PEER REVIEW IN DESIGN: UNDERSTANDING THE IMPACT OF COLLABORATION ON THE REVIEW PROCESS AND STUDENT PERCEPTION

by

Mahender Arjun Mandala

Bachelor of Engineering (Mechanical), Osmania University, 2008

Master of Science (Mechanical), University of Kansas, 2011

Submitted to the Graduate Faculty of

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2016

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCE

This dissertation was presented

by

Mahender Arjun Mandala

It was defended on

July 22, 2016

and approved by

Mary Goldberg, Assistant Professor, Rehabilitation Science and Technology

Christian Schunn, Professor, Psychology

Steven Dow, Assistant Professor, Cognitive Science, UC San Diego

Dissertation Advisor: Jonathan Pearlman, Assistant Professor, Rehabilitation Science and Technology

PEER REVIEW IN DESIGN: UNDERSTANDING THE IMPACT OF COLLABORATION ON THE REVIEW PROCESS AND STUDENT PERCEPTION

Mahender Arjun Mandala, PhD

University of Pittsburgh, 2016

A cornerstone of design and design education is frequent situated feedback. With increasing class sizes, and shrinking financial and human resources, providing rich feedback to students becomes increasingly difficult. In the field of writing, web-based peer review—the process of utilizing equal status learners within a class to provide feedback to each other on their work using networked computing systems—has been shown to be a reliable and valid source of feedback in addition to improving student learning.

Designers communicate in myriad ways, using the many languages of design and combining visual and descriptive information. This complex discourse of design intent makes peer reviews by design students ambiguous and often not helpful to the receivers of this feedback. Furthermore, engaging students in the review process itself is often difficult. Teams can complement individual diversity and may assist novice designers collectively resolve complex task. However, teams often incur production losses and may be impacted by individual biases. In the current work, we look at utilizing a collaborative team of reviewers, working collectively and synchronously, in generating web based peer reviews in a sophomore engineering design class.

Students participated in a cross-over design, conducting peer reviews as *individuals* and *collaborative teams* in parallel sequences. Raters coded the feedback generated on the basis of their appropriateness and accuracy. Self-report surveys and passive observation of teams

conducting reviews captured student opinion on the process, its value, and the contrasting experience they had conducting team and individual reviews.

We found team reviews generated better quality feedback in comparison to individual reviews. Furthermore, students preferred conducting reviews in teams, finding the process ‘fun’ and engaging. We observed several learning benefits of using collaboration in reviewing including improved understanding of the assessment criteria, roles, expectations, and increased team reflection. These results provide insight into how to improve the review process for instructors and researchers, and forms a basis for future research work in this area.

With respect to facilitating peer review process in design based classrooms, we also present recommendations for creating effective review system design and implementation in classroom supported by research and practical experience.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Over the past five years, I have received tremendous support and encouragement from a great number of individuals. First, I would like to express sincere gratitude to my doctoral advisor, Dr. Jonathan Pearlman. His guidance has made this a thoughtful and rewarding journey. Second, I would like to thank my dissertation committee of Drs. Christian Schunn, Steven Dow, and Mary Goldberg for their support over the past two years, helping me move my research from conception to completion, inspiring me to explore new ideas, and encouraging me to keep pushing forward. None of this work could have been accomplished without the support of Drs. William Clark and Irene Mena, who so trustingly offered me their classrooms for experimentation. Equally important were the contributions by the student teams working with me on Peerval: Mayank Sardana and Sushma Ananda. I would like to thank Philip Brooks, for mentoring me as I attempted to interview over a hundred instructors in design education.

During my data collection and writing, Chandu Bolisetti and Sujith Rapolu spent countless hours listening to me talk about my research. Sara Munera and Jorge Candiotti provided needed encouragement. Krishna Samavedam helped me through the final few months as I struggled to juggle life and completing my dissertation. Lastly, thanks to the HERL family, who have contributed in some part or another in this project and have made my work a thoroughly enjoyable experience.

I would like to thank my parents, along with Dometi and Krishnan family, for their unconditional love and support. Thank you for doing what you do, every single day! Special thanks to my partner Shilpa Krishnan who enthusiastically guided me through the writing process and anchored my life in the most stressful of times. My friends, here in Pittsburgh and back home in India, thank you for making my stay away from home easier.

The work presented in Chapter 2.0 was supported National Science Foundation's Innovation Corps for Learning program (Grant #1547954).

DEDICATION

To my parents, Arjun and Indira Mandala, for their unwavering support, countless sacrifices, and encouragement over the years.

1.0 OVERVIEW

There is a renewed and sustained focus on improving engineering education reflected by the exhaustive work done by Accreditation Board for Engineering And Technology (ABET). The primary focus of the recommendation for engineering education includes concentrating on what students learn, and improving student achievement by providing them with real world experiences, rather than focusing on what material is being taught (Engineering Accreditation Commission, 2014). In this direction, engineering design is being increasingly recognized and utilized as a vehicle for change (Altman, Dym, Hurwitz, & Wesner, 2012; Dym, Agogino, Eris, Frey, & Leifer, 2005).

Yet, there exists a large gap between what is considered effective design education and current instructor practices in classes. Researchers and educators have been discussing the various issues in design education for several decades (Altman et al., 2012; Briggs, 2012; Schön, 1987). Dym et al. (2005), in their seminal review of 'Engineering Design Thinking, Teaching and Learning', lay out the issues and complexities surrounding design education. The outstanding issues in design education can be distilled to two major aspects: pedagogy that provides authentic design experience and sustainable assessment practices that mimic the creative and iterative nature of design. In this dissertation work, I focus on sustainable assessment within design education, specifically scaling and improving a central design activity—design critique. Design critiques bring peers and experts together in a design discourse, covering feedback provision, ideation, learning, and understanding.

With increasing inclusion of DBL classrooms in engineering curricula, and a steady rise in enrollment over the years (Yoder, 2014) outpacing resources necessary to engage in meaningful pedagogy, design instructors are often left dealing with managing effort-centric non-traditional classrooms. To understand the current state of affairs in design classrooms, I examine instructional and feedback provision strategies in classrooms within engineering and related fields. This interview-based qualitative study is described in Chapter 2.0 of this dissertation. Web-based peer reviews are increasingly being used across writing and computer science, and have been shown to be reliable, valid, and in many cases, beneficial to student learning (K. Cho & Schunn, 2003b; Gibbs & Simpson, 2004; N.-F. Liu & Carless, 2006; Topping, 1998) Implementing web-based peer reviews in design is challenging. Engineering design crosses multiple domains of knowledge and skills, which at any given instance an individual peer reviewer may not fully possess. Recent work in web-based peer reviews in design classrooms has shown that peers can generate open-ended feedback that is of low quality (C. Kulkarni et al., 2013; Farshid Marbouti, Cardella, & Diefes-Dux, 2014). This prior research motivated the work described in chapter 3.0 which focuses on peer review structuring method that includes using a collaborative team of reviewers working collectively on reviews, and examine its impact on the feedback quality. One outstanding issue in web-based peer review has been student engagement and participation in the process. In chapter 4.0 , student opinion of the peer review process are examined, focusing on the novel structuring methods used, to gain an understanding on what aspects improve student engagement and participation in the process. Finally, a culmination of experience from conducting peer reviews in various design classes and interviewing potential end-users, along with previous research work in the field, is put to use in developing newer computing systems that improve and enhance peer review facilitation in design based classes. Chapter 5.0 , puts forth recommendations for effective peer review systems, along with best practices in implementing them in a classroom.

Nomenclature used in this dissertation:

DBL: Design based learning

Formative assessment: Assessment focused on student learning; information is provided to a learner to scaffold and modify his or her thinking in order to improve learning.

Summative assessment: Assessment focused on program outcomes; student learning is evaluated through grades or marks comparing it to a standard or benchmark.

STEM: Science, Technology, Engineering, and Math.

Peer evaluation: Utilizing students to evaluate their peers on their performance.

Peer review: Utilizing students to provide feedback (summative or formative) to their peers on their work. Used interchangeably with peer feedback, peer critique, and peer assessment, in this dissertation.

2.0 A REPORT ON STATE OF ENGINEERING DESIGN EDUCATION

2.1 INTRODUCTION

Design is often considered a cornerstone of engineering education and profession, and yet, it remains a pedagogical challenge to teach students (Dym et al., 2005). Over several years, educators have experimented with novel pedagogical methodologies to enhance their efforts in training future designers, and in the process have encountered the intricacies of imparting effective design thinking to students (Dym et al., 2005). Design thinking—the cognitive processes that designers evoke during the design process—charts a convoluted, and often ambiguous, route alternating between the domains of knowledge containing facts and truths, and concepts, which do have such true value and are characterized by uncertainty (Hatchuel & Weil, 2003). In recent years, design education has shifted towards a problem based learning model, with real world (or mimicking real world) design projects that provide students with a firsthand design experience. Furthermore, the role of educators has also dramatically changed from a knowledge disseminator to that of a coach (Dym et al., 2005; Dym, Sheppard, & Wesner, 2001). Nonetheless, a key aspect of design education remains the use of situated and frequent feedback that scaffolds and nurtures student learning, and performance.

Feedback, especially focused on providing information to improve learning (also known as formative assessment), plays a crucial role in all-round development of students as independent learners. It impacts students beyond the classroom and into professional practice (Ferguson, 2011). Unfortunately, feedback provision is a resource intensive task that does not scale well with class size. Traditional design education relies on a more intimate cooperative learning environment, epitomized by the studio model (Schön, 1987). Feedback within the studio model is multifaceted and involves both the experts as well as peers. Design education, within the engineering context, has traditionally followed a much subdued role to its arts and architecture counterparts. And although, educators across engineering recognize the importance of feedback and the significance of studio model, the top-down institutional involvement needed to accomplish such a merger between traditional and non-traditional methods impedes any meaningful attempts. In recent years, educators have begun including peer review in some form—feedback on peer presentations or a few assignments, and more formally throughout the course—within engineering design classes. Apart from case studies on stellar examples of using peer reviews in literature, there are very few studies that have looked at practical real-world usage of such assessment activities and perceptions instructors and administrators have on them.

As can be seen, impactful design education requires a coordinated effort from educators, administrators, and the institutions that host them, in navigating the complexities of non-traditional pedagogy, human and capital resources, and outdated faculty incentives (Todd & Magleby, 2004). In this chapter, we examine the practices of engineering design educators across a sample of higher education institutions within the United States. Specifically, we focus on unearthing the strategies used by instructors in structuring their design classes, provisioning feedback to their students, and any barriers faced or strategies used in accomplishing their tasks. Furthermore, we present instructor perspectives on using peer reviews within classrooms.

2.2 RELATED WORK

There are several leading research topics currently being pursued by the design education research community including, cognitive models of learning, design pedagogy, processes, and activities, to name a few (Atman, Eris, McDonnell, Cardella, & Borgford-Parnell, 2014; Dym et al., 2005). Research on all these fronts reveal the intricacies involved in achieving impactful design training. Engineering design process is a complex cognitive and social process (Dym et al., 2005), characterized by ambiguity (Barley, Leonardi, & Bailey, 2012), iteration and negotiation (Schön, 1983), and shaped by the designer's own ongoing construction and application of knowledge (Atman et al., 2007; Dym et al., 2005). Furthermore, unlike traditional experiences in science and mathematics, where problems typically have a finite number of solutions that can be fact checked, design problems require a more divergent approach that explores the multiple solutions that coexist. Adding to this, professional designers often allude to the "fail fast and iterate often" mantra—counter to the expectations of most students, who are typically used to being rewarded for a unique and correct solution. Consequently, student designers may not necessarily possess the experience, technical breadth, and/or aptitude in navigating a multi-solution problem, signifying the role of design educators as coaches.

Effectively developing students' design thinking abilities requires creative classroom practices including, utilizing experiential practices such as problem-based learning, providing appropriate and timely feedback, and encouraging reflexive skills (Frascara & Noël, 2012; Schön, 1987). Elements of these practices make design education, a finance and human resource intensive activity, which, with increasing student enrollment (Yoder, 2014) hampers long-term sustainability.

2.2.1 The quest for best practices

In recent years, design educators, researchers, and practitioners, have come together to address challenges faced in design education and to share their thoughts on how to better educate future designers and engineers. One such confluence of likeminded individuals occurs biennially at the Mudd Design Workshops (MDW), hosted by the Center for Design Education of Harvey Mudd College. This workshop series (latest MDW IX, 2015) has generated several important discussion topics and commitments from its participating members to prioritize and improve design pedagogy. From the start, it seemed clear that design education requires a complete overhaul—including refocusing on *coaching* over *teaching* as a methodology and addressing *grading* and *learning* in new ways (Dym et al., 2001). Over following years, several of the concepts that surfaced in such conferences, have been implemented in classrooms and guidelines developed on what constitutes good design education (Dym et al., 2005).

Today, it is widely accepted that design education is most effective when using a project based approach, with hands-on experiences that enable students to use and sharpen their design thinking skills (Dym et al., 2005). Such an approach is critically—and necessarily—served by formative feedback that helps student designers identify gaps in their learning and performance and make amends to maximize them. Enabling students to reflect on their learning and experiences can boost the permanency of information and skills acquired, it also helps situate the feedback and keep the big picture in view (Briggs, 2012; Rogers, 2001). Yet, engaging and training students to critically reflect on their work or learning is in itself a pedagogical challenge. One potential way to increase student reflection is to use peer critiques, where students provide each other feedback typically using a rubric. Conducting peer reviews inevitably induces self-assessment within reviewers (D. Boud, Cohen, & Sampson, 1999) in addition to enhancing student learning (Nancy Falchikov & Blythman, 2001). Unfortunately, a majority of assessment and pedagogical methods currently employed typically involve instructor to student knowledge transfer, with formal peer-peer learning securing a distant second place.

The use of formal peer review or critiques are more common in design education in arts and architecture, where the culture and expectations of the field have been molded around the studio practice (Dannels, 2005; Dutton, 1987; Gray, 2013). In fact, attendants from early workshops at MDW advocated the use of studio style pedagogy in engineering design (Dym, Wesner, & Winner, 2003), recognizing its impact on multiple dimensions of student experience and learning (Dannels, 2005). Yet, studio based pedagogy in engineering design remains as distant as before. A primary issue with studio based class is scalability—requiring increasing human, capital, and temporal resources. Furthermore, design educators from fields other than arts and architecture seldom have the same cultural and social experiences of relying on peers for feedback and as a source of learning. This is a known limitation in the field with researchers and educators exploring ways to bridge the practices across design fields (Dym et al., 2001; C. Kulkarni et al., 2013; Reimer & Douglas, 2003; Tomayko, 1991, 1996).

Literature in the field is scattered with case studies of good design education practices (Dym, 2012; Reimer & Douglas, 2003; Vasana & Ritzhaupt, 2009), however, they seldom seem to scale well to other institutions, or even remain sustainable within the host institutions over a long term. With design being increasingly recognized as an important activity—one that should be pervasive across several courses including traditional math and sciences—it is pertinent that the research community examine the current practices in the field and adapt and prioritize their work to benefit the larger needs of design community and pedagogy. In this line, the current chapter examines a sample of design educators and their practices across engineering design and human-computer interaction fields within eastern United States. This chapter highlights the following:

RQ1. What are typical feedback provisioning strategies used by design educators in fields other than arts and architecture?

RQ2. What are the issues these individuals face in fulfilling their goals as design educators?

RQ3. What were the participant perceptions with regards to using peer-peer learning methods such as peer review of student work?

2.3 METHODS

2.3.1 Researcher role and study setting

This study emerged from data collected as part of National Science Foundation's I-Corps for learning program in summer of 2015. Over 100 interviews were conducted with stakeholders to evaluate the sustainability and scalability of authors' (MM, JP, and MG) web-based peer review tool in the course of the 7-week program.

After reviewing the data, a subsection were considered suitable to generate a report on the questions described above, and subsequently analyzed for this study (sample size of N=39). This selection was based on whether participants interviewed were instructors of design in a higher education field other than arts and architecture. The primary author (MM) led the NSF I-Corp team as an *Entrepreneurial Lead* (leading efforts to investigate the landscape surrounding the innovation) and conducted a majority of the data collection (N=37/39). Author MG and PB (acknowledgement) conducted the other two interviews included in the study. The interviews spanned several institutions across United States (largely on the east coast) covering numerous types: teaching- vs. research-focused institutions, private vs. public, large vs. small etc.

2.3.2 Interviews

Data in this study is sourced from qualitative semi-structured interviews. The goals of the interview included: a) examine how instructors' provide feedback within their design classes (are there notable strategies), b) what are the pains or highlights of the current feedback provision methods utilized, c) what is the instructor and administrator perception of peer reviews, and d) how do administrators play a role in facilitating design classes. These goals were based on the intent of the interview in understanding the key market segments for a peer review tool being developed for design based classrooms. Semi-structured interviews allowed us to extend our exploration of emerging view points and gain further insights into the participants' workflow and perceptions. Furthermore, probing questions were used to better understand participant responses. Interviews were conducted within the participants' own work setting either in-person or through video conference and lasted anywhere from 30 – 60 minutes.

The interviews took the form of a casual conversation and focused on participants' work. Participants were advised on the purpose of the interview, including the use of data to generate a report on strategies used within the classrooms. Data was collected in the form of handwritten notes, while some interviews were audio recorded with participant consent.

2.3.3 Participants

Participants were recruited using a snowball technique, and represented a range of design instructors (tenured, tenure-track, non-tenure track, adjunct, and teaching assistants), some with service experience as administrators (dean, chair, faculty facilitator etc.). Table 2-1 describes the demographics of the participants who were included in the study.

Table 2-1. Participant rank at institution. Affiliation indicates whether participants were employed by research or teaching focused institutions.

Rank	**N (affiliation)**
Tenured	21 (14 research, 7 teaching)
Tenure-track	7 (research)
Visiting	1 (research)
Adjunct	7 (6 research, 1 teaching)
Teaching Assistant or Lecturer	3 (research)
Total	39 (28 faculty, 11 admin and faculty)

2.3.4 Data analysis

The interviewer compiled the interview notes (transcribed when needed) no later than one day after the interviews. Data were coded using MAXQDA (software for qualitative data analysis, 1989-2016, VERBI Software – Consult – Sozialforschung GmbH, Berlin, Germany). Themes emerged through an inductive coding methodology and by constantly comparing data. An undergraduate researcher cleaned (acknowledgement, EC) and organized data for analyses. The primary author (MM) coded all the data presented in this chapter.

2.4 KEY FINDINGS AND DISCUSSION

This chapter examines the pressing issues in facilitating design classes along with the practices and strategies employed to resolve the issues, by instructors in higher education. Prior to delving into the details and nuances of this chapter, it is important to situate the findings and provide a description of the classroom experiences that forms the basis of faculty opinions and workflow. This chapter does not focus on any specific design class—major or year in school. However, class room experiences shared below cover the breadth of major project-based design classes taught within engineering, computer science, and human-computer interaction fields. Several instructors reported teaching multiple design classes over preceding few academic years, often highlighting the best practices that worked and sometimes failed. Overall, five major themes were uncovered:

1. Class room structure, organization and issues (39/39; 100%)
2. Faculty expectations and perceptions of students entering design class (24/39; 62%)
3. Difficulties in formative assessment (15/39; 38%)
4. Peer reviews positively viewed, yet not widely used (26/39; 67%)
5. Faculty incentives in design education (25/39; 64%).

Below we describe each of these major themes and their significance in details.

2.4.1 Class structure, organization and issues

Classroom size varied widely. Single section classes contained anywhere from 12-75 students, with a few exceptions where students in some classes exceeded 200. Where possible, large design classes were split into sections of 50-100 students. Dividing classes into multiple sections helps manage larger enrollment and physical space limitations. It also allows a more intimate setting—as is possible with such numbers—for the instructors and their assigned teams to gain a level of mutual empathy and understanding. However, it requires increased human resources with additional instructors per section, while concurrently introducing variations in student experiences, grades, and instructor engagement across the sections.

One novel approach noted to control some of the variation was to use a core teaching team (typically one to three instructors) for all the sections, with teaching assistants (TA) or other instructors leading individual sections as mentors or coaches (also known as section-in-charge). The core teaching team handled lectures and overall course facilitation. Grading of final presentations or other similar major milestone assignments were completed either exclusively by the core teaching team or in collaboration with the section-in-charge. In-class experiments, individual assignment grades, and mentoring teams through projects remained in the domain of section-in-charge.

2.4.2 Faculty expectations and perceptions of students entering design classes

A majority of the instructors interviewed described three major issues with incoming students: poor communication skills, narrow or fixed perspective, and avoiding risk in their design process. Several instructors shared their frustration with students miscommunicating or not understanding faculty instruction, goals, deadlines etc. Even within teams, faculty find that students poorly share information—often delegating work to each other and working as independent units within the team, oblivious to potential learning moments their team members encounter. An instructor, pointing out the low written communication skills of her students, describes her predicament in assessing their design work, *"…their writing is so bad that I cannot gauge if they were learning correctly or just do not know how to communicate. My strategy is to have very little writing assignments (twice per semester) and more creative design or sketching assignments."* Strategies such as the one mentioned previously, seemed to be the trend, with many instructors focusing more on oral presentations and structured assignments in lieu of traditional design reports.

Instructors note their struggles with design fixation in students and especially conservative approaches that students often follow, as one instructor cites, *"...projects in early terms overly constraints and creates a design fixation. Students do not think out of the box."* Adding to this instructors also face difficulties in structuring classes to encourage exploring design solution space, for example, an instructor concludes, *"…I have struggled to create a class where students take risks…they need more structure, more instruction. If I give them white space and ask them to create a design, none do or succeed."* A key issues lies in the difficulty faculty face in creating opportunities for iterative design and weaving-in diverse perspectives—often requiring a complete course redesign and increased scaffolding with formative support. Another integral experience of design is failure. Instructors find students do not possess the skills to handle failure in good spirits, and avoid taking risks that may lead to failure. Effective designers allude to the "fail fast and iterate often" philosophy—a philosophy that seems difficult to implement in classrooms. Students are attuned to viewing failure as an expression of their performance and not as an integral part of design. In the end, failure by itself, only creates learning opportunities which need to be seized upon and utilized by the instructors to engage students in the design discourse, and to seek and iterate on diverse ideas.

Instructors were also concerned with decreasing student participation in classes, many declaring *"students are not as engaged as they used to be."* When inquired further, instructors often stated this was due to: increasing use of personal digital devices in class, inability of the students to view the big picture, and difficulty creating and participating in a social community within the course. It comes as no surprise that the digital world is pervasive and surrounds everyday lives of most higher education students. Yet, classrooms examined in this chapter were surprisingly devoid of technology—barring the use of learning management system which was often used strictly for communication and grade archiving. There remains a large opportunity for technology that can better integrate into classrooms and engage students beyond receiving information.

2.4.3 The difficulties in formative assessment within design education

Participants included in this chapter were acutely aware of the significance of frequent and detailed feedback in supplementing and improving student performance and learning. Yet, this task was considered a major pain-point in their weekly workflow, largely due to the structure of incentives designed to engage and justify faculty effort in teaching (described in details in section 2.4.5). Even without the helpful incentives, the effort centric nature of grading and providing timely feedback to a large number of unique design problems and/or solutions that students (and teams) further pushes faculty members to their practical limits. Providing timely feedback, when it matters most to the students, is often at odds with generating detailed and constructive feedback. Furthermore, students often do not receive the type of feedback they seek, as one instructor notes: *"with so many teams, it takes a lot of time to give feedback, yet students want more detail, especially if it is criticism. There is not enough time to bolster my feedback and get it done within a week* [one week was considered timely]*."* Another instructor remarking on the current design education set up at their institution (large research focused private school) states: "*in the current set up, sadly, not every student receives the feedback they should be receiving ... lot of them get sufficiently detailed feedback at the final capstone presentation—and find out why was everyone mean to them* [sic]*".*

Increasing class sizes and reducing resources perceptibly impacts feedback provision. The authors note several institutions where feedback, specifically directed to the unique needs of project teams, was rarely provided. At one large research focused public institution, there was little to no instructor feedback provided at freshman level design class (class enrollment of 300-350). Overwhelmed by the sheer number of teams, instructors at this institution resolved to focus on building team and social skills at freshman level, while refocusing their energies in scaffolding work at senior and capstone design classes. At another similar notable institution, resources and faculty focus were shifted to freshman design classes, where close to 8 instructors (1 faculty, 7 TA) interacted weekly with students in person. An instructor at the institution concludes, *"I was able to wrangle the department to give me so many TA's* [at freshman level]... *at junior and senior level design classes there are not many left* [TA] *and there is relatively little time where we meet specifically with the teams".* Utilizing TA's seems to be an obvious choice in reducing faculty burden, however, instructors who receive such support are often hesitant to involve TA's in deeply engaging roles. Faculty, in this report, instead requested TA's to accomplish grading technical assignments, or rubric based grading of low stake deliverables. In few instances in our examination, TA's did receive a larger role and were expected to provide "*light weight design"* feedback—keeping teams on track and setting realistic expectations. As novice instructors, feedback provision at the level of detail and volume that students' desire, can be quite challenging. A student instructor tasked with providing presentation feedback at a large private research focused university describes her experience: *"It is challenging to grade these presentations all day ... I have mental blocks and do not provide enough feedback to my students".*

In view of the constraints, instructors chose to simplify assignments for grading or completely eliminated them. In other cases, instructors used team presentations to provide directive public feedback, hoping to passively impact other teams present in the class. Overall, the instructors recruited in this report remained largely concerned with feedback provision, and perceived to be providing less or infrequent feedback to their students.

2.4.4 Peer reviews were positively viewed yet not formally implemented

Several instructors seemed to be moving towards using some form of communal feedback methodology such as peer reviews—where peers play an active role in feedback provision—primarily as a countermeasure to decreasing use of formative feedback. At the same time, it was evident from our interviews that instructors value the learning opportunity such peer engagement presents while also simultaneously improving student critiquing skills—skillsets which many believed students do not possess enough of today. The most preferred situation for use of peer reviews reported was in project presentations, where instructors often solicited feedback from students in class. Peers either provided written paper-based or oral feedback. Faculty cited the inconsistent participation, lack of student engagement (as one instructor points out, *"students did* *not care to provide feedback to all teams… they were simply preparing for their turn to present"*), and increased facilitation efforts, as impediments in formalizing its use throughout the course. As an alternative, instructors used discussion boards within their classroom learning management systems, wikis, Facebook posts, or blogs, to generate peer discussion, supplementing the feedback as needed. This methodology allows for easier facilitation compared to presentation feedback, while also making it easier to archive the feedback. However, as noted in presentation feedback, some teams did not receive enough feedback and discussions often devolved off track, often requiring some level of moderation from the instructors.

In support of increasing the use of peer reviews in classrooms, instructors noted several beneficial aspects. A few commonly cited aspects include: improvement in quality of student work as a result of displaying their work to their peers, importance of students providing prudent and meaningful critique and handling ambiguous or critical feedback maturely, and the multi-perspective feedback that peer reviews generate. We found several instructors stating *"we know peer learning is beneficial", "it* [peer review] *is one thing we don't do enough of and I think it is important",* and *"I am not doing it* [peer review] *currently, but I wish I was."* Peer reviews were often not formally implemented, i.e., as an integral pedagogical activity, primarily because of the concern instructors had with the effort needed to facilitate the process and with student participation. One of the two instructors who used peer review process formally in the past described their experience, *"It [peer review facilitation] took a lot of faculty time to set up and was a pain to use [software tool used]. Additionally, student think that grading is not their job but that of teachers."* Another instructor who attempted to use peer reviews opined, *"It's* [peer review] *just a pain… no easy way to do this. The logistics are difficult whether we use LMS or a specific tool."* It was evident from the discussions that instructors had limited awareness of the peer review literature or state of the art—signaling the difficulties of translating research to practice. Instructors also questioned the capacity of students to provide feedback that was helpful in the context of design based learning, as one faculty member concludes, *"... I hesitate whether they* [students] *have capability to give feedback. It requires more expertise than they really have. So ability is a concern to me."* Similar apprehensions were raised by other instructors alluding to the variability of peer projects and associated domain knowledge that is needed to provide valuable feedback.

2.4.5 The state of faculty incentives in design education

Faculty incentives and their impact on the overall practice was a prominent and overarching theme witnessed in our interviews. It had a clear influence on every aspect of instructor and classroom practice, right down to the use of tools such as peer reviews. The practices and incentive structure differed most notably between research focused and teaching focused institutions. Tenured and tenure-track faculty in research focused institutions where expected to split their time equally between research, service, and teaching. In reality, most faculty mentioned spending all their time on research, followed by service and teaching. A second year tenure-track instructor in a large private research focused institution justifies their focus on research over teaching, *"What is my incentive to be a good teacher? They are pretty minimal…some of my worst teachers have gone on to get a tenure…even the actual class instruction and feedback provision is affected because I want more grant proposals in, get more research money and prove myself."* Furthermore, in some institutions instructors often received lesser teaching credits when classes were not typical lecture-type making it difficult to justify spending more time and effort. The problems multiply in multi-section classes or classes with multiple co-instructors. In many institutions, faculty members were often unpaid (or received lower teaching credits) for their time assisting or mentoring teams, making it difficult to ensure consistent feedback and mentoring across sections. Adjunct faculty—who were hired specifically to teach—were under different pressures when facilitating design based classes. Most often, it was their availability on campus that stymied their efforts in consulting with students outside class and in feedback provision. Overall, low or non-existent faculty incentives in research focused institutions have impacted the effort and time faculty spend on teaching—especially affecting resource intensive courses such as design focused project-based classes. A faculty member at a large research focused private school concludes, *"…it is lack of time on* [sic] *faculty to keep up with what's out there* [new pedagogies, tools etc.]*… also very little incentive to make changes to the course."*

Not surprisingly, teaching focused institutions had lesser issues with faculty incentives to teach. Most instructors at such institutions were expected to spend close to 60% of their time in teaching, with rest spread out over research and service. Importantly, tenure requirements were directly tied to teacher ratings and student recommendations. Class sizes in the teaching focused institutions covered in this chapter were often in the range that was considered manageable by most faculty members—typically 8 teams of 3-4 students per team. Instructors in teaching institutions interviewed, were interested in peer learning to enhance their current pedagogy. And like their peers in research focused universities, were unaware of tools and practices available that would help them implement peer learning activities in class.

2.4.6 Study limitations and future work

Considering the sample size, original intent of the data collection, the varied interview protocol, and researcher role and associated biases, we make no claims that the findings of this chapter are fully representative of the range of practices and experiences of design faculty and administrators—they reflect the original context of the data collected (*I-Corp for learning* customer discovery) and should not be viewed as generalizable across the field. Yet, the outcomes of this chapter provide a strong basis for future exploration of this subject. Future work could involve creating follow up surveys and detailed interview scripts to systematically explore and dig deeper into the themes uncovered in this chapter.

2.5 CONCLUSION

As illustrated in the chapter, faculty incentives and practical resource limits play a major role in determining the level of engagement faculty have with their classes—most notably impacting feedback provision. Schools that prioritize research over teaching were associated with low faculty incentives to engage in effort-centric courses such as design based classes. These results are in line with previous work in the field (Todd & Magleby, 2004). We noted evidence of decrease in formative assessment, an issue being widely discussed in higher education (Carless, Salter, Yang, & Lam, 2011; D. J. Nicol & Macfarlane-Dick, 2006). And found that faculty members valued peer-peer learning but were specifically concerned with low student design prowess and critiquing skills.

It is disheartening to see disconnect between what is considered effective design education and actual classroom practices. As novice designers, it is expected that students do not fully possess the necessary design skills to explore the design space, frame the problem, and work towards the best compromise—the burden of engaging students squarely falls on the instruction and course design. Effective design education is complex and requires commitment from faculty, administrators, and institutions. Instructors in our report were intimately aware of this, but faced a challenging landscape, riddled with non-existent incentives and resource constraints.

2.5.1 The promise of peer review

Several instructors in this chapter, already understood the benefits of engaging peers in the assessment process. In fields such as writing, formal web-based peer reviews have become a common practice, yielding largely beneficial learning and performance outcomes (Nancy Falchikov & Blythman, 2001; Gielen, Peeters, Dochy, Onghena, & Struyven, 2010; Li, Liu, & Steckelberg, 2010; E. Z.-F. Liu & Lin, 2007; N.-F. Liu & Carless, 2006; Robinson, 2001; Wooley, Was, Schunn, & Dalton, 2008; Xiao & Lucking, 2008). More recently, design educators have begun examining web-based peer review system to manage the scale issue (C. Kulkarni et al., 2013; Tinapple, Olson, & Sadauskas, 2013).

Well-designed peer review has the potential to resolve several outstanding issues described in this chapter—

- Provide diverse, multi-perspective, and timely feedback catered to individual or team assignments needs (K. Cho, 2004).

- Encourage effective communication skills in creating meaningful reports for assessment as well as crafting useful feedback.
- Allow reflective practices, and enable students to engage deeply with the assessment criteria and course goals (Topping, 1998).
- Allow instructors to utilize open ended assignments, frequently (C. Kulkarni et al., 2013).
- Allow iteration supported by feedback (K. Cho & Schunn, 2003a).
- Allow students to witness multiple design approaches and associated problems as viewed in their peers' work (Tinapple et al., 2013).
- Create a course community where cooperative learning takes center stage and students view a positive interdependence with their peers (Tinapple et al., 2013).

Peer reviews, by themselves, cannot solve all the problems that currently plague design education. Furthermore, their effective implementation requires commitment from the faculty members in redesigning the courses to fully utilize review structure, facilitate and troubleshoot issues that arise from using systems to run the peer reviews, and train students to provide good feedback. There are several outstanding issues to examine in using peer reviews in design based classrooms (recent related work is cited in parenthesis):

- How can novice designers be trained or scaffolded to provide feedback that is beneficial to their peers? (Farshid Marbouti et al., 2014)
- How can we engage students in the process? (Neubaum, Wichmann, Eimler, & Kramer, 2014; D. Nicol, Thomson, & Breslin, 2014)
- Are there methods that can further improve the learning outcomes of participating in peer reviews? (Gielen et al., 2010; N.-F. Liu & Carless, 2006)

- How can computing systems better serve peer reviews in design based classrooms? (C. E. Kulkarni, Bernstein, & Klemmer, 2015; Tinapple et al., 2013)

Nevertheless, the future of design education, and education in general, will increasingly depend on more customized learning experiences, and student involved construction of knowledge. In this line, it is important to understand the context and utility of implementing novel pedagogical tools, and examine ways to make these tools engaging and impactful.

2.6 ACKNOWLEDGEMENT

The authors would like to thank the National Science Foundation's Innovation Corps for Learning program (Grant #1547954) for supporting the data collection efforts which forms the basis of this chapter. We would also like to thank Erin Cole (Department of Mechanical Engineering, Brown University, Providence, RI) for her work in organizing and preparing interview data for qualitative analysis, and Philip Brooks (Entrepreneur-in-Residence, Innovation Institute, University of Pittsburgh, Pittsburgh, PA) for his help with recruiting and conducting some of the interviews, and mentoring the team at I-Corp L program.

3.0 COLLABORATIVE TEAM PEER REVIEW GENERATION IMPROVES FEEDBACK QUALITY

3.1 INTRODUCTION

A core element of design and design education is situated and frequent feedback (Fitch, 2016; Tinapple et al., 2013). As design instructors deal with a faculty rewards system that does not incentivize non-traditional teaching (Dym et al., 2003; Todd & Magleby, 2004), constraining budgets and widening class room sizes, a common casualty is feedback provision. It simply is not feasible to provide careful feedback to large numbers of teams, each with a unique open-ended problem under an increasing temporal, human and financial resource pressure. Increasingly, web-based peer-to-peer feedback has emerged as an alternative to instructor feedback with a potential to scale well and keep in pace with increasing class size, while creating newer avenues for student learning (K. Cho, 2004; Nancy Falchikov & Goldfinch, 2000; C. Kulkarni et al., 2013; Tinapple et al., 2013; Topping, 1998). However, relatively little is known about how to effectively structure peer review for design-based classes, and more importantly how to mimic and maintain the natural learning environment of design reviews and critiques (Cardella, Buzzanell, Cummings, Tolbert, & Zoltowski, 2014).

In studio critique sessions, as is commonly seen in arts and design, students benefit from the feedback, however harsh it may seem, from equal status peers as well as experts (Dannels & Martin, 2008; Reimer & Douglas, 2003). These sessions are highly interactive with peers and instructors engaging in a free-flowing conversation, bringing in multiple perspectives, and often building on each other's assessment (Dannels, 2005). Such evolving critique benefits the reviewees by reducing redundant feedback and tapping into deeper collective knowledge of the group. Further, providing critiques also benefit reviewers, who learn from viewing the critique in action (Dannels, 2005).

The rich interactions of studio critique are not replicated in the current standard of web-based peer feedback, where it is typically individual reviewers who provide feedback, insulated from their co-reviewers' opinions. Consequently, the reviewers miss out on opportunities to discuss with their peers their misunderstandings or support their technical limitations, and also learn from others' critiques as they concurrently evaluate the work. As equal status learners in class, individual student peers may not possess all the necessary skills or the design experience to effectively review open ended creative problems that are part of most design based learning classes. In fact, the nature of the problems being tackled in these classes often necessitates the use of a team of individuals who bring together a range of technical skills and subject knowledge, and who work collaboratively on achieving the project goals. Would a similar collaborative team of individuals working together on a design or project review generate better feedback?

Zhu et al (2014), explored the use of a collaborative team of reviewers in a crowd-sourced environment along with individual reviewers and an aggregate of individual reviewers. They found that such collaborative teams of reviewers working synchronously and collaboratively produced more useful feedback than individual reviewers, which aligned closely with expert feedback, and had increased internal consistency. Additionally, the aggregate feedback from individual reviewers outperformed the collaborative team of reviewers by a nominal margin. These results are promising in that they make the case for exploring collaborative reviewing strategies that benefit and improve upon the peer review processes as is implemented today.

In this study, we explore the use of collaborative team of peer reviewers in an engineering design classroom and examine their impact on quality of feedback generated and the mechanisms that yield this feedback. In crossover experiment, 287 engineering students participated in two peer review assignments. Students, in this sophomore-level introduction to mechanical design class, worked in teams designing a physical product or service. They submitted their design log books for peer reviewing in both the assignments, additionally submitting a project video in the second assignment. These logbook were iteratively updated to include more details. Students conducted the peer reviews under three sequences of review structures between the two assignments – individual review to individual review (control), individual review to collaborative team review, and collaborative team review to individual review. In the collaborative team review condition, students worked in their own project teams, reviewing their peers' work together (collocated) and generating a single team review.

The study measured comment quality as feedback that was accurate and improved the project grade when implemented, and whether there was a net positive or negative sentiment exhibited in the feedback (both deductively coded). A self-report assessment measured student demographics and perceived effort conducting the reviews Section 1.01(a)(i)Appendix A. Furthermore, collaborative teams were passively observed to document the learning and feedback generation processes evoked as reviewers worked through their assessments.

The results of this study could significantly impact how peer reviews are structured in design based classes, and form the basis for developing future collaborative peer management systems.

3.2 RELATED WORK

Research on formal peer review in classrooms dates back to more than three decades (D. J. Boud & Holmes, 1981; Nancy Falchikov, 1986), and the last decade has especially focused on web-based peer review, with its affordances of structure, easy of delivery, and anonymity. In these years, studies have explored the impact of peer review on student learning (Nancy Falchikov & Blythman, 2001; Li et al., 2010), effectiveness of the peer feedback generated (K. Cho, 2004; Ekoniak, Scanlon, & Mohammadi-Aragh, 2013; Topping, 1998), and the student experience of participating in peer reviews (Kaufman & Schunn, 2011; D. Nicol et al., 2014). Moreover, researchers have also compared students and instructors on their scoring and feedback (Nancy Falchikov & Goldfinch, 2000; Hovardas, Tsivitanidou, & Zacharia, 2014; Patchan, Charney, & Schunn, 2009). These studies have shown that peer review is generally reliable, generates more feedback for the students, and has a beneficial impact on student learning.

The recent interest and growth in peer review research are largely associated with advancements in technology that have allowed effortless facilitation of peer reviews (K. Cho & Schunn, 2004; Robinson, 2001) and have made them massively scalable (C. Kulkarni et al., 2013). Moreover, the call for improvement in assessment with increased inclusion of students in the process have further made the case for making peer review an integral part of the pedagogy (Cross & Steadman, 1996; Dym et al., 2005).

3.2.1 Web-based peer review in the domain of design

In design, peer review is not a novel or uncommon activity. Studio based peer and instructor critique have been central to design students' training for over a century. As a primary pedagogical tool in design, studios provide a natural multifaceted learning environment, where students not only develop their design, communication, and reflexive skills (Schön, 1987), but also socialize into the professional values, culture, and expectations of the field (Dannels, 2005). In addition to several benefits, studios serve a dual purpose of providing immediate situated feedback to the designer and supporting assessment of their work. It is only natural that this persuasive appeal of studio based pedagogy has drawn it into creative fields outside the traditional arts and architecture (Reimer & Douglas, 2003; Tomayko, 1991).

The engaging environment of studio is sustainable in small class sizes, where such an interaction can be masterfully managed by the instructor. As class size increases, the facilitation of studio critique becomes a constraining factor in the process. In order to remedy the scale issue, researchers have looked for inspiration in the peer review research and tools developed in the fields of writing and computer science. Tinapple and colleagues developed and implemented a peer review tool for "large creative classroom", their value proposition was peer based public ranking of student projects in class and revealing the identities of the anonymous reviewers and authors, at the end of review phase (Tinapple et al., 2013). They found students socialized into a tighter community and supported each other's work when using their tool in class. Similarly, Kulkarni and colleagues, scaled peer reviewing to a massive open online course with over 400 students and emphasized speedier feedback through novel reviewer matching algorithm (C. E. Kulkarni et al., 2015). The increased speed resulted in performance (measured in grades) improvement in the receiver of feedback and encouraged iteration.

3.2.2 Strategies for organizing the web-based review process

Aside from the impact on learning, peer reviews produce a large volume of feedback for the reviewees. Despite several benefits of peer feedback, there are several outstanding issues that require further examination by researchers. One primary issue is student engagement and participation in the process. Students are often apprehensive of peer feedback (Kaufman & Schunn, 2011). Their apprehension largely stems from ambiguity in feedback received (Cardella et al., 2014) and perception of lack of expertise of reviewers (N.-F. Liu & Carless, 2006). These circumstances create a negative cycle where students provide less helpful or low quality feedback, strengthening their notion that peer reviewers are unreliable, and thus reducing their engagement and participation in future peer review cycles. Researchers in the field of writing have grappled with improving student engagement in the process for over a decade. Studies have looked at instructing peers on providing feedback (Gielen et al., 2010) including showing exemplar snippets of feedback (Sadler, 2002), using a training module to calibrate their marking (Robinson, 2001), using carefully crafted rubrics (Yuan et al., 2016), or creating a more conducive course environment (D. Boud, 2000).

Recent work in the domain of design found that nearly half of the freeform feedback from peers contained only praise or encouragement and lacked any suggestions for improvement or refinement (C. E. Kulkarni et al., 2015), this mimics a similar outcome in a study collecting peer feedback on engineering projects (Vasana & Ritzhaupt, 2009). Another study comparing educators' and students' feedback on engineering design work, found that while educators dug deeper into design problems, students often focused on pointing out communication problems in the documents (F Marbouti, Diefes-Dux, & Cardella, 2015). The authors conclude that students may not understand the expectations of their role as reviewers, not have design suggestions to give, or are unable to—or choosing not to—engage deeply.

Written design communication requires the designer to articulate their ideas and process in both a visual and descriptive fashion that enables the reviewer to form a coherent understanding of the designers' intentions. Additionally, designers and reviewers need to be well versed with the many languages of design used in its communication (Atman, Kilgore, & McKenna, 2008; Dym et al., 2005). We posit some of the ambiguity instilled in peer feedback in design stems from peers' lack of understanding of the designers' intent. Furthermore, engineering design crosses multiple domains of knowledge and skills, which at any given instance an individual peer reviewer may not fully possess. So how can one structure the reviews to engage reviewers more deeply into design issues?

Once again, we look at the studio critique model for inspiration. An often overlooked aspect of studio critique is the collaborative atmosphere of review generation. Peers do not review the work in a vacuum, working instead collaboratively with others in constructing their feedback. Such a collaborative team review process may be particularly beneficial in an online peer review set up, where peers often work with a passive design document and attempt to construct an understanding of the design intent. The success of prior basic research on collaborative reviews (Zhu et al., 2014) adds credence to this idea. However, it is not yet clear how these issues will tradeoff in a real course context, evaluating complex objects and also involving social issues that may be less prevalent in an anonymous online research study. Thus, we test the hypothesis that:

H1: Collaborative team of reviewers will generate better feedback than individual reviewers.

However, collaboration in a team also has its limitations. In addition to increased coordination costs, teams may exhibit "group polarization," a phenomenon in which groups exhibit judgment closely resembling their individual biases rather than the "truth" (Myers & Lamm, 1976). These issues were detected in the work done by Zhu et al., (2014); however, the collaborative team reviewers nonetheless outperformed the individual reviewer in all measures. And although, collaboration may yield better quality feedback, if students find this collective process requires increased effort on their part, they may not fully accept or engage in the process. Thus, we test the hypothesis that:

H2: Student perception of effort required to generate feedback in collaborative teams will be greater than feedback generation as an individual.

3.3 METHODS

The study described below was conducted in a classroom within the School of Engineering at a large public university, where web-based peer review was used in the past. The classroom allowed us to experience an authentic implementation of peer review process and its integration into the syllabus. The study design and execution were deemed to meet the educational strategies exemption by the Institutional Review Board.

3.3.1 Course structure

The study took place within a course titled Introduction to Mechanical Design, a sophomore level introductory course on basic mechanical engineering design and product development process. The course consisted of several in-class lectures, computer-aided design labs and assignments, two team-based design projects carried out through the duration of the course, and no final examination test. The two design projects were assigned 40% of the total grade in the class and were conducted in sequence. Peer reviews and the current research work were part of the first design project, named Design Project 1 that the teams worked on for the majority of the semester. Students elected their own teams, which were constrained to contain exactly five members (with a few exceptions of four-member teams). Students with no team preference were randomly assigned to instructor-generated teams. This team membership remained fixed for both the projects and through to the end of the course.

Due to high enrollment, the course involved two sections of students and two faculty members. Student teams followed the same schedule, syllabus and instruction material in both the sections, with some teams even sharing members across sections.

3.3.1.1 Design project 1

Student elected projects, which focused on new product development or improvement of a physical product or system, were vetted and approved by the instructors. These projects ranged from design and development of a novel wheelchair user umbrella to re-designing of snack boxes for ease of use.

Teams were required to document their work in a design log book, specifically including client statement, their hypotheses, initial user discovery, idea generation, preliminary designs, initial prototype, final design and communication. Additionally, teams created a 5-minute video that contained a summary of their work in a narrative format, and had the liberty to make the video creative, and showcase their prototypes or simulations. Teams were expected to ideate and design the product or system to a level where their designs could be readily fabricated, however, a physical prototype was optional.

Teams participated in peer review twice during the semester. They submitted a logbook for the first review, and added a video along with the logbook for the second review. These two reviews also included instructor grading of the submitted artifacts, with an additional final submission made only to the instructors.

3.3.2 Study design

The study used a 3 x 2 cross-over design of three sequences of intervention over two review cycles for two assignments inside the reviewed Design Project, as described in Table 3-1. Student teams were randomly grouped into one of the three sequences (I-I, T-I, and I-T) prior to the first peer review assignment. The sequence I-I involved individual peer reviews at both the time points, whereas sequences I-T and T-I involved switching between individual and collaborative team reviews.

Table 3-1. Crossover design used in this study.

Condition	Assignment 1	Assignment 2
Sequence **T-I**	Collaborative **T**eam Review	**I**ndividual Review
Sequence **I-T**	**I**ndividual Review	Collaborative **T**eam Review
Sequence **I-I** (control)	**I**ndividual Review	**I**ndividual Review

3.3.3 Participants

Although the peer review process was a class requirement for all students, responding to the survey questionnaire, participating in collaborative peer review and utilizing any of the facilities accorded to participating in the research were not required. Students who participated in all aspects of the research were awarded 2 bonus points out of 100 over the final grade. Those who did not participate in the research were eligible to receive these 2 bonus points by completing a one page reflection, focused on one aspect of the course (e.g. peer review, design project submission, CAD labs, etc.) and present an alternative, along with an explanation for why they feel that would be more effective. Participants received no financial compensation for their time.

All but 25 students agreed to participate in research of the N_{total}=287 students. There were N_{teams}=58 teams: 20 in Sequence I-T, 20 in sequence T-I, and 18 in sequence I-I. Students affiliated with the department of Mechanical Engineering & Material Science made up the largest major in class (79%), followed by students from Bioengineering (19%) and Electrical & Computer Engineering (2%). A majority of students were sophomores (61%) followed by juniors (21%), seniors (14%), and 5^{th} year seniors (4%).

3.3.4 Review structuring

Collaborative peer reviews were performed by already existing project teams. While the assignment submission deadlines and content remained the same across all student teams, the collaborative teams were instructed to meet together, to discuss and generate a single peer review. These teams were given the option to reserve a multimedia room within the school of engineering as a meeting location for their reviewing tasks. Individual reviewers conducted the peer reviews independently.

All reviewers, individual and collaborative team as a whole, were assigned two projects for review in each peer review assignment and used the same instructor developed reviewing rubrics (Appendix B). The rubrics differed for assignment 1 and 2, mirroring the evolution of the logbooks as the designs were revised and refined. Furthermore, the whole review process was double blind, with both the providers and receivers of feedback remaining anonymous throughout the process.

3.3.5 Peer review management

Peerceptiv (Panther Learning Systems Inc., Pittsburgh, PA), also known as SWoRD, is a web-based peer review tool (Figure 3-1) and largely used in the writing assessment field (K. Cho & Schunn, 2007; Schunn, Godley, & DeMartino, 2016). A research version of SWoRD, which allows for increased customization and access to experimental features, was used in this class. Research personnel assisted in setting up, managing and troubleshooting the system for the entire class.

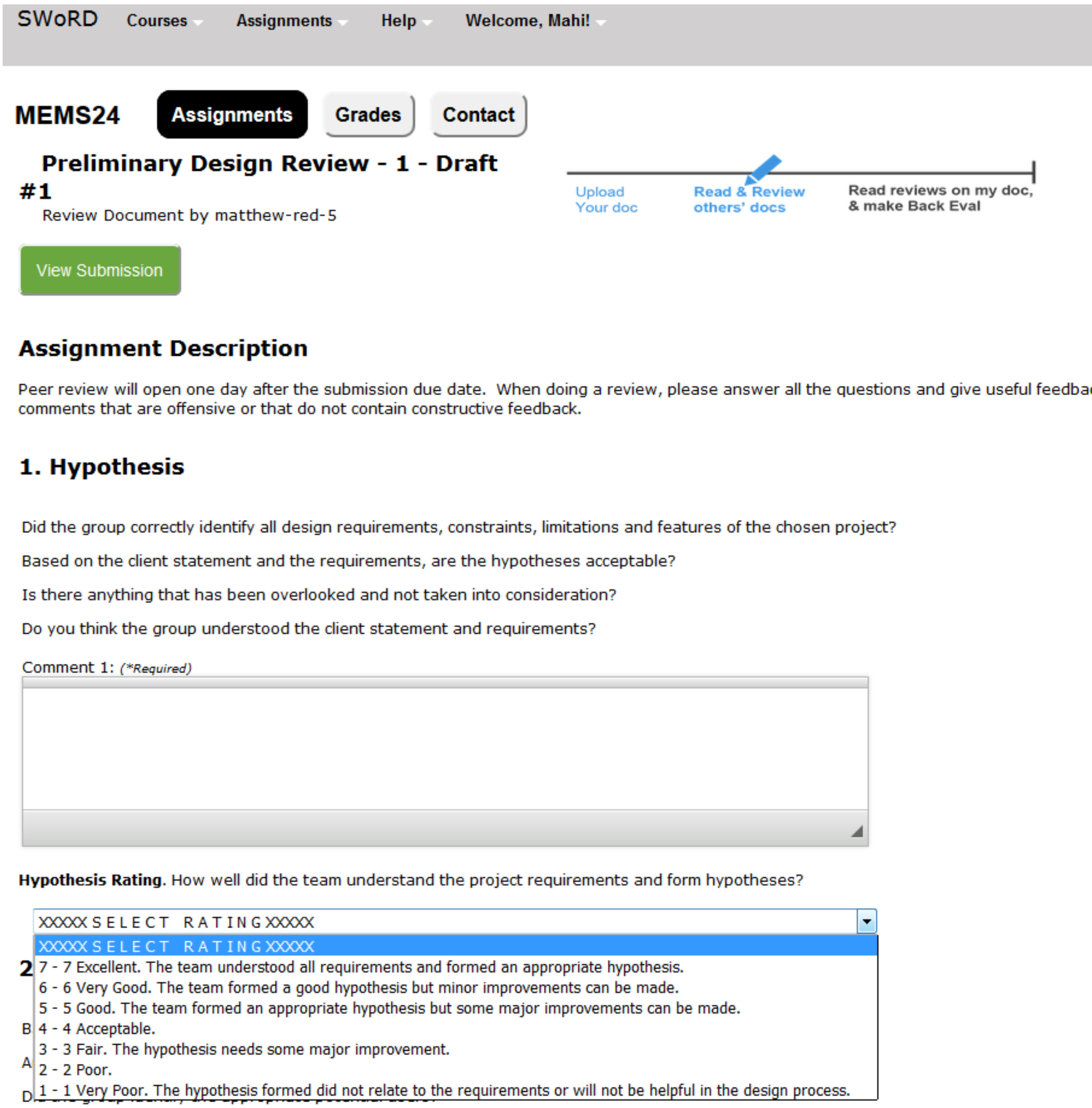

Figure 3-1. Peerceptiv (a.k.a. SWoRD) user interface for reviewing documents.

In order to support collaborative team reviews, custom randomization code was implemented outside the system. The randomization code was written in MATLAB (Release 2012b, The MathWorks, Inc., Natick, Massachusetts, United States) and assigned the same two random reviews to each member of the team, while allocating the rest individually. When students were requested to conduct a collaborative team review, each student in the team received the same two projects to review, with any one member of the team providing the actual feedback and scoring.

3.3.6 Dependent measures and data collection

There were three primary sources of data–survey questionnaires, peer review feedback text and ratings, and field notes from observations of teams conducting collaborative team reviews. Teams conducting collaborative team reviews were asked to allow study personnel to passively observe their reviewing process. A convenience sample of 13 teams out of the 40 teams who were randomized into the collaborative team review intervention across both assignments were observed (6 teams in the first peer review assignment, 7 teams in the second). No individual reviewers were observed. All observations were carried out by the primary author (MM), who described the intent of the observations along with explicit statement on confidentiality of the record. No audio recordings or images were captured during these observations.

3.3.6.1 Feedback quality and sentiment

Feedback quality was measured by accuracy and appropriateness of feedback. Two independent raters (authors WC and IM, also the instructors of the class) rated the feedback on a gradient scale (see code book in Table 3-2) referencing the project logbooks and videos. Raters assigned a code based on whether feedback when implemented in the associated projects would yield a grade change. Feedback was rated per dimension (following the rubrics described above in section 3.3.3). A mean of these scores was used to reflect the overall quality of feedback per reviewer (be it an individual or a team) and empirically ranged from -0.2 to +2.0. Additionally, proportion of high quality feedback (feedback with a score of +2 points) per reviewer was calculated.

Feedback sentiment was characterized as positive, negative or neutral and coded in a similar fashion as quality (Table 3-2) per dimension. A net sentiment score was then calculated and converted once again into an ordinal score reflecting the net sentiment per reviewer.

Instructor rating depended on whether feedback when implemented improved the project grade. To reduce effects of noise from coding, we analyzed the data at the level of comment quality aggregated across dimensions. The quality rating reliability was calculated using Cronbach's alpha. During the initial training, raters had a relatively high reliability in their ratings (Cronbach's $\alpha = 0.76$). These values remained similar post training (Cronbach's $\alpha = 0.72$). Disagreements between raters, were resolved through in-person discussions moderated by author MM.

Table 3-2. Coding schema used to code the open ended feedback.

Quality	**Score**
Increases grade by a grade point or more.	+2
Improves work but does not increase score by a whole grade point.	+1
Does not impact the score.	0
Negatively impacts the score.	-1
Sentiment	**Score**
Positive	+1
Neutral	0
Negative	-1

3.3.6.2 Time spent and effort required

Survey questions enquired about the time and effort individuals spent to complete the reviews. Time was self-reported in units of minutes. Effort was calculated using the NASA TLX (Hart, 2006; Hart & Staveland, 1988), a multidimensional subjective workload assessment questionnaire of end-user workload on a given human-machine interaction. It uses six sub-scales: mental demands, physical demands, temporal demands, own performance, effort, and frustration" (Hart & Staveland, 1988).

3.3.7 Statistical methods and analyses

Group comparisons were conducted using *t*-test or Mann-Whitney *U* test for parametric and non-parametric data, respectively. Effect sizes were represented by standardized mean difference in the form of Cohen's *d*, and were appropriately adjusted for the type of analysis (Ivarsson, Andersen, Johnson, & Lindwall, 2013). The measure of variability is reported with the mean in terms of either standard deviation (denoted by SD) or standard error (denoted by SE) as appropriate.

3.3.7.1 Analysis of fieldnotes

Fieldnotes captured team behavior in the collaborative team review setting. They were analyzed using an inductive framework that resolves data into themes and consequently assists in drawing conclusions (Goetz & LeCompte, 1984). Student names were replaced with pseudonyms. All field notes were coded by primary author, MM.

3.3.7.2 Test for carry-over effects

A recommended strategy to test for carry-over effects is to conduct a mixed model analysis of variance (ANOVA) using repeated measures of time and intervention, and between group's measures for sequence (Portney & Watkins, 2000). Additionally, Grizzle recommends setting a significant threshold of $p<0.1$ (Grizzle, 1965) to test the carry-over effects, represented by the interaction effect of sequence and time in mixed model ANOVA. To avoid issues of non-independence across review cycles, the data were analyzed per assignment, collapsing sequence subgrouping and rather focusing on individual vs. team review for that assignment (e.g., collapsing I-T and I-I in the first assignment data, and T-I and I-I in the second assignment data).

3.4 RESULTS

For practical reasons, only students who completed all three surveys (pre-course, post-assignment 1, and post-assignment 2), N=117, were utilized for data analyses. The selected sample had a similar grade make up (M= 94, SD=4) as the overall class population (M=92, SD=7).

3.4.1 Collaborative team reviewers produced better quality feedback and were more negative than individual reviewers

Reviewers in a collaborative team generated significantly higher quality feedback (N = 16, M = 0.89, SD=0.5, proportion of high quality feedback per reviewer = 24%) compared to individual reviewers (N = 76, M = 0.55, SD=0.43, proportion of high quality feedback per reviewer =

13%), $U = 374.00$, $z = -2.44$, $p = 0.015$, Cohen's $d = 0.53$. Individual reviewers generated slightly more positive feedback (67% positive, 7% negative) compared to collaborative team reviewers (50% positive, 13% negative), however, the difference was not statistically significant, $U = 504.00$, $z = -1.44$, $p = 0.151$, Cohen's $d = 0.30$. See Figure 3-2.

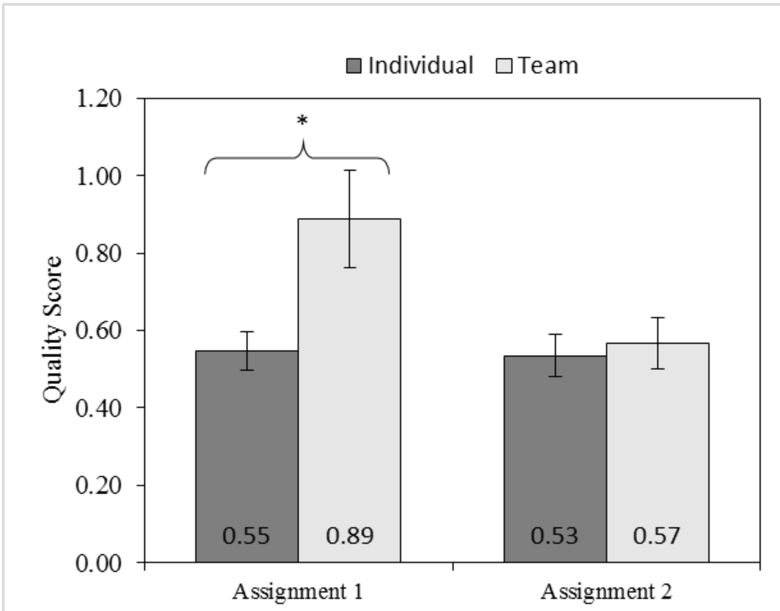

Figure 3-2. The mean quality scores (and SE bars) of feedback generated by individual and collaborative team reviewers in assignment 1 and assignment 2. * denotes statistically significant comparison.

Overall, quality scores for all reviewers dropped from a mean score of 0.63, SD = 0.47 for assignment 1 to 0.55, SD = 0.40 for assignment 2. Looking only at assignment 2 (Figure 3-2), there were no statistically significant differences detected between the quality of review generated by collaborative team review (N = 15, M = 0.57, SD = 0.26, proportion of high quality feedback per reviewer = 10%) and that generated by individual reviewers (N = 72, M = 0.54, SD = 0.46, proportion of high quality feedback per reviewer = 14%), $U = 489.00$, $z = -0.58$, $p = 0.56$, although the direction of differences was the same. Sentiment analysis showed that individual reviewers were slightly more positive (75% positive, 6% negative, compared to collaborative team reviewers (60% positive, 0% negative), however, the difference was not statistically significant, $U = 471.00$, $z = -0.99$, $p = 0.320$.

3.4.2 Collaborative team reviewers spent more time than individual reviewers on the reviewing tasks, yet stated similar effort conducting the reviews in both styles.

Students in the collaborative review teams spent significantly more time (N = 39, M = 116, SD = 50 min) compared to individual reviewers (N = 73, M = 93, SD = 46 min), $U = 1035.500$, $z = -2.392$, $p = 0.017$, Cohen's $d = 0.46$. See Figure 3-3. Similarly, in assignment 2, students in the collaborative review teams spent significantly more time (N = 41, M = 112, SD = 64 min) compared to individual reviewers (N = 72, M = 84, SD = 47 min), $U = 1013.500$, $z = -2.811$, $p = 0.005$, Cohen's $d = 0.55$.

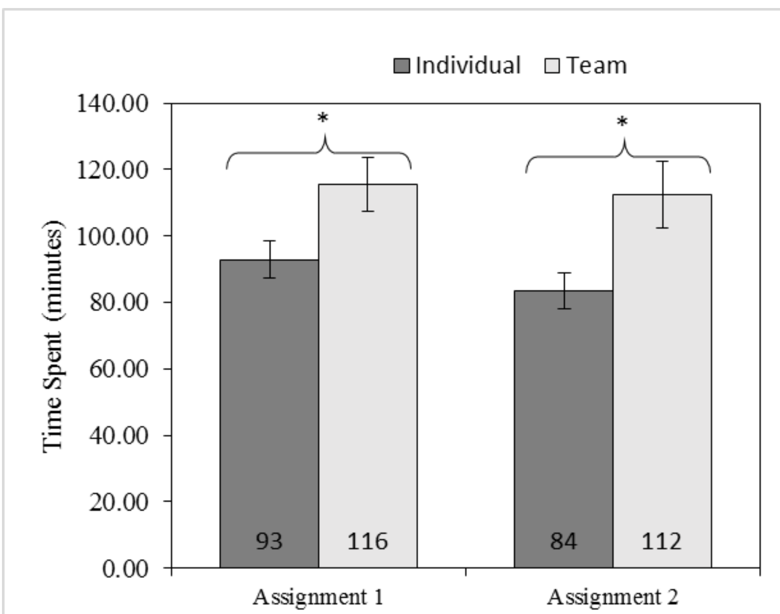

Figure 3-3. The mean time spent (and SE Bars) on generating feedback by individual and collaborative team reviewers in assignment 1 and 2. * denotes statistically significant comparison.

Contrary to hypothesis H2, students in both the assignments perceived similar task effort required in completing both the individual and collaborative team based reviews. In assignment 1, there were no statistical differences detected in the perceived effort required to complete the reviews as reported by students in individual review group (TLX, M = 57%, SD = 12%), and collaborative team review group (TLX, M = 54%, SD = 12%), $t(113) = 1.05$, $p = 0.296$. Similarly, in assignment 2, no statistically significant difference were detected between the individual review group (TLX, M = 56, SD = 12%) and collaborative team review group (TLX, M = 58, SD = 12%), $t(111) = -0.81$, $p = 0.422$.

3.4.3 Collaborative team reviewers seemed engaged and had more 'fun' reviewing

Team observations revealed supporting information that triangulates many of the findings stated above. Many teams utilized the multimedia room reserved for them. Other teams completed the reviews in a meeting room of the library, an empty classroom equipped with computers, or in a public seating area in the lobby. Almost all the teams used a laptop per member, with some teams using printed logbooks and multimedia projector in addition to laptops, for the review and discussion. Across the two assigned projects to review, observed teams spent approximately 48 minutes on the first review, and approximately 40 minutes on the second review.

Overall, students in both the peer review assignments seemed to be quite engaged, with every member of the team participating in the process. As teams continued working on the reviews, they spent more time than they allotted for the review, resulting in some members of the team having to skip a part of the review to make it to their next appointment. In several such instances, students seemed reluctant to leave and miss out on the "fun." For example, while observing a team in review assignment 2, one member left the team towards the end of the first project assigned for review, but called in immediately from his cell phone to inquire about the second project. The fieldnotes stated, "…John was told, you have missed a good one, he stays on speakerphone listening to the team discussion", and concluded, "John is back, and his teammates are showing him the project video and design models… laughter ensues." Such events repeated across several teams, as another excerpt from the fieldnotes of a different team in assignment 1 stated, "Josie has a quiz in 30 mins, she says this is fun and doesn't want to go. Team spends time discussing her predicament, stating she can make it back quickly." Laughter, encouraged by sarcasm, were evident as teams critiqued their peers' work. Not all teams stayed light hearted. One team in particular, seemed quite frustrated with the amount of work required in the class, but, remained engaged in the review process.

Teams spent a significant amount of the review time on determining appropriateness and accuracy of work done, and in explaining the projects or aspects of the projects to each other. They used their own team knowledge to construct an understanding of the projects, with some teams using web-searches, or textbook and class notes to supplement their analysis. An extract from the fieldnotes describes how such interactions were carried out: "the team discusses mechanism of the shaving device, Sid explains the features while Raavi tries to grasp them, some interject with questions, with cascading explanations being providing by members as they begin to understand the design…Raavi states she wouldn't have understood any of that." In addition to working on understanding the projects, teams spent time discussing on the review expectations and their role, clarifying what rubric items mean, and instructor expectations about the project work.

Students collaboratively generated the feedback and voted on the scores to be assigned. This aspect of the process, often times, turned into a deep debate, where consensus was frequently reached by bringing up examples about their own work, or clarifying what is being assessed. For example, one team observed in the second review assignment, spent several minutes debating on the score to be assigned, “Amanda disagrees and reminds the team how their own hard work resulted in a prototype similar to the reviewees, concluding with what a fair assessment should be.”

3.5 DISCUSSION

In this study, we explored the benefit of structuring peer reviews to include a collaborative team review generation element in them, in addition to individual reviewers, on quality of feedback generated and sentiment expressed. We found that collaborative team of reviewers generally produced higher quality feedback compared to feedback generated by individuals. Furthermore, individual reviewers seemed slightly more positive than collaborative team reviewers in their reviews. Surprisingly, there were no differences in student’s perceived effort needed to complete tasks under both, the individual and collaborative team review condition, although they spent more time in the latter condition. However, it is likely that the perception of effort depends on the motivation students have for the tasks rather than on the actual number of hours invested in the task (D. J. Nicol & Macfarlane-Dick, 2006).

The weaker effects in assignment 2 may be due to the complex cross-over design (Portney & Watkins, 2000), with students setting expectations for how to review in assignment 1 or due to changing class conditions across the semester. One significant contributor to student engagement in the review process is the value assigned to the effort invested by the reviewers by the receivers of feedback (Neubaum et al., 2014). In the current study, due to class schedules, grades were assigned by the instructors on the document submitted for peer review. Thus, students did not receive enough time to implement feedback received in a meaningful way. The anticipated lack of attention paid to their feedback may have reduced the effort invested by peers in generating deeper impactful feedback in assignment 2. Additionally, assignment 2 peer review was closer to the end of design project 1, during which time students were already transitioning to work on design project 2. This shifting of student focus can be explained by the notion of a "hidden curriculum". Unlike the formal curriculum established by the instructor, there exists a hidden curriculum that once discovered by students, allows them to efficiently allocate resources to improve their course performance (Snyder, 1971). It can be inferred that focusing on design project 2 at this time point would have maximized the students' grade to effort ratio, thereby further eroding student effort in completing design project 1 peer reviews. Nevertheless, the overall findings of this study lends support to using collaborative team reviews.

3.5.1 Why did collaborative teams generate higher quality feedback?

We offer two explanations for why collective generation of feedback as a team yielded improved feedback quality. First, students in collaborative teams were able to clarify their understanding of the project, the assessment requirements, and expectations of their role as reviewers, resulting in increased accuracy of the feedback. Second, collaborative team reviewers seemed engaged, found peer review fun, and spent more time than individual reviewers, further enhancing their feedback generations efforts. Individual reviewers on the other hand, had to independently form an understanding of all aspects of the review process and the design intent of their peers, and given their novice status in the field of design, it might have made reviewing these open ended problems accurately, a challenge. Nevertheless, individual reviews cannot be completely dismissed. Because collaborative teams are sometimes plagued with issues such as group production losses and potential "group polarization", individual reviewers can examine a larger quantity of projects (i.e., three individual reviewers, each reviewing two random projects would generate six reviews, to generate the same amount of reviews, a team of three individuals would need to review six projects). The current study did not examine the effects of production losses as a result of working in a team. However, it found that the perceived effort was similar in both collaborative team and individual review condition, i.e., students found both activities similarly taxing.

As another concern, an aggregate of individual reviews could still yield better feedback (Reily, Finnerty, & Terveen, 2009; Zhu et al., 2014). However, this aggregation may not work well when individual reviewers do not accurately understand the project at hand or when the project evokes contrasting views. Assuming that collaborative team of reviewers produce higher quality feedback, it may be that an aggregate feedback compiled from a combination of individual and collaborative team feedback will further improve overall feedback quality.

3.5.2 Do collaborative team reviews impact student learning?

This study did not explicitly measure learning impact of the structure of peer review. However, observation of the collaborative team reviewers revealed a cooperative learning atmosphere, where students freely exchanged their thoughts and ideas about their peers' work, while also engaging in self-assessment. In certain cases, teams worked out the design problem of the projects under review, framing the problem, scoping out issues and ideas, and considering what the outcomes would be if they were to work on the problem. Such a level of active engagement with their peers' projects may provide additional design experiences to the team, and since the review teams were also project teams, it creates opportunities for the teams to work together on a new task, socialize, and collectively understand the assessment used in the class. Furthermore, it could be inferred from team observations that some teams delegated work within their project and only had a higher-level understanding of the activities their teammates carried out. As teams reviewed their peers' work and assessed their own work, many within the team gained a deeper perspective of different tasks their team members completed and their impact on assessment. This level of team assessment may perhaps help students recognize the positive interdependence that should exist within the team, and improve their collective efficacy (a team's belief about its own capabilities to work together; see Lent, Schmidt, Schmidt, and Pertmer (2002)). In addition to learning, the quality of feedback generated has been shown to exert a significant positive influence on reviewers' own performance on subsequent iteration of the assignment (Althauser & Darnall, 2001) as well as impact student perception of the peer review process (Kaufman & Schunn, 2011). Nonetheless, formally addressing the learning component of collaborative team peer review remains future work.

3.5.3 Limitations and future work

The study was conducted in an authentic peer review implementation within a large engineering class, and the contextual details could have influenced the obtained results. For example, properly devised rubrics play an important role in helping novice reviewers create meaningful feedback (Yuan et al., 2016). Anecdotal evidence suggests that the rubrics utilized in this class, in some instances, were worded in a way that caused some students to incorrectly understand the rubric objectives.

The coding methodology used for quantifying the quality of feedback focused on only one aspect of the feedback: accuracy. Although this simplistic approach provides a reasonable comparison between the groups, there exists several other criteria that could be assigned to feedback, to create a thorough understanding of the various ways feedback generated by a collaborative team and individual reviewer differs. Dannels and Martin (2008) explored the typology (categorization) of feedback in studio critique ranging from novice to experts. This work suggests that feedback, at least in the studio setting, is composed of following types in decreasing order of frequency, "*judgement, process oriented, brainstorming, interpretation, direct recommendation, investigation, free association, comparison, and identity invoking.*" An aspect of future research work would be to develop and build a design oriented peer feedback typology, and subsequently answer further questions that arise

As mentioned earlier, future work could quantify the benefits of collaborative team peer review structure on the reviewers' learning, motivation, and performance. In the current work, project teams were utilized to conduct the collaborative team reviews as well, subsequent research could look at the benefits of such structuring on the team and explore the use of review only teams that are created independent of the project teams.

A limitation of the current design of collaborative team reviews is that the reviews were required to be conducted in-person. Teams were usually collocated, received access to multimedia room, and in the process incurred some effort in organization. Future work could look into examining the differences that exist between virtual and in-person collaborative review generation, and developing computer systems that facilitate such virtual collaboration.

Finally, the role of study personnel in data collection and analysis could induce potential bias in reporting the results in this study. Primary author MM, moderated rater disagreements when rating the feedback comments on quality and sentiment, conducted and analyzed passive observations of collaborative teams reviewing, and analyzed all survey data. Nevertheless, care has been taken to systematically analyze data and report the results in an objective fashion.

3.6 CONCLUSION

Engineering design is inherently a team process and thrives on diversity—diversity of skills, knowledge, and ideas. Collaborative team peer review brings this diversity to the generation of feedback. We investigated the impact such structuring of peer reviews has on the quality of feedback generated and student perception of effort required. We found that collaborative teams generate better feedback, and students find them no different than individual reviews in the effort required to accomplish them. These results provide alternative review structuring mechanisms for instructors and researchers, and forms a basis for future research work in this area.

4.0 STUDENT OPINION ON PEER REVIEWS

4.1 INTRODUCTION

Advancements in networked computing systems have drastically reduced the effort needed to facilitate peer review of classroom assignments and projects. Classroom based peer reviews have been utilized in writing (K. Cho & Schunn, 2003a), engineering (Farshid Marbouti et al., 2014), human-computer interaction (C. Kulkarni et al., 2013), computer science (Reily et al., 2009), and continue to be implemented in many different settings. The impact of peer reviews on student learning (N Falchikov, 1998; Nancy Falchikov & Blythman, 2001; Li et al., 2010; D. Nicol et al., 2014) and on reducing teachers' grading burden have made them increasingly popular. Furthermore, it is well known that the student engagement with the assessment criteria utilized in class plays a direct role in their performance on the assigned task (Maclellan, 2001). Nonetheless, the success or failure of peer reviews largely depend on one significant stakeholder – students participating in the process. Accordingly, recent research work has focused on understanding student perception of the peer review process including, the values they assign to it (Ertmer et al., 2010) and the cognitive processes evoked in performing the review tasks (D. Nicol et al., 2014).

There are several criteria that influence student perception of the process, including – lack of clarity with regards to their role as reviewer or with the assessment criteria, the additional work required to complete the reviews (N.-F. Liu & Carless, 2006), value attached to the feedback by the receivers, and incentives attached with the process (Neubaum et al., 2014). In general, students seem to be apprehensive about peer review process, however, those with past experience in the process show less negative perception than those with no such experience (Wen & Tsai, 2006). Another related aspect of student perception is participation and engagement in the process. Relatively few studies have examined incentives to improve student engagement with the process. Nonetheless, it is important to gauge student perception of this important pedagogical activity so as to improve the process and the student engagement with it.

In the current study, we look at the impact of an alternate peer review structure on student perception of and engagement with peer review. In addition to utilizing the current standard of peer review, i.e., individual random reviewers elected per assignment or project, we implemented a new structure of review that included co-located teams of students collaboratively reviewing a single project or assignment. This study was part of a larger work conducted in a large sophomore engineering design classroom where the impact of such structuring on quality of feedback generated was evaluated (refer Chapter 3.0 for more details). Collaborative teams of peer reviewers were found to generate better quality feedback than individual reviewers and reported similar effort required between both.

Students in this class participated in the peer review process two times over the course of the semester. The focus of these peer reviews were on the logbooks that were used by the engineering design teams as they worked on a unique product or service design problem. Over the course of the semester, student teams of 4-5 worked together on problem scoping, ideation, concept selection and detailed modeling of their product or service, in some cases, even a prototype. In a crossover design, 40 teams out of the 58 in class were randomized into two sequences of review structuring in two sequential peer review assignments: individual to team and team to individual, with the rest acting as the control (individual to individual). The project teams were also utilized as review teams, and no review-only teams were formed or utilized. Grades on the project work were not determined by peers, however, completing the review task did account for 25% of the assignment grade (assigned by the instructor). The study measured student perception of the peer review process through self-report surveys and passive observations. Open ended survey questions and field notes from observation were inductively coded.

The results of this study could significantly impact how future peer review processes in class are structured, and provide a new pathway for researchers to explore in the field.

4.2 RELATED WORK

What determines student learning in a classroom? Of the several determinants, not surprisingly, the most dominating influence is of the assessment methodology used, even overshadowing teaching (Crooks, 1988; Miller & Parlett, 1974; Snyder, 1971). The assessment methodology deeply impacts the way students perceive course content and set their goals, their engagement in class and the type of learning students undertake (surface- vs. deep-learning) (Gibbs & Simpson, 2004). Furthermore, assessment methodologies such as formative assessment, can potentially refocus students' attention on learning from just performing well in the course (Black & Wiliam, 1998; Hattie & Timperley, 2007). However, not all formative methodologies have a positive influence and researchers often state several criteria that need to be met for it to be beneficial. These criteria state that for feedback to be beneficial it should be – frequent and detailed, focused on actions under students' control, available when needed, clear, appropriate, understandable, and help clarify what is a good performance (Gibbs & Simpson, 2004; D. J. Nicol & Macfarlane-Dick, 2006). With larger classroom sizes, an increasingly common occurrence, the criteria listed above makes formative assessment an increasingly resource intensive methodology burdening the instructors, reducing its usage and depriving students of a central aspect of engaging pedagogy (Ballantyne, Hughes, & Mylonas, 2002; Gibbs & Simpson, 2004; D. J. Nicol & Macfarlane-Dick, 2006).

Peer review, the process of utilizing peers to assess each other's work in generating summative, formative or both types of evaluation, has been shown to be valid and reliable alternative assessment methodology, while also increasing student engagement in the assessment and creating newer avenues for learning (Nancy Falchikov & Goldfinch, 2000; N.-F. Liu & Carless, 2006; Patchan et al., 2009; Topping, 1998). With the advancement in computer systems, the facilitation of peer review no longer remains a significant impediment in the process, giving rise to its extensive adaptation in several domains, most notably in writing (Gielen et al., 2010; Topping, 1998) and more recently in computer science (Reily et al., 2009; Trivedi, Kar, & Patterson-McNeill, 2003). Peer review systems run on the premise that students are actively participating in the process. If the reviewers of peer work do not invest effort in generating meaningful feedback, the receivers of feedback do not benefit, in turn demotivating them from investing effort in reviewing others' work in the future (Kaufman & Schunn, 2011). Furthermore, both lose out on the critical elements of learning in the peer review process – deeper engagement with the assessment criteria, inherent self-assessment, learning from peers' work, and timely and adequate feedback. Therefore, it is quite important that educators and researchers remain aware of student perceptions of peer reviews, and recognize their role as a significant stakeholder in the process.

4.2.1 Student attitude towards peer review

For students, peer review often tends to be an acquired taste. Studies in the field of writing have highlighted the reluctance students have about the peer review process. Student apprehension towards peer review stems from their concern with the evaluative abilities of their peers (Kaufman & Schunn, 2011), their view that the review process is an additional course burden with no explicit value attached to it (N.-F. Liu & Carless, 2006), and when summative assessment is incorporated into peer review, students question the reliability and fairness of their peer markings and become increasingly apprehensive of the whole process (Kaufman & Schunn, 2011; N.-F. Liu & Carless, 2006). Additionally, the way peer review is introduced in class by the educators and their attitude towards it shapes students' initial perception and attitude (N.-F. Liu & Carless, 2006; D. Nicol et al., 2014). Some of these issues are easily addressable, e.g., instructors can create a conducive environment for peer reviews in class; while other issues remain difficult to resolve, e.g., how to incentivize participation. Neubaum et al (2014), manipulated incentives structure for peer reviews in an online course, with reviewers receiving ratings on their evaluation from receivers of feedback, access to assignment solution on review completion, or no direct incentive. They found incentive type impacted the feedback content (receiving ratings improved specificity in feedback) but not feedback participation. Kaufman et al (2011) in their examination of negative perceptions student held against peer reviews, recommend improving training of students in feedback provision as a way to improve their feedback quality, and consequently their perceptions. In the end, the more exposure students have to useful peer feedback, and the process itself, seems to improve students perception of the process (Wen & Tsai, 2006).

Web-based peer reviews in design are in its nascency, however, it is not novel nor uncommon for design education to encompass some form of public critique. In fact, for over a century, the cornerstone of design education has been the design studio (Dannels, 2005) where instructors and peers often engage in open public discussion of student work. Yet, research on student perspectives on these web-based peer review process in design is scarce. Nonetheless, the issue of student engagement in the process remains a fertile investigative grounds for peer review research in the domain of design.

4.2.2 Structuring peer review and its impact on student attitude

A majority of the peer review research work described so far employ the standard review process – students submit individual or team assignments, are randomly assigned several of these peer documents for review, complete the review individually, implement feedback and revise, and receive a grade for completing these tasks. The way the peer review process is implemented and utilized, has been shown to significantly impact the way students both perceive the process and engage with it, e.g., by improving the speed at which feedback was generated and received, Kulkarni et al (2015) found students to better engage in the process, similarly, by creating a virtual community environment Tinapple et al (2013) found students increasingly appreciated their peer feedback.

The current work, supplements research on structuring peer reviews to include a collaborative team review arrangement to improve review quality (Chapter 3.0), by exploring and describing student perception of such structuring. In the sections below, we describe our findings from an active engineering design classroom utilizing peer reviews.

4.3 METHODS

The student opinions described in this study were based on student experience of peer reviews in a large engineering design classroom within the School of Engineering at a public university in the United States. Peer review of student projects was an integral part of the class syllabus. The study design and execution were deemed exempt under category "Evaluation of Educational Strategies, Curricula, or Classroom Management Methods" (IRB# PRO15060428) by the University Institutional Review Board.

4.3.1 Study design and participants

The introductory class utilized in this study focused on providing engineering design and product development experience to sophomore engineering students. In addition to course lectures and assignments, students participated in two team-based design projects through the duration of this course, with membership in the teams remaining constant throughout the semester. Furthermore, students were free to pick their teams, and those who had no reservations about team affiliation were randomly entered into teams created by the instructors of the course. The course consisted of two sections, each led by an instructor. However, the course content, material, and schedule remained the same, in some instances some teams even shared members across both sections. The design projects in class, were sequential, with a few weeks of overlap at the end of the design project 1. Peer review was exclusively utilized for design project 1 and was conducted over two time points in the course, Teams submitted their design logbooks for review in both the peer review assignments, additionally submitting a product video in assignment 2.

The experiment manipulated the structuring of the peer review process, with the addition of a collaborative team of reviewers. Reviewers assigned to the collaborative team review condition, worked in their own project teams, to collectively and synchronously generate a single review of a project. Additionally, collaborative team reviewers remained co-located during the review process and were offered multimedia room reservations in case teams needed them. In a crossover design between the two assignments, randomly elected peer reviewers were assigned to conduct individual reviews (current standard) in assignment 1 followed by collaborative team review in assignment 2, while another set of reviewers were assigned to complete the reverse sequence. Each reviewer (as a team or individually) received 2 randomly selected projects to review, with the whole process remaining double-blinded for the reviewers and receivers of feedback.

4.3.2 Peer review management

Peer reviews were facilitated through an online web-based system, Peerceptiv (Panther Learning Systems Inc., Pittsburgh, PA), also known as SWoRD (K. Cho & Schunn, 2007). While SWoRD supports individual reviewers, it does not have built in support for a team of reviewers working together. In order to accomplish this, each member of the collaborative team review, received the same two projects to review. Only one member from such a team was required to enter the feedback into the system. Custom code was utilized outside SWoRD to accommodate the team randomization and assignment.

4.3.3 Data collection and analysis

Student perception of the process was captured through self-report surveys delivered over the web through Qualtrics (Qualtrics, Provo, UT). The survey questions enquired student opinion on the process, and the contrast that existed between a collaborative team and individual review (see Appendix A). These surveys were conducted at the beginning of the course, at the end of assignment 1 and at the end of assignment 2. Additionally, a sample of the teams conducting a collaborative team review were passively observed by research personnel and documented in the form of field notes.

Descriptive analysis was conducted on quantitative survey data. Binomial tests were used to compare the observed frequencies of dichotomous variables, with the default probability parameter set at 0.5. The open ended responses from surveys and the field notes were analyzed using an inductive framework, resolving information into themes consequently drawing on conclusions. All statistical tests were conducted on SPSS 23.0 (IBM Corp., Armonk, NY).

4.4 RESULTS

Although the class size across both the section combined was 287, there were N=117 (41%) who completed all three surveys. Forty out of the total 58 teams in the class were randomized into the crossover design, conducting collaborative team peer reviews at least once in the two peer review assignment (N=68 students completed the surveys from these 40 teams). Consequent data analyses included only the sample who completed the surveys.

Overall, students seemed to value both providing feedback to and receiving feedback from multiple peers. Furthermore, students preferred conducting reviews as a collaborative team over reviewing individually. In the sections below we describe the results in detail.

4.4.1 Students valued providing feedback to their peers

The majority of students in the course had previous experience providing peer feedback on presentations (74%), assignments (67%), and project work (75%). After participating in peer reviews in the class, students agreed that reviewing helped them see weakness in their own work (76% agree, 8% disagree), helped improve their own work (79% agree, 5% disagree) and helped them learn from seeing their peers' work (80% agree, 8% disagree) (see Figure 4-1). Interestingly, majority of reviewers felt they gave valuable feedback to their peers (80% agree, 3% disagree).

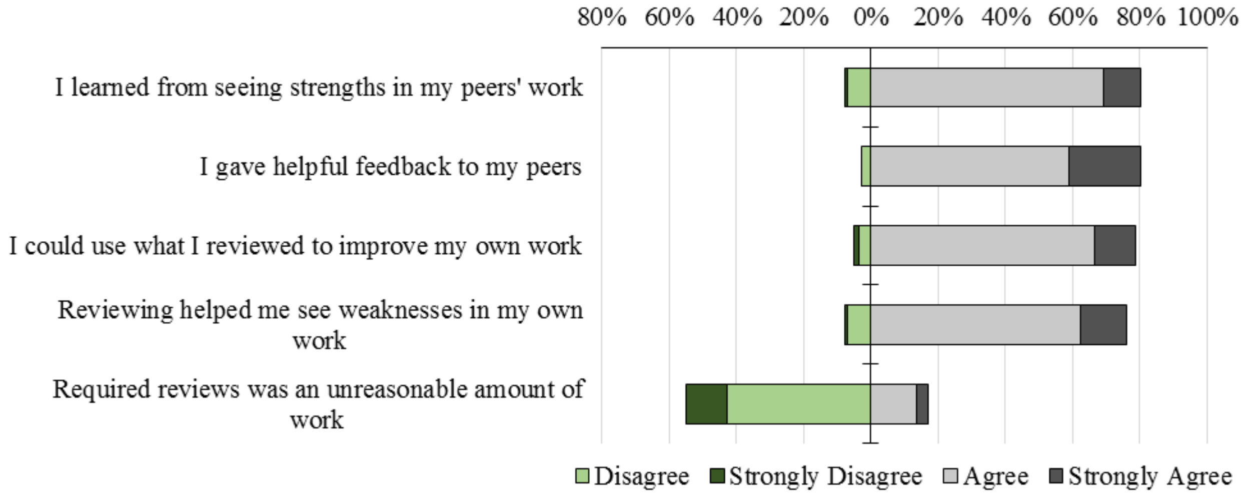

Figure 4-1. Student opinion on reviewing their peers' work.

4.4.2 Students found peer feedback helpful in revising their work

Students agreed that peer feedback they received was helpful (61% agree, 6% disagree), and that this feedback was generally used to correct minor technical aspects of work (66% agree, 13% disagree) (see Figure 4-2). Furthermore, nearly 67% agreed (and 9% disagreed) that peer feedback led to revisions. Another notable result indicates that students were not sure what to do when feedback from peers contradicted each other's (48% agree, 20% disagree).

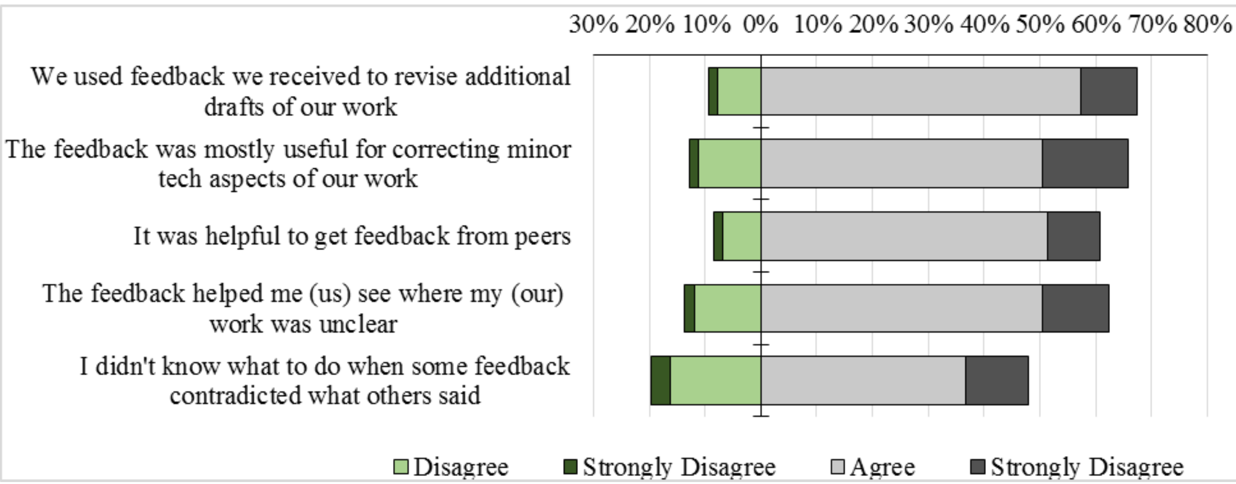

Figure 4-2. Student opinion on feedback from peers.

4.4.3 Collaborative team review vs. individual review

When students were asked to contrast the experience they had conducting reviews as a team and as individuals, contrary to our expectations, students were mostly split on which structure made the review easier to complete (39% agreed individual review were easy, while 38% disagreed, N=68; see Figure 4-1). However, a majority believed that the feedback generated from team reviews was of better quality (70% agree, 12% disagree) than that from individual reviews. Student opinion expressed in open ended survey questions (N=25), captured this sentiment insightfully. For example comments like, "*The only thing I didn't like about peer reviewing together was having to get the group in all one spot* [sic], *at one time to review. It was logistically challenging"*, and, *"It took a longer time to do the team review than the individual review, and it was also harder to pick a time to do it since we all needed to be together in the same room…"* highlight the reasons for the increased effort required to complete team reviews. On the other hand, responses such as, *"Team was better to see other's points of view that you may not have thought of on your own. It also incorporated more perspectives to provide better overall feedback."*, and *"…it* [team review] *was helpful to ask questions about the reports intentions and it was more enjoyable"* demonstrate why students believed team reviews generated higher quality feedback. Students largely agreed that providing feedback as part of a collaborative team was fun (66% agree, 11% disagree; Figure 4-3), as is well described by student comments, *"Team reviews were more fun because you could socialize in between making reviews…", "… I do think we did a better job overall reviewing as a group. It was definitely more fun"*

Another notable result related to student opinion was that providing feedback as a team made them better reviewers (57% agree, 18% disagree, Figure 4-3). Student comments provide a rationale on why collaborative team reviews made them better reviewers, e.g., one student reported, *"Reviewing as a team allowed me to consider other aspects that my teammates brought up that I didn't necessarily think of"*, while another reported, *"I like working as a team. We can bounce ideas off each other, they can find things that I couldn't find that was missing in others work, helped me learn about how to be a better reviewer."*

Overall, students seemed to prefer collaborative team reviews over individual reviews over a variety of reasons. In students' view, *"Working as a team was better because we could all* *collaborate on the responses instead of just having one point of view."*, and, *"Team reviewing is a more genuine review approach. It adds a degree of accountability not achievable with solo* [individual] *reviewing."* The last comment, brings up an interesting perspective on collaborative team reviews – a focus on accountability to provide good feedback.

We anticipated group production losses in the collaborative team reviewers, however, it did not seem to effect the quality of feedback generated (Chapter 3.0) or overall student opinion of the process. A few students did note the issues that arise when working in a team, e.g., *"Team review takes way too long. Too much arguing/making decisions on minor details. Individual review takes significantly less time (one hour individually vs 2.5 hours as a group)"*, and, *"We stressfully made time and arranged an inconvenient meeting to do the team review … the same results would've come from the individual review and in a quicker and more efficient manner."*

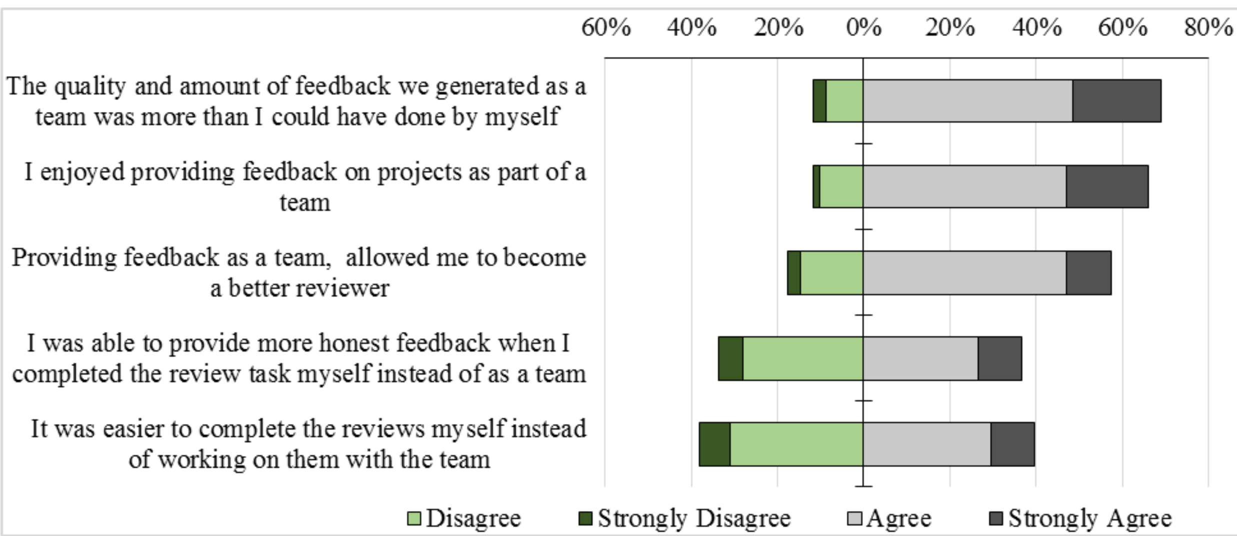

Figure 4-3. Students' opinion contrasting their experiences in collaborative team reviews and individual reviews.

Finally, the survey enquired student preference between individual and collaborative team review. They were asked to base their choice on their own personal learning from completing the review, their anticipated performance in the review task, and effort required to complete the review. In all three cases, students overwhelmingly picked the collaborative team reviews (see Figure 4-4). The percentage of students who picked between the two choices (team or individual) were significantly different from 50% for performance (binomial test p = 0.002, N = 68) and effort (binomial test p = 0.005, N = 68). The results trended towards students picking collaborative team over individual review based on their anticipated learning from the task, however, this result was not significantly different from the random probability of picking either set at 50% (binomial test p = 0.114, N = 68). While these results support the aforementioned student opinion, one surprising outcome here relates to the students choice based on effort required to complete the review, as seen in Figure 4-3. Despite the fact that students noted the extra time and coordination needed to complete the collaborative team review, they preferred it over individual reviewing (68% team review, 32% individual review).

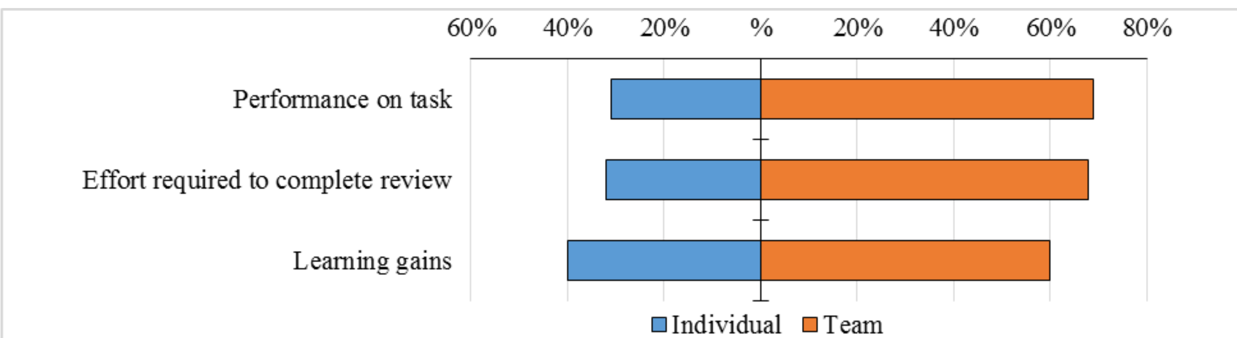

Figure 4-4. Poll showing the type of review (individual or collaborative team) students would like to be part of based on their performance, effort needed of them, and their anticipated learning.

4.4.4 Observations of collaborative team reviews

Thirteen teams were randomly picked for observations with an option to opt out. These teams were passively observed – and archived in the form of field notes – while they completed their collaborative team reviews. Six teams were observed in the first peer review assignment and seven teams in the second. None of the individual reviewers were observed in this study.

Collaborative team reviewers seemed quite engaged in the review process (see Chapter 3.0 for more details). In line with the survey results noted above, students valued reviewing their peers' work. For example, the following excerpt from the field notes: *"Raphael says, 'let's split and review the papers'. Several others disagree and mention that this has really helped them look at their own projects, and they want to give each project their full consideration"* [Team 26], attests to the value students assigned to the process. Furthermore, it highlights the reviewer accountability enforced by collaborative team review.

Teams discussed each project thoroughly, and seemed to be enjoying the process of reviewing. At the end of the observations in the second peer review assignment (after the cross over), 4 out of 7 teams specifically called out collaborative team reviews to be a much better experience than completing the reviews on their own as noted in the field notes: *"Group [sic] review was better, it helped us build off of each other and provide better feedback"* [Team 46], "*This was more fun. We were bored doing it [peer review 1] alone…"* [Team 9] Furthermore, one member from Team 5 remarked, *"If I did this alone, I would keep saying, 'what the heck is this?', and mark them down".* The last comment brings to focus the instructive aspect of the collaborative team peer review, where students were able to discuss projects, review criteria, and reexamine their roles and expectations in class with their team members. Another notable observation related to students' view of project documents to review. Students were not only excited to watch video presentation of their peer work, but also looked forward to viewing them to better understand the work that was done. In several situations, teams skipped over the logbooks to view the product videos first.

Overall, much of what was observed supports the trends and themes uncovered in self-reported student opinion on the collaborative and individual team reviews, mentioned earlier.

4.5 DISCUSSION

In this study we explored student perceptions of web-based peer review in an engineering design classroom. We manipulated the peer review structure by utilizing a collaborative team of reviewers working together and synchronously generating a single review, in addition to using individual reviewers. We found students favored collaborative team reviews over the current standard of individual reviewers. Furthermore, they perceived both providing and receiving feedback as valuable.

4.5.1 Students positively viewed and valued the peer review process

Earlier literature on student perception of peer review indicate students hold a negative attitude towards it (Kaufman & Schunn, 2011; N.-F. Liu & Carless, 2006; Smith, Cooper, & Lancaster, 2002), in stark contrast we noted a more positive outlook from students in the current study. Majority of the students in the class had some experience with peer reviewing in the past. It is known that when students have positive experience with peer reviews, they often carry with them a positive perception of the process (Wen & Tsai, 2006). Additionally, we posit the traditional expectations of the field of design, inherent ambiguity of design problems, and collective navigation of the problem space by students in teams, results in students recognizing positive interdependence needed within and across their course community. Such recognition by the students consequently supports a more open acceptance of the formal peer review process.

Earlier work in the field of writing found that producers of feedback achieve higher gains in performance and learning than receivers of this feedback (K. Cho & MacArthur, 2011; Y. H. Cho & Cho, 2011). While we did not look at the contrast in learning and performance improvement due to producing versus receiving feedback, student opinion here seems to point to the inherent value of reviewing. Furthermore, it was interesting to note that nearly all of the reviewers believed they produced valuable feedback, while a majority of them felt feedback received was useful to correct only minor or technical errors. First, this finding points to reviewers' lack of understanding on what constitutes valuable feedback. In the current class, instructors did not spend any class time on training students on generating good quality feedback. Second, the act of reviewing may itself have resulted in teams examining their own work, potentially restricting the usefulness of feedback received to technical errors.

In general, students seemed satisfied with the feedback they received from their peers, and suggested this led to revision on their work. A prominent finding in this regard was that students were unsure what to do when they received contradictory feedback. It is expected that peer feedback can, at times, be ambiguous or contradictory. Ambiguity in design fields is not uncommon, and capacity to tolerate this ambiguity is often considered an important skillset associated with good designers (Dym et al., 2005). Ambiguity can often lead to divergence of ideas and problem framing (Hatchuel & Weil, 2003). In peer feedback, ambiguous or contradictory feedback can sensitize students to the multiplicity of their peers' reaction to their work – but it can also create confusion. Furthermore, peer feedback requested in the current study did not solely focus on design problems – it also required peers to examine the merits of students following the assignment protocol. As such, some of the contradictory feedback could devalue its usage by students. Nonetheless, future work could address the role of ambiguity in peer feedback in the domain of design.

An interesting observation was that students seemed to prefer watching video presentation of their peer work over reading logbooks. One simple reason could be that video presentation were short (~5 minutes) and provided an opportunity to the students to be creative (and often humorous) in presenting their problem and solution, unlike the logbooks that were written in a more professional and "expected" fashion. Additionally, multimedia presentations provide a much richer contextual environment where design intent can be more readily understood by the peers. Given the choice between reading a text and still image based logbook and a richer descriptive video, it seems natural that students picked watching video presentations.

4.5.2 Why did students prefer collaborative team reviews?

We expected collaborative team reviews to increase student burden, and some of the results indicate this to be true. However, we also found that the reviewing process in collaborative teams was considered more engaging and fun. It is possible that although students spent more time and effort on the review, the increased engagement and association of fun with the task, may have skewed their preference to collaborative team reviews. Additionally, students in collaborative team reviews were able to share their understanding of the work under review and their role as reviewers, share their frustration and excitement with their teammates, and view multiple perspectives on the same work, most of which they were unable to do so when conducting the review independently. Furthermore, reviewing together in project teams, allowed team members to socialize under reduced pressure, on a non-critical non-project task. This reduced pressure on the team may have negated any experience the teams had with free riders who are typically encountered in teams. Students also seemed to recognize the impact collaborative team review had on their learning, which they opined also made them better reviewers. Therefore, on the whole, given the contrast students picked working together as a team on these reviews.

An interesting observation from collaborative team reviews was its impact on reviewer accountability. We posit that the social aspect of reviewing in a team, creates a new expectation that each member of the team should carefully consider and fully participate in the review discussion and feedback generation. This is perhaps an important outcome of using collaborative team reviews, in that, it ensures increased participation in the peer review process across the class undermining the digital distancing that is often created in an anonymous review process.

4.5.3 Limitations and future work

Although self-report surveys provide rich meaningful data, observations and other similar ethnographic techniques (interviews, focus groups etc.) can reveal even more detailed aspects of the phenomenon under investigation. In the current study, due to practical limitations, only collaborative team reviews were observed. Future work could incorporate surveys triangulated with data from other ethnographic sources to capture the whole range of activities students participate in.

A major issue encountered by the students in collaborative team reviews was the administrative tasks associated with organizing the review themselves. Students, in this study, were requested to co-locate and generate a single review per team, which required team members to coordinate a common time to meet. Future work could explore using asynchronous and synchronous virtual collaboration on review generation. Such virtual collaboration has the potential to significantly reduce some of the administrative overhead encountered in real world collaboration. Consequently, research could examine the impact of user interface and experience of such a virtual collaborative review generation tool.

A minority who preferred individual reviews, brought out some of the major issues that could plague collaborative team review. The most notable issue was the potential for teams to engage in prolonged discussion over minor details in the project under review. Although the increased time on task could benefit the team and its members in developing a better understanding of the assessment and their roles both as a reviewer and as members of the team, it can also distract reviewers from providing crucial feedback. Future work could examine aspects of the assignment where students engage in meaningful discussion, and build on developing guides and rubrics specific to collaborative team reviews.

In this study we did not focus on feedback ambiguity and its impact, or specific locations within feedback where such ambiguity exists – it remains an aspect of future work. Additionally, future work could examine the role of artifacts under review on student engagement and impact on feedback generated.

Finally, the role of study personnel in data collection and analysis could induce potential bias in reporting the results in this study. Primary author MM, analyzed all survey data and coded all open ended survey data. Nevertheless, care has been taken to systematically analyze data and report the results in an objective fashion.

4.6 CONCLUSION

Students are an important stakeholder in the peer review process. In the current study, we examined engineering design students' perception of the peer review process. In addition to using individual reviewers, we utilized collective review generation through collaborative team of reviewers. We found that students valued peer feedback, and had a positive perception of the process. Furthermore, students preferred conducting the reviews in collaborative teams, finding them both engaging as well as beneficial to their learning. Collaborative team reviews also seemed to encourage participation. These results add on to previous work on student perceptions of peer reviews, and provides alternative mechanisms to improve student engagement in the process.

4.7 ACKNOWLEDGEMENT

The authors would like to thank Dr. William Clark and Dr. Irene Mena, for their assistance in conducting this research work.

5.0 IMPLEMENTING PEER REVIEWS IN DESIGN BASED CLASSROOMS

5.1 INTRODUCTION

In the past two decades, an increasingly popular recourse for the lack of formative assessment in education has been the use of peers to provide feedback (D. Boud et al., 1999; Nancy Falchikov, 1995; Nancy Falchikov & Blythman, 2001; Nancy Falchikov & Goldfinch, 2000; Nulty, 2011; Topping, 1998; Williams, He, Elger, & Schumacher, 2007). The driving force for use of peer review is the impact it has on multiple dimensions of student learning in addition to lowering the teacher burden. Under certain conditions, peer review has shown to decrease feedback time and increase the amount of feedback generated (Nancy Falchikov, 2013; N.-F. Liu & Carless, 2006; Topping, 1998), promote deeper learning and a stronger understanding of the assessment objectives (Ballantyne et al., 2002; Vickerman, 2009), encourage reflective and critical thinking skills (N Falchikov, 1998; C. Kulkarni et al., 2013; Tinapple et al., 2013), and improve student confidence and encourage learning from observing their peers (Dannels & Martin, 2008; Race, 1998). Advancements in web technologies in recent years have shifted the administration of peer assessment to an online environment, making its implementation in a classroom setting easier compared to earlier methods, additionally, providing opportunities for anonymization of the participants, ability to take the assessment work beyond the brick-and-mortar walls, and improving data analytics. Yet, in a recent examination of classroom practices in design based classes in fields other than arts and architecture, only 2/39 interviewed used formal peer review in class, with 1 of them discontinuing its use (for more details review Chapter 2.0 of this dissertation). The primary reason for hesitancy instructors have in implementing peer reviews is lack of information on how to facilitate such reviews in practice. In this regard, one goal of this chapter is to report best practices to effectively implement peer reviews within current design based classrooms. We base our recommendations on literature and our experience in implementing peer reviews in an engineering design classroom.

A majority of the peer review web tools available today, were developed with a focus on writing (K. Cho & Schunn, 2004; Robinson, 2001), and have not been successfully ported to be used in design based classrooms. Recent work by Tinapple et al (2013) and Kulkarni et al (2015) focusses on design and creative classes, but are yet to be widely available or accessed. Furthermore, our previous work (refer Chapter 3.0 and 4.0) showcases novel peer review structuring methods that could further enhance the peer review experience and outcomes. There are no tools currently available that can support structuring beyond the typical standard of using individual reviewers. As such, we describe characteristics of a novel peer review tool titled, Peerval, that can support new peer review structuring and includes features that enhance the peer reviewing experience for both instructors and students alike. Thus, the second goal of this paper is to describe Peerval, its significant characteristics and implementation methodology, along with future work.

5.2 GUIDELINES ON IMPLEMENTING PEER REVIEWS IN DESIGN CLASSROOMS

Peer review literature typically showcases unique cases of implementation, which are often led by instructors well-versed with peer review literature or expert researchers supporting and facilitating reviews in classrooms. There are several examples of peer review being used within classrooms, both on-land and online (C. Kulkarni et al., 2013; Tinapple et al., 2013; Vasana & Ritzhaupt, 2009). However, these well-crafted exemplars often do not provide enough practical information for instructors to recognize the modifications needed in their course design and pedagogy. The following is an attempt to collect best practices from various sources, along with authors' own experiences implementing the research work described in this dissertation over the past two years.

5.2.1 Classroom set up and environment that encourages positive dependence on peer reviews

Implementing peer reviews, requires a certain buy-in from the instructors of the course. The value instructors attach to peer reviews, can directly impact students' initial perception with the review task. Positive experiences students have with peer reviews encourages and improves future participation in the process (Kaufman & Schunn, 2011; Wen & Tsai, 2006). Furthermore, it is recommended that the course environment creates a positive dependence on peers reviews (D. Boud, 2000). Prior to assigning peer reviews to students, it is imperative that the instructors discuss the rationale for use of the review process (N.-F. Liu & Carless, 2006), establishing a positive relationship between peers in class, and focusing on peer reviews as a source of feedback to iterate and improve work rather than as a grade.

Peer reviews are best used to help students iterate their work. Therefore, it is appropriate to utilize them more than once in a classroom. For example, in the engineering design course (described in Chapter 3.0), the authors used peer reviews over three time points (over the course of a semester) prior to final submission of the project work for instructor grading, with each sequence of reviews lasting one week. Instructors should note the effort peer reviews require from students and plan their assignments and other deliverables around them.

5.2.2 Using grades and collaboration in reviews to enhance student engagement and participation in peer reviews

Assigning a grade to reviewing can enhance students participation in the process (N.-F. Liu & Carless, 2006), while showcasing the value receivers of feedback assign to the feedback from their peers, increases the amount of details reviewers provide in their feedback (Neubaum et al., 2014). Systems such as Peerceptiv (Panther Learning Systems, Inc., Pittsburgh, PA) have built in support for receivers of feedback to assess the helpfulness of feedback, and in turn showcase the value they assign to it (K. Cho & Schunn, 2004).

Results from Chapter 3.0 and 4.0 , in this dissertation, highlight the multidimensional benefits of using collaborative team of reviewers in assessing creative assignments (specifically, design projects). Collaborative team reviewers produce better quality (appropriate and accurate) feedback, compared to individual reviews. Furthermore, students overwhelmingly prefer conducting reviews as a team, calling it more “fun”. Implementation of such reviews will require “off-label” usage of current peer review tools, as currently no tool supports collaborative review generation. The readers are directed to review Chapter 3.0 , in order to get a deeper understanding of the implementation and set up of peer reviews for collaborative assessment.

We recommend utilizing collaborative peer reviews for the first few review assignments or whenever a new rubric is implemented. In our work, collaborative reviewers seemed to benefit reviewers better understand assessment criteria and engage in discussions that help clarify their roles and expectations. Furthermore, when assignments are complex to assess, collaborative team reviews may help reviewers provide feedback with increased accuracy. While collaborative reviews can be used at any course juncture, a practical limitation of using this type of reviews is the increased facilitation costs and the lower volume of reviews generated (e.g. five individual reviewers, each reviewing two randomly elected unique projects generate ten reviews, utilizing a collaborative team of five reviewers, would require this team to review ten projects). In sections below on Peerval, we describe design of a peer review tool that natively supports collaborative team reviews and may reduce the facilitation costs.

5.2.3 Clearly defining assessment criteria and including both ratings and open feedback

One important benefit of peer review is that students gain a clearer understanding of the assessment criteria utilized in classrooms. It is known that student engagement with the assessment criteria utilized in class plays a direct role in their performance on the assigned task (Maclellan, 2001). However, for students to engage in impactful peer reviews it is important they have a clear initial understanding of their roles as reviewers as well as the criteria to be utilized to assess their peers. A commonly used tool to scaffold student feedback is the use of rubrics. A well-crafted rubric can help students provide feedback that is as good as expert feedback (Yuan et al., 2016). However, creating a meaningful rubric is often challenging (Andrade, 2005).

Rubrics are often always work-in-progress, and instructors should utilize feedback from students to constantly iterate and improve it.

When peer review focuses on summative aspects of peer review, student become increasingly hesitant with the process (Kaufman & Schunn, 2011; D. Nicol et al., 2014). This does not necessarily mean that students should not rate their peer work, on the contrary, using a rating scale increases the number of explanations that students provide in support of their ratings (Hicks, Fraser, Desai, & Klemmer, 2015). Thus, we recommend using a rating scale with associated open ended feedback prompts.

5.2.4 Using assignment types that enhance impact of peer reviews

Designers use a variety of communication methods and design languages to share information and solicit feedback on. Most design classes in engineering design conclude with a written report and an oral presentation. Intermediate assignments often require submissions in the form text, graphic, or other relevant formats. Peer review can be used for either individual assignments, or used iteratively to improve students' main project work.

In our implementation of peer reviews in an engineering classroom, students submitted their design logbooks containing text, sketches, and graphics, describing the design process and evolution of designs. Students also submitted short videos describing their design thinking and showcasing physical or virtual prototypes. Students seemed increasingly excited to view and assess the videos. The videos were also viewed as a way to clarify the reviewers understanding of the project.

We recommend using a combination of structured written reports and video providing an overview, along with prototype demonstration. Written reports encourage communication skills that are otherwise easily overlooked in non-writing specific courses, while video, allows students to showcase their creativity and get them excited to participate in the reviews.

5.2.5 Constantly iterating classroom implementation through student-centered feedback

Dym et al (2005) in their seminal review of engineering design thinking and teaching recommend instrumenting "the curriculum-as-laboratory" to help support improving the quality of design pedagogy. This recommendation still stands true today. Instructors, especially when using novel methods such as peer reviews, should systematically collect and utilize student feedback to improve their peer review implementation. Feedback can be collected as part of survey questionnaires, student reflections, interviews, focus groups, or a combination of these methods. This valuable data can support optimizing review types, rubrics, systems used to facilitate reviews, as well as the incentives used to increase student participation.

At a minimum, we recommend utilizing web-based surveys to enquire student perception of the rubrics utilized, and barriers or issues that impede their full participation in the review process.

5.2.6 Guidelines supported by research

The peer review implementation guidelines described above are not exhaustive. There are several additional practices that have been found to enhance the process and its impact including training students to provide better feedback (N.-F. Liu & Carless, 2006; Robinson, 2001)

5.3 COMPUTER SUPPORTED IMPLEMENTATION OF PEER REVIEWS

The tools and processes used to facilitate peer review play a significant role in its success—student engagement in the reviews, quality of peer interaction including that of feedback generated, and instructor engagement as well as sustainability of practice. Some of the best practices described in the section above cannot be efficiently implemented with currently available peer review tools. In order to address these limitations, we are developing a novel peer review tool tailored to augment design-based pedagogy, titled Peerval. Utilizing information collected from design instructors (Chapter 2.0), experience amassed designing and testing collaborative peer review structure (Chapter 3.0) and student perspective on peer review in design based learning (Chapter 4.0), the design specifications and characteristics of the system were developed. The following section describes in detail the features that set Peerval apart from other review tools, concluding with future work that will further enhance its utility. As a work-in-progress, some features and suggested user-interface may significantly change as additional user feedback—generated by instructors and students—is incorporated in the prototype.

5.3.1 Web-based application, easy to access across multiple device, robust and secure

Peerval is being developed as a web-based application, supporting most modern internet browsers. It will be accessible through multiple devices, however, the system will be optimized for screen sizes over 10” diagonal. Peerval development will use modern industry standard encryption methods to keep user data secure and safe. The owner of a course (facilitator, typically an instructor) will be able to pick a cloud storage location (viz., Google Drive, Box, Dropbox or custom Peerval solution[1]) for all course affiliated information to be stored and archived for future retrieval. This option alleviates intellectual property related issues that some design-based classes face by allowing information to be stored in a known and supported location. Peerval will support asynchronous review generation tools, for reviewers to work at their own pace. Additionally, Peerval will also support collaborative peer review generation by a team of reviewers in both synchronous and asynchronous modes. Reviews are double-blinded—reviewers and content generators remain anonymous throughout the peer review process.

5.3.2 Manage assignments, simply and effectively

Peerval will be designed to keep the technology threshold to a minimum. The primary focus of this system is assignment management, and the interface and functionality will focus on simplifying the user experience. Contextually aware help will provide support as is needed to the user. Figure 5-1 describes the anticipated features designed to assist instructors create rich assignments in Peerval.

Peerval will not focus on grade generation, however, reviewers are expected to provide both a numerical score and open-ended feedback as part of their reviews. It has been shown that utilizing a score increases the positivity and the number of explanations reviewers provide in their feedback, compared to those who are asked to provide open feedback only (Hicks et al., 2015). The scores generated will be utilized to identify students (or teams) who receive

[1]Cloud services can be accessed at www.google.com/drive, www.box.com, and www.dropbox.com. Peerval uses Amazon Web Services' (aws.amazon.com) static storage S3.

conflicting reviews or critical feedback—allowing instructors to focus their attention on a select few in the class.

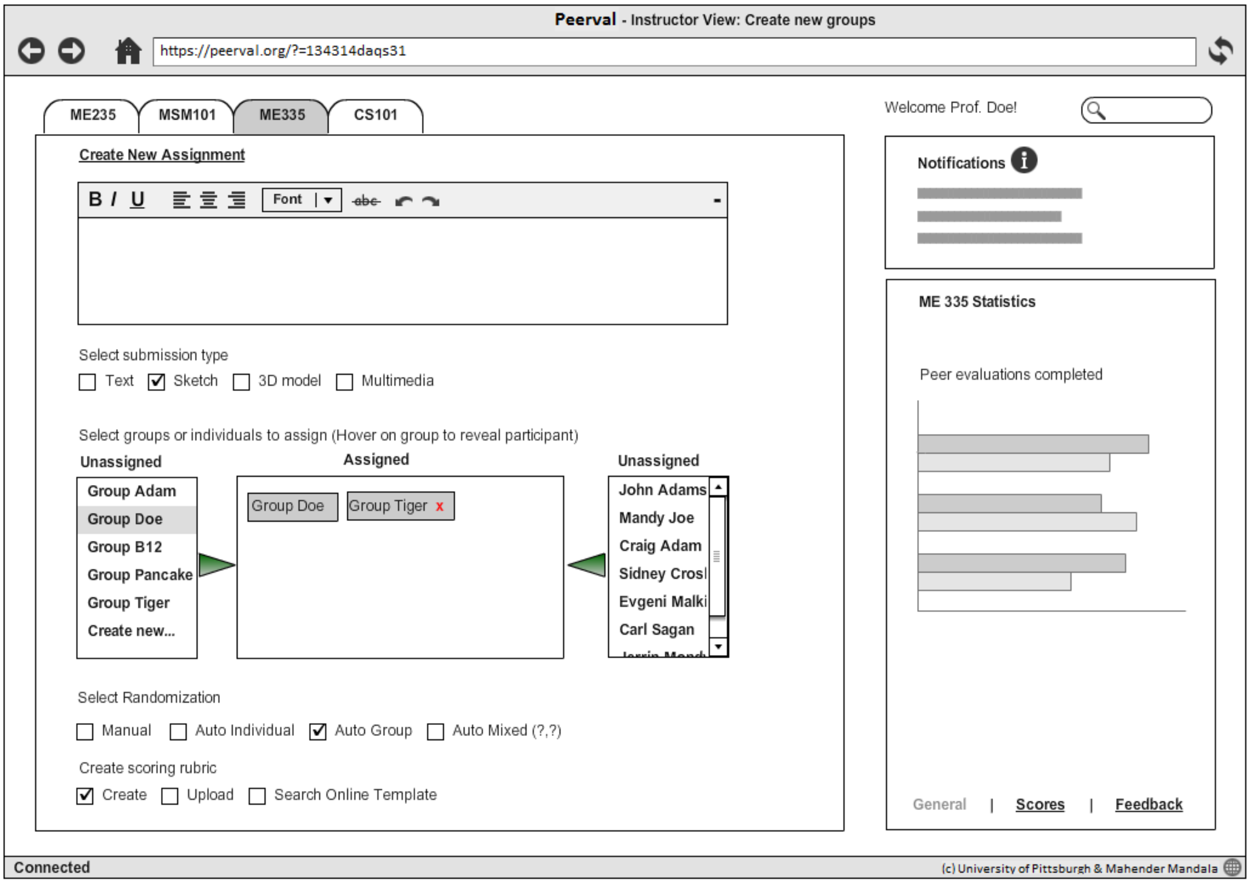

Figure 5-1. Peerval wire-frame showcasing features that support instructors in creating new assignments.

5.3.3 Native support for multiple file formats encountered in design

Peerval will support several digital formats—popular text and multimedia formats, along with STL file format[2] to represent three-dimensional models—allowing a single location for students to upload their design project deliverables. Furthermore, this will allow the system to present all relevant information to the reviewers in a single location as well.

5.3.4 Reviewer tools that support collaboration and reviewer feedback

Peerval will provide several relevant reviewer tools, many of which are currently under development. Some of these tools have been specifically designed to support collaboration in review generation, an approach that can enhance the review quality and engage students in the process (chapter 3.0 and 4.0).

Each peer review task will require a rubric that showcases the scoring system and directs specific open ended feedback. The rubrics can be multi-dimensional and have hierarchy within the dimensions. For example, rubric items can have multiple questions, grouped into different topics or areas of focus. Rubrics play an important role in feedback quality. Well-crafted rubrics have the potential to allow reviewers generate expert-like feedback (Yuan et al., 2016).

Peerval will support reviewer collaboration using tools to discuss, generate, and curate the review. These tools will allow for asynchronous or synchronous collaboration. Reviewers

[2] STL stands for STereoLithography. It is widely used in 3D printing and computer aided manufacturing. STL files describe the surface geometry of a three dimensional object. This file format is supported by most major computer supported design packages.

will be able to utilize a private text chat to discuss the review (viewable by team and instructors) and collaboratively edit their feedback into the review pane. Ratings will be polled across the review team with final choice of rating left to the team. Figure 5-2 describes envisioned collaborative review features.

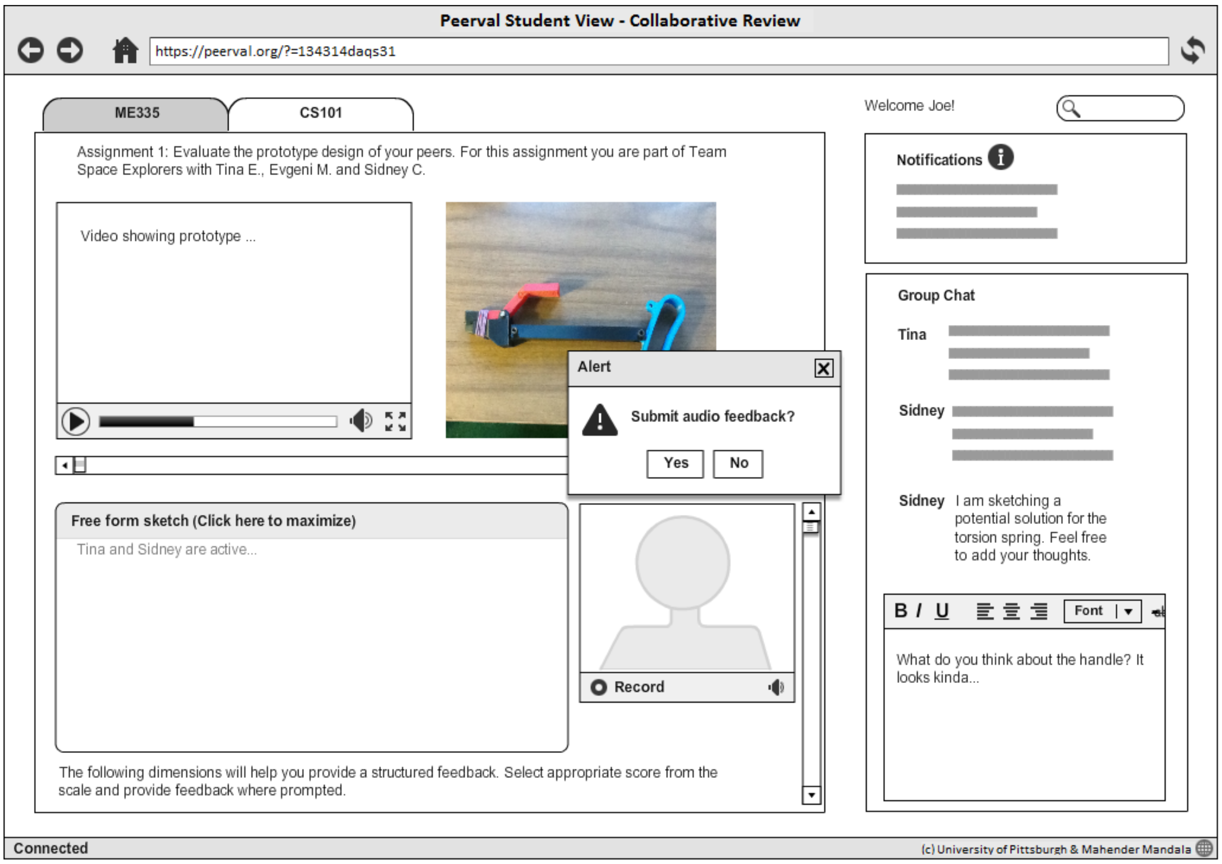

Figure 5-2. Peerval wire-frame showcasing collaborative review features.

Engaging students in the peer review process is paramount to its success. Although assigning grades to peer review can increase participation (N.-F. Liu & Carless, 2006), the experience students have with the peer review system can also play an important role in engaging them. Positive experience, helps change student's perception peer reviewing (Kaufman & Schunn, 2011) further engaging them in the process. One way to increase the positive experience is for students to receive valuable and helpful feedback consistently. When peer reviewers believe their feedback will be valued, they provide more detailed feedback (Neubaum et al., 2014). Peerval will allow receivers of feedback to quickly provide a "thumbs up" or "thumbs down" for helpfulness of the review (indicating to the reviewers the value assigned to their feedback), and "flag" the review for reasons that can be considered inappropriate. When feedback is flagged, the system will alert any teaching assistant or instructors, designated in the course set up, to conduct a review of the of the flagged feedback. Students will also be able to provide open feedback to the reviewers and let them know rationale behind their reviewer-feedback.

5.3.5 Analytics designed to help optimize instructor involvement

Peerval is being designed to capture several key logs from the users. These will include time spent on reviews, time spent on site visiting specific sections, chats and discussions through collaboration, and system flagged individuals. Peerval will flag students (or teams) who receive conflicting or low feedback, reviewers who are flagged by community (described above), and students (or teams) who request consultation from within the review system. Using specific alerts, instructors will be able to manage their course effort in providing feedback, or carrying out interventions with only those students or teams that need them.

5.3.6 Future of Peerval development

Peerval is a work-in-progress system, with the ultimate goals of managing assignments, tracking student progress through their academic careers, and providing data analytics to help instructors optimize their efforts in class. The following features will be implemented in the next phase of development:

5.3.6.1 Requesting specific and customized feedback

Students submitting their work, will be able to request specific feedback from their reviewers in addition to instructor provided rubric. Peerval will provide two ways to request this feedback: additional questions and ratings, or, selecting to upload two-five items for comparative feedback using default or student generated criteria. Allowing students to create their own assessment criteria has the potential to improve student engagement and participation, while also helping them receive specific feedback they are interested in. Furthermore, comparative feedback may allow students to receive feedback on parallel prototypes or converge on a few ideas from multiple solutions with the help of their peers. Figure 5-3 showcases one of several review viewing features in Peerval.

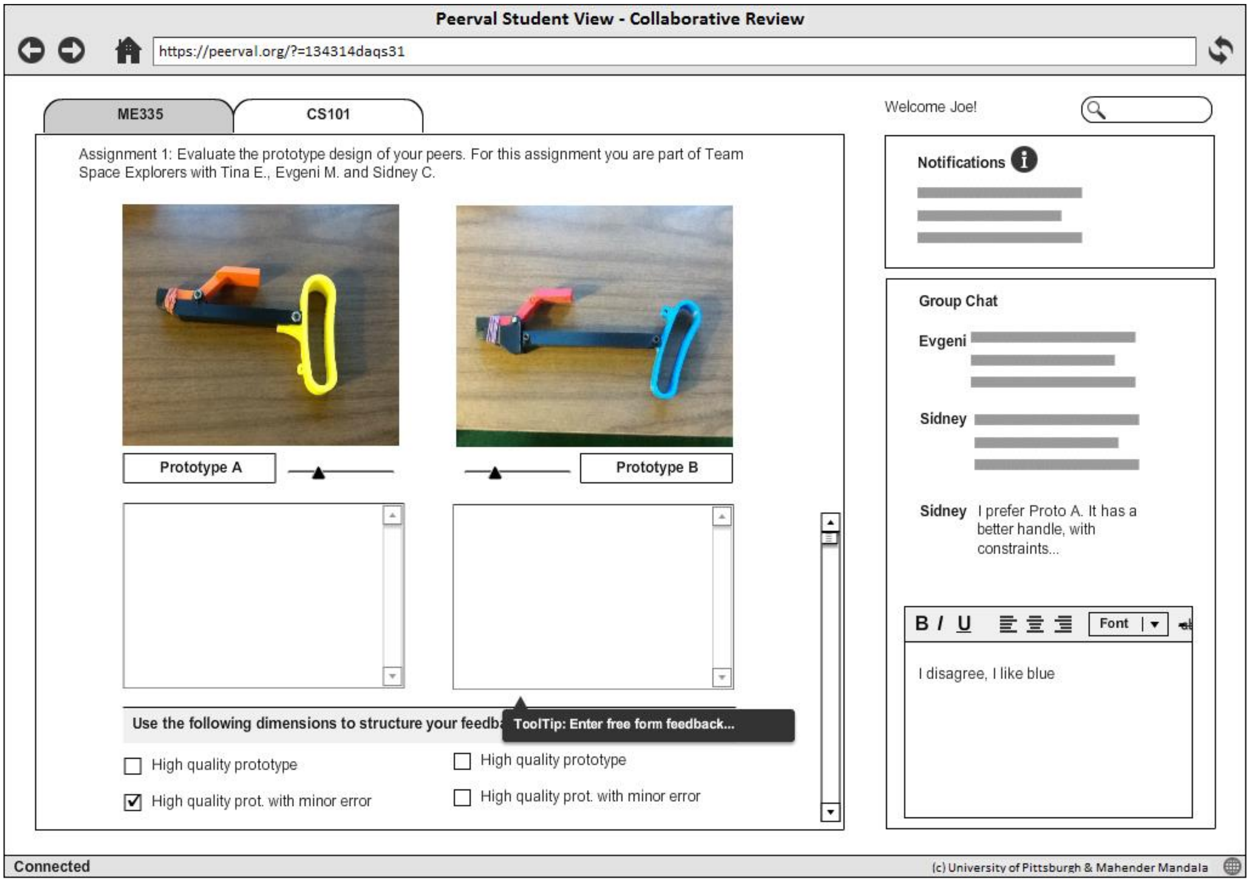

Figure 5-3. Peerval wire-frame showing reviewer view of student generated review criteria.

5.3.6.2 Machine learning and optimizing reviews

Peerval is being designed to collect a large amount usage data, including time spent on individual review questions, the helpfulness of reviews on certain rubric dimensions, dimensions that yield conflicting feedback, historical ratings of reviewers on specific dimensions etc. This information will be utilized to create a "learning system" that actively modifies the rubrics reviewers view, or provide feedback to the instructors on deploying rubrics in future classes. Optimizing rubric questions will help alleviate student burden and increase feedback that students perceive valuable.

5.3.6.3 Engaging students with game like incentives

Reviewers in Peerval, will receive community points for conducting reviews (receive bonus points for helpful reviews). Students will be assigned a minimum number of required reviews by the instructors. As the reviews are completed, those student submissions that receive conflicting reviews or less number of reviews than estimated per submission, will be made available for review by any reviewer in the system. Over time, students will be upgraded through levels, each providing a set of benefits to the reviewers: bonus grades, ability to monitor and comment on other reviews, ability to consult with teaching team on best practices in classes, etc. Furthermore, reviewers will be able to pick specific rubric elements they are interested in providing feedback on, after they complete their required minimum reviews.

Using Peerval over the academic career of students, could help keep them engaged and involved in community activities that benefit their peers. Additionally, community points accrued by users could potentially help determine overall engagement. Those that reach a large enough threshold, could be labeled “lead users” (the reader is directed to review Von Hipple (1986)) of the system and recruited to improve the assessment criteria or other review related items in the course.

5.3.6.4 Feature creep and limitations of research based product development

Developing products focused solely on research ideals can lead to feature creep, i.e. expansive addition of features deviating from the original scope of the product mission, and disconnect from practical guidelines. Including several features can extend the learning curve for tool usage among the primary stakeholder, instructors. Without instructor buy-in, such a tool could quickly add to the peer review facilitation burden and eventually lead to abandoned. The future development of Peerval and addition of features described above will be systematically evaluated with feedback from students and instructors at each juncture.

Nevertheless, basic research in the field will enable development of innovative features. Which when properly vetted and prioritized through a user-centered design approach, will increase the impact of Peerval on both the instructor and student, and consequently encourage adoption and widespread use of peer reviews in design-based learning classes.

5.4 CONCLUSION

In the first half of this chapter, we described the best practices in implementing peer reviews, developed as part of deploying peer reviews in an engineering design classroom and examining contemporary literature. The second half of the chapter described the system characteristic of a new peer review web tool, Peerval. This system supports collaborative peer review (where teams of students work collectively to generate a single collaborative review) and provides tools to enhance instructors' efficiency in managing design classes and feedback.

Peer reviews by themselves are not a panacea to all instructor and classroom problems. Effective implementation of peer reviews requires faculty buy-in, intuitive facilitation tools, creating a favorable course climate, and re-examining assessment criteria. The work presented in this chapter supports the implementation of formal peer review methodology within design-based learning classrooms. Furthermore, it forms a basis of support for future work in translating peer review research to classroom and development of computer supported peer review tools.

5.5 ACKNOWLEDGEMENT

The authors would like to thank Dr. William Clark and Dr. Irene Mena, for their assistance in conducting this research work.

6.0 CONCLUSION AND DISSERTATION CONTRIBUTIONS

This dissertation presented the experimental examination of a collaborative peer review generation methodology with a goal to enhance the use and impact of peer reviews in classrooms. This work is situated in design-based learning, a field naturally supported by cooperative learning, formative feedback, and peer critiques. Design education continues to remain behind the curve in the use of technology to support, enhance, and even scale, the critical aspects of its pedagogy described above.

A major issue that continues to impede effective design education is faculty incentive, which often determines the level of engagement faculty have with their classes, and most noticeably impacting feedback provision in classes. This issue largely impacted faculty at research-focused university, where tenure incentives favors research over teaching. Nevertheless, faculty interviewed in this work (N=39), positively viewed peer reviews yet were hesitant to implement them citing the difficulties in facilitating peer reviews and concerns with the value of feedback generated using such reviews. The complexity of design assignments and projects eroded faculty confidence in using peer reviews, while the potential learning opportunities that reviewing provides in addition to widening the diversity of viewpoints in feedback, encouraged their usage. Overall, faculty felt they were providing less frequent, detailed, and multi-perspective feedback to students in their design classes. These findings support the use of peer reviews to counter the issues raised by faculty, ensuring such reviews produce helpful feedback for students while reducing instructor burden.

Collaborative peer review structuring, where teams of collocated students collectively and synchronously review peer projects, addressed both the poor peer feedback quality in design, and student participation and engagement in peer reviews when implemented in a sophomore level engineering design classroom (N=287). Collaborative teams, meeting in person, generated significantly better quality feedback, found the reviews no more taxing than current standard of independent individual reviewing, and resulted in an overwhelmingly positive student perception of the process. Furthermore, students believed that team reviews made them better individual reviewers in subsequent reviewers. Students cited the group coordination costs as a major limitation of using team reviews. These results support examining collaboration in peer reviews in the context of its impact on student learning and performance.

In support of translating the peer review research to design-based learning classrooms, several recommendations have been put for the in this dissertation work. To summarize, effective peer review implementation requires:

1. Classroom set up and environment
    a. Instructor commitment and positive attitude towards peer reviews.
    b. Setting clear expectations with the process.
    c. Creating a positive dependence on peer reviews within the course.
    d. Focusing on reviews as a source of feedback to iterate and not as grades.
    e. Providing adequate time for students to implement feedback.
2. Using grades and collaborative reviews as incentives
    a. Graded on completing review task (minimum of 15% of total grade).

    b. Using collaborative team reviews on complex assignments, novice reviewers, or when new assignment or rubrics are used.
3. Clearly defining assessment criteria
    a. Involving students in developing rubrics and other criteria.
    b. Using ratings with open ended feedback comments.
4. Choose assignment types that benefit peer reviews
    a. Should encourage iteration.
    b. Include writing and creative multimedia.
5. Constantly iterate peer review implementation
    a. Gather feedback from students using surveys, reflections, interviews etc.

Finally, in order to advance the use of computer supported peer review, a set of design criterion for a novel peer review system, Peerval, were presented. Peerval is being designed bottom up for design-based learning classrooms with built-in support for virtual collaboration in review generation and design artifacts that could enhance the review process (e.g. ability to display 3D models, video, and text within the same review panel).

6.1 DISSERTATION CONTRIBUTION

The contributions of this dissertation work span several dimensions and provide a base for new line of peer review research. The following details the significance of the work presented:

1. Highlights the challenges instructors and design-education, in general, face today.
2. Peer review strategy of using a collaborative team of reviewers to engage and enhance participation, and improve feedback quality.

3. Methodological approaches to study feedback using coding scheme to measure impact on performance and NASA TLX survey tool to measure peer review task related student effort.
4. Guidelines that instructors can use to implement peer reviews in design-based classrooms.
5. Computer system design to accommodate team reviews and enhance peer review process in general.

6.1.1 Generalizability and impact beyond engineering design

The current work focused on using peer reviews in the context of design-based classrooms. However, there are significant implications of the work that have the potential to impact the use of peer reviews beyond the current context. Design poses complex challenges, mixing fact-based domain knowledge and creative conceptualization that necessitates the use of diverse perspectives and collaboration in teams. It seems a natural fit to the use of collaborative team of reviewers in any peer review process utilized in evaluating design. Contemporary web-based peer review processes, in general, often end up adding complexities to any review task by reducing communication to a passive written document, using confusing or unclear assessment criteria, and using hesitant students attempting to unravel the expectations of the process and their role. As such, it can be deduced, based on the results reported in this dissertation, that a collaborative review approach could enhance the review experience in other domains as well, especially when reviewers have little to no experience in reviewing. Early basic research work by Zhu and colleagues (2014), in which a virtual crowd of reviewers working together performed better than individual reviewers on several reviewing tasks including mathematics problems an writing, adds support to this hypothesis.

It is important to note that students had an overwhelmingly positive experience with collaborative team reviews, even with the added burden of conducting the reviews in-person collectively. Furthermore, students picked collaborative team reviews over individual reviews when given a choice. These results indicate the potential of such structuring on engaging students and, potentially, increase their participation in the process.

6.2 FUTUREWORK

This dissertation presents several opportunities for future research and development in enhancing web-based peer reviews. The following summarizes potential future research work:

6.2.1 Impact of collaboration in review generation on feedback typology

In the current work, feedback was examined on quality and appropriateness. Quality was determined by an expert (instructor of the class) and based on its estimated impact on grade when implemented. As such, the quality of feedback as viewed by the student or team receiving this feedback could vary significantly. Future work could compare quality of feedback from two perspectives: student and expert. Understanding the similarities and dissimilarities between the two, can enhance researcher understanding of what constitutes helpful feedback, and potentially uncover strategies to further improve the quality and helpfulness of feedback generated by peers and experts.

Additionally, feedback can be examined under different lenses, i.e. categories beyond a singular quality metric (Dannels & Martin, 2008; Farshid Marbouti et al., 2014; Patchan et al., 2009). Examining feedback under different perspectives can highlight the impact of collaboration, and consequently, help better understand the processes that are invoked when students review their peers' work under different review structures.

6.2.2 Can virtual collaboration mimic in-person collaborative review generation?

Students in this study were requested to conduct collaborative team reviews in-person, collectively. Collocation of review team may have provided several beneficial impacts on the review generated including allowing students to engage in oral and visual communication with their team members, and improve focus on review tasks. However, meeting in-person creates additional burden on students with varying schedules and pressures. Virtual collaboration has the potential to resolve some of these issues. Additionally, virtual collaboration can help identify individual contribution within a team review through digital archiving of discussion, feedback generation etc. Future work could look at virtual collaboration in reviewing compared to in-person collaboration. Findings could help identify areas for improvement in every aspect of the peer review process, including establishing design criterion for computing systems to effectively enhance peer review process.

6.2.3 Determining review artifacts that improve feedback quality and helpfulness

Design uses many communication methods, spanning sketches, models, text, and other multimedia. It is currently unclear which combination of artifacts evokes high quality and helpful feedback. Future work could look at experimentally comparing several artifacts, noting the feedback is evokes, and student perspectives on their usage. The outcome of this work has the potential to impact structuring of peer reviews, course design, and computing system design. Furthermore, the use of various feedback provision methods, e.g. written, audio, or annotated sketch, could be examined to determine the best fit methods that yield improved outcomes in using peer reviews online.

6.2.4 Other review structuring methodologies and their impact

There are additional outstanding issues that could form the basis for future research work, including:

1. Incentive structures and their impact on student engagement, participation, and learning.
2. Assessment criteria design, specifically using students to develop the criteria.
3. Frequency and content of peer reviews and its impact on iteration in design.
4. Validation of NASA TLX tool to measure student effort within the context of peer reviews.

APPENDIX A

SURVEY QUESTIONNAIRE

A.1 PRE-COURSE SURVEY

Thank you for participating in this research study of evaluating peer evaluation in design based learning classes. This study is being conducted by the University of Pittsburgh. The following survey will provide us with important information about your view of, and experience with, design based learning classes, peer evaluation and some general information about you. Your participation is voluntary and you can withdraw at any time. Your answers will be kept strictly confidential. The data collected from this survey will be stored in a secure data facility, and results of the survey will be reported in summary or statistical form only; no individuals will be identified. To minimize any risk of breach of confidentiality, the survey data will be coded and your identifiable information removed. There are no direct benefits to you. Students who complete the surveys and participate in the study will have the opportunity to receive 2 points/100 extra credit. Students who choose not to participate in the study can still be eligible to receive 2 points/100 extra credit by completing a 1 page reflection, focused on one aspect of the course (e.g. peer review, design project submission, CAD labs, etc.) and present an alternative, along with an explanation for why they feel that would be more effective. This study is being directed by Professor Jonathan Pearlman at the University of Pittsburgh, who can be reached at 412-822-3700 if you have any questions. This survey will take about 10 minutes to complete.

1. Please enter your name (first last) : [This will ensure you receive the extra credit at the end of the course]
2. The following questions deal with your experience with peer feedback process. Peer feedback is the process of using peers from the same class to provide feedback to each other on assignments or presentations using a rubric created by the class instructor.
3. Prior to this class, I participated in peer-evaluation, where I provided feedback to my classmates and received feedback from my classmates on… (please select all that apply)

	Once in a class	Twice or more in the same class	None/Never
Presentations			
Assignments			
Project Work			

4. Please share your opinion about receiving feedback from your peers in the past

	Strongly Disagree	Disagree	Neither Agree	Agree	Strongly Agree
I got more feedback from peers than I would have from a teacher.					
It was helpful to get feedback from peers					

It was helpful to get feedback from multiple people					
The feedback I received was generally helpful					
The feedback pointed out ways for me to improve my ideas					
The feedback I received was often wrong or not useful					
I think it is more helpful to receive feedback only from the teacher					

5. Please share your opinion about providing feedback to your peers in the past

	Strongly Disagree	Disagree	Neither Agree	Agree	Strongly Agree
I could use what I reviewed to improve my own work					
I learned from seeing strengths in my peers' work					
Reviewing helped me see weaknesses in my own work					
I thought reviewing as many items as my teacher asked me to, was an unreasonable amount of work.					

6. Considering your previous experience with peer evaluation, how interested/excited are you in having your work reviewed by your peers and also providing feedback on your peers' work in this class?

- Not at all interested or excited
- Somewhat interested or excited
- Highly interested or excited

7. Considering your previous experience with peer evaluation, for each of the following statements, please indicate how true it is for you:

	Not at all true (1)	2	3	Somewhat true (4)	5	6	Very True (7)
I enjoyed doing this activity very much							
I think I am pretty good at this activity							
I put a lot of effort into this							
I felt pressured while doing these reviews							
I did this activity because I had no choice							
I think doing this activity is useful for my personal growth and learning							

8. I have used peer assessment software tools in the past … (example: PeerMark, SWoRD, Arrow, etc.)

- Yes [Please mention the name of the tool as best as you remember it] ____________________
- No

9. Please use the space below to clarify any of your responses or share information with the researchers:

The following questions deal with your experience and feelings about engineering design. Please complete all items on the next four set of questions.

10. Rate your degree of CONFIDENCE (i.e. belief in your current ability) to perform the following tasks by recording a number from 0 to 100. (0=cannot do at all; 50=moderately can do; 100=highly certain can do)

	0	10	20	30	40	50	60	70	80	90	100
conduct engineering design											
identify a design need											
research a design need											
develop design solutions											
select the best possible design											

construct a prototype											
evaluate and test a design											
communicate a design											
redesign											

11. Rate how MOTIVATED you would be to perform the following tasks by recording a number from 0 to 100. (0=not motivated; 50=moderately motivated; 100=highly motivated)

	0	10	20	30	40	50	60	70	80	90	100
conduct engineering design											
identify a design need											
research a design need											
develop design solutions											
select the best possible design											
construct a prototype											
evaluate and test a design											
communicate a design											
redesign											

12. Rate your degree of ANXIETY (how apprehensive you would be) in performing the following tasks by recording a number from 0 to 100. (0=not anxious at all; 50=moderately anxious; 100=highly anxious)

	0	10	20	30	40	50	60	70	80	90	100
conduct engineering design											
identify a design need											
research a design need											
develop design solutions											
select the best possible design											
construct a prototype											
evaluate and test a design											
communicate a design											
redesign											

Thank you for being awesome! This is the last page...The following questions will help the researchers understand your background.

13. What is your age as of September 1, 2015 in years? (e.g., 28):

14. What is your race/ethnicity?

- American Indian or Alaskan Native

- Asian
- Black or African American
- Hispanic or Latino
- Native Hawaiian or Pacific Islander
- White
- Prefer not to answer/Don’t know

15. What type of high school did you attend?

- Public district school (not a magnet school)
- Public magnet school
- Charter school
- Private school (independent, parochial or proprietary)
- Home school
- Prefer not to answer/Don’t know
- Other: ____________________

16. In my K-12 years… (please select all that apply)

- I have taken only core or required STEM courses
- I have taken AP/Honors or other specialized STEM courses
- I was part of a science club after school or on the weekends
- Prefer not to answer

17. What is your focus area (major), as of September 1, 2015?

- Bioengineering
- Chemical & Petroleum Engineering
- Civil & Environmental Engineering
- Electrical & Computer Engineering
- Industrial Engineering
- Mechanical Engineering & Material Science
- Undecided
- Prefer not to answer
- Other (please specify): ____________________

18. In which year of the program are you at, as of September 1, 2015?

- Freshman
- Sophomore
- Junior
- Senior
- 5th year senior
- Not yet enrolled
- Prefer not to answer

Thank you for completing this survey! Please use the text box below to clarify your responses or provide additional information to the researchers.

A.2 POST-ASSIGNMENT 1 SURVEY

Thank you for participating in this research study of evaluating peer evaluation in design based learning classes. This study is being conducted by the University of Pittsburgh. The following survey will provide us with important information about your view of, and experience with, design based learning classes, peer evaluation and some general information about you. Your participation is voluntary and you can withdraw at any time. Your answers will be kept strictly confidential. The data collected from this survey will be stored in a secure data facility, and results of the survey will be reported in summary or statistical form only; no individuals will be identified. To minimize any risk of breach of confidentiality, the survey data will be coded and your identifiable information removed. There are no direct benefits to you. Students who complete the surveys and participate in the study will have the opportunity to receive 2 points/100 extra credit. Students who choose not to participate in the study can still be eligible to receive 2 points/100 extra credit by completing a 1 page reflection, focused on one aspect of the course (e.g. Peer review, design project submission, CAD labs, etc.) and present an alternative, along with an explanation for why they feel that would be more effective. This study is being directed by Professor Jonathan Pearlman at the University of Pittsburgh, who can be reached at 412-822-3700 if you have any questions. This survey will take about 10 minutes to complete.

1. Please enter your name (first last) : [This will ensure you receive the extra credit at the end of the course]

NASA-TLX is a subjective workload assessment tool. NASA-TLX is a multi-dimensional rating procedure that derives an overall workload score based on a weighted average of ratings on six sub-scales. Scales of this sort are extremely useful but their utility suffers from the tendency people have to interpret them in individual ways. The evaluation you are about to perform is a technique that has been developed by NASA. The procedure is simple: You will be presented with a series of pairs of rating scale titles and asked to choose which of the items was more important to your experience of workload in the tasks that you just performed (peer evaluation).Mental Demand: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching etc.)? Physical Demand: How much physical activity was required (e.g., pushing, controlling, activating etc.)? Was the peer review task easy or demanding, strenuous or slack? Temporal (time) Demand: How much time pressure did you feel due to the rate or pace at which the peer review tasks or task elements occurred? Was it slow, leisurely, rapid, frantic? Performance: How successful do you think you were in accomplishing the goals of the peer review tasks? Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance? Frustration Level: How insecure, discouraged, irritated, stressed, content, relaxed and complacent did you feel during the task? Please ignore the technical difficulties of using SWORD and focus on the peer review task itself.

2. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (1/15)

- Effort or Performance

3. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (2/15)

- Temporal (time) Demand or Frustration

4. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (3/15)

- Temporal (time) Demand or Effort

5. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (4/15)

- Physical Demand or Frustration

6. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (5/15)

- Performance or Frustration

7. Click on the scale title that represents the more important contributor to workload for the peer review task you performed.(6/15)

- Physical Demand or Temporal (time) Demand

8. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (7/15)

- Physical Demand or Performance

9. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (8/15)

- Temporal (time) Demand or Mental Demand

10. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (9/15)

- Frustration or Effort

11. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (10/15)

- Performance or Mental Demand

12. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (11/15)

- Temporal (time) Demand or Performance

13. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (12/15)

- Effort or Mental Demand

14. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (13/15)

- Mental Demand or Physical Demand

15. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (14/15)

- Effort or Physical Demand

16. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (15/15)

- Mental Demand or Frustration

17. Rating Sheet - please slide the pointer to the left or right representing low or high respectively.

______ Mental Demand

______ Physical Demand

______ Temporal Demand

______ Performance

______ Effort

______ Frustration

Focusing on the peer review task you were assigned, please answer the following questions:

18. Overall, how much time did you spend on all of the review tasks? (enter HH:MM ; e.g., 01:30 for one hour and 30 minutes)

19. How interested/excited are you in having your work reviewed by your peers and also providing feedback on your peers' work in this class?

- Not at all interested or excited
- Somewhat interested or excited
- Highly interested or excited

For each of the following statements, please indicate how true it is for you, using the scale provided:

	Not at all true (1)	2	3	Somewhat true	5	6	Very true (7)
I enjoyed doing these reviews very much							
I think I am pretty good at evaluating my peers' work							
After working at this activity for a while, I felt pretty competent							
I put a lot of effort into these reviews							
It was important to me to do well at this activity							
I am satisfied with my performance at this activity							
I felt pressured while doing these reviews							
I did these reviews because I had no choice							
I think doing this activity is useful for my personal growth and learning							
I think this is an important activity							

20. How was your experience working with your team? (select one closest to your experience)

- Team worked very well, everyone participated and added value to the work
- Team mostly worked well, everyone participated most of the time
- Team worked well sometimes, individual participation varied
- Team had some difficulty working together, some members had little to no participation

- Team had a lot of difficulties working together, most members provided no value addition to the team

Additional Comments? Please use the box below to provide any additional comments or feedback!

Thank you for helping us better this class and the peer-review process!

A.3 POST-ASSIGNMENT 2 SURVEY (FINAL)

Thank you for participating in this research study of evaluating peer evaluation in design based learning classes. This study is being conducted by the University of Pittsburgh. The following survey will provide us with important information about your view of, and experience with, design based learning classes, peer evaluation and some general information about you. Your participation is voluntary and you can withdraw at any time. Your answers will be kept strictly confidential. The data collected from this survey will be stored in a secure data facility, and results of the survey will be reported in summary or statistical form only; no individuals will be identified. To minimize any risk of breach of confidentiality, the survey data will be coded and your identifiable information removed. There are no direct benefits to you. Students who complete the surveys and participate in the study will have the opportunity to receive 2 points/100 extra credit. Students who choose not to participate in the study can still be eligible to receive 2 points/100 extra credit by completing a 1 page reflection, focused on one aspect of the course (e.g. peer review, design project submission, CAD labs, etc.) and present an alternative, along with an explanation for why they feel that would be more effective. This study is being directed by Professor Jonathan Pearlman at the University of Pittsburgh, who can be reached at 412-822-3700 if you have any questions. This survey will take about 10 minutes to complete.

1. Please enter your name (first last) : [This will ensure you receive the extra credit at the end of the course]

The following questions deal with your experience and feelings about engineering design. Please complete all items on the next three set of questions.

2. Rate your degree of CONFIDENCE (i.e. belief in your current ability) to perform the following tasks by recording a number from 0 to 100. (0=cannot do at all; 50=moderately can do; 100=highly certain can do)

	0	10	20	30	40	50	60	70	80	90	100
conduct engineering design											
identify a design need											
research a design need											
develop design solutions											
select the best possible design											
construct a prototype											
evaluate and test a design											
communicate a design											
redesign											

3. Rate how MOTIVATED you would be to perform the following tasks by recording a number from 0 to 100. (0=not motivated; 50=moderately motivated; 100=highly motivated)

	0	10	20	30	40	50	60	70	80	90	100
conduct engineering design											
identify a design need											
research a design need											
develop design solutions											
select the best possible design											
construct a prototype											
evaluate and test a design											
communicate a design											
redesign											

4. Rate your degree of ANXIETY (how apprehensive you would be) in performing the following tasks by recording a number from 0 to 100. (0=not anxious at all; 50=moderately anxious; 100=highly anxious)

	0	10	20	30	40	50	60	70	80	90	100
conduct engineering design											
identify a design need											
research a design need											
develop design solutions											

select the best possible design											
construct a prototype											
evaluate and test a design											
communicate a design											
redesign											

NASA-TLX is a subjective workload assessment tool. NASA-TLX is a multi-dimensional rating procedure that derives an overall workload score based on a weighted average of ratings on six sub-scales. Scales of this sort are extremely useful but their utility suffers from the tendency people have to interpret them in individual ways. The evaluation you are about to perform is a technique that has been developed by NASA. The procedure is simple: You will be presented with a series of pairs of rating scale titles and asked to choose which of the items was more important to your experience of workload in the tasks that you just performed (peer evaluation).Mental Demand: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching etc.)? Physical Demand: How much physical activity was required (e.g., pushing, controlling, activating etc.)? Was the peer review task easy or demanding, strenuous or slack? Temporal (time) Demand: How much time pressure did you feel due to the rate or pace at which the peer review tasks or task elements occurred? Was it slow, leisurely, rapid, frantic? Performance: How successful do you think you were in accomplishing the goals of the peer review tasks? Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance? Frustration Level: How insecure, discouraged, irritated, stressed, content, relaxed and complacent did you feel during the task? Please ignore the technical difficulties of using SWORD and focus on the peer review task itself.

5. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (1/15)

- Effort or Performance

6. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (2/15)

7. Temporal (time) Demand or Frustration

Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (3/15)

- Temporal (time) Demand or Effort

8. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (4/15)

- Physical Demand or Frustration

9. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (5/15)

- Performance or Frustration

10. Click on the scale title that represents the more important contributor to workload for the peer review task you performed.(6/15)

- Physical Demand or Temporal (time) Demand

11. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (7/15)

- Physical Demand or Performance

12. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (8/15)

- Temporal (time) Demand or Mental Demand

13. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (9/15)

- Frustration or Effort

14. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (10/15)

- Performance or Mental Demand

15. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (11/15)

- Temporal (time) Demand or Performance

16. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (12/15)

- Effort or Mental Demand

17. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (13/15)

- Mental Demand or Physical Demand

18. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (14/15)

- Effort or Physical Demand

19. Click on the scale title that represents the more important contributor to workload for the peer review task you performed. (15/15)

- Mental Demand or Frustration

20. Rating Sheet - please slide the pointer to the left or right representing low or high respectively.

   ______ Mental Demand

   ______ Physical Demand

   ______ Temporal Demand

   ______ Performance

   ______ Effort

   ______ Frustration

Focusing on the Peer Review - 2 you were assigned recently, please answer the following questions:

21. Overall, how much time did you spend on all of the review tasks? (enter HH:MM ; e.g., 01:30 for one hour and 30 minutes)
22. How interested/excited are you in having your work reviewed by your peers and also providing feedback on your peers' work in this class?

- Not at all interested or excited
- Somewhat interested or excited
- Highly interested or excited

For each of the following statements, please indicate how true it is for you, using the scale provided:

	Not at all true (1)	2	3	Somewhat true	5	6	Very true (7)
I enjoyed doing these reviews very much							

I think I am pretty good at evaluating my peers' work							
After working at this activity for a while, I felt pretty competent							
I put a lot of effort into these reviews							
It was important to me to do well at this activity							
I am satisfied with my performance at this activity							
I felt pressured while doing these reviews							
I did these reviews because I had no choice							
I think doing this activity is useful for my personal growth and learning							
I think this is an important activity							

Considering your experience over all, please answer the following questions that deal with peer review assignments.

23. Please share your opinion about receiving feedback from your peers in this class

	Strongly Disagree	Disagree	Neither Agree	Agree	Strongly Agree
We got more feedback from peers than we would have from a teacher					
It was helpful to get feedback from peers					
The feedback helped me (us) see where my (our) work was unclear to my (our) audience					

It was helpful not to know the reviewers' real names					
It was helpful that the reviewers didn't know my (or teams') real name					
It was helpful to get feedback from multiple people					
The feedback we received was generally helpful					
I didn't know what to do when some feedback contradicted what others said					
There was too much feedback to use it all					
The feedback pointed out ways for me (us) to improve my (our) ideas					
We used feedback we received to revise additional drafts of our work.					
The feedback was mostly useful for correcting minor technical aspects of our work (not ideas)					
The feedback I received was often wrong or not useful					
I think it is more helpful to receive feedback only from the teacher					

24. Please share your opinion about providing feedback to your peers in this class

	Strongly	Disagree	Neither	Agree	Strongly
I could use what I reviewed to improve my own work					
I learned from seeing strengths in my peers' work					

Reviewing helped me see weaknesses in my own work					
I understood what I was supposed to give feedback on					
I understood how to rate/score the work of my peers					
I gave helpful feedback to my peers					
It was hard to give criticism in a nice way					
I thought reviewing as many items as my teacher asked me to, was an unreasonable amount of work					

25. Did you or your team participate in the team review research, where you were asked to meet together and complete the reviews as a team?

- Yes
- No

Please answer the following questions related to your experience reviewing the projects in two different settings – as a team and individually.

26. Please share your opinion about providing feedback as a collaborative team vs individually

	Strongly Disagree	Disagree	Neither Agree	Agree	Strongly Agree
I enjoyed providing feedback on projects as part of a team					
I was able to provide more honest feedback when I completed the					

review task myself instead of as a team					
Providing feedback as a team, allowed me to become a better reviewer					
It was easier to complete the reviews myself instead of working on them with the team					
The quality and amount of feedback we generated as a team was more than I could have done by myself					

27. Please describe any strategy you used to complete the review tasks (e.g., reading peer work before meeting with the team, reviewing all assigned documents and then providing feedback at the same time etc.):

- Individually
- As a team

28. You have completed two review tasks both as a team and as individuals (peer review 1 and peer review 2), comparing the two tasks, which one do you prefer based on:

	Team Review	Individual Review
Effort required to complete the review		
Your performance on the review task		
Your personal learning gain		

29. Optional comments on team vs individual review experience:

30. In general, how was your experience working with your team? (select one closest to your experience)

- Team worked very well, everyone participated and added value to the work
- Team mostly worked well, everyone participated most of the time
- Team worked well sometimes, individual participation varied
- Team had some difficulty working together, some members had little to no participation
- Team had a lot of difficulties working together, most members provided no value addition to the team

31. Additional Comments? Please use the box below to provide any additional comments or feedback!

Thank you for helping us better this class and the peer-review process!

APPENDIX B

PEER REVIEW RUBRIC

B.1 ASSIGNMENT 1 RUBRIC

B.1.1 Hypothesis

Base your feedback on the following questions/prompts: Did the group correctly identify all design requirements, constraints, limitations and features of the chosen project? Based on the client statement and the requirements, are the hypotheses acceptable? Is there anything that has been overlooked and not taken into consideration? Do you think the group understood the client statement and requirements?

Rate the prototype in terms of how well did the team understand the project requirements and form hypotheses?

7 - Excellent. The team understood all requirements and formed an appropriate hypothesis.

6 - Very Good. The team formed a good hypothesis but minor improvements can be made.

5 - Good. The team formed an appropriate hypothesis but some major improvements can be made.

4 - Acceptable.

3 - Fair. The hypothesis needs some major improvement.

2 - Poor

1 - Very Poor. The hypothesis formed did not relate to the requirements or will not be helpful in the design process.

B.1.2 User Discovery

Base your feedback on the following questions/prompts: Based on the hypotheses and client statement, did the group ask appropriate questions to the potential users? Are the questions well thought and clear for the interviewees to understand? Did the group identify the appropriate potential users? Did you identify anyone in the list of people interviewed who are most likely not potential users?

Rate the team on how well do you think the team did in initial user discovery?

7 - Excellent. The team identified the potential users correctly and asked excellent questions, and captured the "user story".

6 – Very Good. The team identified the potential users correctly and asked some good questions. However, the team did not ask enough questions.

5 - Good. Potential users were identified correctly and appropriate questions were asked. Some questions were irrelevant and can be constructed better.

4 - Acceptable.

3 - Fair. Some of the users identified may not be potential users and questions were irrelevant or not helpful for the design process.

2 - Poor.

1 - Very Poor. The team did not identify the correct group of users and asked poorly constructed and irrelevant questions.

B.1.3 Idea Generation

Base your feedback on the following questions/prompts: Did the group re-write the client statement based on the interviews? Are the changes appropriate? Did the group demonstrate an appropriate level of maturity in interpreting and understanding user feedback? When generating ideas, did the team sufficiently discuss and take into account all of their user feedback when coming up with ideas? Were multiple methods used to generate ideas? Did the team seek additional solutions beyond the first potential concepts?

Rate how well the team used brainstorming and other techniques to generate a rich set of possible design solutions and features?

7 - Excellent. The team brainstormed well using all the user input and came up with a variety of great solutions and features that addressed all project requirements.

6 - Very Good. The team brainstormed well and designed great solutions and features that meet all the project requirements. User input however was under-utilized.

5 - Good. The team brainstormed well and designed some good solutions. Some minor improvements can be made and some user input might have been wrongly interpreted or neglected.

4 - Acceptable.

3 - Fair. Some good solutions and some bad solutions. Solutions did not have a great variety and user input was not used at all.

2 - Poor.

1 - Very Poor. The team did not brainstorm well, created bad solutions and did not understand nor use user input at all. Solutions did not show any creativity and variation.

B.1.4 Preliminary Designs

Base your feedback on the following questions/prompts: Were the candidate solutions reasonable attempts at meeting the problem statement? Were they distinct enough from each other to represent fundamentally different ideas? Did they demonstrate an appropriate level of technical maturity? Could you understand the design intent? Were they adequately evaluated against each other?

Rate how well the designs presented represent different and meaningful attempts at solving the problem:

7 - Excellent work with multiple, distinct approaches to solving the problem.

6 - Very good work. Distinct concepts.

5 - Minor shortcomings in either the quality or creativity of the design candidates, but still good quality work.

4 - Shortcomings in creativity, distinctiveness of solutions, or addressing the problem statement.

3 - Fair.

2 - Major shortcomings in the quality of the candidate designs. Appropriate level of design sophistication not conveyed.

1 - Very major shortcoming with candidate designs that lack technical maturity level in either design generation or presentation.

B.1.5 Presentation of Designs

Base your feedback on the following questions/prompts: Did the level of detail and images used in presenting the chosen design adequately convey the design intent? Did the models emphasize the important parts of the design? Do you feel that another engineer could pick up the concept based on this presentation?

Rate how well did the sketches convey the design intent:

7 - Excellent. Appealing visual design; highly informative graphics, appropriately placed, and easy to interpret.

6 - Very good. Clearly conveyed concepts that are easy to understand.

5 - Good. Sketches are informative but the quality and understanding could be improved.

4 - Reasonable sketches, but 2 or more are difficult to understand.

3 - Fair. Several design features cannot be seen or understood clearly.

2 - Poor.

1 - Very poor. Sketches are not clear and design intent cannot be understood.

B.2 ASSIGNMENT 2 RUBRIC

B.2.1 Initial Prototype

Base your feedback on the following questions/prompts: Was it easy to understand the design function from the prototype? Did the team put a reasonable amount of effort into creating the prototype? Is the prototype a good representation of the design concept(s), and does it capture all of the design requirements?

Rate the prototype in terms of how well it meets minimum requirements (CAD model) and conveys the team' design ideas to the user:

7 - Excellent: The prototype is a high fidelity representation of the design idea (e.g. working physical mock-up or polished CAD model and/or animation) and is easy to understand

6 - Very Good: Reasonable representation that is above basic expectations (e.g. CAD model plus clay or cardboard mock-up) that enables me to understand the design concepts.

5 - Good: The prototype is a good representation of the design ideas (e.g. representative CAD model).

4 - Below Average: The prototype is a CAD model, but more detail would have been helpful.

3 - Fair: The prototype is a CAD model (or other prototype form) that leaves me wondering about some of the design features. I am not quite sure how the design works.

2 - Poor: The prototype is less than a CAD model and only weakly conveys the design concepts.

1 - Missing/Omitted: There was no prototype shown in the design log book.

B.2.2 User Discovery with Initial Prototype

Base your feedback on the following questions/prompts: Based on the hypotheses and client statement, did the group ask appropriate questions to the potential users to a) get assessment of their design from the Users' perspective and b) gather information to guide the final design choices? (Did the group identify the appropriate potential users? Are the questions well thought and clear for the interviewees to understand? From the questions, can the team get stories of (how, when, where, why) the user may or may not use the device?)

Rate the team's use of discovery to obtain information to guide the final design:

7 - Excellent: The team identified potential users correctly, asked excellent questions that captured the "user story" with respect to their device. In the process, they gathered valuable information for refining the design.

6 – Very good.

5 - Good: The team identified potential users correctly and asked good questions.

4 - Below Average: Potential users were mostly identified correctly and questions were asked, but the questions could have been better directed to obtain user stories with respect to this design.

3 - Fair: Some of the users identified may not be potential users and questions were irrelevant or not helpful for the design process.

2 - Poor: The team did not identify the correct group of users and/or asked poorly constructed and irrelevant questions.

1 - Missing/Omitted: There was no User Discovery with Initial Prototype shown in the design log book.

B.2.3 Evolution of Design

Base your feedback on the following questions/prompts: How well did the team's final design evolve during the project. Consider: a) the preliminary designs – is the final design a clear evolution that incorporates the best of those ideas (or moves in a new direction that better addresses the problem statement)? and b) user feedback -- does the final design incorporate user feedback?

Rate the progress of the design from preliminary concepts to final design.

7 - Excellent: The final design is a clear evolution from the preliminary designs (or the design is a completely new and better concept than any of the preliminary designs) and clearly incorporates features that directly address user feedback.

6 - Very good.

5 - Good: The final design has reasonable components of evolution from the preliminary design set and shows evidence of incorporation of user feedback.

4 - Below Average: The design is basically the same as one of the preliminary designs with small improvements.

3 - Fair.

2 - Poor: Little to no evolution has been shown. The design ideas have not changed from preliminary designs, no incorporation of features is shown, and no user feedback has been used.

1 - Missing/Omitted: No final design is shown in the design log book.

B.2.4 Presentation of Designs

Base your feedback on the following questions/prompts: Watch the video and evaluate its creativity and how well it conveys: a) the need for the device or product, and who the prospective users will be, b) the specific design requirements, and c) the final design, including how it meets the requirements.

Rate the video on its creativity and effectiveness in presenting the design problem and solution:

7 - Excellent: The video is very creative and does an outstanding job of explaining the need for the design, the users, and what the specific design requirements are, and it clearly demonstrates how the final design meets those requirements.

6 - Very good.

5 - Good: The video is good, covers all required points, but not what I would consider very creative.

4 - Below Average: The video is lacking in one key area (need for the design, the users, specific design requirements, or demonstration of how the final design meets those requirements.)

3 - Fair: The video is lacking in two key areas. Note that a very creative video may fall into this category if it does not address the necessary points.

2 - Poor: The video is lacking in three or more key areas.

1 - Missing/Omitted: No video is provided.

BIBLIOGRAPHY

Althauser, R., & Darnall, K. (2001). Enhancing critical reading and writing through peer reviews: An exploration of assisted performance. *Teaching Sociology*, 23-35.

Altman, A., Dym, C. L., Hurwitz, R., & Wesner, J. W. (2012). The key ideas of mdw viii: A summary.

Andrade, H. G. (2005). Teaching with rubrics: The good, the bad, and the ugly. *College teaching, 53*(1), 27-31.

Atman, C. J., Adams, R. S., Cardella, M. E., Turns, J., Mosborg, S., & Saleem, J. (2007). Engineering design processes: A comparison of students and expert practitioners. *Journal of Engineering Education, 96*(4), 359.

Atman, C. J., Eris, O., McDonnell, J., Cardella, M. E., & Borgford-Parnell, J. (2014). Engineering design education: Research, practice and examples that link the two *Cambridge handbook of engineering education research*

Atman, C. J., Kilgore, D., & McKenna, A. (2008). Characterizing design learning: A mixed-methods study of engineering designers' use of language. *Journal of Engineering Education, 97*(3), 309-326.

Ballantyne, R., Hughes, K., & Mylonas, A. (2002). Developing procedures for implementing peer assessment in large classes using an action research process. *Assessment & Evaluation in Higher Education, 27*(5), 427-441.

Barley, W. C., Leonardi, P. M., & Bailey, D. E. (2012). Engineering objects for collaboration: Strategies of ambiguity and clarity at knowledge boundaries. *Human communication research, 38*(3), 280-308.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in education, 5*(1), 7-74.

Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in continuing education, 22*(2), 151-167.

Boud, D., Cohen, R., & Sampson, J. (1999). Peer learning and assessment. *Assessment & Evaluation in Higher Education, 24*(4), 413-426.

Boud, D. J., & Holmes, W. H. (1981). Self and peer marking in an undergraduate engineering course. *Education, IEEE Transactions on, 24*(4), 267-274.

Briggs, C. (2012). *Recent progress in design education.* Paper presented at the 50th AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition.

Cardella, M. E., Buzzanell, P., Cummings, A., Tolbert, D., & Zoltowski, C. B. (2014). A tale of two design contexts: Quantitative and qualitative explorations of student-instructor interactions amidst ambiguity.

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in higher education, 36*(4), 395-407.

Cho, K. (2004). *When multi-peers give better advice than an expert: The type and impact of feedback given by students and an expert on student writing.* (3149958 Ph.D.), University of Pittsburgh, Ann Arbor. ProQuest Dissertations & Theses Full Text database.

Cho, K., & MacArthur, C. (2011). Learning by reviewing. *Journal of Educational Psychology, 103*(1), 73.

Cho, K., & Schunn, C. (2003a). *Battling the tyranny of the thousands with a sword: Scaffolded writing and rewriting in the discipline.* Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications.

Cho, K., & Schunn, C. (2003b). *Validity and reliability of peer assessments with a missing data estimation technique.* Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications.

Cho, K., & Schunn, C. D. (2004). *The sword is mightier than the pen: Scaffolded writing and rewriting in the discipline.* Paper presented at the Advanced Learning Technologies, 2004. Proceedings. IEEE International Conference on.

Cho, K., & Schunn, C. D. (2007). Scaffolded writing and rewriting in the discipline: A web-based reciprocal peer review system. *Computers & Education, 48*(3), 409-426.

Cho, Y. H., & Cho, K. (2011). Peer reviewers learn from giving comments. *Instructional Science, 39*(5), 629-643.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438-481.

Cross, K. P., & Steadman, M. H. (1996). *Classroom research: Implementing the scholarship of teaching*: Jossey-Bass San Francisco.

Dannels, D. P. (2005). Performing tribal rituals: A genre analysis of "crits" in design studios. *Communication Education, 54*(2), 136-160.

Dannels, D. P., & Martin, K. N. (2008). Critiquing critiques a genre analysis of feedback across novice to expert design studios. *Journal of Business and Technical Communication, 22*(2), 135-159.

Dutton, T. A. (1987). Design and studio pedagogy. *Journal of architectural education, 41*(1), 16-25.

Dym, C. L. (2012). Engineering design@ hmc: A testament to innovation and entrepreneurship. *International Journal of Engineering Education, 28*(2), 240.

Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education, 94*(1), 103-120.

Dym, C. L., Sheppard, S. D., & Wesner, J. W. (2001). A report on mudd design workshop ii:" Designing design education for the 21st century". *Journal of Engineering Education, 90*(3), 291.

Dym, C. L., Wesner, J. W., & Winner, L. (2003). Social dimensions of engineering design: Observations from mudd design workshop iii. *Journal of Engineering Education, 92*(1), 105-107.

Ekoniak, M., Scanlon, M. J., & Mohammadi-Aragh, M. J. (2013, 23-26 Oct. 2013). *Improving student writing through multiple peer feedback.* Paper presented at the Frontiers in Education Conference, 2013 IEEE.

ABET. (2014). Criteria for accrediting engineering programs: Effective for reviews during the 2015-16 accreditation cycle. Baltimore, MD: Engineering Accreditation Commission

Ertmer, P. A., Richardson, J. C., Lehman, J. D., Newby, T. J., Xe, C., Mong, C., & Sadaf, A. (2010). Peer feedback in a large undergraduate blended course: Perceptions of value and learning. *Journal of Educational Computing Research, 43*(1), 67-88. doi: 10.2190/EC.43.1.e

Falchikov, N. (1986). Product comparisons and process benefits of collaborative peer group and self assessments. *Assessment and Evaluation in Higher Education, 11*(2), 146-166.

Falchikov, N. (1995). Peer feedback marking: Developing peer assessment. *Programmed Learning, 32*(2), 175-187.

Falchikov, N. (1998). Involving students in feedback and assessment. *SEDA PAPER*, 9-22.

Falchikov, N. (2013). *Improving assessment through student involvement: Practical solutions for aiding learning in higher and further education*: Routledge.

Falchikov, N., & Blythman, M. (2001). *Learning together: Peer tutoring in higher education*. London: Routledge-Falmer.

Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research, 70*(3), 287-322.

Ferguson, P. (2011). Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education, 36*(1), 51-62.

Fitch, S. (2016). Art critiques: A guide. *Studies in Art Education, 57*(2), 185-187.

Frascara, J., & Noël, G. (2012). What's missing in design education today? *Visible Language, 46*(1/2), 36.

Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and teaching in higher education, 1*(1), 3-31.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction, 20*(4), 304-315. doi: http://dx.doi.org/10.1016/j.learninstruc.2009.08.007

Goetz, J. P., & LeCompte, M. D. (1984). *Ethnography and qualitative design in educational research* (Vol. 19): Academic Press Orlando, FL.

Gray, C. M. (2013). Informal peer critique and the negotiation of habitus in a design studio. *Art, Design & Communication in Higher Education, 12*(2), 195-209.

Grizzle, J. E. (1965). The two-period change-over design and its use in clinical trials. *biometrics*, 467-480.

Hart, S. G. (2006). *Nasa-task load index (nasa-tlx); 20 years later*. Paper presented at the Proceedings of the human factors and ergonomics society annual meeting.

Hart, S. G., & Staveland, L. E. (1988). Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology, 52*, 139-183.

Hatchuel, A., & Weil, B. (2003). *A new approach of innovative design: An introduction to ck theory*. Paper presented at the DS 31: Proceedings of ICED 03, the 14th International Conference on Engineering Design, Stockholm.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112.

Hicks, C. M., Fraser, C. A., Desai, P., & Klemmer, S. (2015). *Do numeric ratings impact peer reviewers?* Paper presented at the Proceedings of the Second (2015) ACM Conference on Learning@ Scale.

Hovardas, T., Tsivitanidou, O. E., & Zacharia, Z. C. (2014). Peer versus expert feedback: An investigation of the quality of peer feedback among secondary school students. *Computers & Education, 71*(0), 133-152.

Ivarsson, A., Andersen, M. B., Johnson, U., & Lindwall, M. (2013). To adjust or not adjust: Nonparametric effect sizes, confidence intervals, and real-world meaning. *Psychology of Sport and Exercise, 14*(1), 97-102.

Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: Their origin and impact on revision work. *Instructional Science, 39*(3), 387-406.

Kulkarni, C., Wei, K. P., Le, H., Chia, D., Papadopoulos, K., Cheng, J., . . . Klemmer, S. R. (2013). Peer and self assessment in massive online classes *Design thinking research* (pp. 131-168): Springer.

Kulkarni, C. E., Bernstein, M. S., & Klemmer, S. R. (2015). *Peerstudio: Rapid peer feedback emphasizes revision and improves performance.* Paper presented at the Proceedings of the Second (2015) ACM Conference on Learning@ Scale.

Lent, R. W., Schmidt, L., Schmidt, J., & Pertmer, G. (2002). Exploration of collective efficacy beliefs in student project teams: Implications for student and team outcomes. *age, 7*, 1.

Li, L., Liu, X., & Steckelberg, A. L. (2010). Assessor or assessee: How student learning improves by giving and receiving peer feedback. *British Journal of Educational Technology, 41*(3), 525-536. doi: 10.1111/j.1467-8535.2009.00968.x

Liu, E. Z.-F., & Lin, S. S. J. (2007). Relationship between peer feedback, cognitive and metacognitive strategies and achievement in networked peer assessment. *British Journal of Educational Technology, 38*(6), 1122-1125. doi: 10.1111/j.1467-8535.2007.00702.x

Liu, N.-F., & Carless, D. (2006). Peer feedback: The learning element of peer assessment. *Teaching in Higher education, 11*(3), 279-290.

Maclellan, E. (2001). Assessment for learning: The differing perceptions of tutors and students. *Assessment & Evaluation in Higher Education, 26*(4), 307-318.

Marbouti, F., Cardella, M. E., & Diefes-Dux, H. A. (2014). *Student peer feedback on design.* Paper presented at the Frontiers in Education Conference (FIE), 2014 IEEE.

Marbouti, F., Diefes-Dux, H., & Cardella, M. (2015). *Students' and engineering educators' feedback on design.* Paper presented at the American Society for Engineering Education Annual Conference, Seattle, WA.

Miller, C. M., & Parlett, M. (1974). Up to the mark: A study of the examination game.

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin, 83*(4), 602.

Neubaum, G., Wichmann, A., Eimler, S. C., & Kramer, N. C. (2014). *Investigating incentives for students to provide peer feedback in a semi-open online course: An experimental study*. Paper presented at the Proceedings of The International Symposium on Open Collaboration, Berlin, Germany.

Nicol, D., Thomson, A., & Breslin, C. (2014). Rethinking feedback practices in higher education: A peer review perspective. *Assessment & Evaluation in Higher Education, 39*(1), 102-122. doi: 10.1080/02602938.2013.795518

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in higher education, 31*(2), 199-218.

Nulty, D. D. (2011). Peer and self-assessment in the first year of university. *Assessment & Evaluation in Higher Education, 36*(5), 493-507.

Patchan, M. M., Charney, D., & Schunn, C. D. (2009). A validation study of students' end comments: Comparing comments by students, a writing instructor, and a content instructor. *Journal of Writing Research, 1*(2), 124-152.

Portney, L. G., & Watkins, M. P. (2000). *Foundations of clinical research: Applications to practice* (Vol. 2): Prentice Hall Upper Saddle River, NJ.

Race, P. (1998). Practical pointers on peer assessment. *SEDA PAPER*, 113-122.

Reily, K., Finnerty, P. L., & Terveen, L. (2009). *Two peers are better than one: Aggregating peer reviews for computing assignments is surprisingly accurate*. Paper presented at the Proceedings of the ACM 2009 international conference on Supporting group work, Sanibel Island, Florida, USA.

Reimer, Y. J., & Douglas, S. A. (2003). Teaching hci design with the studio approach. *Computer science education, 13*(3), 191-205.

Robinson, R. (2001). Calibrated peer review™: An application to increase student reading & writing skills. *The American Biology Teacher, 63*(7), 474-480.

Rogers, R. R. (2001). Reflection in higher education: A concept analysis. *Innovative Higher Education, 26*(1), 37-57. doi: 10.1023/a:1010986404527

Sadler, R. (2002). Ah!... So that's 'quality'. In P. Schwarts & G. Webb (Eds.), *Assessment: Case studies, experience and practice from higher education*. London: Kogan Page.

Schön, D. A. (1983). *The reflective practitioner: How professionals think in action* (Vol. 5126): Basic books.

Schön, D. A. (1987). Educating the reflective practitioner: Toward a new design for teaching and learning in the professions. *San Francisco*.

Schunn, C., Godley, A., & DeMartino, S. (2016). The reliability and validity of peer review of writing in high school ap english classes. *Journal of Adolescent & Adult Literacy*.

Smith, H., Cooper, A., & Lancaster, L. (2002). Improving the quality of undergraduate peer assessment: A case for student and staff development. *Innovations in Education and Teaching International, 39*(1), 71-81.

Snyder, B. R. (1971). *The hidden curriculum*. Cambridge, MA: MIT Press.

Tinapple, D., Olson, L., & Sadauskas, J. (2013). Critviz: Web-based software supporting peer critique in large creative classrooms. *Bulletin of the IEEE Technical Committee on Learning Technology, 15*(1), 29.

Todd, R. H., & Magleby, S. P. (2004). Evaluation and rewards for faculty involved in engineering design education. *International Journal of Engineering Education, 20*(3), 333-340.

Tomayko, J. E. (1991). *Teaching software development in a studio environment.* Paper presented at the ACM SIGCSE Bulletin.

Tomayko, J. E. (1996). *Carnegie mellon's software development studio: A five year retrospective.* Paper presented at the Software Engineering Education, 1996. Proceedings., Ninth Conference on.

Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249-276.

Trivedi, A., Kar, D. C., & Patterson-McNeill, H. (2003). Automatic assignment management and peer evaluation. *J. Comput. Sci. Coll., 18*(4), 30-37.

Vasana, S., & Ritzhaupt, A. D. (2009). A case study of a method for hybrid peer-evaluation in engineering education. *World Transactions on Engineering and Technology Education, 7*(1), 34-40.

Vickerman, P. (2009). Student perspectives on formative peer assessment: An attempt to deepen learning? *Assessment & Evaluation in Higher Education, 34*(2), 221-230.

Von Hippel, E. (1986). Lead users: A source of novel product concepts. *Management Science, 32*(7), 791-805.

Wen, M. L., & Tsai, C.-C. (2006). University students' perceptions of and attitudes toward (online) peer assessment. *Higher Education, 51*(1), 27-44.

Williams, B. C., He, B. B., Elger, D. F., & Schumacher, B. E. (2007). Peer evaluation as a motivator for improved team performance in bio/ag engineering design classes. *Engineering Education, 23*(4), 698-704.

Wooley, R., Was, C., Schunn, C. D., & Dalton, D. (2008). *The effects of feedback elaboration on the giver of feedback.* Paper presented at the 30th Annual Meeting of the Cognitive Science Society.

Xiao, Y., & Lucking, R. (2008). The impact of two types of peer assessment on students' performance and satisfaction within a wiki environment. *The Internet and Higher Education, 11*(3–4), 186-193. doi: http://dx.doi.org/10.1016/j.iheduc.2008.06.005

Yoder, B. L. (2014). Engineering by the numbers *Profiles of engineering and engineering technoloy colleges* (pp. 11-47). Washington DC: ASEE.

Yuan, A., Luther, K., Krause, M., Vennix, S. I., Dow, S. P., & Hartmann, B. (2016). *Almost an expert: The effects of rubrics and expertise on perceived value of crowdsourced design critiques*. Paper presented at the Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing, San Francisco, California, USA.

Zhu, H., Dow, S. P., Kraut, R. E., & Kittur, A. (2014). *Reviewing versus doing: Learning and performance in crowd assessment*. Paper presented at the Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing, Baltimore, Maryland, USA.