ABSTRACT


Performance assessment is a common method of determining proficiency and what students can do with that knowledge. Students in engineering design graphics courses engage in performance tasks, such as creating technical sketches or solid computer models of parts, and instructors must determine how well students can execute tasks aligned with the course objectives. The extant literature contains documented changes in the objectives taught in the classes, skills required for industry, and methods of assessing students’ proficiencies in the desired skills. Documented approaches to assessing student performance in engineering design graphics courses are presented and used for further investigation. This study examines the current performance assessment practices utilized in post-secondary introduction to engineering design graphics (EDG) courses in a two-part study.

A web-based survey was developed, distributed, and employed to investigate course performance objectives, the importance of performance assessment, type of work assessed, and performance practices in introductory EDG courses. Responses from current introductory EDG instructors provided insights into the current practices and provided participants for a follow-up study that investigated the reliability of current performance assessment methods in an introductory EDG course.

Three example projects of different quality were randomly selected from existing student portfolios. The projects were stratified by the projects’ grades that were provided by the original course instructor. They were then sent to current instructors at other institutions, along with an existing performance assessment task and assessment instrument. Scores
provided by the instructors were analyzed with Kruskal-Wallis and post-hoc Dunn’s tests to
determine if there was a difference in the scores given by eight introductory EDG instructors.
Inter-rater reliability, project, and type of work were measured to examine the consistency of
ratings provided by the participants. Results of the study were consistent with other
investigations of inter-rater reliability for rubrics in technical graphics courses. There was
some overall reliability but still room for improvement in the area of consistency for
technical graphics performance assessments methods.

by

Kevin Gregory Sutton

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

Technology Education

Raleigh, North Carolina

2018

APPROVED BY:

Dr. Aaron C. Clark  
Co-Chair of Advisory Committee

Dr. Cameron D. Denson  
Co-Chair of Advisory Committee

Dr. James Bartlett

Dr. Brian Matthews
DEDICATION

This “whatever you call that thing, project, you are working on” is dedicated to the fond and laughter filled memories of my dear friend Elaine Brewer Horne (1942-2015).
BIOGRAPHY

Kevin Gregory Sutton is the youngest of three children June 29, 1990, in Boone, North Carolina. He graduated from Watauga High School in 2008 and then attended North Carolina State University, where he earned his Bachelors of Science degree in Technology Education. He received his MAEd in Curriculum and Instruction from Virginia Tech, along with a graduate certificate in Integrative STEM Education. Upon completing his master’s degree, Kevin returned to Raleigh, NC, where he married Brittany Lauren Patterson and worked on his Doctorate of Education at NC State in Technology Education.
ACKNOWLEDGMENTS

Saying that there are several people whom I need to thank for helping me through the process of finishing this step in my education would be an understatement. Between support, encouragement, wisdom, and love, there are several people that I would like to recognize for their efforts. These people helped guide, push, or accompany me during my educational career.

Chronologically, Jeff and Linda Sutton are the first people that I must thank for getting me started on the path towards my education. Their value of education was always evident and their unwavering support for following my interests began my educational process. I greatly appreciate their support and encouragement to dream big and follow my education.

Following my education has not only directed my life goals, but I am also thankful for the people that I have met along the way. Most of all is the person that I am fortunate enough to spend my life with as we raise children and navigate life’s obstacles together. Brittany, your love and encouragement have helped me through many challenging times and turned an average life into an adventure. As you know, my life would be boring without you.

To my children, Kennedy Grace and Landon David, your arrival has brought great joy to my life. Your smiles and laughter take away all cares in the world, while your tiny but powerful hugs remove all stress. I want to thank you for your sweet encouragement and for sharing cookies with your Dada. I wonder what wonderful things you will do.

Also in the family realm, I would like to thank my favorite brother, Michael, and favorite sister, Kimberlee, for their encouragement. I aspire to return the unwavering support you have shown me throughout all of your life pursuits.
Aside from my family, I also need to thank the following academic mentors and committee members:

First, I would like to thank my committee chairs, Dr. Aaron C. Clark and Dr. Cameron Denson, for their significant commitment to my completion of this degree with their time and efforts. I am thankful for having the opportunity to work on other projects and being pushed to further the field beyond simply adding another dissertation to the literature. Your support and “friendly” recommendations have made a significant impact on where I currently am and will go in my career.

For their gracious dedication and time committed to thorough examination and questioning of my work, I would like to express my gratitude to my committee members Dr. James Bartlett and Dr. Brian Matthews. Your insight and efforts have pushed me through to help complete this process.

Special recognition and gratitude is expressed to Dr. Jeremy V. Ernst, Dr. Alice Scales, and Dr. Joe R. Busby for their encouragement and gracious assistance to find opportunities to continue my academic journey. Your efforts throughout my educational journey have greatly assisted my process. I have greatly enjoyed working beside you professionally and greatly value your trust and friendship.

Finally, I would like to thank my peers I have shared numerous courses and experiences with. I have met many wonderful people throughout my courses and want to specifically recognize Michael Grubbs, Laura Doyle, Tyler Love, Anita Deck, Daniel Kelly, Nolan Fahrer, Dominick Manusos, and Daniel Bates. Learning is a collaborative effort, and each of you has greatly enhanced my experiences over the past few years. I look forward to numerous years of learning and life experiences.
# TABLE OF CONTENTS

LIST OF TABLES ......................................................................................................................... ix

LIST OF FIGURES ...................................................................................................................... x

CHAPTER 1: INTRODUCTION ...................................................................................................... 1

Introduction ................................................................................................................................. 1

Technical Graphics ..................................................................................................................... 3

Rationale ..................................................................................................................................... 5

Statement of Problem .................................................................................................................. 7

Research Questions ..................................................................................................................... 7

Methodology ............................................................................................................................... 7

Limitations and Assumptions ....................................................................................................... 11

Definition of Terms .................................................................................................................... 11

Summary .................................................................................................................................... 14

CHAPTER 2: LITERATURE REVIEW .......................................................................................... 16

Introduction ................................................................................................................................. 16

Curriculum .................................................................................................................................. 16

Student Learning Objectives ....................................................................................................... 17

Performance Assessment ........................................................................................................... 22

Rubrics ....................................................................................................................................... 23

Technical Graphics Curriculum and Student Learning Objectives .............................................. 26

Performance Assessment in Technical Graphics ......................................................................... 32

Challenges of Assessing Student Performance in Engineering Design Graphics Courses .......... 33

Computer Automated Grading of Technical Graphics Assignments ......................................... 36

Rubrics for Grading of Engineering Graphics Assignments ....................................................... 39

Investigating Performance Assessment Methods in Technical Graphics .................................... 41

Summary .................................................................................................................................... 42

CHAPTER 3: METHODOLOGY .................................................................................................. 44

Introduction ................................................................................................................................. 44

Survey ....................................................................................................................................... 44

Procedures ................................................................................................................................. 46

Population and Sampling ........................................................................................................... 47

Survey Administration ............................................................................................................... 48

Reliability Study ....................................................................................................................... 51
Hypotheses .................................................................................................................. 54
Participants .................................................................................................................. 54
Methodology ................................................................................................................. 55
Kruskal-Wallis and Dunn Test ..................................................................................... 57
Inter-rater Reliability .................................................................................................... 58
Study Limitations ......................................................................................................... 61
Summary ....................................................................................................................... 63
CHAPTER 4: RESULTS .................................................................................................. 64
Introduction .................................................................................................................... 64
Research Questions ....................................................................................................... 64
Survey Population ......................................................................................................... 65
Fundamental Technical Graphics Course Enrollment .................................................. 66
Fundamental Technical Graphics Student Learning Objectives .................................... 66
Fundamental Technical Graphics Performance Assessment ......................................... 68
Follow-Up Study Participants ...................................................................................... 71
Ability to Differentiate Between Submissions ............................................................... 72
Correlation of Scores by Raters .................................................................................... 76
Interrater Agreement ..................................................................................................... 79
Summary ....................................................................................................................... 85
CHAPTER 5: DISCUSSION ............................................................................................. 87
Introduction .................................................................................................................... 87
Performance Assessment Trends in Technical Graphics Courses ................................ 88
Rubric Development and Validation .......................................................................... 89
Author’s Implications .................................................................................................... 93
Recommendations for Performance Assessment in Technical Graphics ..................... 95
Summary ....................................................................................................................... 98
REFERENCES .................................................................................................................. 99
APPENDICES ................................................................................................................. 113
APPENDIX A: IRB EXEMPT STATUS APPROVAL ..................................................... 114
APPENDIX B: ENGINEERING DESIGN GRAPHICS PERFORMANCE ASSESSMENT SURVEY EMAIL ................................................................. 115
APPENDIX C: ENGINEERING DESIGN GRAPHICS PERFORMANCE ASSESSMENT SURVEY ................................................................. 116
APPENDIX D: REVERSE ENGINEERING DESIGN PROJECT AND ASSESSMENT INSTRUMENT ........................................................................... 121
APPENDIX E: ENGINEERING GRAPHICS PERFORMANCE ASSESSMENT FOLLOW-UP EMAIL ................................................................................................................ 124
APPENDIX F: ENGINEERING GRAPHICS PERFORMANCE ASSESSMENT FOLLOW-UP REMINDER EMAIL ........................................................................ 126
APPENDIX G: ENGINEERING GRAPHICS PERFORMANCE ASSESSMENT ...... 128
APPENDIX H: ALARM CLOCK PROJECT PORTFOLIO ........................................ 137
APPENDIX I: KNIFE PROJECT PORTFOLIO ............................................................. 153
APPENDIX J: MICROPHONE PROJECT PORTFOLIO .............................................. 170
APPENDIX K: RAW SCORES FOR ALARM PROJECT ............................................. 182
APPENDIX L: RAW SCORES FOR KNIFE PROJECT ............................................... 183
APPENDIX M: RAW SCORES FOR MICROPHONE PROJECT ................................. 184
LIST OF TABLES

Table 2.1. Recreated and simplified version of taxonomies for psychomotor learning from original documents. ................................................................. 19

Table 2.2. Recreation of Barr (2012) student learning objectives for technical graphics used with written permission.......................................................... 31

Table 4.1. Academic rank of population ........................................................................ 66

Table 4.2. Course objectives covered in introductory engineering graphics courses. ........ 67

Table 4.3. Responses to whether specific or general rubrics are utilized in introductory engineering graphics courses......................................................... 70

Table 4.4. Responses to whether analytical or holistic rubrics are utilized in introductory engineering graphics courses......................................................... 70

Table 4.5. Results for Dunn's test for total and each category for the project ................. 73

Table 4.6. Spearman’s Correlation Coefficient for all ratings ........................................... 77

Table 4.7. Spearman’s Correlation Coefficient for Alarm Project....................................... 78

Table 4.8. Spearman’s Correlation Coefficient for Knife Project ....................................... 78

Table 4.9. Spearman’s Correlation Coefficient for Microphone Project ........................ 78

Table 4.10. Intra-class Correlation Coefficients (ICC) for Total Sum and Category Sums .. 80

Table 4.11. Intra-class Correlation Coefficients (ICC) for Sketching Criteria ..................... 81

Table 4.12. Intra-class Correlation Coefficients (ICC) for Model Criteria ......................... 82

Table 4.13. Intra-class Correlation Coefficients (ICC) for Drawing Criteria ...................... 83

Table 4.14. Intra-class Correlation Coefficients (ICC) for Assembly Criteria ..................... 84
LIST OF FIGURES

Figure 1.1: Graphical overview of methodology ................................................................. 10
Figure 2.1: Graphic representation the taxonomy of psychomotor learning domain for technical graphics outlined by Clark and Ernst (2010). .............................................. 21
Figure 2.2: Triad of modern technical graphics instruction from Barr (2012). Used with written author’s permission. ................................................................. 30
Figure 3.4: Graphical overview of methodology ................................................................. 51
Figure 4.1: Grading method utilized for different types of work assessed in university introductory engineering graphics courses. ......................................................... 69
Figure 4.2. Performance assessment scores for total projects and individual categories. ..... 74
Figure 4.3. Overall project scores by rater ......................................................................... 76
CHAPTER 1: INTRODUCTION

Introduction

Higher education places heavy emphasis on measuring student achievement of educational outcomes that are not only used for grading, but also research, course improvement, and accreditation to hold institutions accountable for student learning (Deming & Figlio, 2016). Preparing students to enter competitive careers that require not only a solid foundation of knowledge, but also the ability to solve complex problems, perform specific professional performance tasks, and continue life-long learning requires assessment of not only student knowledge, but also the ability to apply knowledge in practical ways through performance assessment (Baartman, Gulikers, & Dijkstra, 2013). The concept is hardly new and has seen numerous names including “authentic” or “alternative” assessment. Institutions of higher education are increasingly utilizing performance assessments to provide outcomes measures of student competency and program effectiveness (Baartman et al., 2006; Cummings, Maddux, & Richmond, 2008; Simon & Forgette-Giroux, 2001) and can be seen in various forms including essays, performance tasks, demonstrations, projects, portfolios, and games or simulated environments (Pearson, 2018).

Performance assessments may have multiple correct answers or variation in quality of response that require reliable measurement that can consistently distinguish between different levels of quality for meaningful comparison (Allen & Knight, 2009). The use of rubrics in higher education has grown with benefits cited, including improved consistency and accuracy of marking, finer differentiation between assignments, facilitated process for feedback, and feedback for course improvement with published studies of rubric use in medical fields, liberal arts, management, education, and technology (Hack, 2015; Loveland 2005).
Judgments made by the individual instructors determine student scores but also involve decision making about how well the students are learning the objectives as well as inform decisions about the student, instructor, and overall course achievement (Hahm, 2014; Khattri, Reeve, & Kane, 1998). Even though the use of a rubric can ameliorate inconsistencies in the scoring process, the use of a rubric alone does not guarantee reliability or meaningful results. Reddy and Andrade (2010) indicate that:

…the value of rubrics in identifying the understandings and skills to be taught and learned, and of providing detailed, criteria-specific feedback to instructors and departments on which of those understandings and skills have been mastered by students and which have not. As a result, the rubrics informed the process of making improvements to courses and instructions. The key to this process, of course, is a clear, valid and reliable rubric, without which the method is useless at best and possibly even misleading. (p. 441)

Rater reliability is the most commonly studied condition of validity in rubric development that looks at the consistency of ratings made by two independent raters, inter-rater reliability, and by the same rater at two different points in time, intra-rater reliability (Moskal & Leydens, 2000). Reliability studies based on the foundation provided by Shrout and Fleiss (1979) provide insight into the consistency of grading processes. Consistency in ratings is necessary to fairly evaluate each student and provide data to support effective instruction of learning outcomes. Processes for developing and refining rubrics to improve reliability exist, such as Allen & Knight (2009), but were not always utilized in developing and implementing performance assessments.
Technical Graphics

Technical graphics is a field that utilizes performance assessment on a large scale in higher education. Thousands of students each year enroll in fundamental technical graphics courses in post-secondary institutions that offer engineering degrees in the United States. North Carolina State University’s enrollment statistics report that 7,518 students have taken the Foundations of Engineering Graphics course in the past 10 years. These courses prepare students with the theory and methods for communicating, recording, and solving problems graphically and are a commonly required course for students enrolled in many engineering majors. Skills learned in these courses, such as interpreting technical drawing information, make students more enticing to future employers (Martinez, 1999).

Over time, curriculum and assessment of student objectives have changed in these courses due to several developments, including emerging technologies, accreditation, and university requirements. Accreditation agencies ensure that university programs produce graduates ready to enter a global workforce. For engineering, ABET requires that graduates are able to communicate effectively and use the techniques, skills, and modern tools required for engineering practice (ABET, 2016a). To be accredited, programs must state what students are expected to know and be able to perform upon graduation, along with assessments of this knowledge to determine the degree outcomes are being attained and utilized for continuous course improvement (ABET, 2016a).

Along with the push for accreditation, advancements in computers and software have changed the tools and practices in the engineering field, which are reflected in the content of introductory engineering design graphics courses. In 2006, Clark and Scales noted the diminishing use of hand drawing tools with a shift towards three-dimensional (3D) solid
modeling using Computer Aided Design (CAD) in programs. Barr (2012) furthered the call from the Engineering Design Graphics Division (EDGD) of the American Society for Engineering Education (ASEE) for students to be able to interpret and create CAD models and engineering drawings.

There have been multiple calls in the graphics literature for additional focus on assessment: “There seems to be a consensus of those EDGD members that reliable tests are needed for measuring gains in knowledge and skills” (Demel, Croft, & Meyers, 2002, para. 11). Clark and Scales (2003) echoed the need for attention to our methods for assessing the content we teach due to the changes we have seen in the past half-century. Another perspective explained, “…the issue of testing whether students can use CAD to create solid models and drawings needs to be addressed” (Demel, Meyers, & Harper, 2004, p. 9). The demand for focus on how to assess student performance was evident.

Students in fundamental technical graphics courses often completed performance tasks, such as reverse engineering design projects at the end of the semester, to demonstrate their ability to meet course outcomes including the ability to interpret and create CAD models, engineering drawings, and assemblies (Barr, Krueger, Wood, & Pirnia, 2014; Branoff, 2007). Students at some universities, such as North Carolina State University, are assessed on their reverse engineering design projects that required students to make sketches, models, plans, and an assembly from an existing multi-part object such as a flashlight to demonstrate multiple different skills taught in the course. Multiple instructors individually rate students’ ability to create technical sketches, 3D solid models, engineering drawings, and assemblies, but the extent the instructors scoring aligns is uncertain. Judgments made by the individual instructors determine student scores but also involve decision making about how
well the students are learning the objectives as well as inform decisions about the student, instructor, and overall course achievement.

Grading methods of these assessments vary throughout the literature with approaches include cognitive interviewing, as seen in Menary, Robinson, & Belfast (2011) and suggested by Branoff & Dobelis (2014); using computer grading schemes (Ault & Fraser, 2013; Baxter & Guerci, 2003; Kirstukas & Amaya-Bower, 2016); and instructor-developed criteria instruments, referred to as rubrics (Barr et al, 2014; Branoff & Dobelis, 2012, 2013, 2014; McInnis, Sobin, Bertozi, & Planchard, 2010). Challenges to effective assessment included the consistency of raters (Branoff, Devine, & Brown, 2016), large class sizes, the time required to grade student work (Goh, Shukri, & Manao, 2013; Kwon & McMains, 2015), and assess design intent (Devine & Laingen, 2013).

Rationale

Fundamental technical graphics courses in post-secondary institutions in the United States have seen significant changes in the content and practice (Clark & Scales, 2006; Barr, 2012) due to the significant changes in technology and policy. The inclusion of constraint based computer aided design (CAD) into the curriculum means that, “Examining print-outs of solid models or drawings is no longer sufficient to determine the correctness of geometry,” (Wiebe, Branoff, & Hartman, 2003, pp. 7). Large class size in these fundamental courses has lead to a variety of approaches to assess student artifacts. These assessments were used to make judgements about student’s proficiency that were a part of the student's grade as well as provide data utilized for instructional improvement. Significant portions of student’s grades were determined by performance assessment and decisions were made with this data (Baldizan & McMullin, 2005; Elrod & Stewart, 2005). This research addressed two issues
related to the assessment of student work in these fundamental graphics courses.

First, discussion of different assessment practices in the literature provide the advantages and disadvantages for each approach but do not explain the extent to which they are being used. The grassroots development of these assessment methods also limit the ability to determine the extent that these methods were utilized. This research helped clarify the type of student artifacts generated and assessment methods utilized in fundamental technical graphics courses. Data related to current assessment practices assisted instructors, course developers, and other researchers in making informed decisions.

Secondly, most courses with large enrollments employed multiple instructors and teaching assistants that were responsible for teaching and assessing students. Knowing the consistency of scores assigned by multiple assessors using these assessments methods, regardless of the instructor or section of the course, informs the extent comparisons can be made between scores (Allen & Knight, 2009). Even though sample rubrics were available as examples for the profession, these instruments need to be investigated to determine their ability to help raters consistently score student work across sections, instructors, or universities. Of all the grading schemes suggested in the technical graphics literature, Branoff et al., (2016) provided the only study that investigates the quality of their rubric to determine the reliability of grading across multiple graders. They suggested that additional reliability studies should be completed to investigate the quality of performance assessment instruments used in these courses. This study furthered their investigation of the consistency in assessing student work in fundamental technical graphics courses and demonstrated a methodology to adapt Allen & Knight’s (2009) steps for rubric validation to investigate performance assessment instruments in the technical graphics field.
Statement of Problem

The intent of this study was to clarify performance assessment practices utilized in fundamental technical graphics courses and investigate the reliability of an existing assessment method. Without this research, the degree of scoring consistency was unknown, and the student assessment data might be potentially useless or misleading for student grades and course improvements.

Research Questions

1) What is the status of performance assessment in fundamental technical graphics courses at postsecondary institutions?
   i) How many students are assessed in these courses?
   ii) What student learning objectives are assessed?
   iii) What type of student work is assessed?
   iv) How is each type of student work assessed?

2) Are course instructors at the post-secondary level able to differentiate between the quality of student work on reverse engineering projects in fundamentals technical graphics courses using an existing assessment instrument?

3) Is there significant correlation between rater scores when using an existing performance assessment instrument?

4) Is there significant inter-rater agreement for criteria of an existing performance assessment instrument?

Methodology

This research first clarified the status of performance assessment to determine course objectives, types of student work assessed, and assessment methods within the Engineering
Design Graphics Division (EDGD), of the American Society for Engineering Education (ASEE), membership. A literature review provided information on the trends and areas of ambiguity surrounding student performance assessment within the membership. Questions were developed about the emerging trends and put in an electronic survey. University faculty from three institutions with expertise in statistics, technical graphics, performance assessment, and survey development with a documented record of expertise in the topic reviewed the survey for content, sequencing, wording, and question type for face validity. After this review, the web-based survey was sent to 200 EDGD members via email. Participation was encouraged by sending weekly follow-up emails for three weeks but were stopped after the responses ceased. The 50 final responses provide data that were analyzed to determine the performance assessment trends that are provided in chapter four.

The second purpose of this study was to use Allen and Knight’s (2009) methodology to investigate the reliability of the performance assessment practices utilized for end-of-course assessment in an introductory technical graphics course as described by Branoff (2004), which is similar to that of Barr, Schmidt, Krueger, and Twu, (2000). The assessment required students to create technical sketches, 3D solid models, engineering drawings, and part assemblies. Three existing student projects were randomly selected that demonstrated three quality levels of work from a set of projects created in a previous semester at North Carolina State University. These projects and instructions for participating in the study were emailed to the participants identified by the initial survey data. All the participants were currently teaching a fundamental technical graphics course utilizing the same proprietary modeling software that the projects were created. Eight of the twelve participants individually assessed each of the projects. The data were analyzed in STATA 14.2 to
investigate the guiding research questions.

The statistical tests were selected based on the population, sample size, use in previous studies for comparison purposes. Each of the tests were non-parametric due to the small sample size and are further discussed in the third chapter. Differences between the mean scores provided for each project and type of work in the project were compared with the Kruskal-Wallis (1952) H test. Dunn’s (1964) test was used to determine where differences existed. Rater correlations were assessed with Spearman’s Correlation Coefficient and instrument reliability was compared with intra-class correlation coefficients, as seen in Branoff et al., (2016). An overview of the methodology can be seen in Figure 1.1 and an in depth rationale found in chapter three.
• Review of literature on learning objectives, engineering graphics learning objectives, performance assessment, assessment practices in engineering graphics courses, rubrics, and rubric reliability.

• Questions developed to collect demographic and course information related to performance assessment in introductory engineering design graphics courses.
• Survey reviewed by university faculty with expertise in technical graphics, surveys, and statistical research methods.
• Reviewers looked at wording, order, and completeness.
• Edits made and approved by all reviewers.

• Three example projects randomly selected from stratified pool of existing student projects.
• All student identifiers removed and files consistently renamed and organized into folders.
• Existing project brief and assessment instrument included with projects in file.

• Procedures and instruments reviewed and approved by NC State IRB.

• Survey sent to active EDGD email list (n=200).
• Follow-up emails sent once a week for three weeks.
• 50 responses with 39 currently teaching an introductory engineering design graphics course at a post secondary institution.

• From the 39 instructors completing the survey, 12 agreed that they would participate in a follow-up study and had the required software to open and examine the student work.
• Each of the 12 qualifying participants were sent the project brief, assessment instrument, and three example projects.
• A follow up of participants was conducted weekly to encourage completion.
• A total of eight participants provided ratings for each of the three example projects.

• Descriptive statistics for survey results.
• Kruskal-Wallis H test for difference of scores by project and category.
• Post-hoc Dunn’s test to determine which projects and categories varied.
• Spearman's Rho and Intra-class Correlation Coefficients calculated for rater consistancy and instrument reliability.

*Figure 1.1: Graphical overview of methodology*
Limitations and Assumptions

Some limitations and assumptions are inherent in the design of this study.

First, the 2016-2017 EDGD directory was used to create a list of email addresses to solicit participants for this study. Not all post-secondary introductory engineering graphics instructors are active members of EDGD, and there is no available directory of all instructors. This study utilized instructors who were active in EDGD, therefore the results reflect EDGD members that were currently teaching fundamental technical graphics courses.

Second, the study assumed that all participants in the follow-up study utilized the provided performance assessment instrument and completed their ratings individually. The instructions provided that these conditions should be met. However, not being present during the rating process due to distance between participants limited the ability to ensure that these directions were followed.

Third, the follow-up study used an example project and three samples from the researcher’s university. These projects were stratified based on the scores provided by the instructor who taught the course and the reliability of these scores is untested. With other instructors the scores from these projects may have been different and placed them in other quality categories. Inter-rater reliability and ability to differentiate between scores may have differed on other performance tasks or with other example projects.

Definition of Terms

Commonly used terms with explicit meaning for the researcher and for the sake of this study are defined as follows:

Agreement: “…is the degree to which scores or ratings are identical (Kottner et al., 2011, p. 668).”
Assembly: “A collection of parts and/or subassemblies that have been put together to make a device or structure that performs a specific function (Lieu & Sorby, 2016, p. G-1).”

Assessment: “…one or more processes that identify, collect, and prepare data to evaluate the attainment of student outcomes. Effective assessment uses relevant direct, indirect, quantitative and qualitative measures as appropriate to the outcome being measured. Appropriate sampling methods may be used as part of an assessment process (ABET, 2016, para. 4).”

CAD: “Computer-aided drawing. The use of computer hardware and software for the purpose of creating, modifying, and storing engineering drawings in an electronic format (Lieu & Sorby, 2016, p. G-2.).” Also, “Computer-aided design: The process by which computers are used to model and analyze designed products (Lieu & Sorby, 2016, p. G-3).”

Constraint-based CAD: “Constraint-based CAD tools create a solid model as a series of features that correspond to operations that would be used to create the physical object. Features can be created dependently or independently of each other with respect to the effects of modifications made to the geometry. The geometry of each feature is controlled by the use of modifiable constraints that allow for the dynamic update of model geometry as the design criteria change. While terminology within the literature varies, ideas of parametric, associative, feature-based, and dimension-driven will be included in this definition.”
(Hartman, 2003, p. 10).”

Criteria: “... (also called traits) are the major dimensions of work that are important
(Arter & Chappuis, 2006, p. 3).”

Drawing: “A collection of images and other detailed graphical specifications intended
to represent physical objects or processes for the purposes of
accurately re-creating those objects or processes (Lieu & Sorby, 2016, G-5).”

Evaluation: “...one or more processes for interpreting the data and evidence
accumulated through assessment processes. Evaluation determines the
extent to which student outcomes are being attained. Evaluation results
in decisions and actions regarding program improvement (ABET, 2016, para. 5).”

Interrater agreement: “…is the degree to which two or more raters achieve identical
results under similar assessment conditions (Kottner et al., 2011, p. 668).”

Interrater reliability: “…is equivalent to the consistency of scores from different
raters across all students (Liao, Hunt, & Chen, 2010, p. 617).”

Performance Assessment: "...Engaging and worthy problems or questions of
importance, in which students must use knowledge to fashion
performances effectively and creatively. The tasks are either replicas
of or analogous to the kinds of problems faced by adult citizens and
consumers or professionals in the field (Wiggins, 1993, p. 229).”

Program educational objectives: “...are broad statements that describe what graduates
are expected to attain within a few years after graduation. Program educational objectives are based on the needs of the program’s constituencies (ABET, 2016, para. 2).”

Rater: “Every person who makes a judgement about a person or object (Kottner et al., 2011, p. 668).”

Rubric: “…a document that articulates the expectations for an assignment by listing the criteria or what counts, and describing levels of quality from excellent to poor,” (Reddy & Andrade, 2010, p. 435).

Student Outcomes: “… describe what students are expected to know and be able to do by the time of graduation. These relate to the knowledge, skills, and behaviors that students acquire as they progress through the program (ABET, 2016a, para. 3).”

Three-dimensional (3D) Modeling: “Mathematical modeling where the appearance, volumetric, and inertial properties of parts, assemblies, or structures are created with the assistance of computers and display devices (Lieu & Sorby, 2016, p. G-15).”

Summary

Determining student learning objectives and assessing student performance are ongoing challenges for fundamental technical graphics course designers and instructors. Multiple approaches to assessing student learning outcomes through performance assessment exist, but the extent of their usage was unknown. The first goal of this research was to gather empirical data to provide an overview of performance assessment practices used in engineering design graphics courses at the post-secondary level by surveying current
instructors. Also, reliable scores on performance assessment are necessary to make meaningful comparisons between the scores (Allen & Knight, 2009) and the documentation scoring reliability in engineering design graphics are limited (Branoff et al., 2016). The second goal of this research looked at the reliability of an existing performance assessment instrument and methods used to assess students on their final reverse engineering design project at the researcher’s university. Surrounding literature, methods for the study, results, and implications follow.
CHAPTER 2: LITERATURE REVIEW

Introduction

Understanding performance assessment practice in the technical graphics field requires looking at the epistemology of learning, curriculum, and performance assessment as a whole, along with how this relates to technical graphics. This chapter provides an overview of the historical context and existing literature surrounding the assessment of student-learning outcomes in the fundamental technical graphics courses that support this study.

Curriculum

Curriculum refers to the knowledge and skills that students are expected to learn throughout an educational sequence, including all content learning objectives, assignments, activities, lessons, and assessments (Curriculum, 2014). More analytically, a curriculum reflects a cultural understanding of what students should know and be able to do to benefit society (Westbury, Hirsch, & Cornbleth, 2015). Researchers and organizations work to determine content and practices that are appropriate to teach students in order to prepare them for industry practice.

One organization that focuses on ensuring that engineers are adequately prepared for industry is The Accreditation Board for Engineering and Technology, Inc. (ABET, 2016b) that provides accreditation for post-secondary institutions in the fields of applied science, computing, engineering, and technology. This organization was founded in “1932 as the Engineers’ Council for Professional Development (ECPD), an engineering professional body dedicated to the education, accreditation, regulation, and professional development of engineering professionals and students in the United States” (ABET, 2016b, para 1). ABET (2016b) suggested that a conglomeration of engineering societies and associations provide
the basis for experts who determine the standards that engineering students should know and be able to do as defined by the Engineering Accreditation Commission.

Of these criteria, the student outcomes are of particular interest for those engaged in curriculum development. The discipline of engineering design graphics is focused on outcomes (g) that require students to be able to communicate effectively and (k) that require students to use the techniques, skills, and modern engineering tools required for engineering practice (Barr, Krueger, & Aanstoos, 2004).

Student Learning Objectives

Gagne (1985) defined learning as a process that changes student dispositions and capabilities that their behavior should reflect. Student learning objectives are the educational goals that determine the knowledge or skills that individuals should possess at the end of a learning sequence.

Multiple learning taxonomies that define the creation and assessment of educational sequences exist (ex. Biggs & Collins, 1982; Bloom et al., 1956; Fink, 2003; and Webb, 2002). Seminal work in the classification of these objectives began with Benjamin S. Bloom and his team that created a framework commonly referred to as “Bloom’s Taxonomy.” Bloom et al. (1956) broke down learning into cognitive, affective, and psychomotor domains. These categories provided a framework in which different types of knowledge is classified and studied.

The cognitive learning domain deals with “recall or recognition of knowledge and the development of intellectual abilities and skills” (Bloom et al., 1956, p. 7). Bloom and Krathwohl (1956) further defined this domain with a hierarchy of nouns, classifying the levels of knowledge in this domain. The framework, originally intended for use by university
examiners, grew to be a common reference for educators, curriculum planners, administrators, researchers, and classroom teachers (Anderson & Sosniak, 1994). Anderson and Krathwohl (2001) made updates to the original work on the cognitive domain to make it more useful to educators and applicable to 21st century learning. This updated framework is the most commonly used cognitive learning taxonomy (Clark & Ernst, 2010; Forehand, 2010). Cognitive learning taxonomies organize what learners know and are able to do with that knowledge.

The affective learning domain pertains to student interests, opinions, emotions, attitudes, and values (Krathwohl, Bloom, & Masia, 1964). Learning in this domain centers on how students become aware, respond, value, compare, and act based on an idea or phenomena. This framework was built from how people receive stimulus, organize their feelings, and act or reorganize their value system. Learning objectives from the affective domain assist in phrasing the desire for students to appreciate the significance of ideas or topics rather than just knowing about them or being able to do them. Finks Taxonomy (2003) crossed from the cognitive domain into the affective domain with the human dimensions and caring levels but is still not as detailed or widely used as the work done by Krathwohl et al. (1964).

Finally, “the psychomotor domain addresses the fact that neither conscious knowledge nor values and attitudes are sufficient to explain effective performance of learned tasks,” (Rovai et al., 2009, p. 8). Developing skills requiring physical movement, manual tasks, or operation of a machine, such as an excavator or computer, fall within the psychomotor domain. Even though Bloom et al. (1956) and his group categorized psychomotor learning, they never disseminated a taxonomy for it as they did for the
cognitive and affective domains. Simpson (1966), Dave (1970), and Harrow (1972) proposed three different taxonomies for the development of psychomotor knowledge that are recreated in a simplified version from their original documents in *Table 2.1*.

| Table 2.1. Recreated and simplified version of taxonomies for psychomotor learning from original documents. |
|---|---|---|---|
| **Level** | **Description** | **Level** | **Description** | **Level** | **Description** |
| Perception | Awareness | Reflex Movement | Involuntary reaction | Imitation | Copy action of another, observe and replicate |
| Set | Readiness | Basic Fundamental Movements | Basic simple movement | Manipulation | Reproduce activity from instruction or memory |
| Guided Response | Attempt | Perceptual Abilities | Basic response | Precision | Execute skill reliably, independent of help |
| Mechanism | Basic proficiency | Physical Abilities | fitness | Articulation | Adapt and integrate expertise to satisfy a non-standard objective |
| Complex Overt Response | Expert proficiency | Skilled Movements | Complex operations | Naturalization | Automated, unconscious mastery of activity and related skills at strategic level |
| Adaptation | Adaptable proficiency | Non-discursive Communication | Meaningfully expressive activity or output |
| Origination | Creative proficiency |

Harrow’s (1972) model is commonly used in the study of athletic ability, dance, in the preparation to give a large speech, human reflexes, responses, and the ability to communicate, but lacks “set” as covered by Simpson (1966). Simpson (1966) accounted for
the fact that people must receive and be prepared to learn a psychomotor skill. This model takes into account all human senses and preparedness, and it lends itself to situations in which people are pushed beyond their comfort zone, with situations such as skydiving or other activities that might involve a personal phobia. For basic adult training, Dave’s (1970) model covers the basic steps from copying an observed or explained action as seen, copying an action with some modification for similar situations, regular use of the skill, and finally mastering a skill by using it naturally in a variety of situations without hesitation. Dave’s taxonomy is the most commonly used taxonomy for this domain (Zollman, 2012).

This research was specifically interested in the assessment of student performance that falls within the psychomotor domain of learning, specifically for fundamental technical graphics courses. While Bloom et al. provided a basis for the assessment of student learning objectives, they encouraged each field to develop their own taxonomy, with specific language and topics appropriate to that field. Clark and Ernst (2010) developed a taxonomy of the psychomotor domain classifies skill development for assessment purposes in the technical graphics field. Dave’s (1970) levels of skill development were the basis of this technical graphics taxonomy of the psychomotor domain, seen in Figure 2.1.
This study focused on the assessment of student skills in fundamental technical graphics courses that fall within Clark and Ernst’s (2010) taxonomy. For fundamental technical graphics courses, precision was the appropriate aim of measurement because it required students to demonstrate skill reliably independent of assistance, which Gagne (1985) constitutes as learning. Imitation and manipulation only reflect the ability to mimic a provided action without regard to independence. Articulation and naturalization were beyond the scope of introductory courses as they occur over time and experience during internship levels and throughout a career (Clark & Ernst, 2010). Overall, Clark and Ernst’s taxonomy was the basis of this research that investigated the methods and reliability of assessing student skills to the level of precision, the third level within Clark and Ernst’s taxonomy, within fundamental technical graphics courses.

*Figure 2.1:* Graphic representation the taxonomy of psychomotor learning domain for technical graphics outlined by Clark and Ernst (2010).
Performance Assessment

Evidence supporting learning requires an outward sign that can be observed in scenarios, such as student behaviors or work (Posner & Rudnitsky, 2006). Assessment is the process (or a series of processes) used to identify, collect, and prepare data regarding the extent that students have attained learning outcomes (ABET, 2016). Assessment data is used for making judgements about the extent knowledge has been learned that can be used for multiple purposes, including enhancing learning through improving instruction, communicating the progress of students, and providing accountability (ABET, 2016; Barrow, 2006; Fahrer, 2009; Herman & Aschbacher, 1992; Linn, 2000; Walvoord, 2010).

Assessments come in multiple forms and variations. Gordan (1998) listed objective paper and pencil items, standardized cognitive assessment, performance assessment, informal observation, essay items, and portfolios as the main categories of classroom assessment and stated that there was not one specific assessment that can measure a student’s knowledge and skill in a particular area. Knowledge is commonly assessed through traditional psychometrically driven testing that contain questions with single correct answers that do not require the demonstration of knowledge but rather the recall of information provided to them by others (Chauncey, 2004; Hanna & Dettmer, 2004). Student learning objectives in the cognitive domain can be achieved through standardized tests that are objective and efficient to grade, but they are decontextualized from professional tasks and skills (Kaslow et al., 2007). These assessments demonstrate knowledge, but not professional competence. Competence is defined by Miller (1990) as “the quality of being functionally adequate, or having sufficient knowledge, judgement, skill, or strength for a particular duty” (p. 563). Measurement of student competence requires measurement of performance on a realistic task
related to professional practice in the form of performance assessment (Baartman et al., 2013; Knight, 2000)

Institutions of higher education are increasingly utilizing performance assessments to provide outcomes measures of student competency and program effectiveness (Baartman et al., 2006; Cummings, Maddux, & Richmond, 2008). Performance assessment, also referred to as “alternative assessment” provide meaningful learning tasks directly related to the learning task that allow students to demonstrate their ability to solve real-world problems and make appropriate decisions (Cummings et al., 2008; Hunt et al., 2000). Performance assessments can be seen in various forms including essays, performance tasks, demonstrations, projects, portfolios, and games or simulated environments (Pearson, 2018) that provide opportunities for students to demonstrate knowledge and skills in a context that results in a tangible product or observable performance. Those products are then rated for proficiency based on a provided set of criteria (Marzano, Pickering, & McTighe, 1993). These ratings provide information about how students are able to apply skills and knowledge to perform a task of professional value (Baartman et al., 2013). While performance assessments provide evidence about student achievement of specific outcomes, grading these assessments takes longer than standardized tests but generate scrutiny for their reliability that has generated a shift towards the use of rubrics.

**Rubrics**

The use of rubrics and the amount of literature discussing rubrics in higher education has increased over the past few decades as the push for learning beyond recall of factual knowledge have swelled (Hack, 2015). Assessment of learning in higher order thinking skills based on learning taxonomies increase in authenticity to professional tasks but are also less
standardized which threatens reliability due to more interpretation and subjectivity in rating (Baartman et al., 2013). Educators commonly turn to rubrics to reduce subjectivity in measuring students’ performance of higher order skills by clarifying and categorizing the intended criteria and levels of skill (Arter & McTighe, 2001; Busching, 1998; Jonsson & Svingby, 2007; Khattri, Reeve, & Kane, 1998; Perlman, 2004). Rubric rating criteria should be directly related to industry practices, so high assessment scores suggest high performance outside of class or in the future workplace (Moskal & Leydens, 2000). Students and teachers make judgements all the time (Turley & Gallagher, 2008), but clearly set definitions of what is wanted and what constitutes quality are intended to shift judgements to a collective rather than the opinions of an individual (Noyes, 1912).

The use of rubrics in higher education has grown with benefits cited, including improved consistency and accuracy of marking, finer differentiation between assignments, facilitated process for feedback, and feedback for course improvement (Hack, 2015; Loveland 2005). To investigate claims of rubric benefits, Jonsson and Svingby (2007) reviewed 75 relevant studies and concluded that rubrics do not facilitate valid judgement on their own but have the potential to promote learning and improve instruction by making expectations and criteria clear.

Rubrics help facilitate performance assessment, but do not come in a one-size-fits-all option. These assessment tools can be categorized as analytical or holistic along with general or task-specific (e.g., Arter & Chappuis, 2006; Herman, Baker, & Linn, 2004; McTighe & Wiggins, 1999; Mertler, 2002; Moskal, 2000; Perlman, 2004). Each type of rubric possesses its own advantages and disadvantages, which makes them appropriate to specific situations.
Analytic rubrics provide clear descriptions of each criterion; where holistic rubrics describe all of the work at the same time, with all criteria being evaluated simultaneously. Holistic rubrics make scoring faster because all criteria are considered simultaneously, but they do not provide diagnostic information to the teacher or specific feedback to the students that informs them of areas for improvement (Brookhart, 2013). Also, rubrics can also be categorized as general or task-specific. General rubrics can be applied to a whole family of tasks, such as technical sketching, while task-specific rubrics describes work that is specific to a single task. According to Brookhart (2013), general rubrics can be used for several tasks or assignments, but take more time to achieve inter-rater reliability.

Turley and Gallagher (2008) suggested that rubric critics do not like the standardization of all work. However, standardization is imperative to make data-driven decisions about learning to use a normed set of practices, such as ANSI standards and a uniform technical language. Performance assessment critics question the reliability of rater judgements, but rubrics can help define assessment criteria and improve the reliability of rater judgements, especially when they are analytic, topic-specific, and accompanied by exemplars and/or rater training (Goodrich, 1996; Jonsson & Svingby, 2007).

Rubric development procedures appear throughout educational literature (ex. Arter, 1993; Pasco County, 1996; San Diego State University, 2017) but Loveland (2005) suggests that guidelines usually follow the eight steps presented below:

First, the teacher should find out how the real world defines quality performance or a product. Second, gather examples of student and expert work that illustrates a range of quality. Third, sort the examples into four to six groupings by quality of performance or product. Fourth, differentiate
within the performance or product specific skills or attributes. Fifth, write
descriptive statements for these attributes. Sixth, within the specific skill or
attributes, write an operational definition at the different levels. Seventh, link
your student and expert examples to the different criteria and levels for
instructional, communication, and professional development purposes. In the
final stage, teachers use the rubric in their classrooms and evaluate the results
(p. 20-21).

Loveland (2005) discussed questions that need to be addressed to assess a rubric’s
quality but as a linear process. Allen and Knight (2009) provided a more in-depth iterative
rubric development process that encourages a collaboration of developers rather than a single
instructor. Their process provides steps to iteratively define and improve the validity and
reliability of a rubric that they encourage for the development and testing of reliability and
usefulness of a rubric. This more in depth approach provides multiple approaches to make
sure that a rubric aids in differencing between different qualities of student work based on
industry standards, but also the consistency of scoring for courses with more than one section
or instructor.

Technical Graphics Curriculum and Student Learning Objectives

Engineers and architects have used geometry, sketching, and drawings as a graphic
language to convey ideas and designs for centuries (Booker, 1963). This language used in the
design process is utilized for communicating, solving problems, visualizing objects, and
conducting analyses (Bertoline et al., 2011). Brunelleschi, Crozet, da Vinci, Durer, Eiffel,
Francesca, Hall, Hood, Monge, Moyer, Palladio, and others contributed to advances in
technical graphics with breakthroughs, such as descriptive geometry, projection theory, and
multi-view drawings to convey concepts (Matthews, 2004). These techniques were adopted by others and passed along through training. Bennett (1926) and French (1918) reported that different training techniques produced the unique drafting and crafting skills that could be seen in different countries. The noticeable difference between the training of these professionals had on their professional practice indicated the importance of how professionals are prepared.

Historically in the United States, the 20th century produced changes within engineering graphics with the formation of graphic organizations. Collaboration between the American National standards Institute (ANSI), the American Society for Engineering Education, the Society of Automotive Engineers, and the American Society of Mechanical Engineers to establish drafting standards began in 1926 (Harris & Meyers, 2007). In 1928, the Engineering Design Graphics Division (EDGD) was formed within the American Society for Engineering Education (ASEE) to promote “teaching, research, discussion, and communication of engineering design graphics” (ASEE, 2016, para. 2). EDGD works to ensure that the content taught to future professionals reflects the constant changes in tools and techniques required for the engineering profession.

Changes in technical graphics instruction can be seen in shifts of practice, such as changes from first-angle projection seen in Crozet’s *Descriptive Geometry* (1821) to third-angle projection presented by Hall in 1902 in his rendition of *Descriptive Geometry*. Additional changes, such as the addition of direct auxiliary or a direct method, where auxiliary planes are projected to show the desired relationships, were included in Moyer’s (1904) *Engineering Descriptive Geometry*. Combining direct method and third-angle
projection is now used in the United States and took place in 1926 with Hood’s *Drawing Descriptive Geometry by the Direct Method* (Hammond et al., 1971; Matthews, 2004).

Changes in engineering graphics came from new methods as previously mentioned along with monumental changes, such as the introduction of the blueprint process to the United States in 1876. Blueprints were created by hand using tools, such as T-squares, triangles, drawing boards, and French curves, during, what Harris & Meyers (2007) describe as, the “Golden Age of Drafting” in the earlier part of the 19th century. Changes to designs were completed manually, and working in ink required projects to be restarted if changes or mistakes were made. Computers and CAD software brought about significant changes in the way that drawings are produced. Massive vacuum tube computers running two-dimensional drawing software, such as Ivan Southerland’s “Sketchpad,” led the shift to digital drafting, and have continually developed to sleeker, more powerful, and personally owned computers that run parametric solid-modeling software that is currently used in industry (Harris & Meyers, 2007).

While some universities attempted to mimic traditional hand drafting methods on the computer with CAD, Bill Ross (personal communication, June 22, 2017) reports that North Carolina State University and Purdue University began to focus on model centered technical graphics courses in the mid-1980s. These courses differed because they taught the students the process of creating digital models of objects that could be modified, how to put them together in assemblies, and how to generate engineering drawings from the models. Topics covered included sketching and text, engineering geometry and modeling, projection theory, detail drawings, part assemblies, working drawings, dimensioning, section views, auxiliary views, and application of part data.
A study by Clark and Scales (2006) found that the general trends for courses in the technical graphics curriculum were seeing a decrease in courses on geometric dimensioning and tolerancing, manual drafting, 2D CAD, 3D non-constrained CAD, and computer-aided manufacturing. Even though there were decreases in many these courses, the study also found there were large gains in courses that included 3D constraint based modeling and animation (Clark & Scales, 2006). Areas that were noted that needed additional investigation included online and distance education instruction, increased emphasis in 3D CAD, and 3D prototyping. The top-rated research interests found were high school outreach and 3D printing, but there was a variety of other areas suggested for additional investigation, including curriculum development, meeting ABET requirements, and assessment of student learning (Clark & Scales 2006).

Graphics literature discusses the content that is included in their fundamental graphics course because many universities only require one semester of graphics training for engineering students (Nozaki et al., 2016). Some universities offer additional courses that allow students to go further in depth, but the content of the introduction course is a large topic:

Common topics in introductory courses tend to be sketching, multiview and pictorial sketching/drawing, sectional views, dimensioning, manufacturing processes, and constraint-based CAD. Final projects in these courses range from design projects where students actually design and build prototypes to reverse engineering projects where students measure and re-engineer existing products (Branoff, 2007, p. 4).
Having a conceptual model helps researchers focus, not only the main elements of a discipline, but also on an understanding of their interrelationships (McGrath et. al., 1991). McGrath et al. (1991) further suggest that, “A conceptual model defines the discipline and gives reason for its existence,” (p. 191). This study used the conceptual model provided by Barr (2012) as the basis of content for fundamental technical graphics courses.

*Figure 2.2:* Triad of modern technical graphics instruction from Barr (2012). Used with written author’s permission.

The literature provides an insight into the topics that are currently included in fundamental technical graphics courses. Several efforts to determine suitable student learning objectives in graphics education include: Barr, 1999; Bertozzi, et al., 2007; Branoff,
Hartman, & Wiebe, 2002; Meyers, 2000; Nozaki et al., 2016; Planchard, 2007; Sadowski & Sorby, 2013; Smith, 2003. While the previously stated studies discuss course content, Barr (2012) provides results from a survey of the EDGD membership regarding common student learning objectives utilized for their graphics courses. Barr’s results from the survey of the field can be seen in table 2.2. The rank was based on a five point Likert scale with five representing strongly agree and one being strongly disagree. Rankings with high scores represent what the membership felt should be the primary learning objectives of their courses.

<table>
<thead>
<tr>
<th>Outcomes</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to create 3-D solid computer models</td>
<td>4.75</td>
</tr>
<tr>
<td>Ability to sketch engineering objects in freehand mode</td>
<td>4.54</td>
</tr>
<tr>
<td>Ability to visualize 3-D solid computer models</td>
<td>4.54</td>
</tr>
<tr>
<td>Ability to create 3-D assemblies of computer models</td>
<td>4.54</td>
</tr>
<tr>
<td>Ability to create dimensions</td>
<td>4.38</td>
</tr>
<tr>
<td>Ability to generate engineering drawings from computer models</td>
<td>4.33</td>
</tr>
<tr>
<td>Ability to create 2-D computer geometry</td>
<td>4.29</td>
</tr>
<tr>
<td>Ability to create section views</td>
<td>4.13</td>
</tr>
<tr>
<td>Ability to perform design projects</td>
<td>4.08</td>
</tr>
<tr>
<td>Ability to analyze 3-D computer models</td>
<td>4.08</td>
</tr>
<tr>
<td>Knowledge of manufacturing and rapid prototyping methods</td>
<td>3.63</td>
</tr>
<tr>
<td>Ability to create presentation graphics</td>
<td>3.46</td>
</tr>
<tr>
<td>Ability to solve traditional descriptive geometry problems</td>
<td>2.75</td>
</tr>
<tr>
<td>Ability to create geometric construction with hand tools</td>
<td>2.71</td>
</tr>
</tbody>
</table>
Survey results provide the student learning objectives that commonly guide fundamental technical graphics courses by defining the skills that students should possess after completing a fundamental technical graphics course. With the goal of making systematic improvements to the technical graphics curriculum, it is imperative for all in the field to know the status of how well students are meeting the learning objectives of the course.

**Performance Assessment in Technical Graphics**

Assessment of student learning outcomes is a prevalent theme throughout technical graphics literature. Nee (1999) suggested that assessment has been one of the most scrutinized areas of education over the past few decades and that has left university graphics faculty working to assess student work and programs to meet the demands of universities and accreditation. Discussions range from what is appropriate to be assessed (e.g., Barr, Krueger, & Aanstoos, 2004; Branoff et al., 2002; Clark & Scales 2003); analysis of specific course outcomes and assessment systems (e.g., Baldizan & McMullin, 2005; Elrod & Stewart, 2005); and several scoring methodologies including criteria assessed and reports of how well students scored on these assessments (e.g., Branoff & Dobelis, 2012, Company, Contero, & Salvador-Herranz, 2013). Along with growing class sizes, the advancement of computers and software have brought challenges for instructors to accurately and efficiently provide feedback and scores for student work.

Multiple assessment approaches including computer automated grading (e.g. Ault & Fraser, 2013; Baxter & Guerci, 2003; Devine & Laingen, 2013; Goh et al., 2013; Hekman & Gordon, 2013; Kirstukas & Amaya-Bower, 2016; Peng, McGary, Johnson, Yalvac, & Ozturk, 2012; Sanna Lamberti, Paravati, & Demartini, 2012), instructors and teaching
assistants using digital files and grading rubrics (e.g., Ault & Fraser, 2013; Branoff et al., 2016), and more qualitative measures including think aloud methods and papers (e.g., Hartman, 2004; Menary et al., 2011) have been pursued and reported. Each of these approaches provides benefits and limitations for use in introductory engineering graphic courses.

**Challenges of Assessing Student Performance in Engineering Design Graphics Courses**

Personal computers and the availability of industry-standard software for students has brought a focus on students learning to develop constraint based solid models using CAD systems. Knowing how to prepare future professionals has led to studies to understand the way that professionals create solid models and trends (e.g., Diwakaran & Johnson, 2012; Hartman, 2004; Johnson, 2009; Johnson & Diwakaran, 2010, 2011).

Instruction of students regarding part modeling strategies to capture design intent continues to evolve, even as the solid modeling systems become more complex and users in industry develop best practices. Intentional exercises to make students aware of the need for careful part planning should include more strategic discussions regarding the uses of part models and alternative methods. Models used for these exercises need to be sufficiently complex to challenge students’ ability to decompose parts and consider alternative modeling strategies. There is no single correct answer regarding part modeling strategies, and students must rely on experience and situational decision making to build robust solid models. (Ault, Bu, Liu, 2014, p. 12)

The process and procedures of creating solid models and assemblies have required changes the content of the introductory engineering graphics courses and have produced the challenge of how to assess student’s models and assemblies. This is a time-consuming
process not only because of the number of students in courses, but also due to the dynamic nature of computer-generated 3D solid models.

Student-created solid models on the computer provide a significant challenge to assess because design intent is built within the model by the way that the model is created and constrained. Branoff (2003) explained that students cannot be assessed by static measures of whether the model looks correct on a printout, but rather that the files must be opened on the computer to examine the geometry and design intent built into the model. Kirstukas (2013) explained that the challenge with evaluating design intent comes from the fact that models may look the same, even though they are modeled differently. Ault, Bu, and Liu (2014) describe design intent as “the intelligence built into the solid model to control the behavior of the part when subjected to changes or alteration,” (p. 2). When the correct geometry and relations are not present within the model, the model will break or produce unexpected results when parameters within the model are altered. There are different terms used such as “model flexing” (Ault & Fraser, 2013) and “dynamic modeling” (Wiebe, Branoff, & Hartman, 2003), but the idea of changing parameters to efficiently make multiple variations or future iterations of a model is a real-world occurrence that demonstrates the robustness of a model. The way models are defined and constrained provide insight about the way designer wishes to be able to modify the design in the future (Hartman, 2005).

Kirstukas (2013) provided a quality example of how different models that appear the same but behave differently and appear quite dissimilar after changing two of the same constraints in each model. One approach to checking students’ design intent is outlined by Devine and Liangen (2013) where they suggested that students be given a part to model and then provide measurements for a distance and face area. Students are then asked to make
changes to several dimensions. Next, students repeat the same measures that they previously checked. With the appropriate constraints, the model should have consistent changes, and the students’ measurements should match with the numbers that they originally checked. Peng et al. (2012) also offered a similar approach.

Both Ault and Fraser (2013) as well as Branoff and Dobelis (2014) suggested that more qualitative approaches to assessing student work would provide additional insight into their reasoning and decisions during the modeling process. Examples of assessment methods aimed at understanding the reasoning behind the modeler’s process and design decisions include Hartman’s (2005) observation of participants modeling, and watching student created videos of modeling, as seen in Chester (2007) and Menary et al. (2011). These approaches give a deeper insight into the design intent by knowing the reason models were created a certain way but require the assessor to go through the modeling process with the student.

Hartman (2005) utilized a think-aloud method where participants modeled a part they were given and were asked to explain their reasoning for each of their actions during the modeling process. These sessions were recorded and then transcribed to be evaluated for common trends. This time-consuming process was intended to look at the reasoning and trends rather than to determine if participants were able to create solid models. Menary et al. (2011) suggested that students could create a video using screen capture software during their modeling process, so they can explain their rationale for their steps or write a report to explain their process. Chester (2007) utilized screenshot videos to assess 34 students modeling skills and investigated the differences between two instructional strategies and students metacognitive abilities while modeling.
While watching students model parts and reading their explanations for decisions they made, the assessor was provided insight into the student’s thought process. However, these qualitative approaches to assessing student-learning objectives are cumbersome and should be used in situations where the process is more important than the product. Insight from these studies provide suggestions for teaching strategies in graphics courses but cannot be the primary performance assessment method utilized for assessing student learning objectives in large classes.

**Computer Automated Grading of Technical Graphics Assignments**

The rising number of students going through engineering graphics courses at universities provides a challenge for faculty to assess student work and provide feedback in a timely fashion (Ault & Fraser, 2013; Branoff & Dobelis, 2014). Opening each file to probe the relationships, geometry, and other features to provide feedback to each student is a cumbersome task (Devine & Laingen 2013; Kirstukas, 2016) that has encouraged exploring automated systems that will provide students with consistent, accurate, and timely feedback. Multiple automated performance assessment approaches have been developed for assessing student-created solid models and engineering drawings.

Advantages, such as reliability and speed, are mentioned in all publications related to automated grading systems for engineering graphics because computers show no bias, consistently execute all programmed commands, and have the ability to compare multiple items simultaneously. Ault and Fraser (2013) reported that each of the 42 projects in their study usually took 5-6 minutes each to grade; whereas the automated grader ran through each of the 42 projects in less than 30 seconds. Similar results were reported by Kirstukas and
Amaya-Bower (2016) with the instructor spending 152 minutes to examine 38 projects that took the computer only a few seconds to score.

Along with speed, another reported advantage of automated grading systems include the ability to check specific feature use that is tedious for humans to check (Sanna et al., 2012). Metadata created with the file, such as time created, can also be compared between users. Digital files can be easily copied by students, but the time the file was created is always stamped on the file; therefore Hekman and Gordon (2013) programed their system to alert the instructor when two files were submitted with the same time stamp to investigate for potential plagiarism. Automatic comparisons of such fine detail that are tedious for humans to check provide significant advantages for instructors.

Literature disseminating the efforts of creating automated grading systems for technical graphics also describe the limitations to their work, which include the need for additional flexibility, more advanced programing, and additional refinement. Technical coding for the assessment of student work in technical graphics courses still requires additional development, such as Ault and Fraser (2013) and Kirstukas (2016), were looking to refine their working algorithms, others, such as Goh et al. (2013), have further to go because their system was a proof of concept that was only able to recognize circles.

A majority of automated systems are designed to work with AutoCAD because it is the most popular engineering graphics software. One of the most extensive efforts for automated grading with DXF files has been completed at San Diego State University with the development process being reported in four master’s theses. Deo (2010) and Jadhav (2010) began with the graphic user interface and initial programing of the software. Ivaturi (2013) and Karna (2013) continued these efforts by refining the coding and moving the automated
system online. The work done at San Diego State University generated interest, but the proprietary system caused the work of Goh et al. (2013), Hekman and Gordon (2013), and Kwon and McMains (2015) to have to start from the ground up. DXF files are used in each of these approaches but are compared to example files through different avenues. Kwon and McMains (2015) used Matlab to compare all of the downloaded files, while Hekman and Gordon (2013) began with a checker that students checked their files against before moving to an email client. Overall, there have been multiple efforts that compared AutoCAD files to a template file, but there is a lack of a widespread or completed method.

AutoCAD is not the only software used in courses that have worked with automated assessment methods. Several 3-D modeling software, including NX, Solidworks™, and Creo, are present in the literature with their own Application Programing Interface (API) that allows users to query created models. Dassalt Systems utilizes a design checker feature to compare models with an example model for their Solidworks™ certification exam, and Baxter and Guerci (2003) described using this feature in their class. Highlights of using the design checker align with the overall speed, as previously mentioned, and ability to select the criteria that the checker searches for in a model.

Outside of Solidworks™, Ault and Fraser (2013) utilized PTC Precision LMS (Learning Management System) to assess student parts created in Creo 1.0, based on the criteria of proper selection of features, placement in the global coordinate system, and correct use of dimensioning. Student work was assessed based off of similarity to comparison instructor generated part files of correct and incorrect solutions. Specifically, in this study, “Precision LMS checked for the presence of two extrude features, one revolve feature, eight holes, and two patterns,” (p. 7). Advantages of the system used in this study include the
grading of 42 student submissions within 30 seconds and the potential to allow students to submit and receive feedback about their modeling procedure without any time required by the instructor. Student use of an unexpected modeling feature yielded low scores because the system was setup to search within another feature and model changes were not tested.

Additionally, Kirstukas (2016) and Kirstukas and Amaya-Bower (2016) describe the development of a system using the API for NX. Where other efforts describe the system, Kirstukas (2016) goes into further detail about his system by including the criteria that the system specifically checks for, an explanation of each criteria, and the number of points to be deducted for each specific error. Hundreds of hours have been invested into this system that is expected to save more time in the future. This is the only system that has been reported to handle multiple student approaches to modeling, even though the author states that it will need to be refined as each submission has the potential to reveal a weakness.

Overall, there are high expectations for the objective nature and speed of automated grading systems for assessment of student work in technical graphics. Several efforts are occurring simultaneously to achieve the goal of providing students with timely and effective feedback. There is no consensus on what the systems should be checking for in a quality model, but the existing literature provides a place to build on and further discusses what constitutes quality work that computer systems can assess.

**Rubrics for Grading of Engineering Graphics Assignments**

The limitations and continued development of automated systems leave a significant portion of performance assessment to instructors and teaching assistants who use rubrics as aid in assessing student work. These rubrics specify both the criteria and competency levels that the student work will be evaluated (Khattri, Reeve, & Kane, 1998). Engineering
programs across the United States utilize rubrics to measure students’ outcomes for, not only course grades, but also for ABET accreditation (e.g., Dahm, 2014; Ossman, 2010). The term rubric is utilized often for assessment of student performance in technical graphics literature, with some providing examples of the scoring categories for the performance assessment utilized in their courses and research. Examples of these include Branoff’s (2004) trip lever and final project rubric, Ault and Frasier’s (2013) pipe flange manual grading rubric, Branoff and Dobelis’ (2012) new rubric, and Barr et al.’s (2000) evaluation sheet. The most comprehensive overview of criteria for assessment of student work is in Company et al.’s (2013) overview where they suggest using rubrics to disseminate quality CAD practices in academia that spread into industry.

Where each of these examples provide details about the criteria evaluated in the student work, they do not define levels of quality for the criteria from excellent to poor, as suggested by Reddy and Andrade (2010). McInnis et. al. (2010) were the only ones to provide a definition of levels for their outlined criteria that follow a one to four scale of: incorrect in virtually all drawings, incorrect in many drawings, incorrect in only a few drawings, and correct in all drawings.

Currently, the evaluation of student performance in technical graphics is conducted utilizing instructor-developed performance assessments that leave questions about their validity and generalizability because holistic grading scales do not define the levels of work within the criteria. Without a clear definition, the number of points awarded may fluctuate greatly between raters. Moskal and Leydens (2000) posited that well-defined score categories should assist in maintaining consistency in scoring regardless of the rater or time rated.
Previous examples offer the performance assessment practices and results of individual authors without addressing the reliability of their methods.

**Investigating Performance Assessment Methods in Technical Graphics**

While some authors are reported their practice for the sake of sharing with others in the field, some studies look to improve performance assessment practices in technical graphics by making comparisons between two different assessment methods or conducting reliability studies to look at correlations between the scores assigned by multiple raters looking at the same example project.

Branoff and Dobelis (2014) compare two rubrics by looking at the scores on the same projects assessed with both rubrics. One rubric accounted for model accuracy and the time to create the model, with penalty points for errors, while the other was composed of categories (without quality indicators) compiled from literature trends and points. Statistical difference existed between the scores using the two different rubrics with larger differences existing within the scores on lower quality work. As the work quality improved, the correlation between the scores with the two different rubrics also increased.

While Branoff and Dobelis (2014) looked at two manual assessment practices, Ault and Fraser (2013) as well as Kirstukas and Bower (2016), discuss the differences between manual and automated assessment of student performance on solid models. Ault and Fraser (2013) mostly noted the increased speed of the automated system but did not report analysis of the scores due to the differences in the criteria used for scoring; however, Kirstukas and Bower (2016) provided a comparison of the scores. Their results revealed that some scores had near perfect correlations between manual and computer assessment, but one-third to one-half of the scores were significantly lower when assessed by the computer.
Dunn (2009) suggested that measurements are subject to error, and human judgement is prone to this type of error. The extent the scores varied informed the reliability of a measure, and the extent judgements can be made based on the scores provided by individual judges (Cohen, 1960; Fleiss, 1971). Branoff et al. (2016) conducted a reliability study to look at the consistency when multiple raters independently scored the same projects. For this reliability study, three raters assessed 10 different projects from two different models (20 total projects). Branoff et al. (2016) reported that the study rubric appears to generate reliable results for the overall rating of the two models, but there was not significant correlation between all the raters on both of the projects. Low correlations on individual categories suggest unaccounted for variance and point to areas that need refinement. Discussion for future work included repeating the study with additional raters or at other universities, along with trying to improve correlations by providing more detailed descriptors for each category, providing specific criteria for point values, and providing training before individually scoring.

This study furthers Branoff et al.’s (2016) investigation critically looking at the reliability of performance assessment instruments utilized in technical graphics courses by expanding to additional types of student work and instructors at multiple different universities. Measures of how consistently raters score student work provide areas for improvement and the extent decisions based on the scores provided by individual instructors can be used for course improvement.

**Summary**

Assessing student learning the psychomotor domain occurs across many disciplines through performance assessment that require instructors and assessors to make decisions
about the extent of student proficiency. Variation in scores and opinions about what constitutes quality work commonly lean towards rubrics to assist in clarifying the criteria for quality work and increase rater reliability (Reddy & Andrade, 2010). Rubric development and validation processes exist to aid in the construction of rubrics and understanding of the extent scores can be compared using that scoring method (Allen & Knight, 2009).

Fundamental technical graphics at the post-secondary level commonly use performance assessments measure student learning, but approaches vary due to changes in curriculum, student learning objectives, and performance assessment practices that are adapting to the advancements of software and technology that have changed professional engineering practices. Multiple individual approaches to assessing student performance have emerged in the technical graphics literature, such as automated systems, observation of students’ performance, and rubrics for assessing artifacts, but the extent of their use is unknown. This investigation aimed to determine current student learning objectives, performance assessment practices, and investigate the reliability of an existing performance assessment practice used in a fundamental technical graphics course through the methods outlined in the following chapter.
CHAPTER 3: METHODOLOGY

Introduction

The intent of this study was to clarify performance assessment practices utilized in fundamental technical graphics courses within the Engineering Design Graphics Division and investigate the reliability of an existing assessment method. This research consisted of a survey and a follow-up reliability study using Allen and Knight’s (2009) methodology for validating rubrics. Plans and procedures for research are intentionally chosen to address the specific needs of the topic that should be influenced by the nature of the problem, the researchers’ personal experience, and the audience (Creswell, 2012). The following chapter provides a detailed explanation of the participants, instruments, and data analysis utilized to answer the guiding questions of this research.

Survey

Surveys are utilized to collect information from individuals, groups, or organizations through interviews or questionnaires (Marsden & Wright, 2010; Trochim, Donnelly, & Arora 2015). Creswell (2009) claims that surveys “…provide quantitative or numeric descriptions of trends, or opinions of a population by studying a sample of that population,” (p. 20) that researchers can use to generalize or make claims about that population. Various approaches to survey methods include contacts in person, over the phone, by mail, or online. Each method provides their own advantages and limitations. This study utilized a web-based questionnaire to collect information about performance assessment practices utilized in fundamental technical graphics courses within the EDGD membership (Babbie, 1990) and determine eligible participants for the follow-up reliability study.
Information regarding the type of work assessed, assessment methods, who assess the work, etc. in fundamental technical graphics courses at postsecondary institutions might have been collected in a few ways with some approaches more appropriate than others. This information could be collected by visiting each university and directly observing the course throughout the semester, by calling each of the universities and talking to each of the technical graphics faculty, mailing surveys to the technical graphics faculty, or sending an electronic survey to the faculty (Creswell, 2009). Couper and Bosnjak (2010) and Umbach (2005) suggest that web-based surveys provide the lowest cost and fastest time frame for collecting information from a spread out population. With these universities spread across the United States and the contact information provided for the participants as email addresses, the decision to use a web survey was appropriate.

Collecting information about the technical graphics community has relied upon web-based questionnaires in the past, with examples provided by Clark and Scales (2000, 2006) and Downs (2009). Clark and Scales (2000) state that the purpose of their study was to, “…identify current trends and issues related to the engineering graphics profession and to see if any conclusions could be drawn to assist graphics educators in making decisions for establishing the direction of growth for institutions,”(p. 24). The initial study conducted by Clark and Scales (2000) utilized the guidelines established by Lyberg, et al. (1997) to develop a survey about trends related to technical graphics. The survey consisted of four parts. The first part focused on course offerings, software use, content areas, and incorporation of new technologies in the classroom. The second part looked at student populations and their needs. The third examined professional development and faculty concerns. Finally, the fourth concerned degrees offered by programs and need for training
technical graphics teachers (Clark & Scales, 2000). Descriptive statistics from the responses provided a lay of the land in technical graphics programs. Clark and Scales (2006) utilized their original (2000) survey with the addition of a fifth section on future research to determine if there were changes in the profession and search for trends in the field. Finally, Downs (2009) also utilized Clark and Scales (2006) survey with the addition of a section on distance learning.

Overall, the benefits of low cost, speed, and prior use of this methodology in similar studies lead to the decision to use a web-based questionnaire to address the first research question regarding the status of performance assessment practices in technical graphics community.

**Procedures**

The survey used to conduct this research (Appendix C) was developed, refined, and administered based on procedures and considerations regarding population, sampling, question, content, bias and administration issues outlined in the literature (Clark & Scales, 2000; Creswell, 2009; Diem, 2002; Fink, 2002; Lyberg et al., 1997; Trochim et al., 2015). The literature revealed variations in the way that student performance is assessed in engineering graphics courses (e.g., Ault & Fraser, 2013; Branoff et al., 2016) and that several challenges to performance assessment exist, such as time, subjectivity, and accuracy (Kirstukas, 2016). Survey questions were developed based upon topics discussed in engineering design graphics performance assessment and rubric literature, including course learning objectives, enrollment, types of student work assessed, assessment methods, and rubric usage questions.
Developing face-validity of the instrument relied upon the method outlined by Clark and Scales (2000). University faculty from three institutions with expertise in statistics, technical graphics, performance assessment, and survey development reviewed the instrument and provided feedback on content, sequencing, wording, and question type. The reviewers responded that the initial draft of the survey was thorough and ready to use other than one raised concern. One reviewer mentioned that there might be something missing from the response list, but could not think of anything missing and suggested that the “other” answer provided would make a space for any responses besides the ones provided. Once the reviewers provided full approval, the researcher moved forward with the web-based survey provided in Appendix C.

In addition to the steps utilized by Clark and Scales (2000), a Flesh-Kinkaid readability test was conducted to address Diem’s (2002) concern for audience language. The Flesch-Kinkaid readability formula ensured that the questions were at an appropriate reading level for the intended participants based on length of sentences in words, average word length in syllables, percent of personal words and personal sentences (Flesch, 1948). This analysis supported that the reading level of the questions was appropriate for high school students and did not exceed the reading ability of the intended population.

**Population and Sampling**

The population of interest for this investigation were technical graphics educators at the post-secondary level teaching introductory courses. Lists and contact information for all of these professionals were not available and lead past studies (ex. Clark & Scales 2000, 2006; Downs 2009) to use a convenience sample of the active members of EDGD to investigate trends in the technical graphics field. EDGD membership is open to ASEE
members that are interested in technical graphics and currently are or have been actively engaged in education or training of engineers or allied professions (EDGD, 2016). This population is interested in research, instruction, and improvements in technical graphics in addition to having a reputation for cooperative work (Downs, 2009).

The researcher used email to solicit participation (as seen in Downs, 2009) using the 2016 EDGD directory of active members. Downs’ (2009) requirement that the participants must be actively teaching a technical graphics course was maintained to ensure information reflected current practice. Two-hundred participants were solicited via email.

**Survey Administration**

Web-based surveys are subject to several aspects of bias including display issues, social desirability, false respondents, ballot stuffing, and non-response (Couper, 2000). The following section discusses these specific issues and steps taken to mitigate these potential areas for bias in the administration of this web-based survey research.

Evans and Mathur (2005) present the challenge of display issues caused by the variety of devices and browsers used to view web-based surveys. Display problems may result in situation that could include all of the answer choices not appearing or overall technical failure that may influence responses or the ability to participate in the study. Downs (2009) ran a check on multiple devices and browsers before distribution of his web-based survey to ensure proper display across platforms. This study followed this example by testing the survey on various computers made by different manufactures, available web-browsers, and mobile devices.

Trochim, et al. (2015) explains that data from surveys should reflect the population’s true values. Besides incentives, Trochim, et al. (2015) propose that survey respondents may
choose responses that reflect opinions others desire rather than their own for social desirability. More simply stated, people might choose responses that they perceive others would prefer rather than what they currently do or believe. Huang (2006) discusses that this fear comes from the potential of their identity being attached or known when responses are reported to others. Social desirability bias was addressed with Trochim, et al.’s suggestion of having the participants choose to be anonymous and by not addressing any topics that deal with personal or illegal issues.

Another area of concern in survey data collection discussed by Trochim, et al. (2015) is false response. False response occurs when individuals outside of the population of interest respond to the survey. This issue was mitigated using Downs’ (2009) approach of sending individualized survey links only to the participants, which would restrict individuals using different email addresses to complete the questionnaire.

In addition to limiting the survey to the intended participants, participants were also limited to one response to prevent ballot stuffing. Ballot stuffing occurs when respondents submit more than one response during the survey research. Multiple responses from a single participant pose a threat to the validity of the survey data by allowing the opinion of one person to have a greater influence on the results (Couper, 2000). Individual links allowed participants to complete the questionnaire only one time to ensure that each participant’s response carried the same weight.

Finally, non-response was the last concern for bias in the web-based survey that was addressed in the administration of this study. Non-response leaves room to question if those who did not answer would have had different opinions than those that did respond (Creswell, 2009). Survey research commonly uses incentives to address this issue by increasing
response rates (Singer & Bossarte, 2006). However, even though incentives do increase response rates, they also undermine the voluntary participation, potentially change people’s motivation for completing the survey, and could create risks that their responses are not consistent with their true values and perspectives (Berry Pevar, & Zander-Contugno, 2008). Oldendick (2012) suggests that reminder emails are an effective method of increasing response rates. Participants that had not completed the survey were sent regular reminder emails to encourage participation three times, as utilized by Downs (2009). Reminders ceased after the response rates slowed dramatically and 50 responses (25.0%) were collected. The number of responses were similar to those collected in previous surveys of the EDGD membership, including Downs (2009) 57 responses (23.4%), Clark and Scales (2000) 71 responses (21%), and Clark and Scales (2006) 51 responses (21%).

Upon completion, the survey closed and the responses were downloaded. Descriptive statistics from the compiled responses were recorded to answer the first research question of this study.
Figure 3.4: Graphical overview of methodology

Reliability Study

- Review of literature on learning objectives, engineering graphics learning objectives, performance assessment, assessment practices in engineering graphics courses, rubrics, and rubric reliability.

- Questions developed to collect demographic and course information related to performance assessment in introductory engineering design graphics courses.
  - Survey reviewed by university faculty with expertise in technical graphics, surveys, and statistical research methods.
  - Reviewers looked at wording, order, and completeness.
  - Edits made and approved by all reviewers.

- Three example projects randomly selected from stratified pool of existing student projects.
  - All student identifiers removed and files consistently renamed and organized into folders.
  - Existing project brief and assessment instrument included with projects in file.

- Procedures and instruments reviewed and approved by NC State IRB.

- Survey sent to active EDGD email list (n=200).
  - Follow-up emails sent once a week for three weeks.
  - 50 responses with 39 currently teaching an introductory engineering design graphics course at a post secondary institution.

- From the 39 instructors completing the survey, 12 agreed that they would participate in a follow-up study and had the required software to open and examine the student work.
  - Each of the 12 qualifying participants were sent the project brief, assessment instrument, and three example projects.
  - A follow up of participants was conducted weekly to encourage completion.
  - A total of eight participants provided ratings for each of the three example projects.

- Descriptive statistics for survey results.
  - Kruskal-Wallis H test for difference of scores by project and category.
  - Post-hoc Dunn’s test to determine which projects and categories varied.
  - Spearman’s Rho and Intra-class Correlation Coefficients calculated for rater consistancy and instrument reliability.
The first part of this study focused on understanding current performance assessment practices in the technical graphics field. Figure 3.4 provides an overview of the study process. The focus of the second part of this study was the reliability of performance assessment measures used to assess student-learning objectives in a fundamental technical graphics course (Appendix D). Determining student proficiency of course learning objectives in fundamental technical graphics courses commonly rely on performance assessments (Branoff, 2004, Barr et al., 2000). Laio, Hunt, and Chen (2010) suggested that the reliability of performance assessments have been criticized and should be investigated to know the extent of agreement between scores provided by different raters. Performance assessment data is only useful to inform decisions and make improvements in the course with reliable scoring (Reddy & Andrade, 2010).

Determining the reliability a performance assessment can be seen throughout different education fields, including language (ex. Allen & Knight, 2009; Stolarova et al., 2014) and medicine (ex. Kottner et al., 2011; Liao et al., 2010). The source of this research began with methodologies for the setup of reliability studies outlined by Shrout and Fliess (1979) and later by McGraw and Wong (1996). A limited number of reliability studies exist within the technical graphics field that relate to the agreement of multiple raters. Comparisons of different methods can be seen in Kirstukas and Amaya-Bower’s (2016) work, but a human rater in this situation produced only one set of the scores. However, Branoff et al., (2016) and Company et al., (2013) provided examples of reliability studies using multiple human raters in technical graphics assignments.

Contero et al. (2013) compared student self-assessment scores to instructor scores using a reliability change index that highlights areas of significant differences between the
scores. However, this study focused more on the design and results of Branoff et al.’s (2016) study of reliability between instructors because the survey data revealed that instructors and teaching assistants utilizing rubrics are the primary source of grading in fundamental technical graphics courses. Three instructors from Illinois State University each rated 10 student submissions for two different models (a total of twenty projects per rater). The projects were randomly selected from stratified groups based on the quality determined by the original rater to provide a sample of different qualities of work. After random selection of the stratified projects, two other instructors from the same university department individually rated each of the twenty projects. Correlations between the raters were compared using Spearman’s correlation coefficient to determine if there were significant correlations between the raters on their overall rating for each model and each rubric criteria for both of the models. Not only were correlations between ratings compared, but also overall agreement using a two-way random intra-class correlation. The researchers reported significant results for the overall evaluation of the two projects, but found no consensus for two of the categories on one project and those same two categories and one other for the other project. Ultimately, the study reported overall reliability using the rubric to evaluate the models, but refinement was needed to increase the reliability of scoring for the individual categories of base/core, end conditions, and base model origin.

Past reliability studies in the technical graphics field, along with guidelines outlined by Allen and Knight (2009) for rubric development and testing, were utilized in developing this research study. The following research questions guided this study to evaluate the existing instrument for assessing students on their reverse engineering design projects:
1. Are raters able to differentiate between the quality of student work on example reverse engineering projects in fundamentals of engineering graphics using the existing assessment instrument?

2. Is there interrater agreement for each example project when using the existing assessment instrument?

**Hypotheses**

The following hypotheses were tested to investigate the current status of performance assessment and the reliability of an existing performance assessment instrument used in fundamental post-secondary technical graphics courses:

- **H₀₁**: The mean score of each sample project will *not* differ from each other when rated by technical graphics educators.
- **H₀₂**: There is not substantial inter-rater reliability among the existing criteria for each of the example projects.
- **Hₐ₁**: The mean score of at least one sample project will differ when rated by technical graphics educators.
- **Hₐ₂**: There is substantial inter-rater reliability among the existing criteria for each of the example projects.

**Participants**

Participants were solicited by email address found in the 2016 EDGD membership directory, responded to the initial survey, and agreed to be contacted for a follow-up study. These participants in the reliability study were randomly selected using a random number generator from this convenience sample of respondents who were currently teaching a fundamental technical graphics course and used the same engineering graphics software as
the researcher’s university. The initial survey indicated that the highest number of instructors teaching a fundamental technical graphics course at a university was eight. For this reason, eight participants were used which also exceeds the number of raters required for a reliability study (Koo & Li, 2016; McGraw & Wong, 1996; Shrout & Fleiss, 1979).

**Methodology**

This section outlines the process and rationale for selecting the sample projects, distribution, directions, and data collection process of this study, which looked at the ability of instructors to use an instrument to differentiate between the quality of student’s work on reverse engineering design challenges as well as the consistency of their ratings using the existing instrument.

Intentional selection of example projects to grade required a pool of existing student projects to make inferences about actual practice, where the subject of each student project is different. One hundred student submissions on the reverse engineering design project for two sections of the course taught at NC State University in the spring 2015 semester were assessed using the existing assessment instrument and compiled by the instructor of the course. Determining whether the instrument is able to differentiate between the quality of work required having raters look at different levels of work to see if they are able to distinguish between those qualities as seen in Branoff et al. (2016). Koo and Li (2016) stress the importance of a heterogeneous sample when possible in a reliability study. Scores from the original course instructor were used to stratify student projects into three categories. The categories were 70-79, 80-89, and 90-100. Categories were not defined below 70 points because there were only a few projects below this threshold, and each of them lacked entire
sections of the projects, including them would result in a rating of 0 from all raters and falsely raise the level of agreement of scores.

Each complete project was assigned a number. The researcher randomly selected one project from each of the stratified categories using an online random number generator. Random sampling provided each of the projects an equal opportunity of being selected and ensured that the projects selected for this study were representative of projects from each of these stratified categories (Creswell, 2012). Three projects were selected based on meeting Walter et al., (1998) and Koo and Li’s (2016) threshold for reliability study design along with respect for the participants time.

All student and institution names were removed from the projects so that identities could not be associated with the work samples. Projects were then assigned names based on the object modeled in the project, which included an alarm clock, microphone, and knife. The physical portfolios for these projects can be seen in appendix H, I, and J respectively. The distinct names were selected to avoid confusion that may cause data entry errors and avoid numbers or letters that are ordinal and may suggest rank of quality or grades.

Participants received directions, a copy of the design challenge with the grading instrument, a digital survey link for submitting ratings, and three digital folders containing the example projects via email (Appendix E). Each project folder contained a PDF copy of the student’s portfolio containing the student’s cover page, technical sketch, engineering drawings, and assembly drawings, along with the digital files for each student’s parts and assemblies in their native file format. Categories and criteria from the assessment instrument were provided to each of the participants.
**Kruskal-Wallis and Dunn Test**

Determining whether there is a difference between groups when there is one nominal variable and one measurement variable commonly uses one-way analysis of variance, but this test requires the assumption of normality (McDonald, 2014). Without clearly being able to demonstrate normality due to the small sample size, the Kruskal-Wallis H test was used to determine if the ratings for each project differed by project overall scores and for section scores (Scales & Petlick, 2004). This non-parametric test extends the Mann-Whitney U test, which compares more than two independent groups, to investigate if there is equality of the means or medians on either a continuous or an ordinal dependent variable (Sprent & Smeeton, 2007).

For this test, Kruskal and Wallis (1952) require the assumption that observations are all independent, all those within a given sample come from a single population, populations are of approximately the same form, and that $F$ and $x^2$ tests assume approximate normality. The methods and assumptions provided by Kruskal-Wallis are still in use for non-parametric statistical analysis with only the addition of the fact that the samples are selected randomly from the population for the second assumption (Hollander, Wolfe, & Chicken, 2013).

The Kruskal-Wallis test faces the same challenge of the one-way ANOVA in that it is an omnibus test and is sufficient for determining if there is a difference between groups, but does not provide data as to which of the populations differ (Dinno, 2015). Dunn (1964) provided test for analyzing which populations differ based on ranks for those that wish to know where differences lie, rather than the fact that there are differences. Ruxton and Beauchamp (2008) claim that the Dunn procedure (as seen in Zar, 1999) is the most common method of pairwise multiple comparisons without making assumptions about normality.
Esposito and Bauer (2017) provide an example of using a Kruskal-Wallace H test to compare rank score with a follow-up Dunn’s test in their study where they compare scores between grade levels for an open-ended performance task without the assumption of normality or equivalent distributions. This study used the methodology that Esposito and Bauer (2017) utilized except looked for differences in the rank scores between projects rather than across grade levels.

Both the Kruskal-Wallis and Dunn’s tests were run in STATA 14.2 (alpha set to .05) to determine if there were statistical differences between the scores provided for each project, type of work in the project, and where any differences occurred. This question addressed Allen and Knight’s (2009) step to determine if an instrument can distinguish differences in the quality of work. The small sample size and the inability to make assumptions about the normality of the population were the main reasons for using non-parametric tests to investigate if there are differences between the mean scores from the participants for each of the projects and subcategories.

**Inter-rater Reliability**

Inter-reliability looks at the consistency of scores from different raters across students (Liao et al., 2010). Knowing the consistency of raters is important because raters commonly have different standards for evaluation (Hoyt, 2000; Myford & Wolfe, 2003) and having multiple raters providing different ratings for the same behavior defeats the purpose of objective measurement (Stemler, 2004). Understanding the reliability of a measurement provides information to practitioners about the value of a measurement (Koo, Li, 2016).

Shrout and Fleiss (1979) stated that a typical interrater reliability study has \( n \) targets rated independently by \( k \) raters. Branoff et al., (2016) began to look at the reliability of
performance assessment within technical graphics with a reliability study, where three instructors each rated 10 submissions for each of two separate CAD models on nine criteria. This study had eight raters from multiple universities independently rate three randomly selected projects on 18 different categories to investigate interrater reliability of the raters using Spearman’s rho and the reliability of instrument or rating process using intra-class correlation coefficients (ICC).

Liao et al., (2010) state that correlation tests, such as Pearson, Kendall’s Tau, and Spearman’s rho, determine the consistency ratings that multiple raters independently assign for the same performance. Correlation coefficients range from -1 to 1 with 0 representing no correlation between ordinal variables. This study used the analysis provided by Spearman (1910) over others for two reasons. First, the use of ranks rather than raw data for a comparison require no assumptions about the distribution of either variables (Sedgwick, 2014). Second, Puth, Neuhauser, and Ruxton (2015) suggest that using the same measure as previous studies allows for an easier comparison. Branoff et al., (2016) provided the only reliability study available in the technical graphics field that uses a correlation test to look at the consistency of ratings across multiple instructors on the same projects. This study used Spearman’s rho run in STATA 14.2 to make comparisons with the previous study.

Stolarova et al., (2014) advocate that Pearson correlations can provide an association of strength that enables comparisons but should not be the only measure of inter-rater reliability. ICC is a widely used reliability index for interrater reliability analysis in multiple fields of research (Koo & Li, 2016; Shieh 2012). ICC measures the degree two or more raters use an instrument to differentiate quality (Liao, et al., 2010; Kottner et al., 2011). Stolarova
et al., (2014) explained that ICC is a quality criterion of an assessment instrument or rating process, and can be regarded as an estimate of instrument reliability under certain conditions.

Multiple forms of ICC exist that can provide different results when applied to the same data (Koo & Li, 2016). Bartko (1976), McGraw and Wong (1996) and Shrout and Fleiss (1979) provided salient definitions for the different forms of ICC, along with conditions for their use. Shrout and Fleiss (1979) initially defined six forms of ICC, accompanied by four more forms, for a total of 10 by McGraw and Wong (1996). Distinguishing features for each form of ICC are based on the model, type, and definition. Landers (2015) and Koo and Li (2016) provide useful guidelines for selecting and reporting ICC and results were consulted for correct analysis of this study.

Model selection for ICC is based on whether or not there are consistent raters for all ratees and if they are a sample or population of raters. The participants for this study were a sample of fundamental technical graphics instructors from EDGD that each assessed the sample projects independently. With these conditions, a two-way random effects model was selected that utilizes a random effects model for raters that essentially controls for rater effects when producing an estimate of reliability (Landers, 2015). This model allows generalizations about reliability that can be applied to any raters who possess the same characteristics of raters in the study (Koo & Li, 2016).

Type selection for ICC determines whether there is interest in single rater measurement or mean of k rater measurements. Situations should use the ‘single rater’ type when ratings from one instructor or teaching assistant is the basis of student scores. Conversely, studies should use the ‘mean of k raters’ type when the mean of scores provided by multiple instructors are the basis of each student’s score (Landers 2011). This study
reported both types for comparison to results obtained by Branoff et al., (2016) and for a comparison of the rating procedures.

The definition selection for ICC is the final consideration that determines whether consistency or absolute agreement between raters is the primary concern. This study was interested in absolute agreement to investigate the extent that different raters assign the same score to the same subject.

Data were analyzed with two-way random effects ICC, absolute, model in STATA 14.2 (with alpha set .05). Chapter four contains the analysis results and comparisons. Thresholds of confidence interval ICC estimate values are arbitrary but provide a rough estimate for comparison. Koo and Li (2016) suggest that the most common thresholds for reliability are poor below 0.5, moderate from 0.5 to 0.75, good from 0.75 to 0.9, and excellent above 0.9.

**Study Limitations**

There are a few noteworthy limitations to this study that should be considered when reviewing and building on this study. Limitations include project selection, sample size, software selection, and the population. The following elaborates on each of these.

First, the projects were randomly selected from a stratified pool that gave each project an equal probability of being selected from their original score based stratified categories. Projects were sorted based on the scores provided by the original instructor without any evidence supporting the reliability of the scoring. The projects may have been in different categories if scored by a different rater but original scores were the only data available for stratification. The projects had variations in different parts of the work samples meaning that some had better technical sketches with lower quality CAD models or other variations in the
quality of the work. The results may have been different in the event that three other projects were selected. Only three projects were evaluated due to the time required to grade each of these in an attempt to be respectful of the volunteer raters’ time.

Second, eight participants completed the follow-up study over time with multiple reminder emails and encouraging phone calls. Levels of agreement can be compared with this population size, but looking at rater trends, such as correlations between rater agreement and years of experience, would require a substantially larger population.

Third, participants were only eligible to participate in the study if they currently use the CAD package Solidworks™ at their institution. Each of the available student work samples were completed in Solidworks™, and the literature suggested that the raters would be required to open the files to examine the models and assemblies for design intent and other modeling practices (Ault & Fraser, 2013; Branoff, 2003; Branoff & Dobelis, 2014; Kirstukas, 2013). Availability and familiarity with this specific software significantly reduced the number of eligible participants, and it is not possible to conclude that the outcomes might have been different using non-native file types and allowing instructors to grade projects using the software of their choice.

Efforts to assess student work in EDG courses are occurring around the world. Literature from European and Asian countries, ex. Contero and Company (2013), Goh et al. (2013), Othman (2007), and Sanna et al. (2012) demonstrate efforts in the direction of efficiently and reliably assessing student performance in EDG courses. This study utilized the EDGD of ASEE that has some outside representation but is heavily comprised of individuals working at universities in the United States. Also, many instructors engage in assessing student performance in EDG courses within the United States that are not members
of the EDGD and their current practice and rating sample are not reflected in this study.

Ultimately, there are multiple groups working to assess student work in fundamental technical graphics courses around the world however, the available EDGD directory was the available sample to work with for this study.

This investigation answers the guiding research questions, and the limitations presented are an acknowledgement of the extent this information portrays. Every research design is subject to limitations. Some limitations also provide a point of direction for further study.

**Summary**

This research utilized the literature and existing statistical methods to answer the guiding research questions. Gathering data to inform professional practice within the technical engineering graphics field required the use of a web-based survey to reach the 2016 EDGD membership. Challenges, rationale, and comparisons to studies using this same method were provided. Additionally, a study was conducted to investigate the reliability of a performance assessment measurement used in a fundamental technical graphics course using steps outlined by Allen and Knight (2009). Eight participants were selected from the survey respondents to grade three existing projects using the existing performance assessment instrument. The scores were analyzed using non-parametric methods to determine if the raters were able to differentiate between the quality of work on the different projects, investigate the correlation of scores provided by the raters, and report rater agreement and instrument reliability. Rationale, supporting literature, and precedent studies were discussed. Results and comparisons of data are found in the following chapter.
CHAPTER 4: RESULTS

Introduction

Results from both a survey of performance assessment trends in introductory engineering graphics courses and follow-up study of existing performance assessments reliability are reported in this chapter, providing data to answer the questions guiding this research. Participant demographic information, along with current performance assessment practices and introduction engineering graphics courses at post-secondary institutions are reported and analyzed. These data reveal information related to course size, performance objective alignment, types of student work assessed, and how student performance is assessed at each institution. This chapter also includes the results of a follow-up investigation into the reliability of a current performance assessment instrument that examines scoring trends across raters, projects, and sections of a common set of projects. Finally, a brief summary of the results are provided in the conclusion.

Research Questions

1) What is the status of performance assessment in fundamental technical graphics courses at postsecondary institutions?
   i) How many students are assessed in these courses?
   ii) What student learning objectives are assessed?
   iii) What type of student work is assessed?
   iv) How is each type of student work assessed?

2) Are course instructors at the post-secondary level able to differentiate between the quality of student work on reverse engineering projects in fundamentals technical graphics courses using an existing assessment instrument?
3) Is there significant correlation between rater scores when using an existing performance assessment instrument?

4) Is there significant inter-rater agreement for criteria of an existing performance assessment instrument?

**Survey Population**

To answer the first research question as well as sub questions, an online survey (Appendix C) was utilized. This survey collected data related to current performance assessment practices utilized in introductory engineering graphics courses. Two hundred invitations to participate in this study were sent via email to members of the Engineering Design Graphics Division (EDGD) of the American Society for Engineering Education (ASEE). Fifty members responded, providing a response rate of 25%. Of the 50 responses, 47 taught undergraduate engineering graphics courses and 39 of those 47 were currently teaching an introductory engineering graphics course at their university. These 39 respondents met the selection criteria established for this study and were included in the following analysis.

Experience is reported as the number of years teaching introductory graphics. The mean experience for this study is 14.59 years (SD = 11.23) and ranged from 1 to 41 years of experience teaching engineering graphics. These instructors currently teach between one and six courses per semester (M=1.81, SD=1.06). The academic rank of the participants ranged from graduate teaching assistants to full professor.

Academic rank data were self-reported and ranged from lecturer to full professor as seen in *Table 4.1*. 
Table 4.1. Academic rank of population.

<table>
<thead>
<tr>
<th>Academic Rank</th>
<th>Percentage</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instructor/Lecturer</td>
<td>23.08%</td>
<td>9</td>
</tr>
<tr>
<td>Teaching Assistant Professor</td>
<td>5.13%</td>
<td>2</td>
</tr>
<tr>
<td>Teaching Associate Professor</td>
<td>5.13%</td>
<td>2</td>
</tr>
<tr>
<td>Assistant Professor</td>
<td>12.82%</td>
<td>5</td>
</tr>
<tr>
<td>Associate Professor</td>
<td>20.51%</td>
<td>8</td>
</tr>
<tr>
<td>Professor</td>
<td>23.08%</td>
<td>9</td>
</tr>
<tr>
<td>Other</td>
<td>10.26%</td>
<td>4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>39</strong></td>
</tr>
</tbody>
</table>

Fundamental Technical Graphics Course Enrollment

The survey asked the respondents about the number of sections and students per section, providing the answer to the first sub question of research question one and providing an understanding of how many students are assessed in fundamental technical graphics courses at universities across the United States. The number of sections varied by the university, ranging from 1 to 30 (\(M=7.76, SD=6.59, n=38\)), and the average number of students per section reported ranged from 15 to 380 (\(M=51.31, SD=61.2, n=39\)). The range in class size and number of sections offered show that some universities offer an introductory engineering graphics course to a small class one time per year, where other universities are offering many large sections every semester to serve thousands of students per year.

Fundamental Technical Graphics Student Learning Objectives

The second sub question for research question one concerned the student learning objectives of the course. Participants selected the objectives that most aligned with their fundamental technical graphics course from the performance objectives for introductory
engineering graphics courses, as defined by Barr (2012). The objectives were placed in descending order by the percentage of participants that indicated that it aligned with their course. This ranking can be seen in Table 4.2.

Table 4.2. Course objectives covered in introductory engineering graphics courses.

<table>
<thead>
<tr>
<th>Objectives</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to create dimensions</td>
<td>94.87%</td>
</tr>
<tr>
<td>Ability to create section views</td>
<td>84.62%</td>
</tr>
<tr>
<td>Ability to sketch engineering objects in freehand mode</td>
<td>79.49%</td>
</tr>
<tr>
<td>Ability to visualize 3-D solid computer models</td>
<td>76.92%</td>
</tr>
<tr>
<td>Ability to create 2-D computer geometry</td>
<td>76.92%</td>
</tr>
<tr>
<td>Ability to create 3-D solid computer models</td>
<td>71.79%</td>
</tr>
<tr>
<td>Ability to generate engineering drawings from computer models</td>
<td>69.23%</td>
</tr>
<tr>
<td>Ability to perform design projects</td>
<td>56.41%</td>
</tr>
<tr>
<td>Ability to create presentation graphics</td>
<td>43.59%</td>
</tr>
<tr>
<td>Ability to analyze 3-D computer models</td>
<td>38.46%</td>
</tr>
<tr>
<td>Ability to solve traditional descriptive geometry problems</td>
<td>35.90%</td>
</tr>
<tr>
<td>Ability to create geometric construction with hand tools</td>
<td>25.64%</td>
</tr>
<tr>
<td>Knowledge of manufacturing and rapid prototyping methods</td>
<td>17.95%</td>
</tr>
<tr>
<td>Other</td>
<td>17.95%</td>
</tr>
</tbody>
</table>

The “other” category allowed participants to write in their own objectives and yielded responses including, field sketching, creating moving assemblies, teamwork, tolerancing, and the ability to read and understand engineering drawings. From the list of objectives, the first twelve use the term ability, which indicates that the students should be able to demonstrate some sort of performance to meet this objective.
Fundamental Technical Graphics Performance Assessment

Instructors indicated that a strong majority (94.87 percent, \( n = 39 \)) of courses require students to use an engineering graphics software as part of their course. When asked: “What percentage of the student's course grade is determined by assignments requiring students to demonstrate technical ability?” the instructors indicated that over half of the student’s course grade is determined by their ability to perform technical tasks (\( M = 65.18, SD = 22.12, n = 39 \)).

The type of technical tasks that the students create in the class were reported in Figure 4.1. Responses indicate that the majority of fundamental courses require students to turn in technical sketches (92.31%), computer generated assemblies (89.74%), computer generated engineering drawings (69.23%), and computer generated 3D models (69.23%). Far fewer courses require students to turn in physical models created by hand (5.13%) and digitally fabricated models (15.38%). Respondents who selected the other option (12.82%) were given an open-ended text box and their responses included: field sketches, open-ended design project deliverables, written reports, presentations, and working drawing packages.

A range of assessment methods discussed in the literature sparked interest in the way that each type of work is assessed. Figure 4.2, below, shows the responses and reveals performance assessment method trends. A majority of the assessment relies on the instructors and teaching assistants. Physical models are always assessed by an instructor or teaching assistant, along with the “other” categories previously listed. Automated systems, peer evaluation, and self-evaluation are only used on computer generated files and technical sketches. Ault and Fraser (2013), Baxter and Guerci (2003), Goh et al. (2013) Hekman and Gordon (2013), Kirstukas (2016), and Kwon and McMains (2015) provide multiple discussions about automated systems and their advantages. However, the survey results
suggest that these systems are not yet widespread with most universities are not utilizing these systems and rely mostly on instructors and teaching assistants for performance assessment grading.

Figure 4.1: Grading method utilized for different types of work assessed in university introductory engineering graphics courses.

Figure 4.1: Grading method utilized for different types of work assessed in university introductory engineering graphics courses.

Manual grading methods commonly rely on rubrics in order to define criteria and specify performance that qualifies for each level or grade. A majority of the participants ($n=39, 79.49\%$) indicated that they utilize rubrics for assessing student work in their course. Among those that indicated that they use rubrics, there is a wide variety of rubric types that are utilized in courses, along with some uncertainty about different types of rubrics. Differences and uncertainty can be seen in the responses regarding whether the rubrics utilized were specific or general. Participants were asked what type of rubric they used and to select all that applied. Responses can be seen in Table 4.3 and indicate that more specific
rubrics ($n=39, 72.97\%$) are utilized than general ($n=39, 48.65\%$), but there are some universities that utilize both types of rubric. Two respondents ($n=39, 5.41\%$) were unsure of whether their rubrics were specific or general.

Table 4.3. Responses to whether specific or general rubrics are utilized in introductory engineering graphics courses.

<table>
<thead>
<tr>
<th>Answer</th>
<th>%</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific (made specifically for a certain assignment)</td>
<td>72.97%</td>
<td>27</td>
</tr>
<tr>
<td>General (can be used for multiple assignments)</td>
<td>48.65%</td>
<td>18</td>
</tr>
<tr>
<td>I am not sure</td>
<td>5.41%</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>37</td>
</tr>
</tbody>
</table>

Variety continued and uncertainty grew when participants were asked whether the rubrics utilized in their courses were analytical or holistic. Table 4-4 below displays the responses that include less overlap in types of rubrics than the previous question but the uncertainty about types of rubrics increased ($n=37, 16.22\%$). From this data, analytical rubrics appear to be more common ($n=37, 56.76\%$) than holistic rubrics ($n=37, 32.43\%$) in introductory engineering graphics courses.

Table 4.4. Responses to whether analytical or holistic rubrics are utilized in introductory engineering graphics courses.

<table>
<thead>
<tr>
<th>Answer</th>
<th>%</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical (guidelines for each level of performance are provided for each criterion individually)</td>
<td>56.76%</td>
<td>21</td>
</tr>
<tr>
<td>Holistic (guidelines for each level of performance are the same for all criterion)</td>
<td>32.43%</td>
<td>12</td>
</tr>
<tr>
<td>I am not sure</td>
<td>16.22%</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>100%</td>
<td>37</td>
</tr>
</tbody>
</table>
Along with the type of rubrics used, participants were asked about other rubric use. Survey results indicated that rater training is provided for using grading rubrics in some fundamental technical graphics courses but that is not always the case. A slight majority of the participants regularly offer training for using their grading rubrics in their courses (n=37, 51.35%). Meanwhile, other participants indicated that training is offered sometimes (n=37, 24.32%), and some do not offer training at all (n=37, 24.32%).

Finally, the survey asked if students were provided with written feedback on their work. A majority of respondents (n=39, 64.10%) indicated that they do provide written feedback on assignments. Only one participant stated that he or she does not provide written feedback on assignments, while the remaining, 33.33%, participants sometimes offer written feedback on assignments.

**Follow-Up Study Participants**

Survey results provided demographic information about performance assessment practice in fundamental technical graphics. The second part of this study was interested in the reliability of performance assessment in the courses. Eight instructors who currently teach introduction engineering graphics courses at post-secondary institutions participated in the follow-up study to investigate the reliability of an existing technical graphics performance instrument utilized to assess final projects. Each of the participants individually assigned scores for each of categories for all three example projects provided. Experience teaching engineering graphics courses ranged from 1 to 41 years for the participants, with an average of 10.25 years of engineering graphics teaching experience. Each of the instructors taught between one and three sections of engineering graphics courses per semester and report having between 25 and 120 students per section.
Ability to Differentiate Between Submissions

Allen and Knight (2009) suggest that a quality rubric is able to differentiate between submissions of different quality. To determine whether the eight raters were able to differentiate between the qualities of the alarm, knife, and microphone projects, the following hypotheses were examined:

\( H_0: \) There is no difference between the scores for each project.

\( H_A: \) There is a difference between at least one of the scores for the sample projects.

A Kruskall-Wallis H test showed that there was a statistically significant difference in total score between sample projects, \( \chi^2(2) = 15.419, \ p = 0.0005 \), with a mean total score of 95.625 for Project Alarm, 70.6875 for Project Microphone, and 72.5 for Project Knife. Evidence supports that we should reject the null hypothesis in favor of the alternative because there is a difference between at least one of the project scores.

To determine which project’s scores were significantly different, a post hoc Dunn’s test revealed that there was a significant difference in the total score for Project Alarm between both Project Microphone, \( \chi^2 = 3.435, \ p = 0.003 \), and Project Knife, \( \chi^2 = 3.364, \ p = 0.0004 \). However, there was not statistical evidence to support a difference between Project Microphone and Project Knife, \( \chi^2 = -0.0708, \ p = 0.4718 \).
Table 4.5. Results for Dunn's test for total and each category for the project.

<table>
<thead>
<tr>
<th>Projects</th>
<th>Chi-Squared</th>
<th>P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alarm and Microphone</td>
<td>3.435</td>
<td>.000**</td>
</tr>
<tr>
<td>Alarm and Knife</td>
<td>3.365</td>
<td>.000**</td>
</tr>
<tr>
<td>Microphone and Knife</td>
<td>-0.071</td>
<td>.472</td>
</tr>
<tr>
<td><strong>Sketch</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alarm and Microphone</td>
<td>2.516</td>
<td>.006*</td>
</tr>
<tr>
<td>Alarm and Knife</td>
<td>4.022</td>
<td>.000**</td>
</tr>
<tr>
<td>Microphone and Knife</td>
<td>1.506</td>
<td>.066</td>
</tr>
<tr>
<td><strong>3D Models</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alarm and Microphone</td>
<td>2.667</td>
<td>.004*</td>
</tr>
<tr>
<td>Alarm and Knife</td>
<td>2.902</td>
<td>.002*</td>
</tr>
<tr>
<td>Microphone and Knife</td>
<td>0.234</td>
<td>.407</td>
</tr>
<tr>
<td><strong>Engineering Drawings</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alarm and Microphone</td>
<td>3.009</td>
<td>.001*</td>
</tr>
<tr>
<td>Alarm and Knife</td>
<td>1.988</td>
<td>.023*</td>
</tr>
<tr>
<td>Microphone and Knife</td>
<td>-1.021</td>
<td>.154</td>
</tr>
<tr>
<td><strong>Assembly</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alarm and Microphone</td>
<td>3.451</td>
<td>.000**</td>
</tr>
<tr>
<td>Alarm and Knife</td>
<td>1.373</td>
<td>.085</td>
</tr>
<tr>
<td>Microphone and Knife</td>
<td>-2.078</td>
<td>.019*</td>
</tr>
</tbody>
</table>

*Note: * Significant at $\alpha = .05$ ** Significant at $\alpha = .001$

The scores indicated that projects had varying quality throughout in different areas of the projects. Even with a higher overall average score, the knife project was the lowest ranked for the sketching portion. Determining where each of the project’s strengths and weaknesses were, according to the raters, required looking at the sketching, engineering drawings, 3D solid models, and assembly scores individually. A Kruskall-Wallis H and post hoc Dunn’s test were run for each of the categories. The statistical results of the test can be seen in Table 4.5 and the scores can be viewed graphically in Figure 4-2. Each category displayed statistical differences in the scores for two of the three pairs of projects but not always the same two projects.
Figure 4.2. Performance assessment scores for total projects and individual categories.
Scores in the sketching section support that raters found the sketching done in the alarm project to be better than the sketching in the microphone project, \( \chi^2 = 2.5163, p = 0.0059 \), and the knife project, \( \chi^2 = 4.0225, p = 0.0000 \). There is not any statistical support for a difference in the scores between the sketching portion of the microphone and knife project, \( \chi^2 = 1.5062, p = 0.0660 \). Scores in the 3D solid models, \( \chi^2 = 2.6672, p = 0.0038 \), \( \chi^2 = 2.9015, p = 0.0019 \), and engineering drawing sections, \( \chi^2 = 3.0087, p = 0.0013 \), \( \chi^2 = 1.9879, p = 0.0234 \), tell similar stories with statistical support showing scores on the alarm clock project were higher than both the microphone and knife project, respectively. The data does not support a statistical difference between the scores assigned for 3D solid models, \( \chi^2 = 0.2343, p = 0.4074 \), and engineering drawings, \( \chi^2 = -1.0208, p = 0.1537 \), for the microphone and knife projects.

The assembly section is the only section that does not place the alarm project scores higher than both of the other projects. The alarm project was rated higher than the microphone project, \( \chi^2 = 3.4508, p = 0.0003 \), on the assembly section. Raters awarded more points to the knife project than the microphone project, \( \chi^2 = -2.0777, p = 0.0189 \). However, there is no evidence to support a difference between the scores assigned to the alarm and knife projects, \( \chi^2 = 1.3731, p = 0.0849 \), on the assembly section.

Scores assigned by the raters demonstrate that the alarm clock project was consistently strong across each of the categories, where the microphone and knife projects had their own respective strengths and weaknesses. The data supports that there is a statistically significant difference between the score for the alarm clock project when compared to both of the other projects. The difference in overall score between the microphone and knife project were not statistically significant. This evidence supports the
rejection of the null hypothesis because at least one mean score was different from the other scores for overall scoring on the sketch, model, drawing, and assembly category scores. This indicates a foundation of the fact that rater scores do differentiate for different qualities of work when using the existing performance assessment instrument.

**Correlation of Scores by Raters**

Research question three investigated the consistency of scores provided by each of the raters for each example project when using the existing assessment instrument. *Figure 4.3* graphically demonstrates the variation of overall project scores provided by each of the raters. The strength of the correlation between each rater were calculated and tested for all scores together and all scores on each project individually using Spearman’s rho in STATA 14.2.

![Overall project scores by rater.](image)

*Figure 4.3. Overall project scores by rater.*
Results as displayed in Table 4.6 indicated that statistically significant correlations exist between each pair of raters when looking at scores for all projects across each category combined. The strongest correlation of all scores was between raters one and eight: 
\[ r(54)=.897, p<.001 \], and the weakest correlation existed between raters seven and eight: 
\[ r(54)=.500, p<.001 \]. Averages of the correlation values suggest that rater one had the highest correlation of scores with all other raters, while rater seven had the lowest. Looking at only the combined scores would support rejection of the null hypothesis because there are statistically significant correlations between each pair of raters. However, determining the consistency of scores provided for each project required testing the correlation of scores for each project individually.

Table 4.6. Spearman’s Correlation Coefficient for all ratings

<table>
<thead>
<tr>
<th></th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
<th>Rater 4</th>
<th>Rater 5</th>
<th>Rater 6</th>
<th>Rater 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater 2</td>
<td>.605**</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 3</td>
<td>.810**</td>
<td>.648**</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 4</td>
<td>.706**</td>
<td>.685**</td>
<td>.753**</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 5</td>
<td>.887**</td>
<td>.719**</td>
<td>.776**</td>
<td>.745**</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 6</td>
<td>.668**</td>
<td>.631**</td>
<td>.632**</td>
<td>.588**</td>
<td>.744**</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td>Rater 7</td>
<td>.619**</td>
<td>.569**</td>
<td>.624**</td>
<td>.625**</td>
<td>.607**</td>
<td>.572**</td>
<td>-----</td>
</tr>
<tr>
<td>Rater 8</td>
<td>.897**</td>
<td>.526**</td>
<td>.796**</td>
<td>.714**</td>
<td>.841**</td>
<td>.568**</td>
<td>.500**</td>
</tr>
</tbody>
</table>

*Note: * Significant at \( \alpha =.05 \) ** Significant at \( \alpha =.001 \)

Tables 4.7, 4.8, and 4.9 display correlation results for the alarm, knife, and microphone projects respectively. All correlation coefficients for the alarm project are statistically significant and higher than correlations on both other projects. Score correlations for the knife project were statistically significant for all pairs of raters except for raters four and seven, \( r(18)=.454, p=.058 \) and raters seven and eight, \( r(18)=.427, p=.077 \). On the
microphone project, twelve rater pairing score correlations were not statistically significant with the lowest being between rater two and seven, \( r(18) = -.074, p = .769 \).

Table 4.7. Spearman’s Correlation Coefficient for Alarm Project

<table>
<thead>
<tr>
<th></th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
<th>Rater 4</th>
<th>Rater 5</th>
<th>Rater 6</th>
<th>Rater 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater 2</td>
<td>.984**</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 3</td>
<td>.963**</td>
<td>.984**</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 4</td>
<td>.907**</td>
<td>.898**</td>
<td>.875**</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 5</td>
<td>.975**</td>
<td>.995**</td>
<td>.977**</td>
<td>.907**</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 6</td>
<td>.675**</td>
<td>.710**</td>
<td>.703**</td>
<td>.698**</td>
<td>.766**</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Rater 7</td>
<td>.896**</td>
<td>.921**</td>
<td>.909**</td>
<td>.890**</td>
<td>.927**</td>
<td>.691**</td>
<td>------</td>
</tr>
<tr>
<td>Rater 8</td>
<td>.972**</td>
<td>.979**</td>
<td>.960**</td>
<td>.898**</td>
<td>.971**</td>
<td>.663**</td>
<td>.886**</td>
</tr>
</tbody>
</table>

Note: * Significant at \( \alpha = .05 \) ** Significant at \( \alpha = .001 \)

Table 4.8. Spearman’s Correlation Coefficient for Knife Project

<table>
<thead>
<tr>
<th></th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
<th>Rater 4</th>
<th>Rater 5</th>
<th>Rater 6</th>
<th>Rater 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater 2</td>
<td>.759**</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 3</td>
<td>.721**</td>
<td>.707**</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 4</td>
<td>.766**</td>
<td>.727**</td>
<td>.750**</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 5</td>
<td>.884**</td>
<td>.656**</td>
<td>.759**</td>
<td>.842**</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 6</td>
<td>.798**</td>
<td>.722**</td>
<td>.663**</td>
<td>.707**</td>
<td>.801**</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Rater 7</td>
<td>.587*</td>
<td>.529*</td>
<td>.542*</td>
<td>.454*</td>
<td>.535*</td>
<td>.690**</td>
<td>------</td>
</tr>
<tr>
<td>Rater 8</td>
<td>.885**</td>
<td>.606**</td>
<td>.753**</td>
<td>.726**</td>
<td>.883**</td>
<td>.670**</td>
<td>.427</td>
</tr>
</tbody>
</table>

Note: * Significant at \( \alpha = .05 \) ** Significant at \( \alpha = .001 \)

Table 4.9. Spearman’s Correlation Coefficient for Microphone Project

<table>
<thead>
<tr>
<th></th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
<th>Rater 4</th>
<th>Rater 5</th>
<th>Rater 6</th>
<th>Rater 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rater 2</td>
<td>.135</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 3</td>
<td>.809**</td>
<td>.246</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 4</td>
<td>.539*</td>
<td>.202</td>
<td>.546*</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 5</td>
<td>.780**</td>
<td>.553*</td>
<td>.615**</td>
<td>.457</td>
<td>------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rater 6</td>
<td>.563*</td>
<td>.486*</td>
<td>.621**</td>
<td>.307</td>
<td>.723**</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>Rater 7</td>
<td>.693**</td>
<td>-.074</td>
<td>.325</td>
<td>.415</td>
<td>.474*</td>
<td>.446</td>
<td>------</td>
</tr>
<tr>
<td>Rater 8</td>
<td>.853**</td>
<td>.105</td>
<td>.742**</td>
<td>.565*</td>
<td>.700**</td>
<td>.420</td>
<td>.432</td>
</tr>
</tbody>
</table>

Note: * Significant at \( \alpha = .05 \) ** Significant at \( \alpha = .001 \)
In summary for the correlation of rater scores, the data supports that there is a significant correlation between scores provided by raters using the existing performance assessment instrument for some but not all projects. Ratings should be consistent between all raters for all projects, but the results indicated that that is not the case. In the case of the alarm project, the data supports that we reject the null hypothesis because significant correlation exists between all rater scores. However, on the microphone and knife projects, significant correlations exist between some but not all rater scores. Raters were consistent across the higher rated, alarm, project, but the consistency of the scores decreased with the lower quality, microphone and knife, projects. This evidence suggest that we should fail to reject the null hypothesis because there are not statistically significant correlations between all scores provided by raters.

**Interrater Agreement**

The fourth research question investigated whether there was an interrater agreement for the scores provided by raters when using the performance assessment instrument. Table 4.10 show results for ICC(2,8) agreement for individual and average measure.
Table 4.10. Intra-class Correlation Coefficients (ICC) for Total Sum and Category Sums

<table>
<thead>
<tr>
<th>Category</th>
<th>Single Measures</th>
<th>Average Measures</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
<th>Value</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>.582</td>
<td>.917</td>
<td>.194</td>
<td>.938</td>
<td>25.350</td>
<td>.000**</td>
</tr>
<tr>
<td>Sketch Sum</td>
<td>.741</td>
<td>.958</td>
<td>.352</td>
<td>.983</td>
<td>48.680</td>
<td>.000**</td>
</tr>
<tr>
<td>Model Sum</td>
<td>.335</td>
<td>.801</td>
<td>.055</td>
<td>.957</td>
<td>9.010</td>
<td>.003*</td>
</tr>
<tr>
<td>Drawing Sum</td>
<td>.414</td>
<td>.849</td>
<td>.088</td>
<td>.968</td>
<td>10.760</td>
<td>.001**</td>
</tr>
<tr>
<td>Assembly Sum</td>
<td>.606</td>
<td>.925</td>
<td>.214</td>
<td>.983</td>
<td>20.380</td>
<td>.000**</td>
</tr>
<tr>
<td>Portfolio</td>
<td>.480</td>
<td>.881</td>
<td>.123</td>
<td>.975</td>
<td>12.660</td>
<td>.001**</td>
</tr>
</tbody>
</table>

Note: * Significant at α = .05  ** Significant at α = .001

The ICC (2,8) agreement analysis revealed significant results for overall scores and for the sum scores for each type of work submitted on the project. Overall, the average measure ICC(2,8) was .917, with a 95% confidence interval between .658 and .998 (F(2,14)=25.350, p<.001. This suggests that 91.7% of the variance in the overall average score across all raters is a true variance. Koo and Li (2016) suggest that even though the overall scores fall into the excellent reliability range, the 95% confidence interval suggests that the reliability could be in the moderate or good range. The average measures for each category all fall within the good to excellent range, but single measures are much lower. The single measure for the sum model (.335), drawing (.414), and portfolio (.480) scores all fall...
within the poor reliability range, and the 95% confidence intervals for all of the single measures suggest that they might all fall in the poor range.

ICC (2,8) agreement was also run for each criteria of the instrument. Tables 4.11, 4.12, 4.13, and 4.14 provide results for the analysis for the sketching, model, drawing, and assembly categories, respectively. The intra-class correlation analysis revealed significant results for twelve of the eighteen individual categories within the instrument.

Table 4.11. Intra-class Correlation Coefficients (ICC) for Sketching Criteria

<table>
<thead>
<tr>
<th>Sketch</th>
<th>Single Measures</th>
<th>Average Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sketch 1</td>
<td>Intraclass Correlations: .685, .946</td>
<td>95% Confidence Interval: .289, .764</td>
</tr>
<tr>
<td>Sketch 2</td>
<td>Intraclass Correlations: .786, .967</td>
<td>95% Confidence Interval: .427, .856</td>
</tr>
<tr>
<td>Sketch 3</td>
<td>Intraclass Correlations: .384, .833</td>
<td>95% Confidence Interval: .071, .379</td>
</tr>
<tr>
<td>Sketch 4</td>
<td>Intraclass Correlations: .749, .960</td>
<td>95% Confidence Interval: .369, .824</td>
</tr>
<tr>
<td>Sketch 5</td>
<td>Intraclass Correlations: .301, .775</td>
<td>95% Confidence Interval: .035, .225</td>
</tr>
</tbody>
</table>

Note: * Significant at α = .05 ** Significant at α = .001

Results for the technical sketch category, seen in Table 4.11, demonstrate that there was significant agreement across all technical sketching criteria. The highest level of agreement on the sketch can be seen in the second category, view alignment, that had an average measures ICC(2,8) of .967, with a 95% confidence interval between .856 and .999 (F(2,14)=35.120, p<.001. The fifth category, appropriate part name, had the lowest level of
agreement among raters on the technical sketch, with an average measures ICC(2,8) of .775, with a 95% confidence interval between .225 and .994 (F(2,14)=6.730, p=.009. For individual measures, all categories except three, line type (.384), and five, appropriate part name (.301), have at least moderate levels of agreement. The average rater scores for the technical sketch categories do have significant agreement for all individual categories but not for individual raters for the line type and appropriate part name categories. Overall, for the sketch, the data supports the rejection of the null hypothesis because there is significant agreement of raters on the sketch categories.

Table 4.12. Intra-class Correlation Coefficients (ICC) for Model Criteria

<table>
<thead>
<tr>
<th>Model</th>
<th>Single Measures</th>
<th>Average Measures</th>
<th>95% Confidence Interval</th>
<th>Ftest with True Value 0 (df1=2 df2=14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>Single Measures</td>
<td>.137</td>
<td>-.026</td>
<td>.903</td>
</tr>
<tr>
<td></td>
<td>Average Measures</td>
<td>.560</td>
<td>-.258</td>
<td>.967</td>
</tr>
<tr>
<td>Model 2</td>
<td>Single Measures</td>
<td>.412</td>
<td>.073</td>
<td>.969</td>
</tr>
<tr>
<td></td>
<td>Average Measures</td>
<td>.849</td>
<td>.387</td>
<td>.996</td>
</tr>
<tr>
<td>Model 3</td>
<td>Single Measures</td>
<td>.010</td>
<td>-.027</td>
<td>.873</td>
</tr>
<tr>
<td></td>
<td>Average Measures</td>
<td>.469</td>
<td>-.265</td>
<td>.982</td>
</tr>
<tr>
<td>Model 4</td>
<td>Single Measures</td>
<td>.047</td>
<td>-.045</td>
<td>.823</td>
</tr>
<tr>
<td></td>
<td>Average Measures</td>
<td>.283</td>
<td>-.528</td>
<td>.974</td>
</tr>
<tr>
<td>Model 5</td>
<td>Single Measures</td>
<td>.065</td>
<td>-.042</td>
<td>.847</td>
</tr>
<tr>
<td></td>
<td>Average Measures</td>
<td>.358</td>
<td>-.479</td>
<td>.978</td>
</tr>
</tbody>
</table>

Note: * Significant at α =.05 ** Significant at α =.001

Category sum analysis, seen in Table 4.12, reveal that the model section had the lowest rater agreement for both average (.801) and single measures (.335) ICC(2,8). The individual category results indicated that there was only significant agreement for the second
category, fully defined sketches (.849). Intra-class correlations fall within the poor category for the model orientation (Model 1), accuracy (Model 3), model procedure (Model 4), and appropriate part name (Model 5) criteria. This evidence supports the null hypothesis that there is no rater agreement for model scores.

Table 4.13. Intra-class Correlation Coefficients (ICC) for Drawing Criteria

<table>
<thead>
<tr>
<th></th>
<th>Intraclass Correlations</th>
<th>95% Confidence Interval</th>
<th>Ftest with True Value 0 (df1=2 df2=14)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower Bound</td>
<td>Upper Bound</td>
</tr>
<tr>
<td><strong>Drawing 1</strong></td>
<td>Single Measures</td>
<td>.343</td>
<td>.051</td>
</tr>
<tr>
<td></td>
<td>Average Measures</td>
<td>.807</td>
<td>.303</td>
</tr>
<tr>
<td><strong>Drawing 2</strong></td>
<td>Single Measures</td>
<td>.281</td>
<td>.010</td>
</tr>
<tr>
<td></td>
<td>Average Measures</td>
<td>.757</td>
<td>.077</td>
</tr>
<tr>
<td><strong>Drawing 3</strong></td>
<td>Single Measures</td>
<td>.278</td>
<td>.037</td>
</tr>
<tr>
<td></td>
<td>Average Measures</td>
<td>.755</td>
<td>.234</td>
</tr>
</tbody>
</table>

* Significant at α =.05 ** Significant at α =.001

Table 4.13 presents the results of the intra-class correlation analysis for individual criteria for the engineering drawing section. For average measures, the instrument reliability fell in the good category for all three criteria. Individual measures fell within poor reliability, but all have significant agreement at α =.05. This evidence supports the rejection of the null hypothesis because there is significant agreement for all individual categories on the engineering drawing assessment criteria.
Table 4.14. Intra-class Correlation Coefficients (ICC) for Assembly Criteria

<table>
<thead>
<tr>
<th>Assembly</th>
<th>Single Measures</th>
<th>Average Measures</th>
<th>95% Confidence Interval</th>
<th>Ftest with True Value 0 (df1=2 df2=14)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intraclass Correlations</td>
<td>Lower Bound</td>
<td>Upper Bound</td>
<td>Value</td>
</tr>
<tr>
<td>Assembly 1</td>
<td>.480</td>
<td>.123</td>
<td>.975</td>
<td>12.660</td>
</tr>
<tr>
<td></td>
<td>.881</td>
<td>.529</td>
<td>.997</td>
<td>12.660</td>
</tr>
<tr>
<td>Assembly 2</td>
<td>.019</td>
<td>-.081</td>
<td>.823</td>
<td>1.200</td>
</tr>
<tr>
<td></td>
<td>.136</td>
<td>-1.490</td>
<td>.974</td>
<td>1.200</td>
</tr>
<tr>
<td>Assembly 3</td>
<td>.256</td>
<td>-.020</td>
<td>.948</td>
<td>3.750</td>
</tr>
<tr>
<td></td>
<td>.734</td>
<td>-.185</td>
<td>.993</td>
<td>3.750</td>
</tr>
<tr>
<td>Assembly 4</td>
<td>.313</td>
<td>.031</td>
<td>.955</td>
<td>6.000</td>
</tr>
<tr>
<td></td>
<td>.784</td>
<td>.202</td>
<td>.994</td>
<td>6.000</td>
</tr>
</tbody>
</table>

Note: * Significant at α = .05 ** Significant at α = .001

Finally, Table 4.14 presents the results of the intra-class correlation analysis for the assembly section. Correct mates/relations (Assembly 1), ICC(2,8)=.881 was the only individual category with significant agreement, at the α = .001 level. Materials (Assembly 3) (.734) and CAD drawing of assembly (Assembly 4) (.784) both had significant rater agreement with α = .05. However, scores for the assembly orientation (Assembly 2) (.136) category did not yield significant rater agreement. The data supports the null hypothesis because there is no significant rater agreement for each individual criteria on the assembly section.

The data supports that there is significant rater agreement for sum scores for the entire projects and category sums. Dropping to the individual scoring criteria supports that there is significant rater agreement for all technical sketch and engineering drawing criteria but not for all criteria in the CAD model and CAD assembly categories.
Summary

Results presented in this chapter demonstrate that generally fundamental technical graphics courses at post-secondary institutions use similar student learning outcomes and commonly assesses student proficiency with performance assessments. The most commonly assessed student artifacts include technical sketches, computer generated engineering drawings, computer generated 3D solid models, and computer generated assemblies. Assessment of these artifacts are mostly graded by instructors and teaching assistants using rubrics. Training is not always provided for raters and levels for scoring criteria are not always defined.

The scores were significantly different for at least one project overall and on each project category. This supports rater’s ability to differentiate between the qualities of student work when using the existing instrument. However, the consistency of this differentiation varies with the quality of the project. There were significant correlations between the scores provided by all raters for the higher quality, alarm clock, project; however, not all rater’s scores had a significant correlation for the other two projects. The data supports that scores provided by multiple instructors can have high correlation on some projects and simultaneously lower correlations between scores on other projects. Even though the raters were able to find differences in the quality of the project, the lack of correlation between judges scores, specifically on the microphone project, suggest that these are not reliable differences.

Finally, ICC analysis results support that there is significant rater agreement when looking at overall score and sum scores for each of the project categories in the event of using average measures. Individual criteria ICC results support that there is a significant rater
agreement on twelve of the eighteen rating criteria for this project, but the other six criteria are subject to low rater agreement and should be further investigated to improve the reliability of the instrument.
CHAPTER 5: DISCUSSION

Introduction

Higher education in the United States is increasingly utilizing performance assessments as a measure the attainment of student learning objectives and provide accountability for students and institutions (Baartman et al., 2006; Baartman, Gulikers, & Dijkstra, 2013; Cummings, Maddux, & Richmond, 2008). Performance assessments are used throughout multiple disciplines (Hack, 2015) including technical graphics. Thousands of students per year take fundamental technical graphics courses that must be assessed based on skills and knowledge for the sake of reporting student progress, program evaluation, and accreditation (ABET, 2016a). Institutions handle the student load by creating multiple sections or larger sections but, ultimately, need to assess the skills of a large number of students. Reliable instruments are required to determine the success of the lessons in the curriculum relating to specific learning objectives. For example, the ability to freehand sketch engineering parts needs to be examined to see how well students are learning to sketch during the course. Without a reliable instrument, there is a possibility for deficiencies in skill to go unnoticed for individual students and even entire programs (Moskal & Leydens, 2000).

This research was conducted as a two-part study. A web-based survey of the EDGD membership was used to investigate the status of performance assessment in fundamental technical graphics courses at postsecondary institutions including the number of students assessed, learning objectives assessed, type of student work assessed, and methods for assessment of performance tasks. Results revealed performance assessment trends, discussed in the next section, and provided participants for the following reliability study.

The reliability study had eight fundamental EDG instructors from various universities
each individually assess three existing reverse engineering projects of various quality with the guidance of the current performance assessment instrument. Example projects were randomly selected from existing projects that were stratified by grades given by the instructor when the project were completed with one project originally receiving an A, 90-100; one receiving a B, 80-89; and one a C, 70-79. Participant scores were analyzed with the Kruskal-Wallis H and post hoc Dunn’s tests to determine if course instructors were able to differentiate between the quality of student work on reverse engineering projects in fundamentals technical graphics courses using the existing assessment instrument. Spearman’s Rho was used to analyze if there was significant correlation between rater scores when using the performance assessment instrument. Finally, intra-class correlation coefficients (ICC(2,8)) were used to test for significant inter-rater agreement for criteria of the existing performance assessment instrument. Results from these tests and implications are discussed in the following rubric development and validation section.

Performance Assessment Trends in Technical Graphics Courses

Across universities, the data suggests that these courses have common objectives. A majority of the participants suggest that their fundamental technical graphics course covers dimensioning, section views, 2-D computer geometry, and 3-D solid computer models, how to sketch engineering objects in freehand mode, visualize 3-D solid computer models, generate engineering drawings from computer models, and perform design projects. These results support that most fundamental technical graphics courses are still currently utilizing the top rated outcomes reported by Barr (2012). This study expanded upon Barr’s (2012) work by also including types of work and methods employed in these courses to measure these learning outcomes. This study found that the types of work assessed in these courses
also shared commonality. Digital and physical models were assessed at a few universities, but the four primary types of student work assessed in fundamental technical graphics courses are technical sketches, computer generated engineering drawings, 3D models, and assemblies. Similarities can be seen across course objectives and types of student work assessed in fundamental technical graphics courses.

Multiple approaches to performance assessment have emerged in technical graphics courses, including computer-automated methods, example Ault & Fraser (2013); manual grading with rubrics or checklists, example Barr et al, (2014); verbal protocol analysis, example Menary, Robinson, & Belfast (2011); peer; self; observation; and adaptive comparative judgement. Each of these approaches is thoroughly described in the literature with their advantages and limitations, but it is difficult to tell the extent of their usage. Even though there are many positives, including the speed and accuracy of automated computer grading, the data suggest that these are not widely utilized in fundamental EDG courses at this time. However, the survey results from this study support that the majority of the performance assessment is completed by instructors and teaching assistants, with peer assessment in a distant third place. A few universities utilize self-assessment and computer automated assessment systems, but a majority of the performance assessment workload falls to manual grading done by instructors and teaching assistants using rubrics. These results support the need for validated rubrics in fundamental EDG courses as they are the primary measure of student achievement.

**Rubric Development and Validation**

Even though this study supports that rubrics are one of the primary measures of student achievement within fundamental EDG courses, the the development and validation of
rubrics are not widespread throughout the literature. Examples of instruments can be seen such as Branoff’s (2004) trip lever and final project rubric, Ault and Frasier’s (2013) pipe flange manual grading rubric, Branoff and Dobelis’ (2012) new rubric, and Barr et al.’s (2000) evaluation sheet, but none of these provide the process or data supporting the reliability or conditions for reliable scoring with these instruments. Branoff et al. (2016) provided one of the first studies in the field to examine the intra-rater reliability of an instrument within the technical graphics field.

This study utilized Branoff et al.’s (2016) recommendations to increase the number of raters along with use raters from multiple universities. This study also expanded upon Branoff et al.’s (2016) study that looked at the grading of 3D CAD models by expanding to other types of student work including technical sketches, engineering drawings, and CAD assemblies in addition to 3D CAD models. The additional types of student work allowed this study to look at each of the primary types of student artifacts assessed in fundamental EDG courses as reported by the EDG instructors in the first part of study. This study expanded upon Branoff et al.’s (2016) methodology by using steps outlined by Allen and Knight (2009) to not only investigate the reliability of the measurement process, but also the ability of the process to differentiate between quality work.

This study tested an existing technical graphics performance assessment instrument to determine if scores differentiated between the quality of work on the example projects and the reliability of scores provided by multiple instructors. The scores provided by eight fundamental technical graphics instructors across multiple universities, who rated three example projects independently, were analyzed three ways to determine the usefulness and reliability of this scoring procedure.
Performance assessment scores are useful in the event that they can differentiate between high and low quality work (Allen & Knight, 2009). Examples such as Branoff and Dobelis’ (2012) comparison of two rubrics demonstrated that the same projects may receive different scores based on the rubric they are graded with and that one of their rubrics did not provide much variation in scores across all of the work while the other rubric did. This study used varying quality example projects, as seen in Allen & Knight’s (2009), to determine if the scores varied with the quality of the work. Results from a Kruskal-Wallis H test and post hoc Dunn’s test found and identified significant differences between the scores for at least one of the projects for overall scores and scores for each of the four types of student work. The data supports that the scores provided by individual raters using this instrument were able to differentiate between the qualities of student work on the example projects and that there was some variety of quality across these projects. This variation was also important for the next part of the study as Koo and Li (2016) stress the importance of a heterogeneous sample when possible for a reliability study.

After investigating the ability to differentiate between quality of work, Allen and Knight (2009) suggested to investigate the reliability of scores provided using the instrument. Methods outlined by Shrout and Fleiss (1979) and used by Branoff et al. (2016) were utilized for a reliability study to answer questions about rater consistency and agreement due to similarities of research design and for easier comparison of results as suggested by Puth, Neuhauser, and Ruxton (2015).

Spearman’s Rho was used to analyze score correlations across raters. The results indicated that there was significant correlation between each pair of instructors for the alarm clock project, but significant correlations did not exist between each pair of raters on the
microphone and knife projects. The significant correlations between each pair of raters on the one project while not always on the others demonstrate that there can be significant correlation between scores on some projects while not on others. These results are similar to Branoff et al’s (2016) that had significant correlation between all pairs of raters for one model while simultaneously not on another model. These results demonstrate the importance investigating rater consistency across different instances the instrument is used to validate the consistency of scoring across each situation in which the instrument is used.

Finally, an ICC(2,8) agreement was performed to investigate the extent raters provided the same score for each project, section, and category. Results from this analysis provide insight about the reliability of scores for each area. ICC average measures were higher than single measures for all areas, similar to results found by Branoff et al. (2016). The data support agreement among scores across raters and was statistically significant for overall score and category scores. Branoff et al. (2016) reported that the study rubric appears to generate reliable results for the overall rating of the two models, but there was not significant correlation between all the raters on both of the projects. Low correlations on individual categories suggest unaccounted for variance and point to areas that need refinement. This study found agreement on individual criteria was significant for all but four categories. Overall, there is statistically significant reliability of total scores, but there is a need to look further at the individual criteria that raters did not agree on to increase instrument reliability.

This study was able to identify areas of weakness in the current instrument and suggest further investigation be done to improve the reliability the specific areas identified with low reliability along with increasing the individual measure of reliability if scores made
by individual instructors independently are going to be used for comparison rather than average scores made by multiple instructors. Andrade and Reddy (2010) along with Branoff et al. (2016) suggested possibilities to improve instrument reliability that include providing more detailed descriptors for each category, providing specific criteria for point values, and providing training before individually scoring. The process utilized in this study demonstrates a method for investigating the reliability of instruments that provide insights about areas of improvement and results that are easily compared to using this analysis.

Author’s Implications

The process of investigating and validating a performance assessment instrument is a lengthy but necessary process establish reliability for assessing student work. This study demonstrated to the researcher that performance assessments should be analyzed even if they are in use, analyzed in depth, and analyzed through a variety of statistical analysis. The following explains how this research lead to these suggestions.

Even though performance assessment is used in many fields to measure student-learning objectives, the reliability of scoring is not always examined or well documented. This study demonstrated the importance of investigating scoring methods to pinpoint areas for improvement rather than accept the methods because they were in use or developed by esteemed colleagues.

Not only should performance assessment practices always be tested, but they should also be tested for each application or variation that they will be utilized for. Looking at overall score correlations may mask the truth as seen in Table 4.6 where each pair of raters scores were statistically significant. Breaking the correlations down into the different projects revealed that the high correlations for the alarm project, seen in Table 4.7, had masked the
low number of significant rater pairs on the microphone project, seen in Table 4.9. A similar situation occurs in the overall ICC results in Table 4.10 where each of the categories had statistically significant agreement, but looking directly into each of the criteria, in Tables 4.11, 4.12, 4.13, showed that there was not significant agreement across each of the criteria. These results support the importance of digging into the different applications to ensure that the scoring is reliable for all situations and criteria.

In addition to levels, a variety of analysis provide a more holistic investigation that can identify different weaknesses that require different approaches to improve. Figure 4.3 demonstrates the overall scores for each of the projects where the final scores have some variation but an overall pattern for all but rater 7. There is a jump in the scores where rater 7 provided scores in the 40s for both the knife and microphone projects while most all of the other raters had provided passing scores. This variation indicates perhaps some intervention or training might be needed improve the correlation of scoring for this rater that seems to view the quality of the projects quite differently than the others. Differences in the number of pairs of raters with significant correlation across the different projects suggests that there might be a difference in the reliability for different quality or subject matter. Using the same part for all students or having a similar complexity is recommended to help improve consistency there. Finally, the intra-class correlations revealed specific criteria with and without significant agreement for the raters. The areas with low agreement could possibly be improved with better wording, specific levels and descriptions, or training. A variety of analysis provides multiple angles for researchers to analyze performance assessment practices for areas of improvement.
This reliability study demonstrates a method for investigating the reliability of performance assessment practices but also emphasizes the need for investigating scoring practices in performance assessment. Overall, this study supports the need for investigating existing performance assessment scoring, investigating reliability in depth across all situations the assessment is utilized, and using a variety of analysis to determine different areas for improvement in practice.

**Recommendations for Performance Assessment in Technical Graphics**

This study builds upon previous investigations of performance assessment in technical graphics courses and creates several follow-up questions and possible future directions. Work demonstrating the potential of computer assessment of student work provides hope to instructors buried under mountains of student work to assess, but the survey data shows that a majority of student work in fundamental technical graphics courses is still completed by instructors and teaching assistants. Information gathered in this study provides evidence for several areas of follow-up investigation and work.

First, review of the literature shows that there are multiple efforts that have been completed or are in the process of developing ways to judge the quality of student work in technical graphics (Ault & Fraser, 2013; Barr et al, 2014; Baxter & Guerci, 2003; Branoff & Dobelis, 2012, 2013, 2014; Branoff & Dobelis, 2014; Kirstukas & Amaya-Bower, 2016; McInnis, Sobin, Bertozzi, & Planchard, 2010; Menary, Robinson, & Belfast, 2011). Challenges to effective assessment included the consistency of raters (Branoff et al., 2016), large class sizes, the time required to grade student work (Goh, Shukri, & Manao, 2013; Kwon & McMains, 2015), and assess design intent (Devine & Laingen, 2013). These efforts represent substantial amounts of effort and knowledge but only provide the view of either a
single instructor or team of instructors and individual universities. Generating a consensus throughout the engineering design graphics community for what constitutes quality work for each of these four main types of student work would be beneficial to the advancement of reliable performance assessment instruments in this field. Survey results indicated technical sketches, computer generated 3D solid models, computer generated engineering drawings, and computer generated assemblies are the four primary type of student artifacts assessed in fundamental technical graphics courses. Consideration and efforts to develop successful assessments for technical graphics courses will require collaborative efforts by universities or researchers to create reliable instruments. Moving back to the earlier steps of Allen & Knights (2009) process for rubric development and validation may prove useful in the development of reliable instruments of professional value.

Participant comments about the existing instrument were not reported in the results because they were not within the scope of this investigation, but they did indicate that opinions differed about the current criteria. Tolerancing and part difficulty are two of the criteria that were mentioned as additions to the instrument in the participant’s comments. The criteria for quality of student work had been developed through multiple efforts. A consensus or ranking of what constitutes quality for each type of student work commonly submitted in fundamental technical graphics courses would be beneficial for instructors and researchers in technical graphics. Knowing these criteria for each type of work would provide opportunity for increased commonality between instruments and allow additional comparisons.

Second, the lack of correlation between all raters and low individual rater agreement supports that consideration should be given to Branoff et al.’s (2016) statement that
“…category correlations could be improved by further developing the rubric to include more detailed descriptors of each category (p. 8).” Additional investigation should look to see if this further definition for categories increases the reliability of scores that fall into the middle ranges that had lower interrater reliability than the higher ranked project. Comparisons should be conducted between instruments that describe each level of quality for each criteria and those that only list the different categories. Differences in reliability may be found for the lower level work between the instruments and there is also a chance for reduced time for student feedback or increased student feedback as the descriptors may provide verbiage explaining their shortcomings without the instructor having to write it out for individual students. Training of raters, for increasing the reliability of performance assessment methods, is also an area in need of study.

Third, the process of finding multiple raters and shared projects that reflect the common type of work assessed in technical graphics can be quite time intensive. This study shows that many people are working to assess the same types of student artifacts, and the development of general instruments might be beneficial for making informed decisions at multiple universities. Validation of assessment instruments is a challenging and time-consuming task, but it is necessary to know the extent to which students are meeting student-learning outcomes in fundamental technical graphics courses that prepare future engineers to communicate, document, and analyze their designs. Comparisons with other instruments, raters, or modifications using the example work, results, and raw data provided here are encouraged for the improvement of assessing student-learning objectives in technical graphics.
Finally, ICC data from this and Branoff et al. (2016) supports the amount of variance that is “real variance” and is always higher for average measures over single measures. When feasible, average scores from multiple raters should be used when significant decisions about student proficiency are being made using this instrument.

Summary

This study surveyed the technical graphics community to clarify objectives, types of student artifacts assessed, and assessment methods utilized in fundamental technical graphics courses at the post-secondary level. This survey verified that rubrics are the primary method of grading the primary artifacts submitted for performance assessments in these university courses and supplied participants for studying the reliability of an existing performance assessment instrument. This study adapts the methodology of Allen and Knight (2009) and builds upon Branoff et al.’s (2016) work to investigate the reliability of a performance assessment instrument for use in fundamental technical graphics courses. Study results aligned with those found by Branoff et al. (2016) and support the notion of Andrade and Reddy (2010) that the use of a rubric does not ensure a reliable measurement.

These results inform an important discussions about performance assessment in general and performance assessment within the technical graphics field. Further studies are needed for the development of performance assessment instruments with empirical support of reliability to measure student and program success. This study provides a rationale for increased attention to rubrics within technical graphics, suggestions for improved reliability in performance assessment measures, and a method of testing performance assessment instruments that can identify areas of weakness and be easily compared to existing reliability studies.
REFERENCES


Ault, H. K., & Fraser, A. (June, 2013). A comparison of manual vs. online grading for solid models. Paper presented at the 2013 American Society for Engineering Education Annual Conference & Exposition, Atlanta, GA.


Branoff, T. J., & Dobelis, M. (June, 2013). The relationship between student’s ability to model objects from assembly drawing information and spatial visualization ability as measured by the PSVT:R and MCT. Paper presented at the 2013 American Society for Engineering Education Annual Conference & Exposition, Atlanta, GA.

Branoff, T., & Dobelis, M. (June, 2014). Relationship between students’ spatial visualization ability and their ability to create 3D constraint-based models from various types of drawings. Paper presented at the 2014 American Society for Engineering Education Annual Conference & Exposition, Indianapolis, IN.


Webb, N. L. (2002). Depth-of-knowledge levels for four content areas. Retrieved from [http://facstaff.wcer.wisc.edu/normw/All%20content%20areas%20%20DOK%20levels%](http://facstaff.wcer.wisc.edu/normw/All%20content%20areas%20%20DOK%20levels%)


APPENDIX A: IRB EXEMPT STATUS APPROVAL

9/22/2016
North Carolina State University Mail - Clark - 9276 - IRB Protocol assigned Exempt status

Kevin Sutton <kgsutton@ncsu.edu>

Clark - 9276 - IRB Protocol assigned Exempt status

IRB Administrative Office <pins_notifications@ncsu.edu>
Reply-To: debra_paxton@ncsu.edu
To: kgsutton@ncsu.edu

Wed, Sep 21, 2016 at 5:02 PM

Date: September 21, 2016
IRB Protocol 9276 has been assigned Exempt status
Title: Performance Assessment in Introductory Post-Secondary Engineering Design Graphics Courses
PI: Clark, Aaron

The research proposal named above has received administrative review and has been approved as exempt from the policy as outlined in the Code of Federal Regulations (Exemption: 46.101. Exempt b.2). Provided that the only participation of the subjects is as described in the proposal narrative, this project is exempt from further review. This approval does not expire, but any changes must be approved by the IRB prior to implementation.

1. This committee complies with requirements found in Title 45 part 46 of The Code of Federal Regulations. For NCSU projects, the Assurance Number is: FWA00003429.
2. Any changes to the protocol and supporting documents must be submitted and approved by the IRB prior to implementation.
3. If any unanticipated problems or adverse events occur, they must be reported to the IRB office within 5 business days by completing and submitting the unanticipated problem form on the IRB website: http://research.ncsu.edu/sparcs/compliance/irb/irb_form.pdf.

Please let us know if you have any questions.

Sincerely,

Deb Paxton
919.515.4514

IRB Administrator
dapaxton@ncsu.edu
NC State IRB Office

Jennie Oststein
919.515.8754
IRB Coordinator
irb-coordinator@ncsu.edu
NC State IRB Office

https://mail.google.com/mail/u/0/?ui=2&ik=115435c374&view=pt&search=inbox&msg=1574e5e5d017b731&simh=1574e5e5d017b731
APPENDIX B: ENGINEERING DESIGN GRAPHICS PERFORMANCE ASSESSMENT SURVEY EMAIL

From: Aaron Clark

Subject: Survey of Performance Assessment Practices in Introductory Engineering Graphics Courses

I am writing on behalf of Kevin Sutton to request your participation in a brief survey of performance assessment practices in introductory engineering graphics courses. Performance assessment provides data for student achievement at an individual and class level. He is interested in collecting demographic information to better portray the current status of performance assessment in introductory courses.

The survey is brief and will only take 10 minutes to complete. Please click on the provided link or copy paste the web address into your browser to access the survey.

Your participation in the survey is entirely voluntary, and all of your responses will be kept confidential. No personally identifiable information will be associated with your responses to any reports of this data. The NC State University Institutional Review Board has approved this survey. Should you have any comments or questions, please feel free to contact Kevin Sutton at kgsutton@ncsu.edu or (919) 515-0221.

Follow this link to the Survey:
$\{l:/SurveyLink?id=Take the survey\}$

Or copy and paste the URL below into your internet browser:
$\{l:/SurveyURL\}$

Follow the link to opt out of future emails:
$\{l:/OptOutLink?id=Click here to unsubscribe\}$

Thank you in advance for your time and cooperation. Your participation is important to this investigation and the improvement of performance assessment practice in our field.

Professionally yours,

Dr. Aaron C. Clark, DTE
APPENDIX C: ENGINEERING DESIGN GRAPHICS PERFORMANCE ASSESSMENT SURVEY

North Carolina State University
INFORMED CONSENT FORM for RESEARCH
Title of Study: Performance Assessment in Introductory post-secondary Engineering Design Graphics Courses
Principal Investigator: Kevin G. Sutton
Faculty Sponsor (if applicable): Aaron C. Clark

What are some general things you should know about research studies?
You are being asked to take part in a research study. Your participation in this study is voluntary. You have the right to be a part of this study, to choose not to participate or to stop participating at any time without penalty. The purpose of research studies is to gain a better understanding of a certain topic or issue.

You are not guaranteed any personal benefits from being in a study. Research studies also may pose risks to those that participate. In this consent form you will find specific details about the research in which you are being asked to participate. If you do not understand something in this form it is your right to ask the researcher for clarification or more information. A copy of this consent form will be provided to you. If at any time you have questions about your participation, do not hesitate to contact the researcher(s) named above.

What is the purpose of this study?
The purpose of the study is to investigate the current performance assessment practices and the validity of a performance assessment instrument utilized in post-secondary engineering design graphics courses.

What will happen if you take part in the study?
If you agree to participate in this study, you will complete a questionnaire about performance assessment practices at your institution. Some participants will be contacted after the survey and requested to participate in a follow-up study. Participants who agree to participate in the follow-up study will be asked to use an assessment instrument to grade three example student projects.

Risks and Benefits
I do not anticipate any risks to you participating in this study other than those encountered in day-to-day life. There will be no direct benefit for participation in the study. However, your participation will benefit the study of performance assessment in engineering graphics courses.

Confidentiality
The information in the study records will be kept confidential to the full extent allowed by law. Data will be stored securely in password-protected programs. No reference will be made in oral or written reports which could link you to the study. Your email address will only be requested in the survey if you consent to be contacted for participation in a follow-up study. Your provided email address will not be distributed or used as any part of the data analysis.

Compensation
You will not receive anything for participating in this study.

What if you have questions about this study?
If you have questions at any time about the study itself or the procedures implemented in this study, you may contact the researcher, Kevin Sutton, at kgssutton@ncsu.edu, or 919.513.0221.

What if you have questions about your rights as a research participant?
If you feel you have not been treated according to the descriptions in this form, or your rights as a participant in research have been violated during the course of this project, you may contact Deb Paxton, Regulatory Compliance Administrator at dapaxton@ncsu.edu or by phone at 1-919-515-4514.
“I have read and understand the above information. I have received a copy of this form. I agree to participate in this study with the understanding that I may choose not to participate or to stop participating at any time without penalty or loss of benefits to which I am otherwise entitled.”

☐ Agree
☐ Disagree

In what type of university do you teach?
☐ Public
☐ Private
☐ For-profit

What is your academic rank?
☐ Instructor/Lecturer
☐ Teaching Assistant Professor
☐ Teaching Associate Professor
☐ Assistant Professor
☐ Associate Professor
☐ Professor
☐ Professor Emeritus
☐ Other ____________________

What is your current tenure status?
☐ Tenured
☐ Tenure-Track
☐ Non Tenure-Track

Do you currently teach undergraduate engineering graphics courses?
☐ Yes
☐ No

How many years have you taught engineering graphics courses?

How many undergraduate engineering graphics courses do you teach per semester?

Does your institution currently offer an introduction Engineering Design Graphics course?
☐ Yes
☐ No

What is the name of the introduction engineering graphics course?
How many sections of your introduction engineering graphics course are offered each academic year?

What is the average number of students enrolled in each section of the introduction engineering graphics course?

How many instructors teach your introduction engineering graphics course during each academic year?

Are students required to use a graphics software as a part of the introductory engineering graphics course?
- Yes
- No

What is the name and version (year) of the primary software used for instruction?

Select the student outcomes that align closest with your introduction engineering graphics course.
- Ability to create 3-D solid computer models
- Ability to sketch engineering objects in freehand mode
- Ability to visualize 3-D solid computer models
- Ability to create dimensions
- Ability to generate engineering drawings from computer models
- Ability to create 2-D computer geometry
- Ability to create section views
- Ability to perform design projects
- Ability to analyze 3-D computer models
- Knowledge of manufacturing and rapid prototyping methods
- Ability to create presentation graphics
- Ability to solve traditional descriptive geometry problems
- Ability to create geometric construction with hand tools
- Other ____________________

Do the students complete any visualization test(s) during the introduction engineering graphics course?
- Yes
- No
- Sometimes

What is/are the name(s) of the visualization instrument(s) used?
What percentage of the student's course grade is determined by assignments requiring students to demonstrate technical ability?
______ Click on slider bar and drag it to the nearest percentage

What type of student work is assessed in your introductory course? (Please select all that apply)
- Technical sketches
- Computer generated assembly models
- Computer generated engineering drawings
- Computer generated 3D models
- Hand made physical models
- Digitally fabricated physical models (ex. 3D printed model)
- Other ____________________

How is each type of work assessed?

In what format are student assignments submitted? (Select all that apply)

Are rubrics utilized for assessing student work?
- Yes
- No

What types of rubrics are utilized? (Select all that apply)
- Specific (made specifically for a certain assignment)
- General (can be used for multiple assignments)
- I am not sure

Are the rubrics analytical or holistic?
- Analytical (guidelines for each level of performance are provided for each criterion individually)
- Holistic (guidelines for each level of performance are the same for all criterion)
- I am not sure

Are raters provided with training or examples for using assessment rubrics?
- Yes
- No
- Sometimes
Are students provided written feedback on their work?
☐ Yes
☐ No
☐ Sometimes

Are you willing to participate in a brief follow-up study?
☐ Yes
☐ No

Please provide your e-mail address for follow-up. (Your email address will be kept confidential used for follow-up contact only. It will not be compared to your responses for analysis)
APPENDIX D: REVERSE ENGINEERING DESIGN PROJECT AND ASSESSMENT INSTRUMENT

REVERSE ENGINEERING PROJECT REQUIREMENTS

Rationale
The concepts covered during the semester are designed to give students the skills that enable them to create an accurate description of a moderately complex design. The term project is intended to provide a culminating experience whereby students have the opportunity to exercise the knowledge, skills, and judgment developed in the course.

Project Requirements

Part 1 - Re-creating an existing design
For the final project, choose a moderately complex assembly of 3-8 parts, and create basic documentation that could be read, interpreted, and understood by manufacturing personnel.
- Become familiar with the parts and assembly by disassembling it in order to gain access to all parts measurements.
- Use calipers to measure and document all of the dimensions required to fully define the parts. Think about the standards for your dimensions (units, precision, and dimensions between mating parts).
- Create 3D solid models, assembly, and CAD documentation via Solidworks as outlined below.

Part 2 - Modifying the existing design
Modify the existing design by modifying/improving an existing part or creating a new part that improves the existing design functionality. Effort should be given so that the new design is purposeful.
- Create a technical sketch of your modified or new part as outlined below.
- Add modified/new part to your assembly.

Project Components (All sketches and CAD drawings should be turned in with a border and completed title block)
This project will consist of five categories: Technical Sketch, 3D Solid Models, CAD Drawings, Assembly, and Project Portfolio.

1) Technical Sketch of the modified/new part (25%) - Create a multi-view sketch of the modified/new part. The multi-view sketch should demonstrate proper multi-view layout, conventional practices, and correct dimensioning technique. A section or auxiliary view may be required to properly describe the part. Revise sketches as needed for a high quality accurate representation.

Evaluation
✓ Accuracy of Views and Annotations - 8 points
✓ View Alignment - 5 points
✓ Unibody - 4 points
✓ Line Quality - 4 points
✓ Appropriate part name - 5 points

Example of a Technical Sketch:

![Example Sketch]
2) 3D Solid Models (30%) - Create 3D Solid Models of each part including the modified/new part. Use the best feature available to create parts most efficiently in Solidworks, save using appropriate part numbers and materials, and submit all electronic files to moodle as directed.

   Evaluation
   ✓ Model Orientation - 4 points
   ✓ Fully Defined Sketches - 7 points
   ✓ Accuracy - 7 points
   ✓ Model Procedure - 7 points
   ✓ Appropriate Part Name - 5 points

3) CAD Drawing of the most complex part (20%) - Create a detail drawing of the most complex part. This part may not be the same part used for the technical sketch. Create a detail part drawing using Solidworks and include all information necessary to manufacture the part. This should include: shape description (orthographic views, sectional views, and auxiliary views); size description (dimensions); and notes, material, etc.

   Evaluation
   ✓ View Alignment - 5 points
   ✓ Linetype (visible lines, hidden lines, center lines, tangent edges, etc.) - 6 points
   ✓ Annotations - 8 points

4) Assembly (20%) - Use Solidworks to create an assembly of all parts represented in their correct position including the modified/new part. In addition, submit a CAD drawing with the best pictorial view of the assembly, a bill of materials with balloons that identify all parts. Submit all electronic files to moodle as directed.

   Evaluation
   ✓ Correct mates/relations between parts - 10 points
   ✓ Appropriate orientation of assembly - 3 points
   ✓ Materials (represents real parts materials) - 4 points
   ✓ CAD Drawing of Assembly with Bill of Materials with Balloons - (best pictorial view is chosen) - 3 points

5) Project Portfolio (5%) - Submit the final project to the instructor in the following format. The components of the report should be organized in the order that they appear below.

Hardcopy Deliverables
1. Title Page (title, name, section, semester, and date)
2. Technical Sketch of the modified/new part
3. CAD Drawing of the most complex part (different than part used for the technical sketch)
4. CAD Drawing of Assembly with Bill of Materials with Balloons - (best pictorial view is chosen)
5. Evaluation Sheet

Digital Deliverable
6. Compressed Folder - Upload to Moodle submission link (All Solidworks files including 3D Solid Models, Assembly, and Drawing)

Note: After uploading compressed folder to Moodle, verify that it is not corrupt by downloading and opening assembly.

Reference Material
1. Information on Solid Modeling - Solidworks Tutorials.
2. Three Dimensional Modeling - Chapter 4 in Bertoline/Wiebe.
3. Shape Description/Drawing Layout - Chapter 5 in Bertoline/Wiebe.
4. Auxiliary Views - Chapter 6 in Bertoline/Wiebe.
5. Sectional Views - Chapter 8 in Bertoline/Wiebe.
7. Fastening Devices and Methods - Chapter 10 in Bertoline/Wiebe.
8. Working Drawings - Chapter 11 in Bertoline/Wiebe
**FINAL PROJECT**  
**EVALUATION SHEET**

**Student Name:**

1) **Technical Sketch of the modified/new part**
   - Accuracy of Views and Annotations - 8 points
   - View Alignment - 4 points
   - Linetype - 4 points
   - Line Quality - 4 points
   - Appropriate part name - 5 points

   **Comments:** ____________________________________________________________________  
   **25 points**

2) **3D Solid Models of all parts**
   - Model Orientation - 4 points
   - Fully Defined Sketches - 7 points
   - Accuracy - 7 points
   - Model Procedure - 7 points
   - Appropriate Part Name - 5 points

   **Comments:** ____________________________________________________________________  
   **30 points**

3) **CAD Drawing of the most complex part**
   - View Alignment - 6 points
   - Linetype (visible lines, hidden lines, center lines, tangent edges, etc.) - 6 points
   - Annotations - 8 points

   **Comments:** ____________________________________________________________________  
   **20 points**

4) **Assembly**
   - Correct mates/relations between parts - 10 points
   - Appropriate orientation of assembly - 3 points
   - Materials (represents real parts materials) - 4 points
   - CAD Drawing of Assembly with Bill of Materials with Balloons - (best pictorial view is chosen) - 3 points

   **Comments:** ____________________________________________________________________  
   **20 points**

5) **Project Portfolio**
   **5 points**

   **Hardcopy Deliverables**
   1. Title Page (title, name, section, semester, and date)
   2. Technical Sketch of the modified/new part
   3. CAD Drawing of the most complex part (different than part used for the technical sketch)
   4. CAD Drawing of Assembly with Bill of Materials with Balloons - (best pictorial view is chosen)
   5. Evaluation Sheet

   **Digital Deliverable**
   6. Compressed Folder - Upload to Moodle submission link (All Solidworks files including 3D Solid Models, Assembly, and Drawing)

   **Note:** After uploading compressed folder to Moodle, verify that it is not corrupt by downloading and opening assembly.

   **Comments:** ____________________________________________________________________

   **TOTAL:** __________
APPENDIX E: ENGINEERING GRAPHICS PERFORMANCE ASSESSMENT FOLLOW-UP EMAIL

From: Kevin Sutton

Subject: Engineering Design Graphics Performance Assessment Follow-Up Study

I would like to thank you for completing the survey and your willingness to assist in a study to investigate the use of a current assessment instrument. You were selected for this study based on your experience and use of SolidWorks 2015-2016.

I am requesting your assistance in assessing three example reverse engineering project portfolios. Your participation in this study is completely voluntary, and all of your responses will be kept confidential. No personally identifiable information will be associated with your responses to any reports of these data. The NC State University Institutional Review Board has approved this survey. Should you have any comments or questions, please feel free to contact me at kgsutton@ncsu.edu or (919) 515-0221.

Instructions:

1. Please begin by reviewing the assignment and assessment instrument attached to this email as a PDF.
2. Three example portfolios are attached to this email in a zipped folder. These portfolios are for research purposes only and not intended for dissemination.
   a. Each project folder contains a PDF copy of the example portfolio containing a cover page, technical sketch, engineering drawings, and assembly drawings along with the digital files for each example's parts and assemblies in their native file format.
3. Please take a look at each of the three example projects before you begin rating.
   a. The example projects folder will need to be downloaded and unzipped.
   b. Opening the files through the SolidWorks program is recommended to reduce file opening errors.
4. After reviewing the project materials by yourself, please open the Qualtrics link below and input your scores and feedback for each example project as if you were assessing students in your class on their submissions.
5. Please do not discuss your ratings with others to avoid influencing perceptions.

Follow this link to the Survey:
$\{l://SurveyLink/?d=Take the survey\}$
Or copy and paste the URL below into your internet browser:
$\{l://SurveyURL\}$
Follow the link to opt out of future emails:
$\text{OptOutLink}d=\text{Click here to unsubscribe}$

Thank you for your time and concern for investigating assessment practices in the field to improve teaching and learning in engineering graphics courses.

Sincerely,

Kevin Sutton
APPENDIX F: ENGINEERING GRAPHICS PERFORMANCE ASSESSMENT
FOLLOW-UP REMINDER EMAIL

To: Unfinished Participants

From: Kevin Sutton

I would like to thank you for completing the performance assessment survey and your willingness to assist in this follow-up study to investigate the use of a current assessment instrument. I hope to catch you before the end of the semester rush. This study is the final part of my dissertation research and critical to the completion of my degree. You were selected for this study based on your experience and use of SolidWorks™. The specific criteria make the sample for this study small and your timely participation is appreciated beyond what I can express with words.

For this follow-up study, I need your help assessing three example reverse engineering project portfolios. Your participation in this study is completely voluntary, and all of your responses will be kept confidential. No personally identifiable information will be associated with your responses to any reports of these data. The NC State University Institutional Review Board has approved this study. Should you have any comments or questions, please feel free to contact me at kgsutton@ncsu.edu or (919) 515-0221.

Instructions:

1. Please begin by reviewing the assignment and assessment instrument attached to this email as a PDF.
2. Three example portfolios are attached to this email in zipped folders. These portfolios are for research purposes only and not intended for dissemination.
3. Each project folder contained a PDF copy of the example portfolio containing a cover page, technical sketch, engineering drawings, and assembly drawings along with the digital files for each example’s parts and assemblies in their native file format.
4. Please take a look at each of the three example projects before you begin rating.
   - The example projects folder will need to be downloaded and unzipped.
   - Opening the files through the SolidWorks program is recommended to reduce file opening errors.
5. After reviewing the project materials by yourself, please open the Qualtrics link below and input your scores and feedback for each example project as if you were assessing students in your class on their submissions.
6. Please do not discuss your ratings with others to avoid influencing perceptions.

Project description and grading sheet
Alarm clock project
Knife project
Microphone project
Follow **this link to the Survey:**
[Take the survey]

Or copy and paste the URL below into your internet browser:
[https://ncsu.qualtrics.com/SE?Q_DL=6roeydSEh3Ncymp_8jmqnKveHBukQ17_MLRP_3forKE1AP80L9l3&Q_CHL=email](https://ncsu.qualtrics.com/SE?Q_DL=6roeydSEh3Ncymp_8jmqnKveHBukQ17_MLRP_3forKE1AP80L9l3&Q_CHL=email)

Thank you in advance for your time and efforts.

Sincerely,
-Kevin Sutton

Follow the link to opt out of future emails:
[Click here to unsubscribe](#)
APPENDIX G: ENGINEERING GRAPHICS PERFORMANCE ASSESSMENT

9/7/2016  Quatrics Survey Software

Alarm Clock

Project Name: Alarm Clock

Type a number in the box to give a score for each of the criteria. Please use a number between 0 and the maximum number of points designated for each criterion.

Technical Sketch of the modified/new part

Accuracy of Views and Annotations - 8 points  0 point:
View Alignment - 4 Points  0 point:
Line Type - 4 points  0 point:
Line Quality - 4 points  0 point:
Appropriate Part Name - 5 points  0 point:
Total  0 point:

Comments:

3D Solid Models of all parts

Model Orientation - 4 points  0 point:
Fully Defined Sketches - 7 points  0 point:
Accuracy - 7 points  0 point:
Model Procedure - 7 points  0 point:
Appropriate Part Name - 5 points  0 point:

Total

Comments:

**CAD Drawing of the most complex part**

View Alignment - 6 points  
0 point

Line type (visible lines, hidden lines, center lines, tangent edges, etc.) - 6 points 
0 point

Annotations - 8 points 
0 point

Total 
0 point

Comments:

**Assembly**

Correct mates/relations between parts - 10 points 
0 point

Appropriate orientation of assembly - 3 points 
0 point

Materials (represents real part materials) - 4 points 
0 point

CAD Drawing of Assembly with Bill of Materials with Balloons (best pictorial view is chosen) - 3 points 
0 point

Total 
0 point
Project Portfolio

Hardcopy Deliverables
1. Title Page (title, name, section, semester, and date)
2. Technical Sketch of the modified/new part
3. CAD Drawing of the most complex part (different than part used for the technical sketch)
4. CAD Drawing of Assembly with Bill of Materials with Balloons - (best pictorial view is chosen)

Digital Deliverable
5. Compressed Folder - Upload to submission link (All Solidworks files including 3D Solid Models, Assembly, and Drawing)

Portfolio - 5 points
Total

Comments:

Pocket Knife
Project Name: Pocket Knife

Type a number in the box to give a score for each of the criteria. Please use a number between 0 and the maximum number of points designated for each criterion.

Technical Sketch of the modified/new part

Accuracy of Views and Annotations - 8 points
View Alignment - 4 Points
Line Type - 4 points
Line Quality - 4 points
Appropriate Part Name - 5 points
Total

Comments:

3D Solid Models of all parts

Model Orientation - 4 points
Fully Defined Sketches - 7 points
Accuracy - 7 points
Model Procedure - 7 points
Appropriate Part Name - 5 points
Total
Comments:

CAD Drawing of the most complex part

View Alignment - 6 points
Line type (visible lines, hidden lines, center lines, tangent edges, etc.) - 6 points
Annotations - 8 points
Total

Comments:

Assembly

Correct mates/relations between parts - 10 points
Appropriate orientation of assembly - 3 points
Materials (represents real part materials) - 4 points
CAD Drawing of Assembly with Bill of Materials with Balloons (best pictorial view is chosen) - 3 points
Total

Comments:
Project Portfolio

Hardcopy Deliverables
1. Title Page (title, name, section, semester, and date)
2. Technical Sketch of the modified/new part
3. CAD Drawing of the most complex part (different than part used for the technical sketch)
4. CAD Drawing of Assembly with Bill of Materials with Balloons - (best pictorial view is chosen)

Digital Deliverable
6. Compressed Folder - Upload to submission link (All Solidworks files including 3D Solid Models, Assembly, and Drawing)

Portfolio - 5 points
Total

Comments:

Microphone

Project Name: Microphone
Type a number in the box to give a score for each of the criteria. Please use a number between 0 and the maximum number of points designated for each criterion.

**Technical Sketch of the modified/new part**

<table>
<thead>
<tr>
<th>Category</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy of Views and Annotations</td>
<td>8</td>
</tr>
<tr>
<td>View Alignment</td>
<td>4</td>
</tr>
<tr>
<td>Line Type</td>
<td>4</td>
</tr>
<tr>
<td>Line Quality</td>
<td>4</td>
</tr>
<tr>
<td>Appropriate Part Name</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>0</td>
</tr>
</tbody>
</table>

Comments:

---

**3D Solid Models of all parts**

<table>
<thead>
<tr>
<th>Category</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model Orientation</td>
<td>4</td>
</tr>
<tr>
<td>Fully Defined Sketches</td>
<td>7</td>
</tr>
<tr>
<td>Accuracy</td>
<td>7</td>
</tr>
<tr>
<td>Model Procedure</td>
<td>7</td>
</tr>
<tr>
<td>Appropriate Part Name</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>0</td>
</tr>
</tbody>
</table>

Comments:
CAD Drawing of the most complex part

- View Alignment - 6 points 0 point:
- Line type (visible lines, hidden lines, center lines, tangent edges, etc.) - 6 points 0 point:
- Annotations - 8 points 0 point:
- Total 0 point:

Comments:

Assembly

- Correct mates/relations between parts - 10 points 0 point:
- Appropriate orientation of assembly - 3 points 0 point:
- Materials (represents real part materials) - 4 points 0 point:
- CAD Drawing of Assembly with Bill of Materials with Balloons (best pictorial view is chosen) - 3 points 0 point:
- Total 0 point:

Comments:
Project Portfolio

Hardcopy Deliverables
1. Title Page (title, name, section, semester, and date)
2. Technical Sketch of the modified/new part
3. CAD Drawing of the most complex part (different than part used for the technical sketch)
4. CAD Drawing of Assembly with Bill of Materials with Balloons - (best pictorial view is chosen)

Digital Deliverable
6. Compressed Folder - Upload to submission link (All Solidworks files including 3D Solid Models, Assembly, and Drawing)

Portfolio - 5 points
Total

0 point:

Comments:

Powered by Qualtrics
APPENDIX H: ALARM CLOCK PROJECT PORTFOLIO

For Research Only - Do Not Disseminate

Alarm Clock

Hidden for confidentiality

04/19/2016
Rationale

Most of the objects used in our daily lives are designed to perform a single task. For this reason, people end up buying many different objects for different purposes. The continuous research process toward the development of objects that bring comfort to everyday life, is essential for any company to succeed in the modern world. The thought behind this project was to redesign an alarm clock in order to add a simple space dedicated as a pen holder. The goal was to have a more functional object - two in one - that uses space more efficiently. After completely disassembling the product and becoming familiar with all its components, it was concluded that the best and more convenient way to achieve this goal was to modify the top case of the existing alarm clock, integrating a small space on the side of it to hold three pens. It would be very useful having an alarm clock and a pen holder combined in one item, providing the comfort of having a more organized nightstand with everything needed in the right place.

Modified Top Case

Original Top Case
### Table of Parts

<table>
<thead>
<tr>
<th>ITEM NO.</th>
<th>PART NAME</th>
<th>MATERIAL</th>
<th>QTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alarm Base</td>
<td>ABS Plastic</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Battery lid</td>
<td>ABS Plastic</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Button</td>
<td>Silicon Rubber</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>Circuit Board</td>
<td>E-Glass Fiber</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Back Screen Protector</td>
<td>Alum. 1060 Alloy</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Screen</td>
<td>Glass</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>Dual Glass Screen</td>
<td>Glass</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>Front Screen Protector</td>
<td>ABS Plastic</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>Screen Base</td>
<td>ABS Plastic</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Screen Support</td>
<td>ABS Plastic</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>Top case new</td>
<td>Alum. 1060 Alloy</td>
<td>1</td>
</tr>
</tbody>
</table>
APPENDIX I: KNIFE PROJECT PORTFOLIO

For Research Only - Do Not Disseminate

Pocketknife

Hidden for confidentiality

4/25/16
For Research Only - Do Not Disseminate
### For Research Only - Do Not Disseminate

<table>
<thead>
<tr>
<th>ITEM NO.</th>
<th>PART NAME</th>
<th>MATERIAL</th>
<th>QTY.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>outer right plate</td>
<td>NYLON 101</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>inner right plate</td>
<td>STAINLESS STEEL(FERRITIC)</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>plate connector</td>
<td>CHROME STAINLESS STEEL</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>inner left plate</td>
<td>STAINLESS STEEL(FERRITIC)</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>outer left plate</td>
<td>NYLON 101</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>plate connector screw</td>
<td>CHROME STAINLESS STEEL</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>washer</td>
<td>CHROME STAINLESS STEEL</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>housing screw</td>
<td>CHROME STAINLESS STEEL</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>housing</td>
<td>CHROME STAINLESS STEEL</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>o-ring</td>
<td>SILICONE RUBBER</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>clip</td>
<td>CHROME STAINLESS STEEL</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>clip screw</td>
<td>CHROME STAINLESS STEEL</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>blade</td>
<td>CHROME STAINLESS STEEL</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td>knub</td>
<td>CHROME STAINLESS STEEL</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>knub screw</td>
<td>CHROME STAINLESS STEEL</td>
<td>1</td>
</tr>
<tr>
<td>ITEM NO.</td>
<td>PART NUMBER</td>
<td>DESCRIPTION</td>
<td>QTY.</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>----------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1</td>
<td></td>
<td>outer right plate</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>inner right plate</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>plate connector</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>inner left plate</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>outer left plate</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>plate connector screw</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>washer</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>housing screw</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>housing</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>o-ring</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>clip</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>clip screw</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>modified blade</td>
<td>1</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>knob</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>knob screw</td>
<td>1</td>
</tr>
</tbody>
</table>

Modified Knife Assembly

See B.O.M.: 1-1

4/25/16

Hidden for confidentiality
For Research Only - Do Not Disseminate

Hidden for confidentiality

SI

Hidden for confidentiality

INNER RIGHT PLATE

STAINLESS STEEL (FERRITIC)

Sheet 1

Dimensions and tolerances are in millimeters. Components are not to scale.

4/25/16

3:2

N/A

1
For Research Only - Do Not Disseminate

Hidden for confidentiality

Hidden for confidentiality

Thickness 1.66 with 4 fillet

PART 1

NYLON 101
For Research Only - Do Not Disseminate

Hidden for confidentiality

1.5

Ø 0.16

Ø 4.16

R 1.17

0.75

5.25

4.05

Ø 2.85

Hidden for confidentiality

SI

Hidden for confidentiality

 voting specifice enfor e e en for confidenility

4/25/16

10:1

N/A

N/A

1

CHROME STAINLESS STEEL
APPENDIX J: MICROPHONE PROJECT PORTFOLIO

For Research Only - Do Not Disseminate

Blue Yeti Microphone

Hidden for confidentiality

April 21, 2016
For Research Only - Do Not Disseminate

Hidden for confidentiality

Dimensions are in millimeters, unless otherwise noted.

Material: Aluminum

PCB Frame

PART #1

4/21/2016

Rev A

1

Hidden for confidentiality
For Research Only - Do Not Disseminate

<table>
<thead>
<tr>
<th>ITEM NO.</th>
<th>PART NUMBER</th>
<th>DESCRIPTION</th>
<th>QTY.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PCB_Frame</td>
<td>Internal Frame</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>PCB_1</td>
<td>Upper PCB</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>PCB_2</td>
<td>Lower PCB</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Screw_Short</td>
<td>PCB Screw</td>
<td>8</td>
</tr>
<tr>
<td>5</td>
<td>Mic_Head</td>
<td>Microphone Container</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Screw_Long</td>
<td>Casing Screw</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>Mic_Body</td>
<td>Aluminum Casing</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>Power_Button</td>
<td>On/Off Switch</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>Volume_Knob</td>
<td>Volume Dial</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>Dial_Long</td>
<td>Pickup Dial</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>Dial_Short</td>
<td>Gain Dial</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>Back_Stop</td>
<td>Braces connectors</td>
<td>1</td>
</tr>
</tbody>
</table>
## APPENDIX K: RAW SCORES FOR ALARM PROJECT

<table>
<thead>
<tr>
<th></th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
<th>Rater 4</th>
<th>Rater 5</th>
<th>Rater 6</th>
<th>Rater 7</th>
<th>Rater 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asketch1</td>
<td>7</td>
<td>8</td>
<td>8</td>
<td>5</td>
<td>8</td>
<td>8</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Asketch2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Asketch3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Asketch4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Asketch5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Amodel1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Amodel2</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Amodel3</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Amodel4</td>
<td>7</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>7</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Amodel5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Adrawing1</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Adrawing2</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Adrawing3</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>6</td>
<td>7</td>
<td>3</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Aassembly1</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>Aassembly2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Aassembly3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Aassembly4</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Aportfolio</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>
APPENDIX L: RAW SCORES FOR KNIFE PROJECT

<table>
<thead>
<tr>
<th></th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
<th>Rater 4</th>
<th>Rater 5</th>
<th>Rater 6</th>
<th>Rater 7</th>
<th>Rater 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ksketch1</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Ksketch2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ksketch3</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Ksketch4</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Ksketch5</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>Kmodel1</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Kmodel2</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Kmodel3</td>
<td>5</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>7</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td>Kmodel4</td>
<td>5</td>
<td>7</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Kmodel5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Kdrawing1</td>
<td>6</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>6</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Kdrawing2</td>
<td>6</td>
<td>5</td>
<td>2</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Kdrawing3</td>
<td>6</td>
<td>5</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Kassembly1</td>
<td>10</td>
<td>10</td>
<td>8</td>
<td>7</td>
<td>10</td>
<td>8</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Kassembly2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Kassembly3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Kassembly4</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Kportfolio</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
</tbody>
</table>
### APPENDIX M: RAW SCORES FOR MICROPHONE PROJECT

<table>
<thead>
<tr>
<th></th>
<th>Rater 1</th>
<th>Rater 2</th>
<th>Rater 3</th>
<th>Rater 4</th>
<th>Rater 5</th>
<th>Rater 6</th>
<th>Rater 7</th>
<th>Rater 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Msketch1</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>8</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Msketch2</td>
<td>3</td>
<td>0</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Msketch3</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Msketch4</td>
<td>2</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Msketch5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Mmodel1</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Mmodel2</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Mmodel3</td>
<td>7</td>
<td>3</td>
<td>7</td>
<td>7</td>
<td>5</td>
<td>7</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Mmodel4</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>7</td>
<td>7</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Mmodel5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Mdrawing1</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Mdrawing2</td>
<td>6</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Mdrawing3</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Massembly1</td>
<td>8</td>
<td>3</td>
<td>8</td>
<td>4</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>8</td>
</tr>
<tr>
<td>Massembly2</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Massembly3</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Massembly4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Mportfolio</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>1.5</td>
<td>4</td>
</tr>
</tbody>
</table>