Are Two Heads Better Than One in CAD? A Comparison of Various CAD Working Styles.

by

Vrushank Shripad Phadnis

B.E., University of Mumbai (2010)
M.E., Massachusetts Institute of Technology (2013)

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2020

Author . . . .
Department of Mechanical Engineering
August 15, 2020

Certified by . . . .
David R. Wallace
Professor of Mechanical Engineering
Thesis Supervisor

Accepted by . . . .
Nicolas Hadjiconstantinou
Professor of Mechanical Engineering
Graduate Officer

Are Two Heads Better Than One in CAD?

A Comparison of Various CAD Working Styles.

by

Vrushank Shripad Phadnis

Submitted to the Department of Mechanical Engineering
on August 15, 2020, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Mechanical Engineering

Abstract

Collaboration in Computer-Aided Design (CAD) has existed since the inception of CAD tools. The established norm in multi-user CAD work has been to use top-down modeling techniques wherein a complex model is divided into sub-assemblies (or part files) for individual designers to work on separately. In this process, designers integrate their work through a check-in/check-out process. This style of collaboration does not change regardless of team sizes, product types, or over time. However, recent cloud-based CAD tools are expected to change this by offering real-time collaboration like Google Docs.

In this research, we are interested in learning the effects of real-time collaboration on designers' work. We draw heavily from software development research where dyadic work is common and is known as 'pair programming'. We use an experimental approach to investigate research questions pertaining to speed and quality of real-time collaboration.

We found that pair work in CAD is not summative. In other words, the work of two designers does not lead to twice the outcome of individuals. This results is contrary to previous real-time CAD collaboration research but consistent with software programming research. However, we also found that the quality of CAD increases in certain pair CAD settings. We observed that sharing control of the CAD software leads to higher quality and parallelizing work leads to worse quality. To elaborate on our results, we reveal specific patterns of participant behaviour based on audio communication and cursor activity.

In summary, we establish foundational knowledge in real-time CAD collaboration research. Through our work, we share insights which can inform practicing engineers that are interested in adopting pair CAD work.

Thesis Supervisor: David R. Wallace
Title: Professor of Mechanical Engineering

Acknowledgments

Prof. David Wallace, thank you for getting the ball rolling. Your vision and experience in CAD research set me up for success. Our conversations on academia, Canada and teaching are cherished. Prof. Alison Olechowski, I am very fortunate to have started working with you when you were a grad student at MIT to now as the principal investigator (PI) of Ready Lab. Being with you through this transition has given me some of the best learning of my PhD. Thank you for taking me along on this journey. Your determination, empathy, and work ethic are inspiring and make Ready Lab the place it is.

To my committee members, our meetings were very much looked forward to. Prof. Warren Seering, thank you for encouraging me to think boldly about post-PhD life. Prof. Maria Yang, thank you for enabling the thriving design research community at MIT.

Prof. David Hardt, your mentorship has been crucial in shaping my career since I first came to MIT. From MEngM to recently working on 2.830x, you have been a role-model. Prof. Amos Winter and Prof. Sangbae Kim, thank you for including me in the fun-filled journey of 2.007. Your passion for mechanical design and care for students are an inspiration. Prof. Michael Golay, thank you for welcoming me into 22.811x, guiding my career choices and sailing. Your perspective on the world's energy crisis is refreshing and I am privileged to be able to disseminate that.

Thank you to all the undergraduate research students that significantly contributed to this work. Kevin Leonardo, Cameron Arnet, Hamza Arshad, Jinxuan (Janice) Zhou, Jennifer Wu and Meaghan Vella, our meetings and conversations were instrumental in framing this research.

My peers from CADLab, Ready Lab and Ideation made enduring grad school super enjoyable. Dr. Joshua Ramos, thank you for having a solution to everything-CADLab-related. Georgia Van de Zande, Dr. Geoff Tsai, Audrey Bosquet, Dr. Larissa Nietner, Dr. Zhuoxuan (Fanny) Li and Teresa Lin, our time together defined my great experience in CADLab. Dr. Qifang Bao and Dr. Edward Burnell,

thank you for your companionship and enriching discussions. James Chen, Gustavo Zucco and Safa Faidi, our adventures in Myhal will be fondly remembered.

Leslie Reagan, you have always amazed me with your continued dedication in supporting MechE grad students. Kirsten, Stephan, Pablo, Ariane and Sauvanne, I could not have asked for a better home-away-from-home in Canada. Dr. Samuel Raymond, Dr. Mojtaba Forghani, Prof. Khalid Jawed, Sujay Bagi, Michael Arnold, David Wang, Shile Ding and Candy Wang, thank you for your invaluable friendship and making my time at MIT meaningful.

Lastly and most importantly, a very big thank you to my family. Baba and Aai, you have been pivotal to every success in life. Thank you for that initial nudge to apply to MIT and then seeing me through it. Sudiksha, thank you for keeping me smiling and happy, all the time. Ulhas Mama, thank you for instilling positivity during my lows at MIT. MamaAai, my success is a result of your nurturing and lessons on perseverance. I dedicate this accomplishment to you and will continue making you proud.

Raksha, thank you for bringing focus and pace to my PhD (and life). I look forward to our adventures.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

1.1.1 New (Virtual) Normal

At the time of writing this thesis, the world around us is going through a major transformation. The pandemic known as the novel coronavirus disease of 2019 (COVID-19) has had a profound impact on our lives that goes beyond the underlying health crisis [5]. As seen in Figure 1-1, 'stay at home' orders across the world have limited our ability to be in physical proximity with each other.

This warrants novel work practises, and in-person collaboration is no longer the standard. Remote work has become the new norm for many [6]. However, working together virtually is not a new concept and has long been considered as an alternative to physical offices [7]. It can be argued that the onset of the pandemic has accelerated its adoption and acceptance.

Remote work offers a number of advantages like saving in travel time, promoting diverse teams, reducing discrimination [8]. Conversely, there are disadvantages like reduced awareness of others' work, decrease in explicit management, vulnerability to mistrust, and the need for complex technical infrastructure [9]. Teams can strive to achieve the right balance by selectively deploying remote work. This requires a thorough understanding of virtual team work and such insights are often estab-

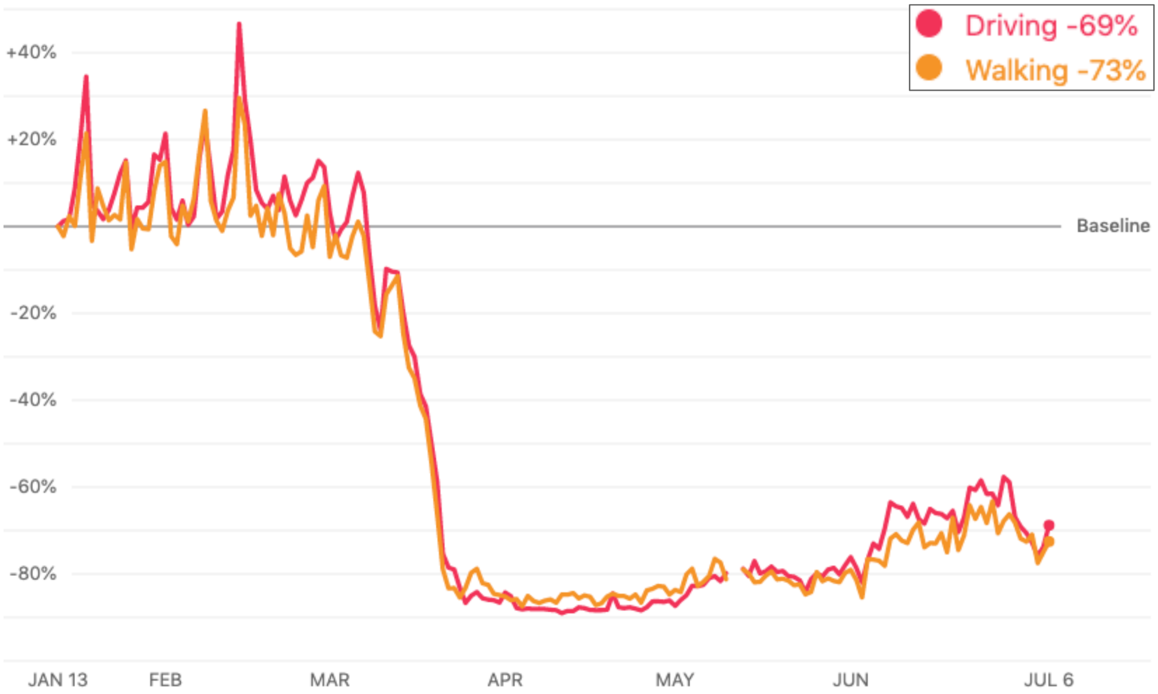


Figure 1-1: Relative mobility change in Mumbai in 2020 due to COVID-19.[1]

lished in research. Through this work, we strive to draw such insights for mechanical engineering by investigating the impact of remote Computer-Aided Design (CAD) collaboration.

1.1.2 Challenges in Innovation of CAD Working Styles

Mechanical engineers may be considered as late adopters of fully virtual work. Some of the hesitation might stem from challenges in replicating the in-person interaction with physical artifacts, in an online environment. For mechanical engineers visual communication and tactile feedback are essential. Recent efforts to use virtual/augmented reality (VR/AR) solutions are touted to provide an enhanced experience that comes close to working with physical objects. These techniques have received limited success in research but show future potential [10]. However, the equipment required for VR/AR applications remains fairly complex to setup and expensive to buy.

Besides tool limitations, the CAD community has to overcome inertia to change from decades of legacy in mechanical engineering. Our habits and working styles are

---

[1] https://www.apple.com/covid19/mobility

established from the long past era; starting with drawing on drafting boards. Some veteran engineers still prefer reviewing 2D part prints over 3D CAD models. Next, master modeling techniques and top-down assembly became popular for collaboration in early CAD tools. This is still the norm. As seen in Figure 1-2, in top-down modeling assembly files are subdivided into lower level sub-assemblies and part files. These part files are then assigned to individual designers to work on separately.

Lastly, the steep learning curve and high cost of procuring new CAD software forces enterprise customers to retain existing software. This risk averse nature of the CAD market may be a deterrent to new entrants from innovating. Conversely, CAD software are very complex and require extensive development time and domain knowledge.

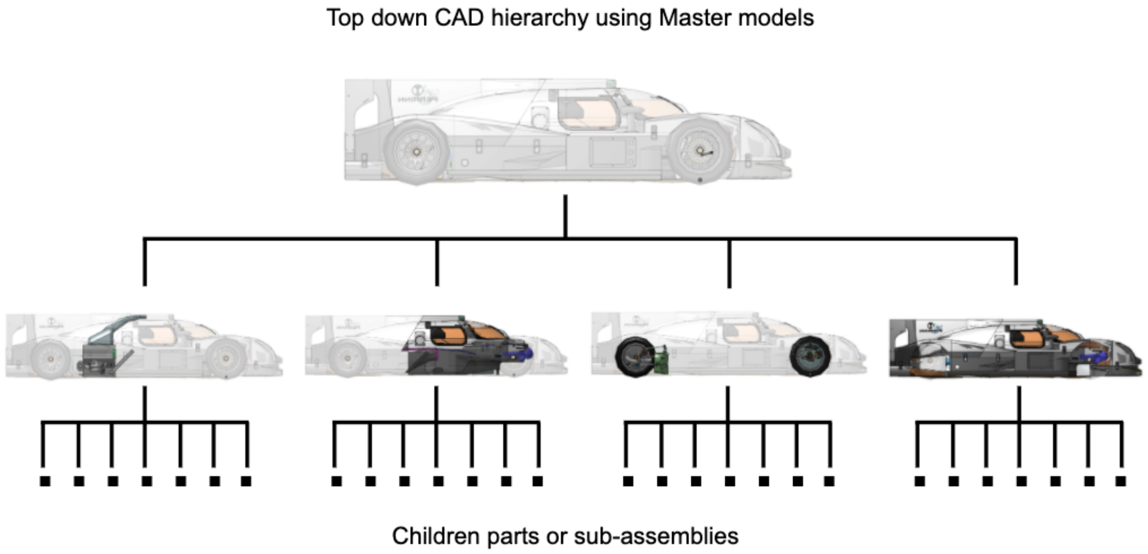
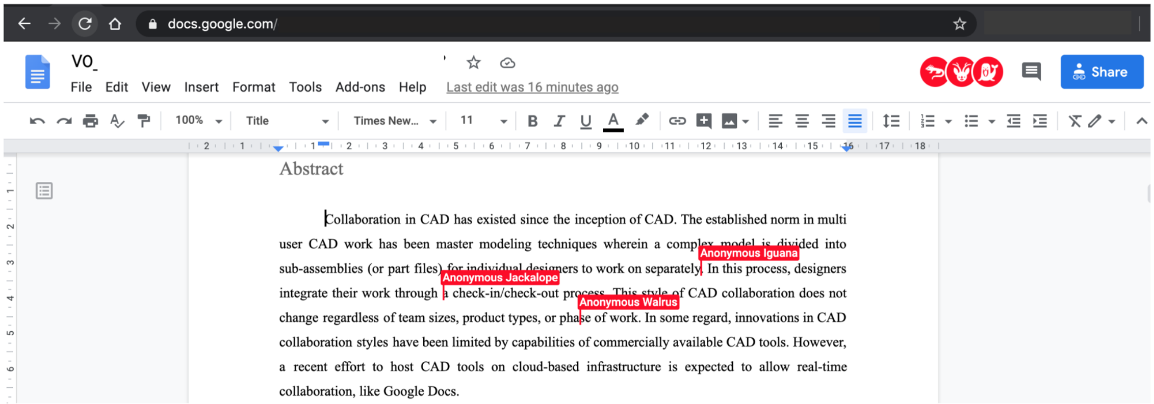


Figure 1-2: Illustration on top-down modelling example showing parent-child relationship. CAD model adapted from the Perrinn project on Onshape.com.

1.1.3 Optimistic Future for Mechanical CAD

We see a varying acceptance of digital technology in the past few generations of people [11]. Receptiveness to innovation in technology is much higher in newer generations. Today's mechanical engineers use a variety of software tools like Slack, Google Docs, MATLAB Online, Asana, *etc.* in addition to traditional tools like CAD. These tools differ immensely from their predecessors. For example, in Figure 1-3, we show four users collaborating on a single text document in real-time using Google Docs. This is vastly different from using email and Microsoft Word to produce the same work. This shift in working style hugely influences a writer's ability [11,12]. In our research, we are interested in understanding the implications of such shifts in working styles; but pertaining to mechanical engineering.



Abstract

Collaboration in CAD has existed since the inception of CAD. The established norm in multi user CAD work has been master modeling techniques wherein a complex model is divided into sub-assemblies (or part files) for individual designers to work on separately. In this process, designers integrate their work through a check-in/check-out process. This style of CAD collaboration does not change regardless of team sizes, product types, or phase of work. In some regard, innovations in CAD collaboration styles have been limited by capabilities of commercially available CAD tools. However, a recent effort to host CAD tools on cloud-based infrastructure is expected to allow real-time collaboration, like Google Docs.

Figure 1-3: Multiple users editing a single text document using Google Docs. [2]

We see a recent uptick in cloud-based offerings from the CAD industry. Tools like SketchUp, Blender and Tinker CAD promise democratization of low fidelity design tools. As they are free of cost and require only a web browser to operate. However, they have limited capabilities and can not be alternatives to traditional CAD software like SOLIDWORKS and Rhinoceros (Rhino).

Cloud-based CAD tools offer an unique advantage by allowing multiple users to access the database at the same time, like in Google Docs. This functionality opens up novel working styles that were not possible in the past. A few of the benefits of using cloud-based CAD tools over traditional CAD tools are outlined in Figure 1-4. Cloud-based cloud use a centralized computing infrastructure for rendering and analysis. This means CAD users can access the CAD file in real-time using a simple web-browser. And there is no need to install a local instance of the CAD software which obviates the need for specific operating system or hardware specifications. All

---

[2] https://docs.google.com

files are securely stored and backed up using enterprise-level software and hardware. Cloud-based storage is also easier to scale and benefits both small and large scale organizations by allowing them to dynamically start/stop subscriptions as needed.

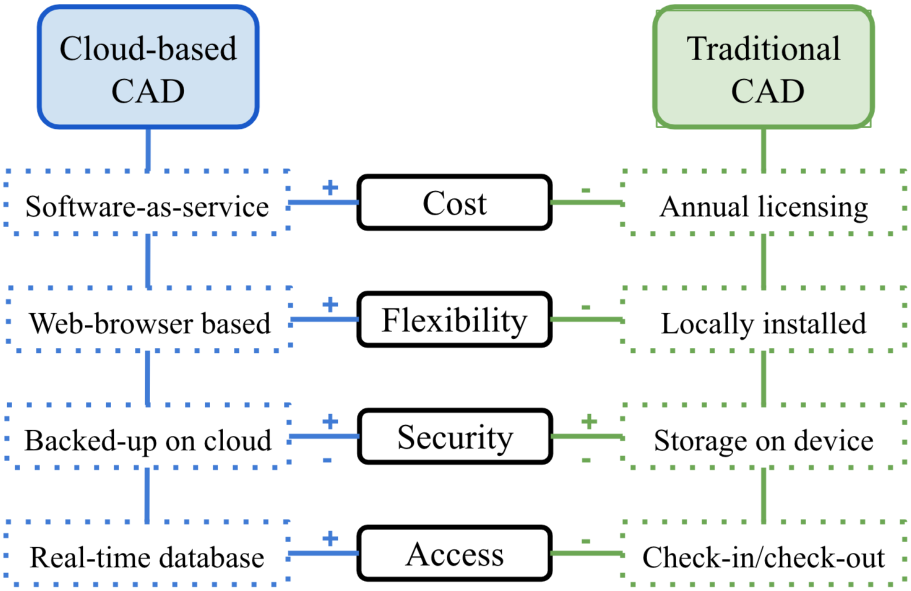


Figure 1-4: Comparison of cloud-based *vs.* traditional CAD.

Figure 1-5 shows some of the current offerings in the CAD industry. Synchronocity in CAD (x-axis) pertains to the ability of a CAD tool to update everyones databases in real-time. In our work we use the words synchronous, cloud-based and real-time CAD interchangeably. Only a few fully-synchronous software are currently available and Onshape is one such example.

Synchronous CAD tools are still sparsely used. This could be partly because of the difficulty in using real-time collaboration. Figure 1-6 shows a summary of industry sentiments on cloud-based CAD. Note that this survey was conducted in 2017 and these sentiments might have changed. However, by 2017 cloud-based CAD tools were certainly commercially available. In the graphic, we see that 50% of the surveyed pool was aware of cloud-based CAD but did not pursue it. One can speculate that is the case because of the large investments made by companies in building their existing CAD infrastructure. However, we think some of these apprehensions might stem from not understanding the use-case for cloud-based CAD. Our investigation of the real-time collaboration features in cloud-based CAD will help designers evaluate

Figure 1-5: Categorization of modern CAD tools.[3]

the relevance of cloud-based CAD in their own work.

1.2 Thesis Outline

The work presented in this thesis is derived and partly reproduced from these publications: [3, 4, 13–16].

Over the past few years, this research has evolved in three distinct phases: literature review, methods development, and execution of experiments. This dissertation is also structured in that order.

1.2.1 Literature Review (Chapter 2)

Prior art looking into implications of real-time CAD collaboration is limited. Firstly, we survey this limited research and also outline topics on the broader theme of collaboration in mechanical engineering. We then present an analogy of real-time collaboration from software development called pair programming. Lastly, we define terminology and nomenclature that is specific to our work.

---

[3] All logos were downloaded from Google Images and are the property of their respective owners.

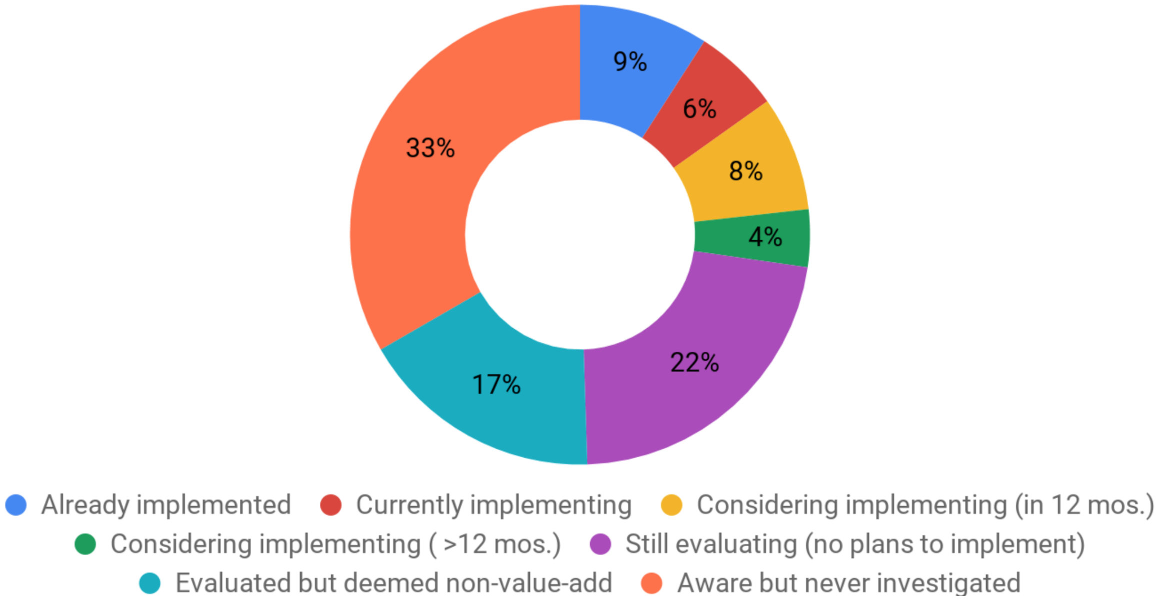


Figure 1-6: Current and future usage plans for cloud-based computer-aided design (CAD) worldwide in 2017. Adapted from [1].

1.2.2 Methods Development (Chapter 3)

Behavioral experiments are complex to set up and need a sophisticated tools. In this section, we present nuances of the hardware and software choices made to develop the experiment toolkit. To gain early insights and validate our toolkit we conducted a preliminary study and the results from the same are presented in this section.

1.2.3 Execution of Experiments (Chapter 4-5)

The final experiments were conducted at the University of Toronto (UofT). As part of this research we had the unique opportunity to set up a new lab space for Ready Lab at UofT. This section describes the final iteration of the experiment and corresponding results.

We begin Chapter 4 with a description of the logistics of running an experiment. Then we define variables of interest and present results. Lastly, in Chapter 5 we draw insights and present guidelines for future engineers interested in adopting pair CAD work.

Chapter 2

Literature Review

2.1 CAD Research

CAD, as a design tool has existed for more than half a century. Some of the foundational work in the field was done at MIT, at CADLab [17]. As a result, there is a rich repository of research topics on CAD tools. In the following sections, we discuss a subset of relevant research.

2.1.1 General Topics on Collaboration

Collaboration has long been a focus within the design research community [18]. Ostergaard and Summers review the literature and put forward a taxonomy of collaborative design activities, consisting of a number of dimensions by which collaborative design is affected: team composition, communication, distribution, design approach, information and nature of the problem. Reflecting on these dimensions, we believe that fully-synchronous CAD presents an unique example of collaborative design work, in particular because of the nature of the problem, distribution of the team, and information. We posit that a new design approach may be necessary to best capture the value-add of the fully-synchronous CAD capabilities.

Another area of relevant research is that which looks at virtual teams. This literature is reviewed by Martins *et al.* [19]. Key implications for this work include that type of technology used by virtual teams is an important input on team effectiveness and efficiency (varies by, for example, richness and novelty), as is task type (varies by, for example, ambiguity and complexity).

We expect that collaboration in computer-aided design could be different from other types of new product development collaboration - the nature of the collaboration is contingent on the work. Therefore, to understand fully synchronous CAD collaboration we need to specifically study work in this context.

2.1.2 Synchronous Collaboration in CAD

Collaboration in CAD has existed in some form since its inception. Like mentioned earlier, top-down assembly and master modeling techniques are the norm in industry and have proven to be efficient at organizing CAD work.

Recently, web-based synchronous CAD has become available commercially at scale. The ability to have multiple designers simultaneously modifying a model from their own workstations, once a dream, is now a reality through cloud-based software offerings [20, 21]. More products are being designed by global teams and high demand for anytime availability of product data is a compelling CAD software requirement. However, the potential collaborative features of synchronous CAD are intriguing as design is a social process that involves multiple people working towards a solution [22]. A synchronous CAD platform hypothetically offers numerous benefits such as: synchronous access, cost effectiveness, higher utilization of resources, and enhanced security [23]. However, there are still questions to be answered surrounding design in the cloud related to usability, security and computational performance.

Moreover, we have yet to understand the effect of these synchronous CAD tools on underlying design processes used at designer (user) level, or team level, and how to best make use of the provided design freedoms. It is plausible that in a fully synchronous collaborative CAD environment there will be an effect on key design metrics like: creativity, communication fluency, design quality, and productivity. Through our research, we aim to elucidate these relationships and present guidelines for best practices for fully-synchronous collaborative CAD work.

A few experiments have aimed to investigate collaborative CAD teams, some even using synchronous collaborative CAD tools. Eves *et al.* present a study of their specific collaborative computer-aided design software, multi-user CAD (MUCAD) which allows synchronous design [24]. Their study is a comparison of the output of four-person teams using MUCAD and traditional CAD (sharing models through email). This study found that use of MUCAD increases the awareness of teammates' activities, and communication between team members. However no statistically significant findings were found related to model quality or team productivity, potentially due to limitations of a small sample size. It is important to note that the MUCAD software in this study was buggy, a major limitation on the industry-like setting for this work and on the quality of the collected data. The authors note anecdotally that multi-user teams appeared to have better interface management, and a better sense of current state of model and what still needed to be done.

Two studies have looked for associations between designer teams or product-type and design behaviors using full-synchronous collaborative CAD tools. Stone *et al.* set out to establish a method to determine the optimal number of designers on a part based on characteristics of the part itself and the architecture of its features [25]. Teams varied in size from one-to-four-person. No statistically significant trends were found. This study was also limited by buggy software. The authors do make an effort to adjust performance to compensate for the impact of bugs, but this is an important limitation of this work to consider. Another study looked at various team-member dynamics, especially communication, during a MUCAD design competition of three-person teams [26]. Teams that encouraged effective forms of communication and teams whose members scored similarly on Purdue Spatial Visualization Test (a test of spatial manipulation ability) performed better than other teams. The authors compared audio recordings and posit that patterns of communication could provide important insight. Though no statistically significant results regarding communication were found, the authors note that anecdotally, high-performing teams tended to communicate less. This study was again limited by bugs in the software.

Holyoak *et al.* present an approach for collaborating on fully-synchronous CAD software [27]. This proposed design process provides guidance based on inputs of design specifications and task distribution (analyzed for dependence *via.* a design structure matrix), aiming to take advantage of the ability for multi-users to work concurrently in parallel. In particular, by analyzing the team for expertise and decision-making authority, the output of the design process is a list of tasks and corresponding personnel groups to accomplish those tasks. The authors claim that their process has the potential to reduce wait time and iterations, and therefore overall design time. The authors present a small data set comparing teams of three who follow this new process to teams of three who iterate through a design, indicating that the multi-user teams following the new process finish design tasks more quickly and with more specifications satisfied. However, these results are not statistically examined for significance, likely due to the small number of trials.

Current research specific to fully synchronous collaborative CAD, while preliminary, makes important observations about trends in behavior and outcomes. These works highlight how little is known about synchronous collaborative CAD when compared to other forms of collaborative design and inspires potential directions for future detailed study.

2.2 Background on Multimodal Methods

Experiments provide a unique advantage to researchers by providing a higher degree of control in defining research questions [28, 29]. Exploratory experiments are especially useful in emerging research fields which do not have an exhaustive pre-existing literature to draw upon; but are also challenging to set up as they need to capture a large array of variables. In Chapter 3, we present an experimental method to answer exploratory research questions about collaborative CAD.

2.2.1 Experimental Toolkits in Design Research

Our work has been influenced by previous research efforts to develop toolkits. Sivanathan *et al.* presented a ubiquitous data capture method for CAD work [30]. Their toolkit facilitated real-time logging of multiple data types like CAD metadata, eye tracking, cursor activity logging, Electroencephalography (EEG), and electrocardiography sensing. The benefits of real time analysis from the framework presented by Sivanathan *et al.* was proposed to be applicable beyond research settings. However, this setup requires a complicated sensor suite and does not lend well to designers' natural way of working. Another toolkit presented by Liu *et al.* used EEG, galvanic skin resistance (GSR)/electrocardiography (ECG) to capture psycho-physiological data. This dataset was used to compute designers' emotions and compare with logs with CAD [31]. A fuzzy model was developed to derive emotions output which was validated in a case study. This approach, like Sivanathan *et al.'s* framework, relied heavily on sensor-based solutions. Nyugen *et al.* present an EEG based toolkit to automate the design protocol analysis methods using microstate analysis [32]. In this work, a model to map EEG signals to events of interest in design is developed. The toolkit is applied in a pilot study setting and its results are presented as a case study. Nyugen *et al.* present a detailed comparison between automated protocol analysis and expert human coders. This work concluded with recommendations on the choice of algorithms in automating design protocols using EEG.

Automated techniques often rely on sensor based data. Techniques like EEG and fMRI have been used in design research as tools to measure participants' cognitive loads [33]. These heavily sensor-based techniques can be intrusive and complex to set up.

Rahman *et al.* present a software based approach to capture design behavioral data. A CAD platform, ENERGY3D was developed to study solar energy systems [34]. This approach is non-intrusive in nature but limited in the variety of variables captured by sensor based toolkits.

Contrary to the multimodal sensor based techniques, Ostergaard *et al.* present an experimental method to study collaboration patterns in design review meetings [35]. In this work, they compare groups *vs.* individuals working through a design review. Data is collected by evaluating participant worksheets post-experiment. The simplicity of the experiment and lack of a sophisticated toolkit limits the amount and variety of data available.

2.2.2 Protocol Analysis Methods

Protocol analysis is an established method in design research and is used to understand hard-to-code designer behaviour [36]. A peculiarity of protocol analysis studies is their small sample sizes. Small data sets are common in protocol analysis because of the time consumed in coding each data set, sometimes taking 10 times longer than the footage being reviewed [36].

Researchers have investigated the effects of virtual *vs.* co-located design processes using protocol analysis methods [37]. Verbalization or think-aloud techniques are common in protocol analysis. For example, Anwar *et al.* modeled cognitive behaviour of designers in the conceptual phase of their work using verbalization [38]. Verbalization techniques have limitations like data validity, steep learning curve, and tasks not being suitable for verbalisation [39].

Another popular protocol analysis method, known as retrospective analysis involves participants recollecting their experience post-experiment. Such retrospective protocol analysis methods are prone to filtering and bias by participants. It is also observed that informal reporting from researchers' notes during the study can lead to data losses. As the research team might not have anticipated all possible events of interest beforehand [40].

2.2.3 Automated Data-Capture Methods

CAD is analogous to other graphics-based interactive software and thus, appropriate to be studied using research methods from the human computer interaction (HCI) community. Techniques like cursor tracking which are a subset of user experience (UX) data capture techniques are ubiquitous in HCI. Cursor tracking algorithms are quick to set up and track spatial location of the cursor/pointer. They can run natively on most operating systems and are less resource intensive compared to more advanced UX analytics software.

In psychology research, cursor tracking is popular and used to to ascertain user behaviour and preferences [41,42]. In design research, eye tracking is more popular [43, 44]. However, traditional eye tracking solutions lose out on UX analytics information like, clicks, scrolls and drag motions. Eye tracking solutions often require specialized hardware and need to be calibrated for each user. Although more accurate than cursor tracking, eye tracking is not perfect. Some eye tracking software triangulate the user gaze location with cursor coordinates to increase accuracy.

Self-reporting of emotions is a common measure of user satisfaction and this can be limiting. Methods like EEG and fMRI are gaining popularity and provide emotion metrics over continuous time [33,45]. But these methods are complicated to set up and require expensive equipment. In that regard, we chose a software based facial image recognition approach. Software based solutions are passive to the participants and may be run in the background. They rely on advanced algorithms to post process face recording footage and generate emotion metrics [46]. More recently, emotion tracking has been of interest within the design research community [31, 47]. Software based emotion tracking is particularly appealing because of its ease of use [14].

In our toolkit, we capture multiple data streams to analyse user behaviour. Such multimodal analysis has been successfully used to investigate human social behaviour [48]. More so, some design researchers have adopted such methods to study CAD work [30]. Multimodal methods provide a good trade off by leveraging the benefits of various data capture techniques. We rely on automated techniques for high frequency (high resolution) data like cursor movements, and emotion expression, but use manual techniques to document complex CAD activity that are hard to automate (or codify).

2.3 Pair Work

Small group work has been studied in research before [49]. Its adoption in software development, known as pair programming has shown potential benefits. In this section, we outline research on pair programming and draw parallels with our work.

2.3.1 Agile Methodology and Pair programming

Extreme programming techniques (XP) were popularized after the release of the Agile manifesto in 2001 [50]. Pair programming, a subset of XP techniques is becoming more common in software development teams. The benefits of pair (dyadic) programming have since been studied in both research and industry settings [51, 52]. The results of these works are mixed but pair programming has shown to impact the speed, quality and engagement of software developers [53]. A better understanding of pair programming methods helps teams to accurately deploy these techniques. For example, one of the benefits of pair programming is its ability to provide an enhanced knowledge-sharing experience [54]. This can immensely improve the on-boarding process of new team members or tasks requiring knowledge transfer between cross-functional stakeholders.

2.3.2 Remote Pair Programming

Traditionally, pair programming sessions strictly adhered to a driver-navigator style collaboration [55]. In such a pair only one participant codes at a time, allowing the second participant to review. Participants share control of a single station and switch roles as seamlessly. Typically, these sessions are held with participants being in physical proximity and sharing the same keyboard and mouse. The premise being, more eyes on the written code makes it higher quality and bug free [56].

However, an increasing number of pair programming sessions are now held remotely. Software collaboration in virtual settings can be categorized into two broad categories, as shown in Figure 2-1. The nuances of each working style will be discussed next.

Traditionally, pair collaboration was achieved using screen sharing tools to maintain the essence of having an increased awareness amongst users [57]. This style of collaboration was introduced in 1987 as What-You-See-Is-What-I-See (WYSIWIS) collaboration and developed as a groupware project called Colab at Xerox [58]. WYSIWIS based tools are still prevalent today and preferred for their increased aware-

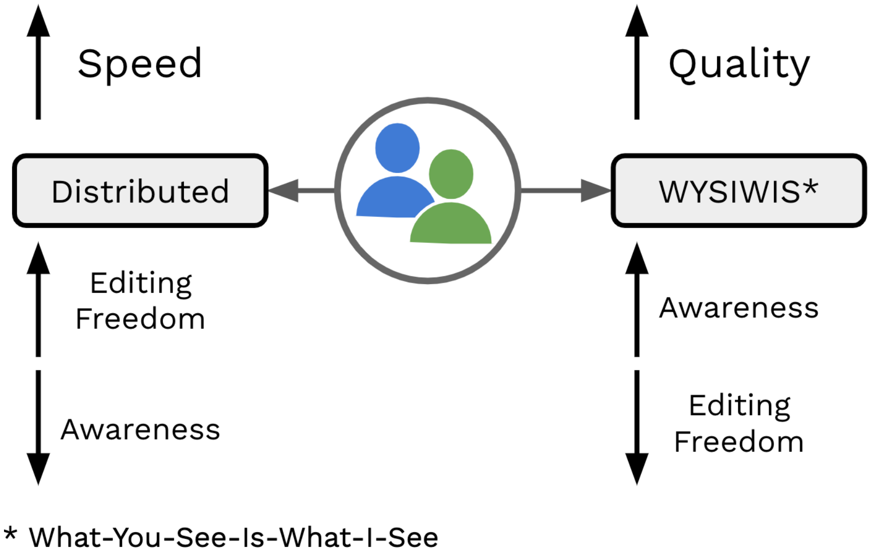


Figure 2-1: Remote pair programming styles.

ness [59–61].

Newer programming compilers can handle multiple code streams at a single time, enabling distributed programming [62, 63]. This style of working can be unstructured and breaks away from the strict driver-navigator dynamic [64]. The added editing freedom allows for parallelization of work and is touted to be faster [52]. However, possibly at the cost of reduction in quality [65]. We suspect that parallel control of the code base reduces awareness amongst participants but shortens development time. Typically, parallel coding setup necessitate advanced tools capable of integrating multiple code streams without conflicts [66].

Similar to cloud-based CAD research, pair programming literature has used speed and quality as primary metrics to compare coder's outcomes [52]. We will build our work on these same comparative metrics: speed and quality.

2.3.3 Introduction to Pair CAD

Preceding software engineering, research on mechanical engineering teams has been long ongoing and is well established [37, 67–72]. Moreso, the tools we use in practise are also mature. Product data management (PDM) is the preferred collaboration platform and is used to aggregate work of multiple individuals [67].

Mechanical design projects follow a progression similar to software development, starting from an ambiguous ideation phase and eventually leading to a structured detailed design outcome [73, 74]. Both disciplines rely on teamwork and in-depth domain knowledge for project success.

Figure 2-2 shows a visual representation of the differences in our proposed CAD working styles. In top-down modelling, multiple individual designers' work is combined into a single assembly file using master modeling techniques. Note that each designer works on a mutually exclusive subset of the CAD assembly.

Cloud-based CAD tools allow multiple designers to work side by side in real-time. We call this style of work pair CAD. The first pair CAD style has both designers working on a single CAD database where they can edit the CAD file in parallel, simultaneously. The second style is akin to screen sharing where both participants share the CAD user-interface (UI) itself but from separate terminals. Only one participant can actively navigate the CAD tool at a time.

Lastly, individual CAD work simply means one designer working on their own CAD file.

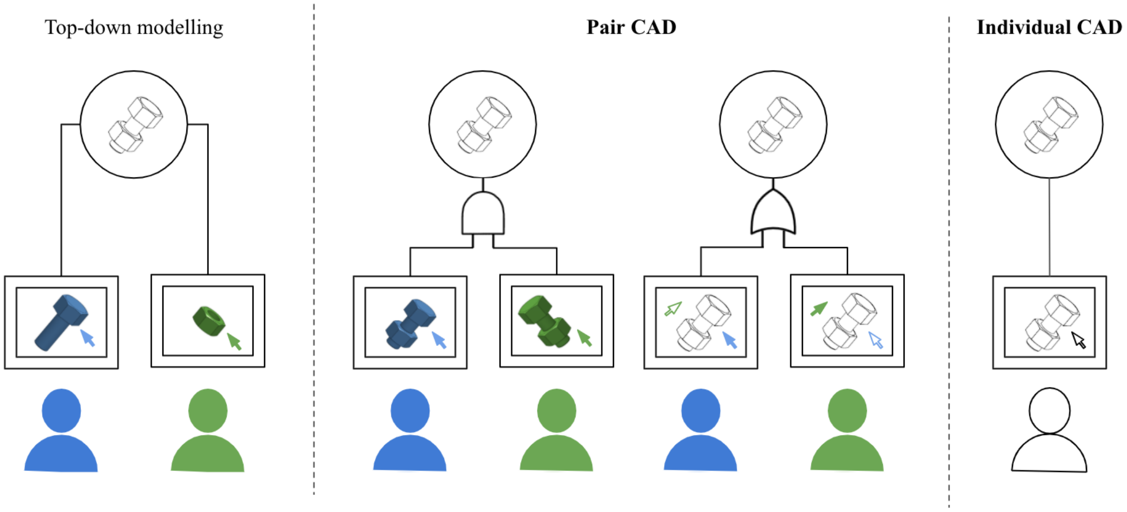


Figure 2-2: Representation of various CAD working styles.

Chapter 3

Development of Multimodal Method And Pilot Study

Our research experiments evolved in two distinct steps. First, metrics were established, measurement toolkit was developed and pilot study was conducted. Then, an improved version of the toolkit was used to carry out the second, larger, iteration of the experiment. Chapter 3 pertains to the first portion of our research experiments. And Chapter 4 presents the second portion.

3.1 Research Strategy

Research tools have evolved over the past few decades. However, the overarching types of research strategies have not changed. Figure 3-1 illustrates an array of such strategies and their inherent trade-offs. It is particularly noteworthy that none of these strategies can solely maximize all three attributes of behavioral research: precision, generalizability, and realism. In this work, we choose lab experiments as they offer most precision in defining abstract research questions. This was important as our work is early-stage and exploratory. We acknowledge the obtrusive nature of lab experiments and strive to minimize this limitation by avoiding sensor- based data generation.

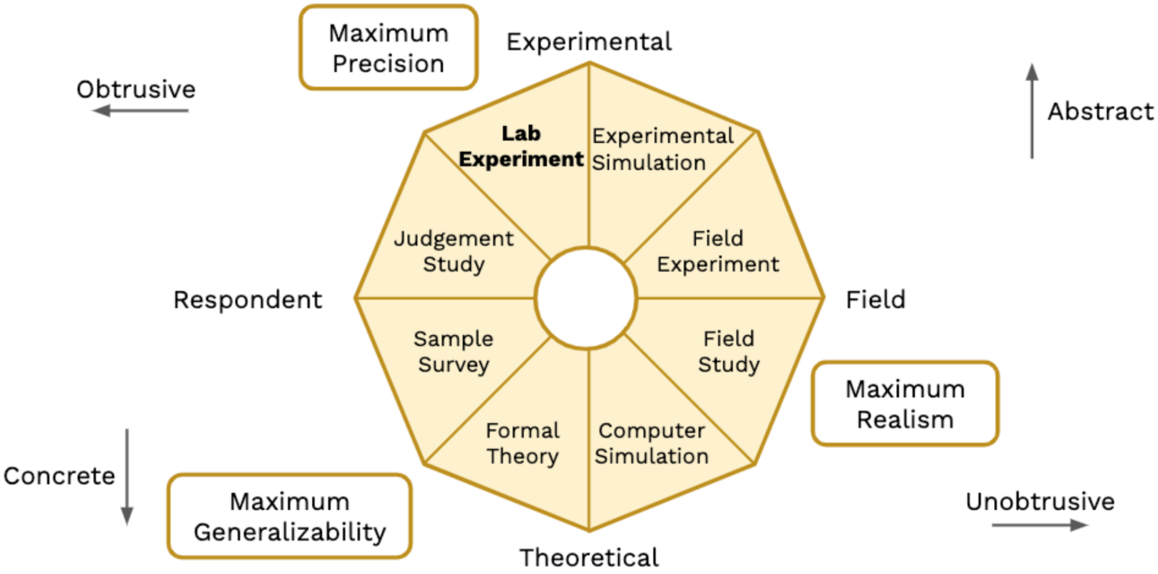


Figure 3-1: Research strategies. Adapted from [2].

3.2 Multimodal Toolkit

This section describes the metrics of interst, hardware specifications, software choices and post processing methods implemented in our multimodal approach.

3.2.1 Metrics of Interest

Understanding CAD collaboration warrants an analysis of both CAD and non-CAD activities. We captured speed and quality attributes which are directly associated with the CAD activity. In addition, non-CAD related metrics like communication, user satisfaction, and user software interaction are useful in explaining the outcome of the CAD activities. A summary of all metrics of interest is seen in Figure 3-2. All metrics were mapped to quantifiable data captured using methods discussed in subsequent sections.

3.2.2 Hardware

A participant workstation in our experiment (Figure 3-3) consisted of the below hardware:

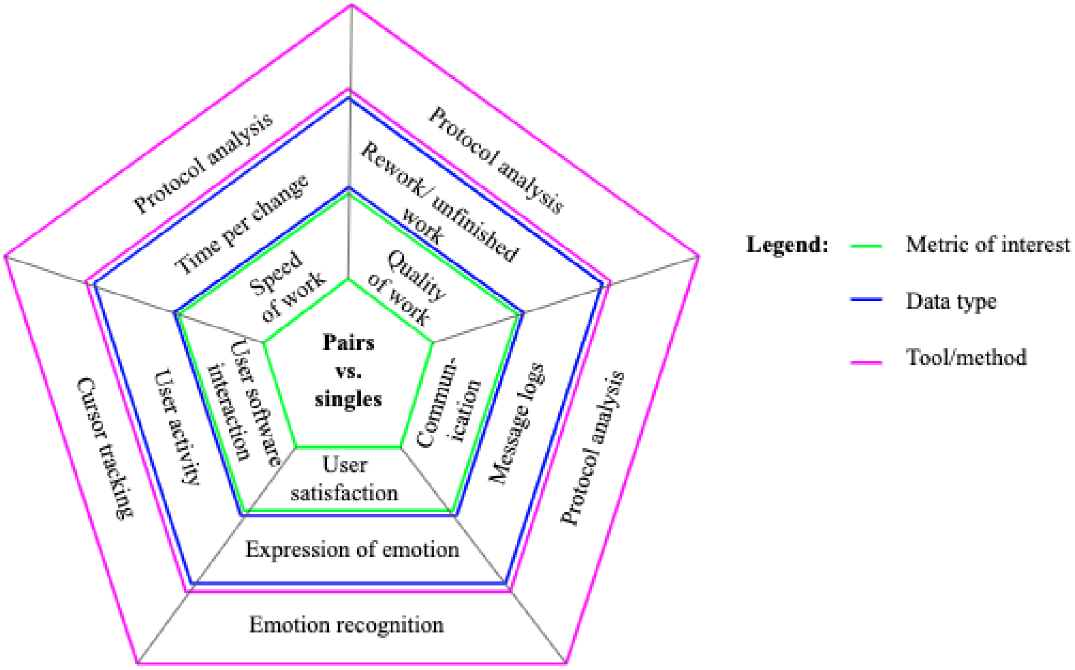


Figure 3-2: Architecture of toolkit showing metrics of interest, data types, and tools.

- Workstations: Mac-based computers running SOLIDWORKS and Onshape. These machines also captured webcamera footage and screen recording footage. Note that the webcamera was inbuilt.
- Monitor: 27" size, 16:9 aspect ratio monitors were used to display the CAD UI.
- Keyboard and mouse: Standard keyboard and mouse.

3.2.3 Data/Information Interconnection

Our toolkit was designed to handle a wide range of operating conditions imposed by the diversity of the CAD working styles. This was partly made possible by using configurable hardware/software. In this section, we provide details on how all stations were setup for the experiments.

We used cloud-based CAD on all collaborative CAD stations. The single working style stations used a local installation of SOLIDWORKS. In addition to participant stations, we used a central command station connected to external hard drives to

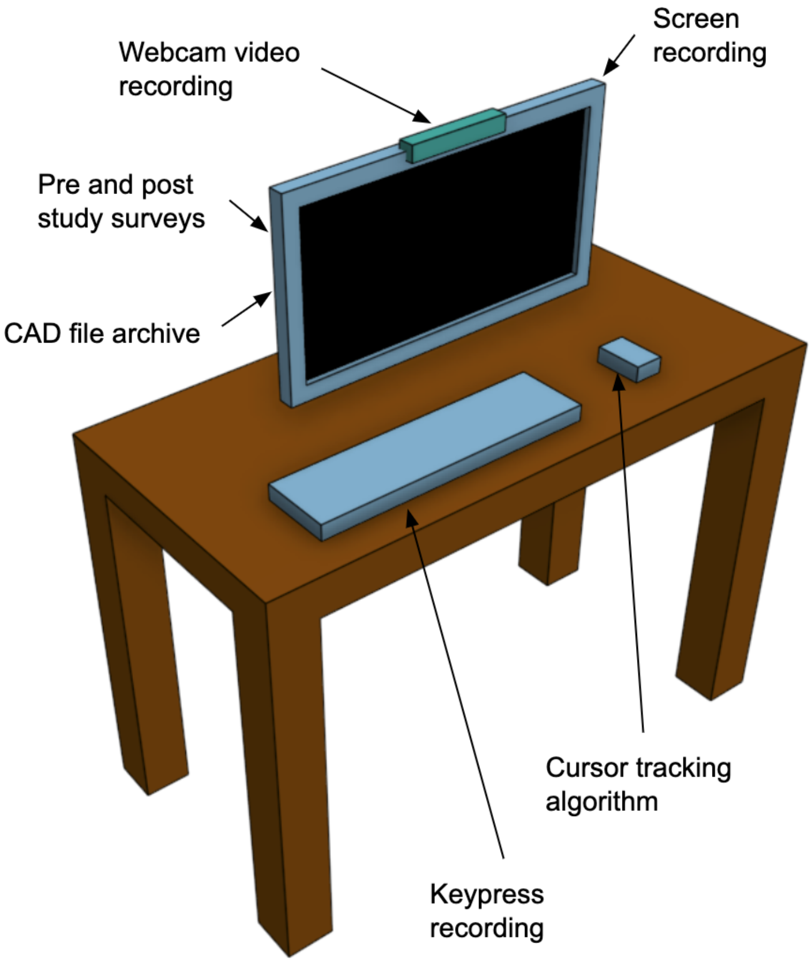


Figure 3-3: Hardware setup used in our experimental toolkit.

archive participant data. Post-experiment, ownership of all CAD files were transferred to the command center. See Figure 3-4 for an illustration of our data flow and interconnections. CAD files provided to pairs were connected online so that each participant could collaborate with their partners.

All data was stored using an anonymous coding scheme in compliance with the ethics board requirements. Video recording data was stored locally on participant stations and then archived post-experiment.

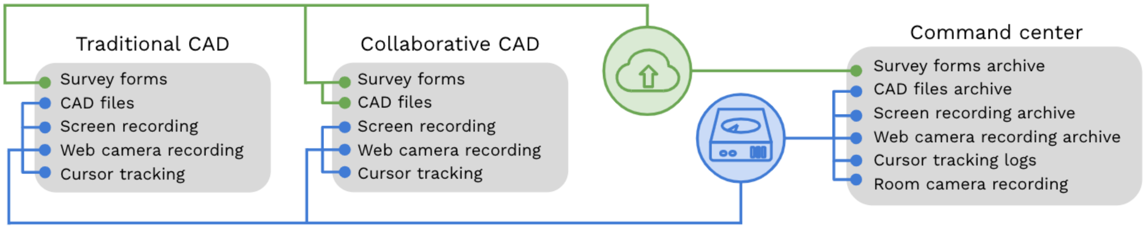


Figure 3-4: Work station layout and data flow diagram.

3.2.4 Software

All data streams captured during our pilot study are tabulated in Table 3.1. A description of each data type follows. Note that some metrics are redundant/overlapping and were used to triangulate results.

Manual Coding

Protocol analysis allows us to codify information that is otherwise hard to capture using automated techniques. A coding schema is central to a protocol analysis. Frameworks like function-behaviour-structure (FBS) are popular in the design research community [75]. However, a CAD-specific coding schema is currently lacking in the literature. We implemented a grounded theory approach to develop a CAD collaboration specific coding schema [76].

The five tracks shown in Figure 3-5 are: geometry modification (CAD work), communication, rotate/3D spin, roll back, and help access. These were derived from watching multiple screen recording videos and identifying patterns in them. All videos

Table 3.1: Summary of data source, post-processing tools, and data type.

Name	Source	Post-processing	Data type
Protocol analysis (manual)	Screen recording	V-Note	Quantitative: Time stamping events: communication, CAD file rotation, model rollback, help menu access
Cursor tracking (automated)	Screen recording	Open CV	Quantitative: 2D position (X/Y location) of cursor
Emotion detection (automated)	Web-camera recording	Affectiva	Quantitative: Percentage of total time an emotion was expressed

were annotated using a video coding software, V-Note. A detailed description of each track in our coding schema follows.

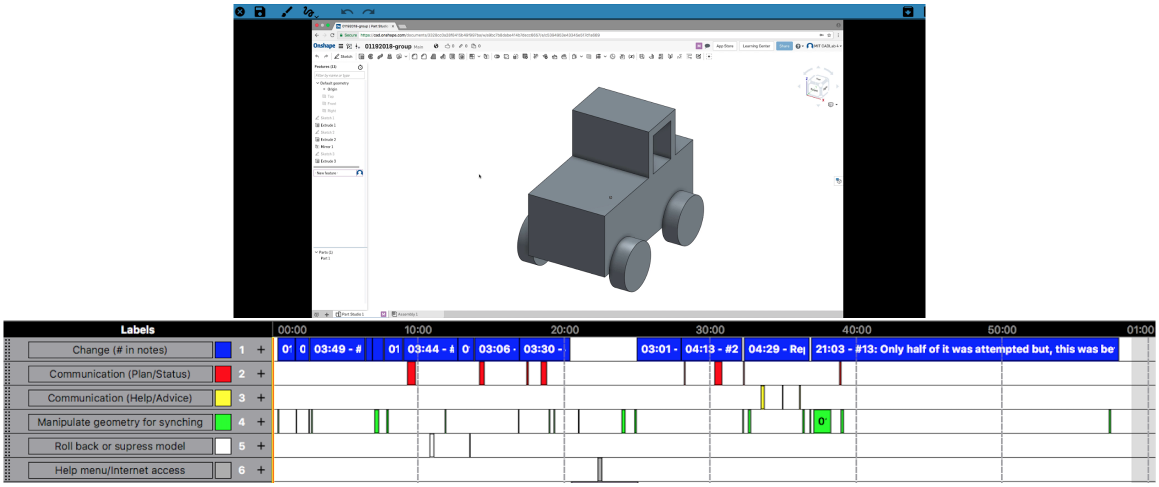


Figure 3-5: Manual coding in progress.

CAD Work

Tracking CAD completion times helped us record the time taken to complete a certain design task. Similar time based metrics have been used by other design researchers to compare different CAD working styles [26]. We included three subcategories into this track: active/productive work, rework, and incomplete/unfinished work. The statistics generated from time spent in each category were used to assess the participants performance in our pilot study.

Communication

Although our setup was capable of capturing written and verbal communication, we chose to limit our participants to written communication for this work. As it is easier to decode and analyse. Communication messages were tagged and marked with time stamps using the manual coding software. We coded messages into two categories: (i) communication to plan/update and (ii) communication to help/advice. Communication between participants was restricted to the design task on hand and participants were instructed not to identify each other in their communications.

Spin/Rotate

This track captured rotating the CAD geometry to seek awareness. We recorded the frequency and time duration of each instance of spinning/rotating of the CAD geometry.

Roll-back or Suppress

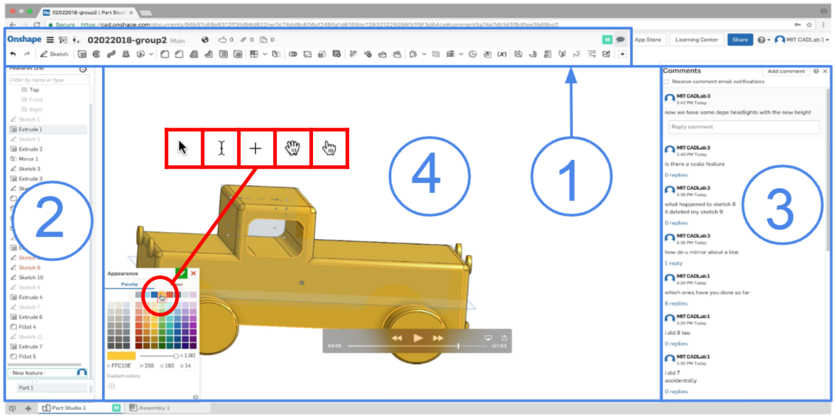
Rolling back the model tree is a common strategy used in CAD and is used to temporarily revert CAD features in the model tree. Each roll-back was documented with the time for which it stayed active and frequency in this rolled back state.

Help Menu Access

We expected our participants to use the software's help menu during the pilot study to assist with information on features. Time spent on referring to the help menu was catalogued in this track. Data generated from the manual coding process was aggregated for each participant and then combined on a per-working-style basis. For paired work, each participants data was calculated separately and then averaged with their partner to compute statistics for the pair.

3.2.5 Cursor Tracking

We built a post-processing object detection algorithm using template matching in OpenCV. This helped us capture cursor locations from screen recording videos. Cursor location data was then categorized in regions of interest as seen in Figure 3-6. The script scanned pixels of an input frame and correlated it to multiple input template images to find a match. This method obviates the need to maintain additional specialized cursor-tracking software during the pilot study. High thresholds were set for the cursor-template matches, and multiple calibration tests were performed. This was an important step in mitigating false detection and ensured repeatability of our cursor tracking results. To further counter false detections, repeated locations for more than 10 seconds were flagged. With our current parameter settings, the algorithm was able to recover more than 85% of cursor locations.



MIT CADLab3
5:42 PM Today
now we have some dope headlights with the new height

MIT CADLab3
5:40 PM Today
is there a scale feature
0 replies

MIT CADLab3
5:38 PM Today
what happened to sketch 8
it deleted my sketch 9
0 replies

MIT CADLab3
5:36 PM Today
how do u mirror about a line
1 reply

MIT CADLab1
5:29 PM Today
which ones have you done so far
8 replies

MIT CADLab1
5:29 PM Today
i did 8 too
0 replies

MIT CADLab1
5:26 PM Today
i did 7
accidentally
0 replies

Figure 3-6: Screen footage divided in regions of interest (ROI). 1: Feature toolbar, 2: Communication, 3: Model tree, 4: CAD window.

Each output log file generated at least 7,000 cursor location entries. Heatmaps were used for qualitative evaluation of cursor tracking data. An example is shown in Figure 3-7.

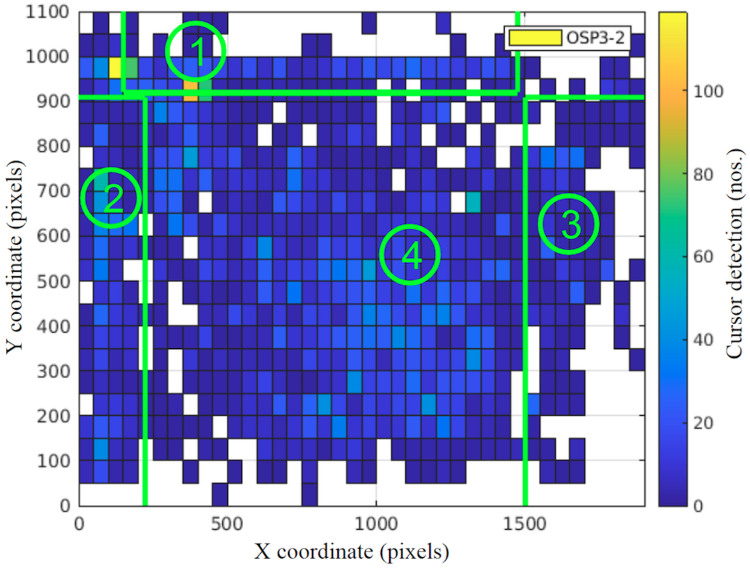


Figure 3-7: Heatmap showing cursor activity for a pair participant.

3.2.6 Emotion Detection

Facial expressions were tracked and processed to calculate emotion metrics for each participant. A tool using the software developer kit (SDK) within Affectiva was built to accept video uploads and generate emotion statistics [14]. The software worked by tracking facial cues and mapping landmark points as shown in Figure 3-8. These maps were then correlated to a library of seven known emotions. And in Figure 3-8 this mapped to the expression of surprise.

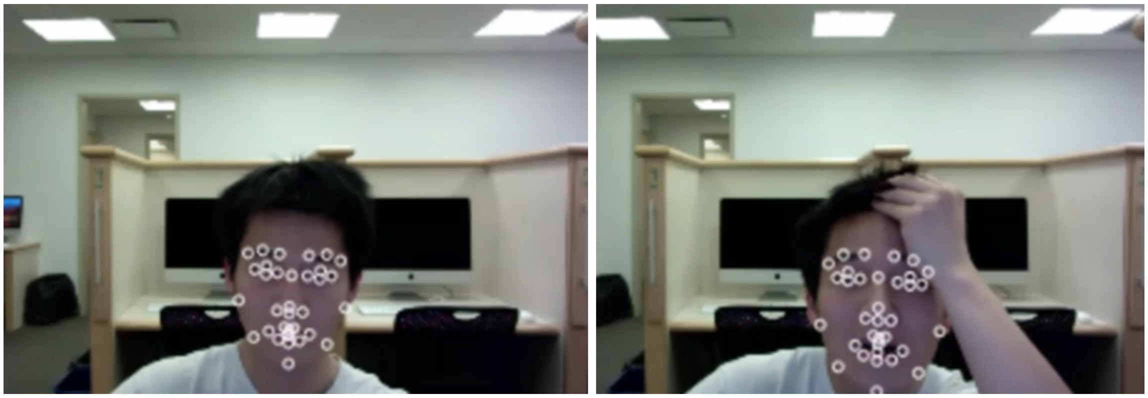


Figure 3-8: Tracking points for the expression of surprise [3].

The SDK output was in the form of a rating from 0-100, signifying the level of expression of a given emotion. A rating of 0 meant the emotion was not present and 100 meant the emotion was fully expressed. Our code sampled data at 2 frames per second (fps) to match the frequency of other software in our toolkit. Overall, we had a 81% detection rate for facial videos processed from the pilot study. An example output corresponding to expression of surprise discussed earlier is plotted in Figure 3-9.

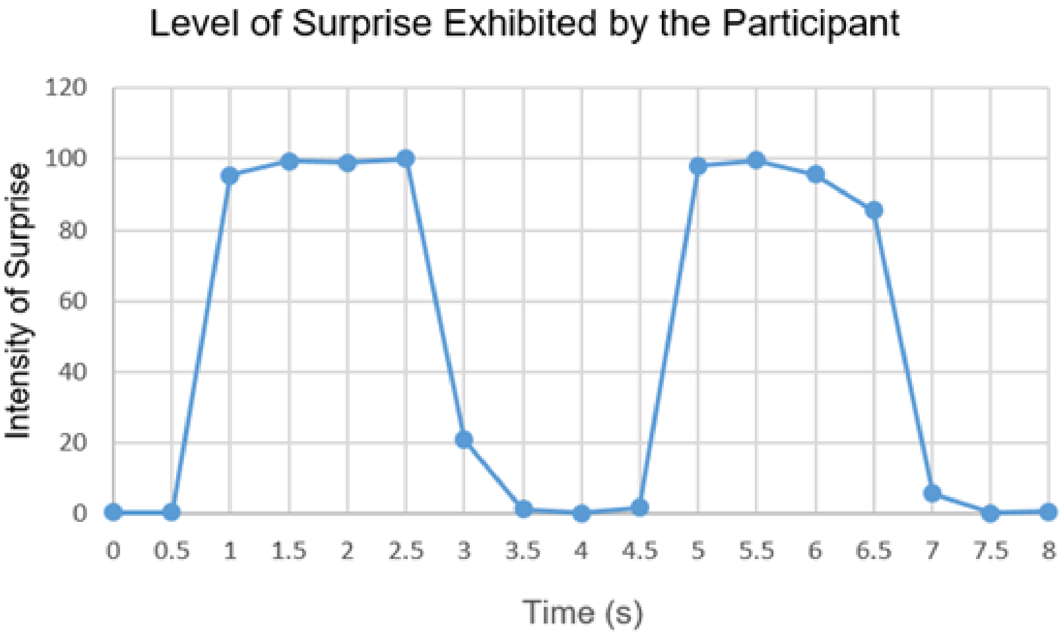


Figure 3-9: Plot showing intensity of surprise [3].

3.3 Pilot Study

Our pilot study goals were to validate the workings of our toolkit and to draw early insights for future work. The pilot study was designed and executed in compliance with the Committee on the Use of Humans as Experimental Subjects (COUHES) at MIT. As part of COUHES requirements, all participant names were anonymized.

3.3.1 Nomenclature

We chose three distinct CAD working styles: single person working in traditional CAD, single person working in synchronous CAD, and lastly a pair working in synchronous CAD. The three working styles help us test all possible modes in which our toolkit would be used in a future study. We were interested in studying pairs *vs.* individual CAD work and chose Onshape (OS) as it supported both working styles. We added a third working style with Solidworks (SW), a traditional CAD package, as benchmark. We adopted the following naming:SWS1 to SWSn for single SOLIDWORKS participants, OSS1 to OSSn for single Onshape participants, OSP1_1 and OSP1_2 for participants working in Onshape pair 1, upto OSPn_1 and OSPn_2 for participants working in Onshape pair n. See Figure 3-10 for a visual representation of the naming scheme.

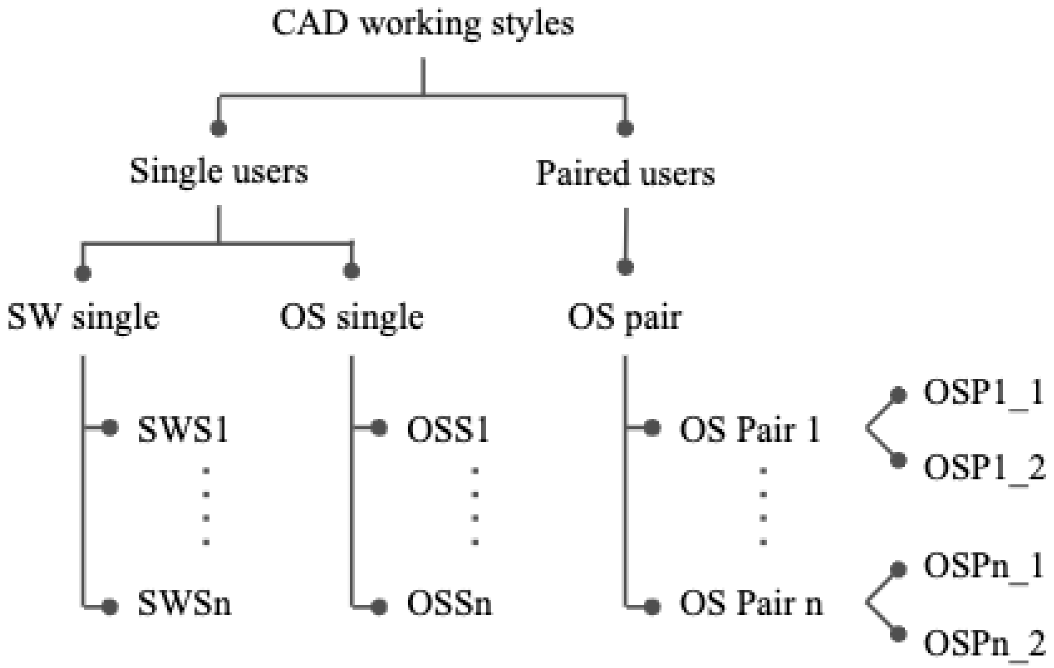


Figure 3-10: Naming scheme of different CAD working styles.

3.3.2 Pilot Study Facility

The experimental setup was built in a pre-existing facility in the MIT Behavioral Research Lab (BRL). Our focus was to be non-intrusive and maintain a design-office-like environment. We kept our setup void of any external cameras or body sensors.

It was important for us that participants were in physical and audio isolation from each other to simulate virtual collaboration. The physical space inside our lab was divided using slide-out partitions. We also added white noise in the environment to cancel out any residual audio. Figure 3-11 shows our experimental setup at the MIT BRL.

Participants were asked not to use personal electronic devices during the pilot study and adequate storage space was provided.

Figure 3-11: Pilot study setup at MIT BRL.

3.3.3 Participants

Our participant pool consisted of 12 students from MIT. It was required that all participants had taken a design class and used CAD for more than a year. A total of 16 people participated in the study but we considered 12 participants data for our analysis. Some data was discarded because of technical glitches or participants lacking prerequisites. On average, our participants had more than 2.5 years of CAD experience. Every participant was given a demo of the unique features in synchronous CAD including its collaborative features.

3.3.4 Design Task

Participants were asked to role-play toy designers working on an early-stage concept of a toy car shown in Figure 3-12.

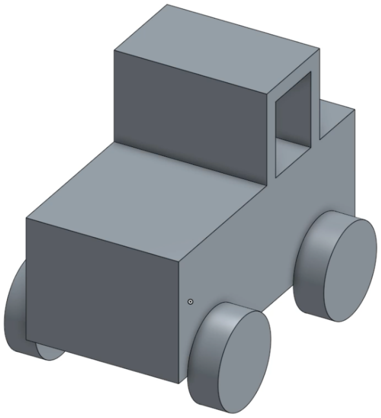


Figure 3-12: Initial CAD file provided to participants.

After reviewing the initial CAD file, participants were tasked with implementing a list of 42 changes that represented feedback from customer review sessions. Examples changes are shown in Table 3.2. We purposely designed our change list to limit participants user-software interaction and perform drafting tasks only. This prescriptive nature of our change list minimized variability in interpretation of the change list and forced participants to spend the majority of their time on CAD modeling. Participants were asked to implement as many changes as possible in 60 mins, with each change awarded a score. The objective was to get the highest score and thus, make as many changes as possible.

3.4 Pilot Study Results and Conclusions

Our pilot study led to 9 CAD files, as shown in Figure 3-13. OS Single and SW single participants' CAD files are not as elaborate as their paired counterparts. This is due

Table 3.2: Examples from change list.

#	Change name	Score
1	Add 3mm fillet overall on the body of the car	1
2	Increase the length of the car by 20%	1
3	Change the diameter of the wheels to 30mm	1
4	Increase width of tires by 50%	1

to fewer changes implemented and this difference is captured more quantitatively in the following sections.

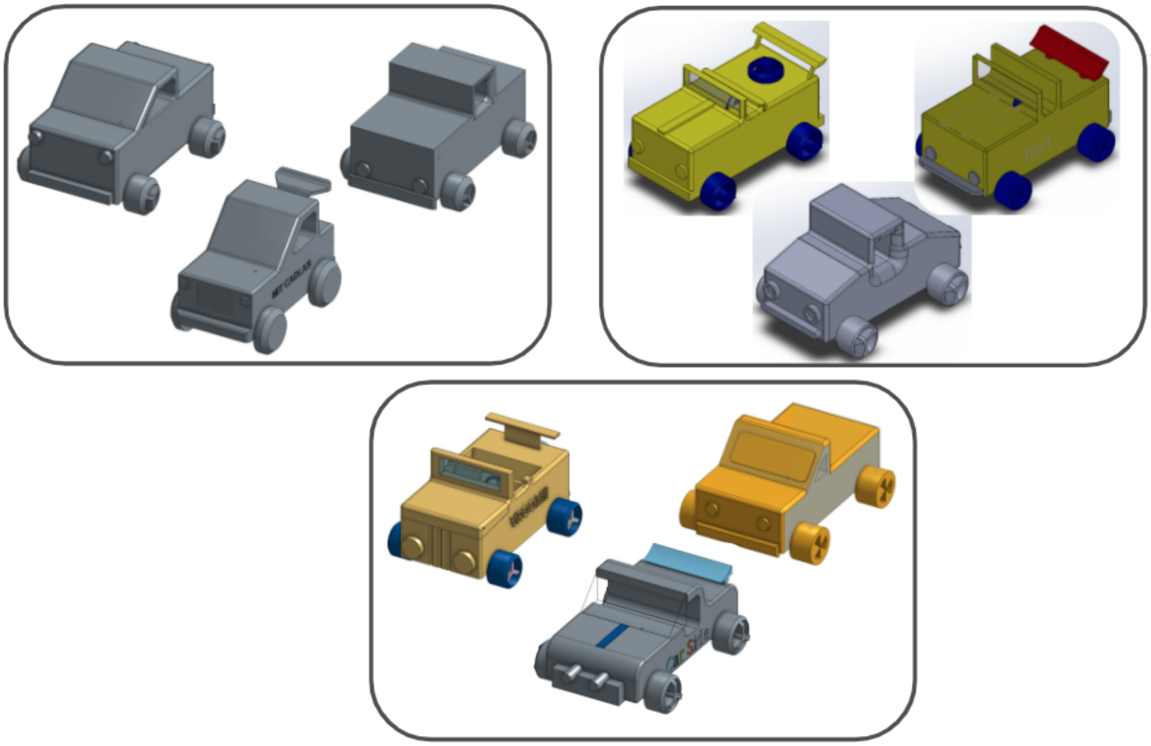


Figure 3-13: Final CAD files (Clockwise): OS Single, SW Single, OS Pairs.

3.4.1 Manual Coding

Inter rater reliability (IRR)

Manual coding of videos relies on judgement of human coders to classify objects into perceived categories. It is thus important to validate a coding schema's repeatability by testing its interpretation by multiple people. Cohen's Kappa is a measure of inter-rater reliability (IRR) and is often used to check the reproducibility of a manual coding schema [77, 78]. A second coder rated 99 samples from the manual coding video repository. Video samples were distributed such that we had at least one representation from each track of our coding schema. These ratings were cross referenced with their original categorization to compute IRR statistics. We achieved 85% agreement between ratings and our IRR assessment resulted in a high Cohen's K statistic of 0.71 suggesting a substantial agreement between raters [79].

Performance Analysis

Participant success in our pilot study was measured by the number of changes completed in 60 minutes. As seen from the first two rows of Table 3.3, paired participants were slower at implementing changes in comparison to individuals. It is noteworthy that the average time taken to complete a change is not directly indicative of the total number of changes implemented. This difference can be explained by analyzing the non-CAD activities of the protocol analysis.

Overall modeling efficiency (OME) is an aggregate metric used to compare relative performance of participants. We adapted this approach from manufacturing literature, where overall equipment efficiency (OEE) is used to compare production efficiency of machines [80–82].

We defined three ratios to compute OME: availability ratio, speed ratio, and quality ratio. Availability ratio captures the total time available to participants to perform CAD related activities. Time spent towards activities like rework, failed changes, and communication were removed from total study time to calculate the available time for each working style.

Next, speed loss ratio was used to capture the relative differences in speed of executing changes. This ratio was calculated by base lining each participant with the fastest participant, SWS1.

And lastly, quality loss ratio accounts for the accuracy with which each working style implements changes. This ratio penalized participants for unfinished or incorrectly implemented changes.

OME was calculated as a product of all ratios, as seen in Equation 3.1

$$OME = \frac{S}{S'} \times \frac{A}{A'} \times \frac{Q}{Q'} \tag{3.1}$$

where,

S = Total number of changes completed by participant

S' = Maximum number of changes completed by any participant, 15 in our case

A = 3600 - time lost performing non-CAD related activities

A' = 3600 (secs)

Q = Number of changes attempted - number of unfinished or incorrect change

Q' = Total number of changes attempted by participant

All ratios and OME values are tabulated in Table 3.3.

Table 3.3: Summary of overall modeling efficiency (OME) calculations.

Working style	SWS1	SWS2	SWS3	OSS1	OSS2	OSS3	OSP1_1	OSP1_2	OSP2_1	OSP2_2	OSP3_1	OSP3_2
No. of changes	15	15	9	13	12	9	9	8.00	6	10	5	8
Ind. speed loss	1.00	1.00	0.60	0.87	0.80	0.60	0.60	0.53	0.40	0.67	0.33	0.53
Avg. speed loss	0.87			0.75			0.51					
	Speed loss SWS			Speed loss OSS			Speed loss OSP					
Qual. loss time	977	361	326	169	336	0	853	1331	0	0	144.00	331
Commun. time	0	0	0	0	0	0	219	156	592	541	725	311
Non-CAD time	977	361	326	169	336	0	1072	1487	592	541	869	641
Ind. availability	0.73	0.90	0.91	0.95	0.91	1.00	0.70	0.59	0.84	0.85	0.76	0.82
Avg. avail. ratio	0.85			0.95			0.76					
	Availability SWS			Availability OSS			Availability OSP					
No. of qual. loss	4	3	2	3	3	0	3	2	0	0	2	2
Ind. qual. loss	0.79	0.83	0.82	0.81	0.80	1.00	0.75	0.80	1.00	1.00	0.71	0.80
Avg. qual. loss	0.81			0.87				0.84				
	Quality loss SWS			Quality loss OSS				Quality loss OSP				
Ind. OME	0.58	0.75	0.45	0.67	0.58	0.60	0.32	0.25	0.33	0.57	0.18	0.35
Avg. OME	0.6			0.6			0.3					
	OME SWS			OME OSS			OME OSP					

3.4.2 Cursor Tracking

Modelling the cursor tracking data proved challenging, given its size and complexity. To capture the time aspect of our data, Markov chain models were built to depict transitions between regions of interest. A model was created for each participant but to simplify our analysis we chose to aggregate data for each working style. This led to two models as shown in Figure 3-14. Each transition state in the below models maps to the regions of interest identified in Figure 3-6.

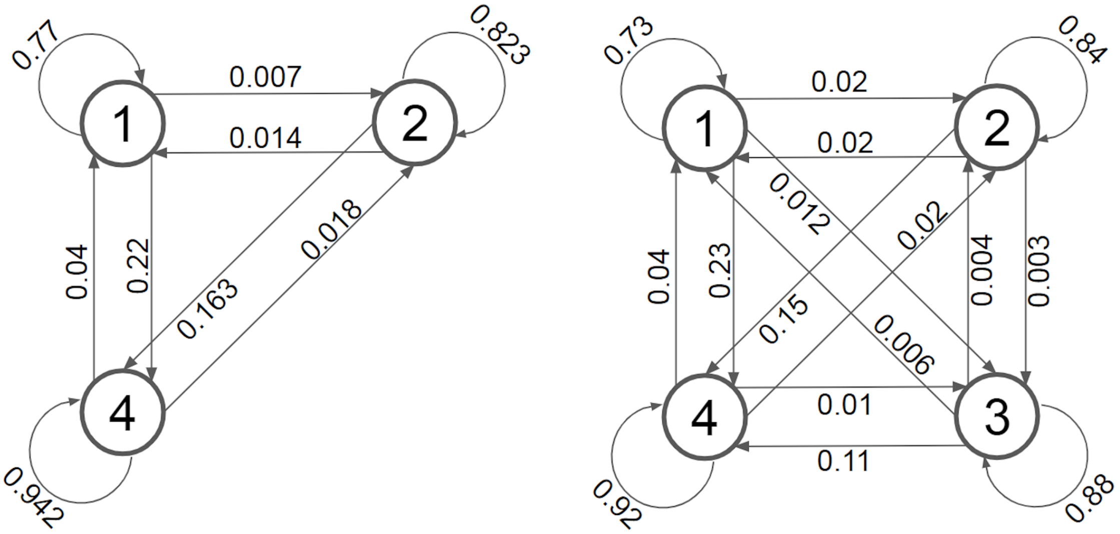


Figure 3-14: (Left) Markov models of single participants and (right) pair participants. Transition states: 1. Feature toolbar, 2. Model tree, 3. Communication, 4. CAD window.

3.4.3 Emotion Detection

Similar to the cursor tracking data, emotion tracking data was computed at a high frequency and then down-sampled. All instances of sustained emotion were catalogued in one of the seven categories in Figure 3-15. Paired participants expressed more emotions than individuals. A more elaborate analysis of this emotion data can be found in Zhou *et al.* 's work [14].

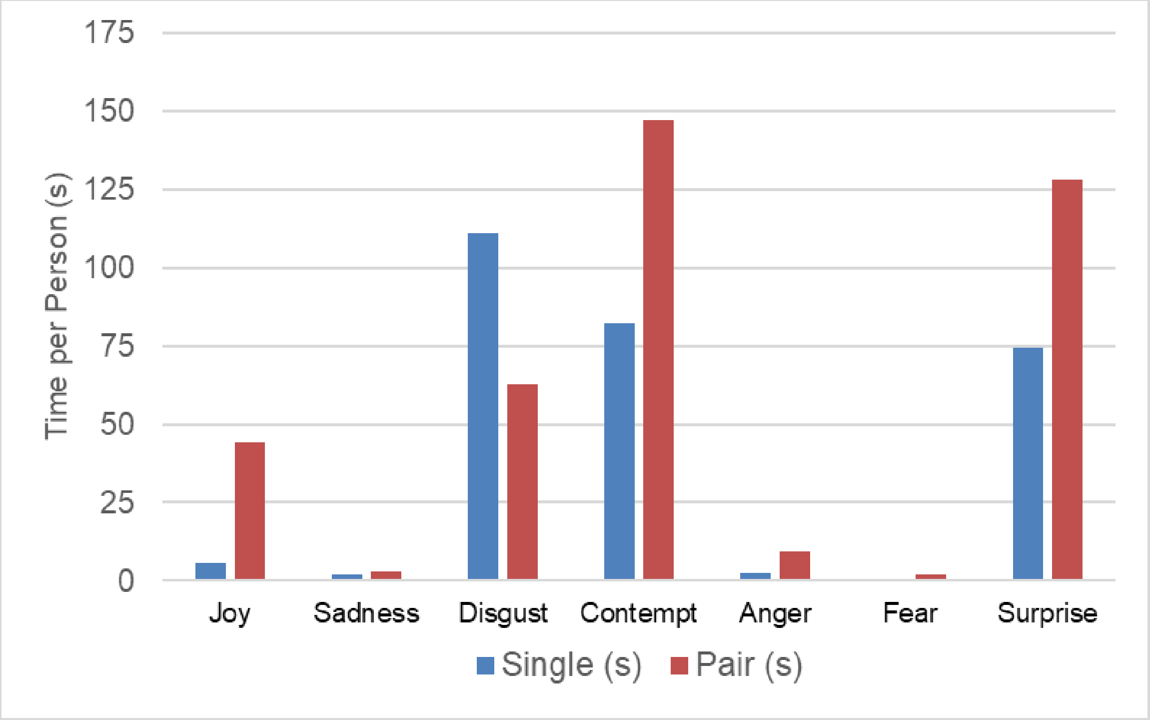


Figure 3-15: Aggregate time metrics for emotions expressed by singles *vs.* pairs [3].

3.4.4 Aggregating all Data Streams

Analysing multiple data streams together can prove challenging and makes it difficult to compare the three working styles comprehensively. In an effort to aggregate data, we created dotted plots as seen in Figure 3-16 (a). These plots show the manual protocol analysis data alongside the cursor tracking data. As the two data-sets were not mutually exclusive, it was difficult to set up a quantitative analysis. We used the dotted plots to derive qualitative understanding of our participants behaviour. For example, in Figure 3-16 (b), it is seen that OSP1_1 and OSP1_2 manipulated the geometry lot more than OSS3.

The dotted plots augment information presented by the protocol analysis and Markov models by adding a time dimension. For example, the transition probabilities shown in Figure 3-14 give us an aggregate view of the participants movement within the CAD environment. But the dotted plots in Figure 3-16 help us understand the distribution of time spent over the timeline of our pilot study. This tells us if the transitions are happening regularly or in spurts.

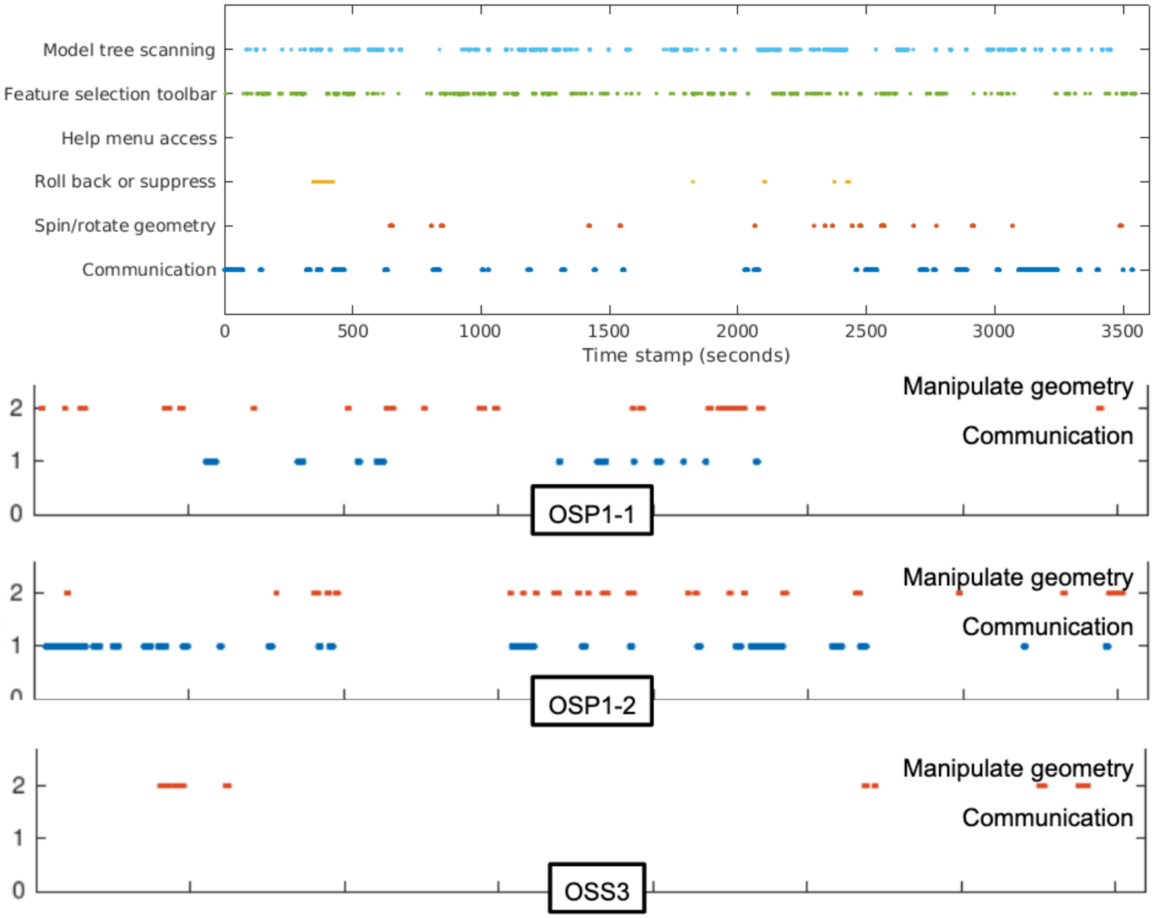


Figure 3-16: (Top) (a) Complete dotted plot for OSP3_1 (bottom) (b) excerpts from participant dotted plot.

3.4.5 Limitations

The small size of our participant pool reduces the confidence in our results. Using students as participants in design research has precedence in the field, but results from such work falls short in their generalisability to real world scenarios. Rigorous training of student participants can help minimize this disconnect and will be implemented in future.

We chose a design task that was drafting focused. This provided greater clarity of the design tasks to participants but limited the scope of our results.

Our toolkit heavily relied on post processing of datasets. This lent convenience in running the experiment, but at the cost of data losses from video post processing algorithms. For example: cursor tracking was implemented for OS pairs and OS single working styles only, as the SWS CAD UI was drastically different in comparison. This posed problematic for our cursor tracking script and it could not deliver consistent results. The comparison between OS pairs and OS singles was adequate to compare individual CAD work against paired work in the pilot study. However, a better coverage of all working styles can be achieved by running cursor tracking natively, a simple extension of our experimental set-up

3.4.6 Conclusions

This chapter outlined a multimodal experimental approach to study CAD work and successfully validated its working in a pilot study setting. The resulting dataset was analysed to create models and representations to study differences in paired CAD work *vs.* individuals.

The hardware and software tools provided consistency and standardization in running experiments. In the pilot study, it was noted that single participants were faster than pairs. This is in agreement with other literature on computer supported collaborative work [83].

More broadly, collaboration is inherent with losses and a good understanding of the underlying causes of such overhead and possible gains is paramount [49].

Chapter 4

Experiment and Results

4.1 Overview

This chapter describes the final version of the experiment which is based on learning from Chapters 1-3. All the major contributions of this research are in the following chapter. We firstly present the revised experiment setup, followed by results pertaining to research questions surrounding speed, quality, communication, UI activity, and user satisfaction.

4.2 Research Framework

4.2.1 Implementation of Pair CAD Working Styles

Figure 4-1 show the two working styles chosen for the experiment. These were first introduced in Chapter 2, Figure 2-2. All participants in the experiment used Onshape as the primary CAD software. Onshape was configured differently and sometimes used alongside additional software, depending on the prescribed working style.

All participants worked on a single part file and not an assembly. This was important as we did not want to provide any spatial compartmentalization and potentially influence the collaboration strategy of pairs. For example, multiple part files in a CAD assembly might lead participants to work on individual parts separately because of the nature of the CAD file.

Parallel CAD participants worked on the same Onshape file, wherein both participants had access to a common database. But they worked on a local instance of the database allowing them to change the CAD file, in parallel. This style of work, like mentioned in Chapeter 2, is analogous to text editing in Google Docs [84]. Parallel CAD pairs had no obvious incentive to look over each others work and this was expected to reduce awareness. Note that Onshape also provides a limited screen sharing option but it is not the natural mode of work of the CAD software.

Shared CAD participants used a screen sharing tool called Use Together in addition to the provided CAD model. This working style was analogous to using video conferencing software like Zoom/Skype. This meant they had to take turns at using the software to work on their CAD database. Each pair participant had their own independent keyboard/mouse control and their own cursor pointer. This style of work was expected to increase awareness amongst pairs as participants had to look over each others work throughout the experiment.

Lastly, the individual CAD working style participants worked by themselves and were representative of traditional CAD use.

4.2.2 Research Model

We compare the three working styles on the basis of speed and quality; shown as RQ1 and RQ2 in Figure 4-2. The results of this comparison were validated by assessing the influence of participants incoming skill level on their pair performance. Lastly, we use supporting questions: S1, S2, and S3 to further elaborate on our primary research questions. The architecture of our research model is based on previous work in pair programming [83].

Below is a list of all our research questions

Primary research questions:

- RQ1: On a per person basis, is Parallel CAD faster than Shared CAD and Individual CAD work?

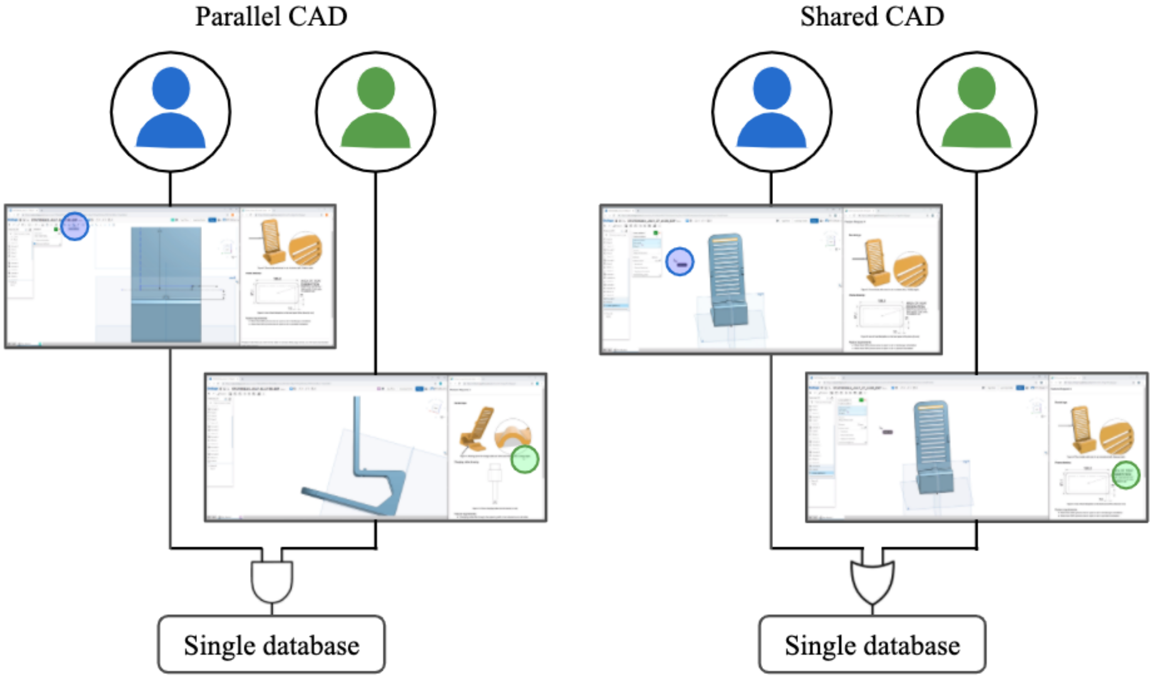


Figure 4-1: Illustration of pair CAD implementation.

- RQ2: Does Shared CAD lead to higher quality work compared to Parallel CAD and Individual CAD?

Supporting research questions:

- S1: Do Shared CAD participants communicate more than Parallel CAD participants?

- S2: Do Parallel CAD participants share work more equally compared to Shared CAD participants?

- S3: Are Parallel CAD and Shared CAD participants more satisfied than Individual CAD participants?

4.3 Experiment Design

As part of this research, we specified and procured new hardware. Our experiment consisted of various phases and this posed unique constraints on the hardware requirements.

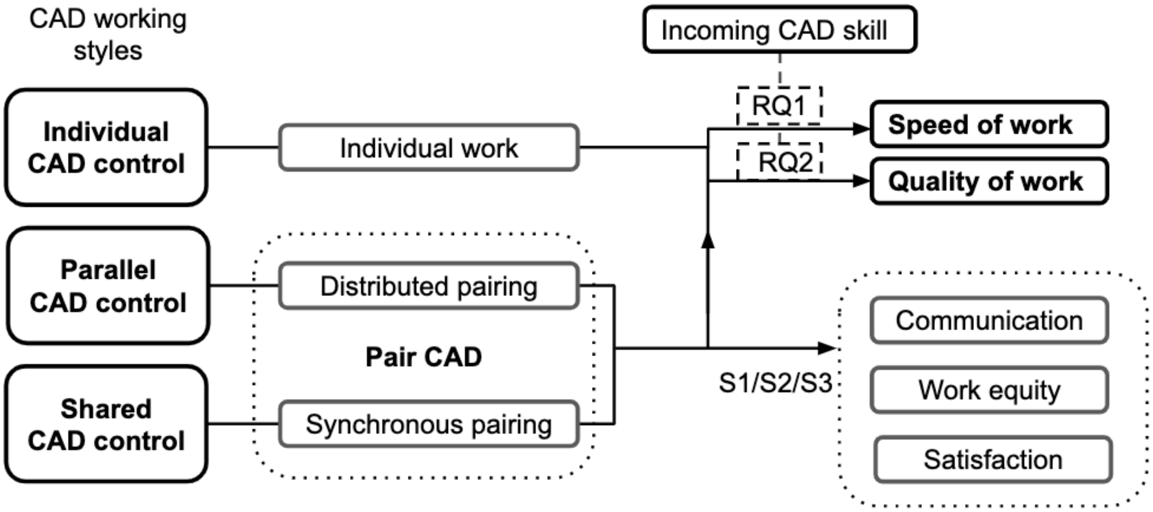


Figure 4-2: Research model.

Before beginning our work, we obtained ethics approval from the Research Ethics Board (REB) at UofT. All participant data presented in this section was code-named to comply with REB stipulations.

4.3.1 Equipment (Hardware and Software)

We designed our lab space to mimic a typical design work environment in industry. This meant we could not use excessive audio video (AV) equipment and chose not to use body mounted sensors like skin conductance and electroencephalogram (EEG). The non-intrusive approach proved successful during our pilot study at MIT BRL and thus, was replicated at UofT. This also helped in simplifying our hardware setup and provided a natural setting for our participants. The details of our hardware are shown in Figure 4-3. All equipment shown highlighted is as below:

1. Audio communication: Over-the-ear headset with an external mic. The setup also had a remote to control volume and mute.

2. Face recording: An external web-camera was used to record participants faces. We recorded footage at 1080p x 60fps with a stable frame rate.

3. Display: Ultra wide 34" monitors were used to mimic a dual screen setup. Our

screen resolutions was 2560 x 1080; enabling us to show the design task and CAD UI alongside each other. See Figure 4-4 for an illustration of the participant dashboard.

4. Input: A standard QWERTY keyboard and a mouse with scroll were used as input devices. Our setup was adaptable for left-handed users.

5. Furniture: A height-adjustable chair and standard height tables were used. After Phase II, portable room dividers were used to reorganize the room into compartments.

6. Computers: Each participant station was capable of running the front end: CAD UI, design task and collaboration tool and back end: recording AV streams and performing UX analysis.

Our minimal hardware setup meant that we had to post-process all data streams to extract the necessary metrics. Below is the list of software used in our experiment.

- UX analytics: A python-based script collected the cursor point locations, button clicks, keyboard typing, and scroll wheel use. This script ran in the background.

- Video data stream: All video data was processed and archived using open broadcaster software (OBS).

- Automation: It was important to maintain the time synchronicity of our data. This was done by starting all software simultaneously and keeping track of any discrepancies. We used a modified version of the cursor tracking python script to initiate all video and UX analysis software at once.

- CAD: A configured version of Onshape was used as our primary cloud-base CAD software. We accessed the CAD UI and all design tasks through Google Chrome. We used Windows 10 as the operating system on all of our computers..

- Audio communication: Participant-to-participant communication was enabled *via.* Google Hangouts and Use Together.

Figure 4-3: (Clockwise) Participant workstations, Phase I setup, Phase II setup, Phase III and IV setup.

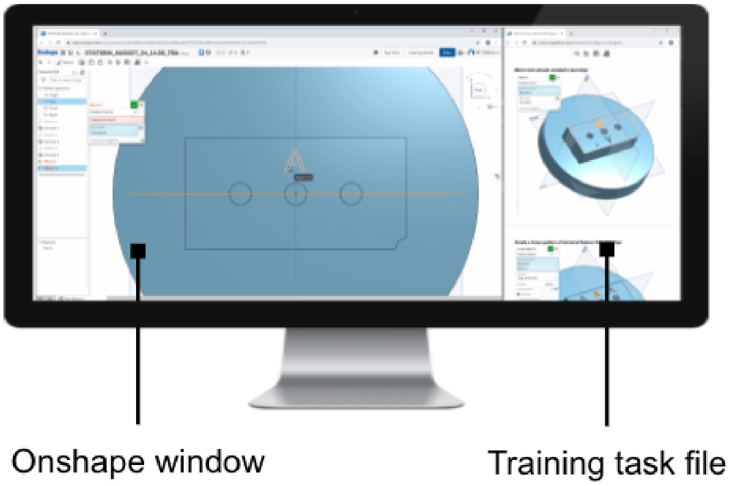


Figure 4-4: Participant dashboard.

4.3.2 Phases of Work

All experiments were conducted in four phases shown in Figure 4-5. A pre-study and post-study survey were used to collect demographic information, feedback and evaluations. To maintain consistency, experiment facilitators adhered to a strict standard operating procedure (SOP). During the experiment, at least two study coordinators were present in the lab. New study coordinators were required to review the SOP and shadow current coordinators, before they were allowed to run experiments on their own.

In Phase I, we trained participants on the provided CAD software. We relied on the participants incoming CAD experience to base our training material. We modified the CAD UI to allow only a subset of all features. This allowed us to cover all available features in the limited training time. This basic feature set was chosen on the basis of relevance to design tasks in our experiment.

The training first focused on 2D drafting, followed by 3D modelling. During the training, a subgroup of features was demoed by the facilitators. And then the participants independently worked through a short exercise that tested their understanding of the demoed feature set. An example instruction page with the training file is shown in Appendix Figure A-1. Participants were encouraged to ask questions during Phase I. Lastly, each Phase I demo was delivered using a script, ensuring a consistent training experience in each run.

We designed a benchmark metric that was better aligned with our design tasks and representative participant performance in study settings. We decided against using the popular Purdue spatial visualisation test - rotation (PSVT-R) to gage our participants CAD skill level [26,85]. PVST-R was too simple and general to give use as a benchmark in our work. We created multiple CAD design tasks in Phase II based on the SOLIDWORKS certification process. The certification is commonly used by industry professionals as attestation of CAD proficiency. Any participants that could not complete a single task in Phase I was asked to abort the experiment.

Phase III introduced CAD collaboration to our participants. Participants were shown a demo of their prescribed collaboration method and then had 5 minutes to work on a modified version of the Phase II design task using the collaboration method. Participants were encouraged to ask questions as they worked through Phase III. Repeating the design task from Phase II helped us keep our participants focus on the collaboration method and not the mechanics of the design task on hand. The experiment room was reorganized using room dividers to isolate participants from each other. This was done to mimic a fully virtual collaboration. To minimize audible noise from multiple pair runs, we induced white noise in the background.

Phase IV was considered the true experiment. In this phase, participants used the prescribed CAD collaboration to work through a series of design tasks. Design tasks were presented in the same order and sequentially. Participants were not allowed to go back to a previous task. Phase IV design tasks were still CAD specific but more open ended than Phase II tasks. This was our longest phase and considering that our participants had spent upward of an hour in lab already, we added a short break before beginning Phase IV. The room layout remained consistent with Phase II and included the white noise.

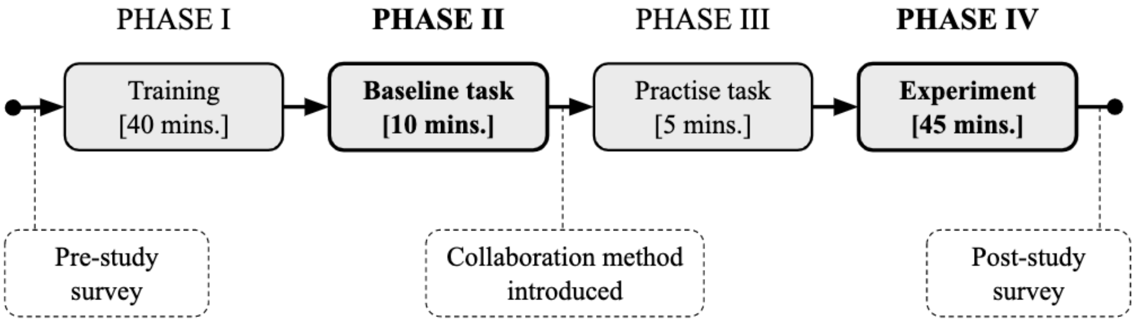


Figure 4-5: Phases of our experiment.

4.3.3 Operationalization Matrix

Table 4.1 shows the mapping of the data collected in the experiment to the research questions from section 4.2.2. Such operationalization matrices are used by other researchers and helped us in planning our experimental setup from the early stages [29, 86].

Table 4.1: Summary of data source, post processing tools, and data.

Construct	Variables	Type of data (qty.)	Data source
Speed of work (RQ1)	Number of CAD tasks completed in Phase IV	Quantitative data (n=37)	Time log sheets and user activity data from Qualtrics survey
Quality of work (RQ2)	Average quality of tasks completed in Phase IV	Quantitative data (n=37)	Rating CAD files using standardized grading rubric [4]
Communication (S1)	Amount (%) of study time spent communicating	Quantitative data (n=23)	Audio trace from web camera recording using open broadcaster software (OBS)
Work equity (S2)	Amount (%) of study time spent interacting with GUI	Quantitative data (n=60)	Custom python script tracking cursor location, clicks, scroll, and keystrokes
Satisfaction (S3)	Evaluating self reported satisfaction scores	Quantitative data (n=60)	Average of Likert scale responses to post study survey questions

4.3.4 Design Task

All design tasks were administered using Qualtrics. This allowed us to collect UX data like form completion times, number of clicks, and computer names. This was useful in precisely calculating Phase IV task completion times. Using Qualtrics, we also ensured that participants progressed through the experiment one task at a time and did not browse through multiple tasks. This unidirectional nature of the design task helped us compare everyone's work consistently.

Phase II design tasks were created to evaluate the CAD drafting skills of our participants. We chose a simple CAD model shown in Figure 4-6 (a) as our starting point. Then, we asked participants to add a feature at a time to eventually end up with the final CAD file. Instruction in Phase II were very prescriptive and structured. We evaluated Phase II work on the basis of speed and quality. Appendix Figure A-2 shows an example Phase II design task page.

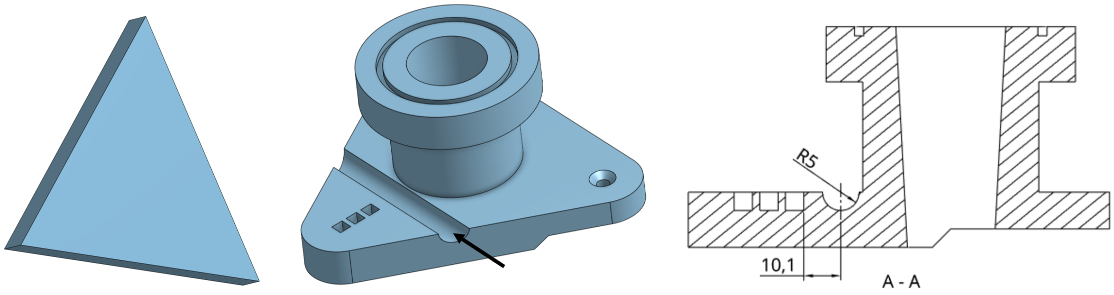


Figure 4-6: (Left) (a) Initial CAD file, (middle) (b) Final CAD file, (right) (c) Instructions (in mm).

In Phase IV, participants worked on a more open-ended task compared to Phase II. Appendix Figure A-3 shows an example Phase IV design task page. In addition to text based description, we included renderings of the requested feature. Before beginning to change the CAD model, participants had to deliberate on the provided design task description.

Phase IV began with an initial CAD file for a phone holder as shown in Figure 4-7 (a).We ensured that Phase IV design tasks were not coupled to each other. This was particularly important for Parallel CAD work.

Figure 4-7: (Left) (a) Initial CAD file, (right) (b) Final CAD file renderings.

4.3.5 Pilot Runs

We ran a total of 15 pilot runs. Recruitment for these runs happened mostly through the research team's network but also included some paid runs. Feedback from each pilot run was debriefed upon and implemented before subsequent pilots. The pilot runs were stopped once no more iterations were required and the research team was happy with the state of our setup. The experience from executing pilot runs served as practise for the research team in running experiments.

4.3.6 Participants

Like in other design research, our participant pool was primarily students [87–89]. For our experiment, we recruited students from UofT, Ryerson University, George Brown College, and Ontario College of Art and Design (OCAD) University. All of which are based in the greater Toronto area (GTA). More than two hundred people expressed interest in participation. Our sign-ups were sourced primarily (50%) from posters (See Appendix Figure C-1, and the remaining from emails, social media posts and others. We accepted the data of 60 participants from a total of 66 experiment runs over Summer 2019, Fall 2019 and Winter 2020. Participants were compensated at an hourly rate of $15 (Canadian Dollars).

A summary of our participant demographics is shown in Table 4.2. On average our participants had 33 months of 3D CAD experience and were 25 years of age. Each working style had participants around the same age and equal amount of CAD experience. We had a sub-optimal gender ratio of 22% (F/M) and this was representative of the demographic of mechanical engineers in the Greater Toronto Area. Lastly, 43.3% of our participants reported English as their second language (ESL) and our participant pool had a diverse group of ethnicities.

In our pairing assignments, we ensured that no two participants knew each other. This ensured that our results were not influenced by any preexisting working relationship. Other than this condition, we randomly allocated participants to their working styles.

Table 4.2: Summary of participants profile.

Name of entry	Individual CAD	Parallel CAD	Shared CAD	Total
Number of participants	14	24	22	60
CAD Experience (months)	33.4	33.4	33.3	33.4
Age (years)	24.5	24.5	25.1	24.7
Gender (F/M) (%)	7.1	37.5	13.6	21.7
ESL (Y/N) (%)	35.7	54.1	36.4	43.3

4.4 Internal Validity

We assumed that a few tenets from pair programming would translate to pair CAD work in mechanical engineering. These assumptions will be tested in the following section. It is to be noted that it was difficult to predict the below outcomes beforehand given the novelty of pair CAD work.

4.4.1 Awareness Assumption

Awareness (as defined in our work) is the ability to understand ones work in the context of your pair's [57]. To assess awareness, we used four Likert scale questions as shown below, where 5 was: "Strongly Agree" and 1 was: "Strongly Disagree". An overall awareness metric was derived by averaging the participants agreement to below prompts.

- *I was able to fully obtain information about the other member of my team (eg. partner's CAD skills, modelling styles, etc.)*
- *I was able to fully obtain information about what the other member of my team was working on*
- *I was able to fully obtain information about how our team's activities will be coordinated*
- *If I continued this study by myself, I'm confident in taking over my partner's work*

Figure 4-8 shows the variation of awareness scores between different working styles. As seen, Shared CAD participants scored higher than Parallel CAD. This result was also found to be statistically significant on an one-tailed, two sample independent paired t-test; which returned a p-value of less than 0.05.

In the post-study survey, we asked participants to identify their main source of awareness. Shared CAD participants reported that they sought awareness equally from observing the CAD model (50.0%) and audio communication (50.0%). Parallel CAD participants primarily used audio communication (75.0%) over observing the CAD file (25.0%) to seek awareness.

4.4.2 Bias From Pre-Existing CAD Experience

The second assumption was that a modified version of the CAD UI and our training will decouple participants performance from their incoming CAD skill. To test the

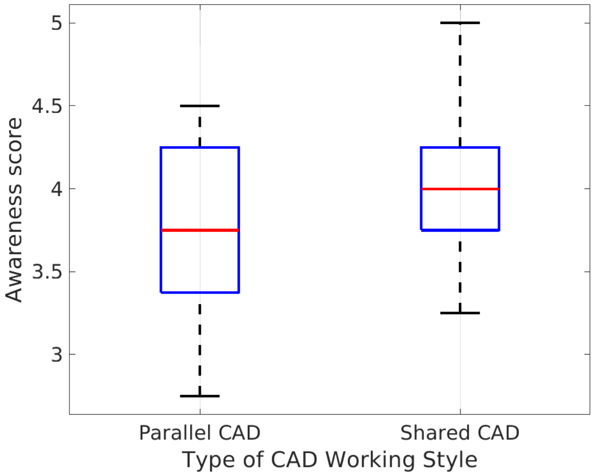


Figure 4-8: Awareness score *vs.* CAD working style.

validity of our assumption, we plotted Phase II data against participants self reported 3D CAD experience. As seen in Figure 4-9 there is no clear correlation between the two variables and an attempt to fit a linear model resulted in a poor R^2 of 0.014. Overall, this analysis gives us confidence in stating that Phase II performance was not dependent on participants incoming CAD skill.

4.5 Defining Speed and Quality Metrics

A summary of Phase IV performance is shown in Table 4.3. Parallel CAD participants implemented the most number of features followed by Shared CAD and Individual CAD. The difference in values was found to be statistically significant in an ANOVA test resulting in a p-value of less than 0.01. On a per person basis, Individual CAD participants executed the most features, followed by Parallel CAD and Shared CAD. This difference was also statistically significant with a p-value of less than 0.01. This distinction in the final outcome of various working styles is seen visually in the Appendix Figures B-1, B-2, B-3 showing the final CAD files for each working style.

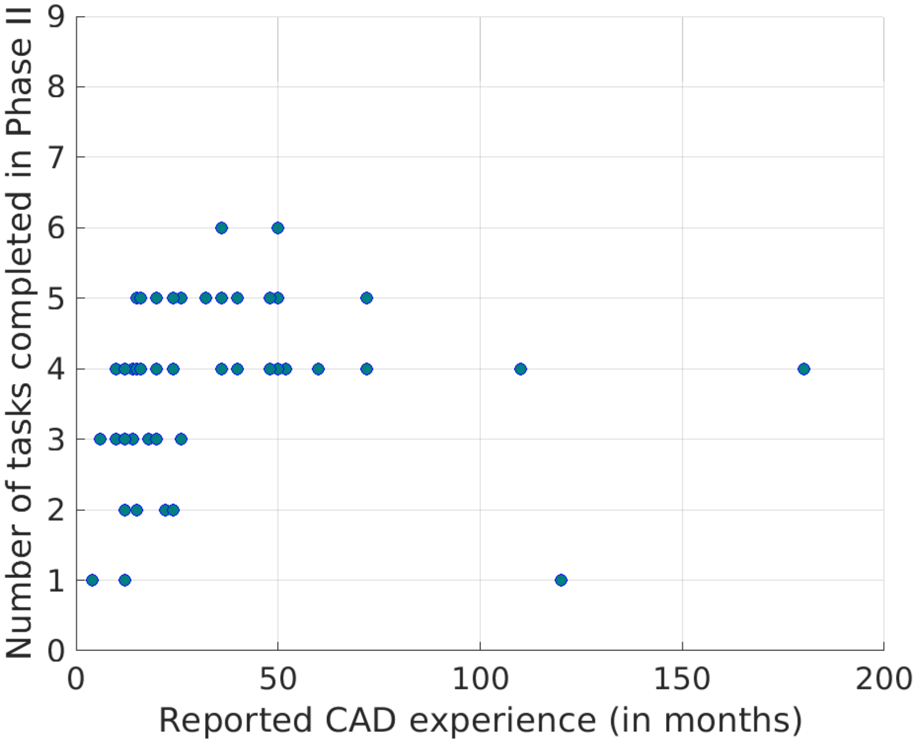


Figure 4-9: Number of Phase II tasks completed *vs.* self-reported CAD experience (in months).

Table 4.3: Summary of participant progress in Phase IV.

	Individual CAD	Parallel CAD	Shared CAD
Average number of tasks completed in Phase IV	4.2	6.3	4.9
Average number of tasks completed per person	4.2	3.15	2.45

4.5.1 Variation Between Design Tasks

Figure 4-10 elaborates on the results presented in Table 4.3. As we can see, the variance in completion times is large. Further the variation between design tasks is not consistent for all working styles. This makes it challenging to use the design task completion times as-is to compare speed of CAD work. In order to level all variances, we normalized the task completion times and derived new speed metrics. The normalization happened at the working style level, which means new speed scores were calculated treating data from each working style separately. The following section is an in-depth explanation of the same.

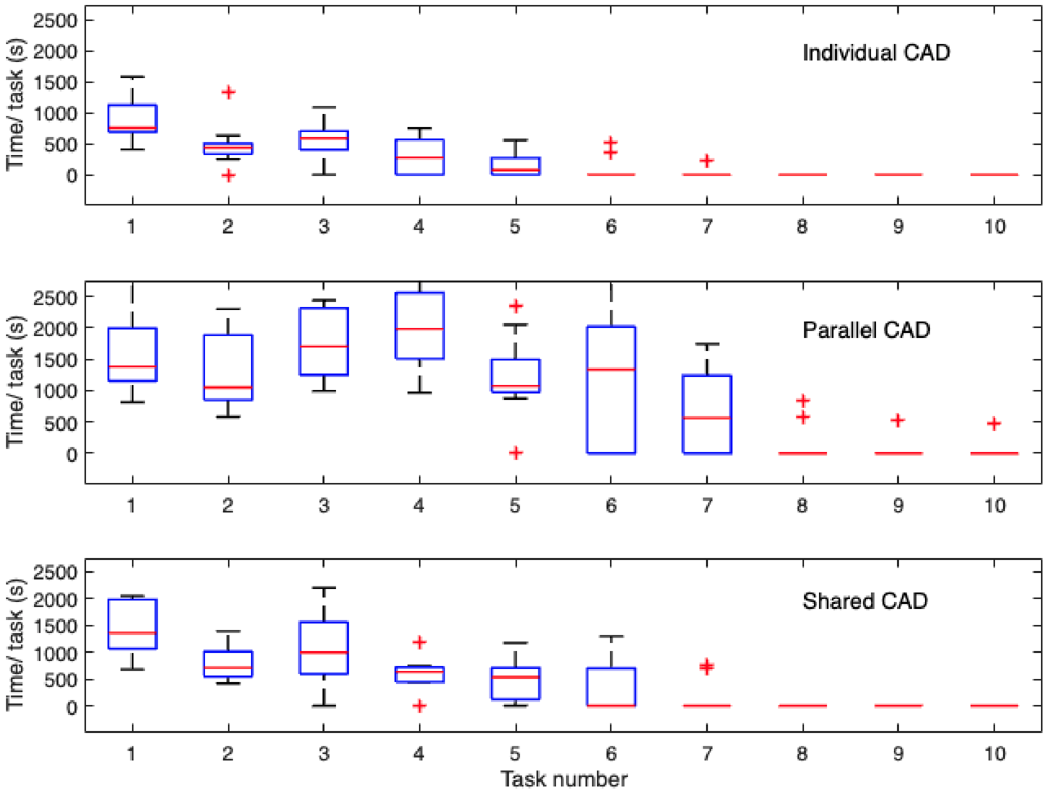


Figure 4-10: Variation in Phase IV task completion time *vs.* working styles.

4.5.2 Phase II and Phase IV Speed

So far we haven't considered the fact that pair CAD work involved two participants. In the below speed metric definition, we calculate CAD speed on a per-person basis and also account for the discrepancy from variation in completion times.

Table 4.4 is an illustration of our method used in calculating Phase IV speed metrics. We first averaged all design tasks times for a given working style. Then we compared all the averages to identify the minimum value. This value was considered as our reference, an indicator of single unit of CAD work for a given working style. Then we divided the rest of the average values with this reference value to calculate a relative CAD work unit for a particular design task. This calculation was repeated for all working styles.

Each participant's Phase IV speed score was then calculated as per Equation 4.1. Phase IV speed score is the cumulative CAD work units collected by completing 'n' number of design tasks and were separately calculated for each working style denoted by 'w'.

$$(Phase\ IV\ speed)_{w,n} = \sum_{i=1}^{n} \frac{(Average\ time)_i}{min\,(Average\ time)_n} \tag{4.1}$$

4.5.3 Phase II and Phase IV Quality

Unlike Phase IV speed, Phase IV quality was calculated as an absolute metric. As shown in Table 4.5, Phase IV quality was derived as an aggregate of ratings from four categories. These categories and their sub-categories were based on prior work by Company *et al.* [90]. Participant scores were calculated using a grading rubric that assessed each design task based on a series of questions. An inter-rater reliability (IRR) evaluation of the grading rubric showed 96% agreement between two coders. This represents "almost perfect agreement" [91]. Quality scores were aggregated using Equation 4.2, wherein scores from each tasks were averaged based on 'n' questions and then averaged over the 'm' number of design tasks completed by a participant. A more detailed description of the Phase IV quality metric can be found in Arshad

Table 4.4: Schematic of Phase IV speed calculations.

Working style	Task numbers			
	1	2	....	n
$Individual\ CAD_1$	.	.		.
.	.	.		.
.	.	.		.
$Individual\ CAD_w$	.	.		.
Average	I_1	I_2	....	I_n
$Parallel\ CAD_1$	.	.		.
.	.	.		.
.	.	.		.
$Parallel\ CAD_w$	.	.		.
Average	P_1	P_2	....	P_n
$Shared\ CAD_1$	.	.		.
.	.	.		.
.	.	.		.
$Shared\ CAD_w$	.	.		.
Average	S_1	S_2	....	S_n

et al.'s publication [4].

$$(Phase\ IV\ quality)_k = \frac{\sum_{i=1}^{m}(\sum_{i=1}^{n} \frac{x_j}{n})}{m} \tag{4.2}$$

Table 4.5: Categories used in Phase IV quality calculations [4].

Metric category	Definition	Indicator
Complete	Replicates drawing accurately	-Replicates size accurately -Replicates shape accurately
Concise	Replication features used (e.g. use of offsets, mirrors)	-Replication features used when available
Consistent	Fully constrained and dimensioned with no new parts	-Fully constrained -Dimensioned in reference to the model
Valid	No failed instances	-No errors in the model tree

4.6 Results

4.6.1 Primary Research Questions

Summary of New Metrics

Phase II speed, Phase II quality, Phase IV speed and Phase IV quality were recalculated for analysis and their corresponding z-score values are used in this section. This was important in order to use the same scale in comparing them.

Phase IV speed and Phase IV quality scores are shown in Figure 4-11. Phase IV speed was highest for Individual CAD participants followed by Parallel CAD and lastly, Shared CAD. This difference was validated by an ANOVA model resulting in a p-value of less than 0.01. Phase IV quality metrics did not show a clear trend but Shared CAD participants scored the highest followed by Individuals and Parallel CAD. This difference was also statistically significant in an ANOVA test with a p-value of 0.03.

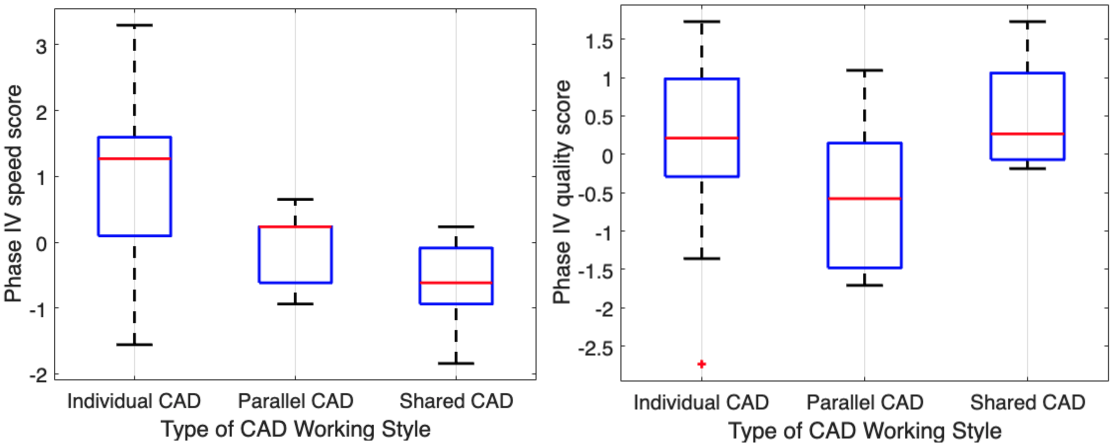


Figure 4-11: (Left) Phase IV speed and (right) Phase IV quality *vs.* working styles.

The following sections further elaborate on our findings on Phase IV speed and quality. We first assess the effect of participants incoming CAD skill, measured in Phase II. Then we assess the influence of Phase IV speed and quality interaction. The raw data used to compute the regression models is in Appendix B: Table B.1, Table B.2 and Table B.3.

Speed: Main Effects

We set up a step-wise linear regression model to test the effect of Phase II performance on Phase IV speed. As seen in Table 4.6, a change in working style had the largest effect on Phase IV speed. It is also noteworthy that Phase II quality was removed from the model results as it proved to be insignificant. We also find that Parallel CAD and Phase IV quality have a sizeable interaction that needs to be reviewed.

A visual representation of main effects on Phase IV speed is shown in Figure 4-12. The range of values in Figure 4-12 illustrate the confidence intervals on each effect

Table 4.6: Results from step-wise regression model with Phase IV speed as dependent variable.

	Estimate	SE	tStat	pValue
Intercept (Individual CAD)	1.05	0.15	7.17	0.00
Phase II speed	-0.48	0.17	-2.90	0.01
Parallel CAD	-1.25	0.20	-6.28	0.00
Shared CAD	-1.93	0.21	-9.19	0.00
Phase IV quality	0.74	0.13	5.73	0.00
Phase II speed : Parallel CAD	0.65	0.20	3.29	0.00
Phase II speed : Shared CAD	0.70	0.21	3.34	0.00
Parallel CAD : Phase IV quality	-1.02	0.18	-5.56	0.00
Shared CAD : Phase IV quality	-0.32	0.23	-1.38	0.17

size. In summary, there is no strong evidence suggesting an influence of Phase II speed or quality on Phase IV speed.

Quality: Main Effects

A step-wise regression model was created for Phase IV quality and the corresponding coefficients are listed in Table 4.7. As we can see, Phase II speed and quality were removed from the results as they were found to be insignificant. The difference between Individual CAD and Shared CAD had the most effect on Phase IV quality. The Phase IV speed and quality interaction was found to be sizable and will be investigated in the next section.

Similar to Figure 4-12, Figure 4-13 shows the main effects for Phase IV quality and their corresponding confidence intervals. We see that a change in working style had the most significant effect, but the wide range of Phase IV speed overlaps with working styles main effect. In summary, there is no strong evidence suggesting an influence of Phase II speed or quality on Phase IV speed. However, the interaction between Phase IV speed and quality warrants further investigation.

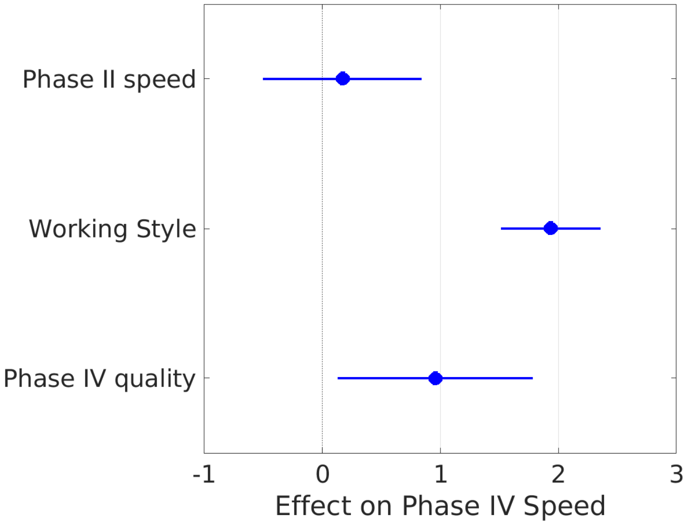


Figure 4-12: Main effects for Phase IV speed.

Table 4.7: Results from step-wise regression model with Phase IV quality as dependent variable.

	Estimate	SE	tStat	pValue
Intercept (Individual CAD)	-0.68	0.27	-2.55	0.01
Parallel CAD	0.06	0.30	0.20	0.84
Shared CAD	1.63	0.35	4.71	0.00
Phase IV speed	0.70	0.17	4.21	0.00
Parallel CAD : Phase IV speed	-1.51	0.32	-4.71	0.00
Shared CAD : Phase IV speed	-0.05	0.31	-0.15	0.88

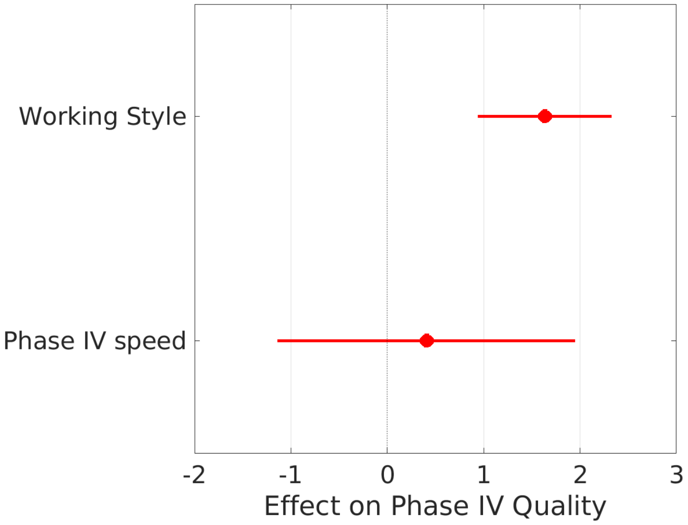


Figure 4-13: Main effects for Phase IV quality.

Phase IV Speed/Quality Interaction

The interaction between Phase IV speed and quality was significant in both regression models and will be reviewed in this section. In Figure 4-14, Effect of Phase IV quality is shown on the X-axis. The goal of this plot is to evaluate the effect of interaction between Phase IV speed and quality on the effect of change in working styles.

In the top-half of Figure 4-14, we see that the main effect of changing working style remains consistent at different Phase IV speeds. Likewise, in the bottom-half we see that for a given working style, the effect of changing Phase IV speed has a different outcome. Parallel CAD results show a decrease in Phase IV quality as Phase IV speed increases, but Shared CAD and Individual CAD show a positive increase in Phase IV quality with an increase in Phase IV speed.

Our findings are further explored in Figure 4-15 which shows the relationship between the raw datapoints for Phase IV speed and quality. The point cluster for Individual CAD and Shared CAD has a positive slope whereas Parallel CAD data points tend towards a negative slope.

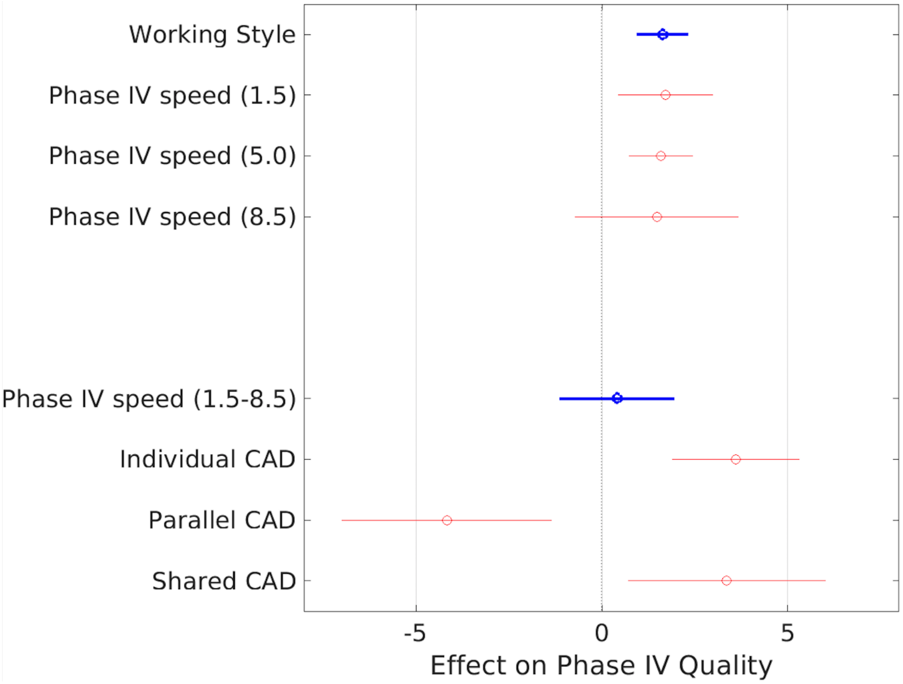


Figure 4-14: Effect of interactions between Phase IV speed and working style on Phase IV quality.

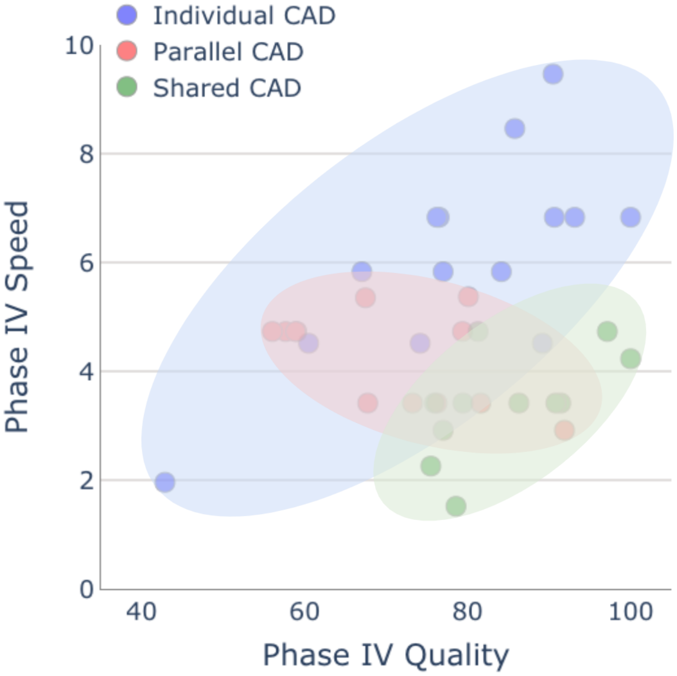


Figure 4-15: Clustering of Phase IV speed *vs.* Phase IV quality data.

4.6.2 Supporting Research Questions

We use supporting questions from section 4.2.2 to elaborate on our results. We used ANOVA and paired t-tests to evaluate participant responses and the results are shown in Table 4.8. The communication data and cursor activity show significant differences, but satisfaction score were consistent between working styles.

Table 4.8: Summary of supporting research questions.

	Parallel CAD	Shared CAD	Individual CAD	p-Value
Communication (% of 45 mins.)	42%	75%	NA	<0.01
Cursor activity (% of 45 mins.)	60%	42%	56%	<0.01
Satisfaction (1-5)	3.95	3.22	4.13	0.059

Communication

To elaborate on our paired t-test results, we explore the level of communication. We plotted the detected audio activity in Figure 4-16. As can be seen, the difference in Parallel CAD and Shared CAD is clear and this validates our results from Table 4.8. Due to technical challenges, the audio data was calculated only for 40 participants which included 10 pairs using Shared CAD and 10 pairs using Parallel CAD.

Cursor Activity

To gage work equity between pairs, we asked participants to rate their responses towards the prompt: *"I contributed more than my partner in this study"*. Responses are shown in Table 4.9. Overall, there is no clear consensus on which working style provided a more equal opportunity to contribute.

In addition to the Likert scale, we used cursor activity to ascertain work distribution between working styles. Figure 4-17 show the aggregate cursor activity for each participant. We see that Individual CAD and Parallel CAD participants show

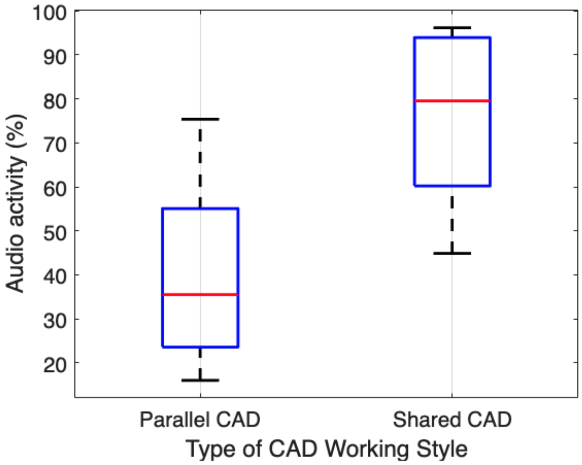


Figure 4-16: Audio activity *vs.* working style.

Table 4.9: Summary of participant responses to work equity prompt.

Response	Parallel CAD	Shared CAD
I did very much more work	4.2%	4.5%
I did somewhat more work	12.5%	13.6%
Equal work	66.7%	59.1%
Partner did somewhat more work	4.2%	22.8%
Partner did very much more work	12.5%	0%

a similar amount of cursor activity. However, Shared CAD participants data shows a compelling behaviour. We see that in each Shared CAD pairing, one participant uses the cursor significantly less compared to their counterpart. This makes the overall cursor activity of Shared CAD participants to be much lower than others. Due to technical challenges, the cursor activity data was calculated only for 50 participants which included 10 pairs using Parallel CAD, 10 pairs using Shared CAD and 10 participants using Individual CAD.

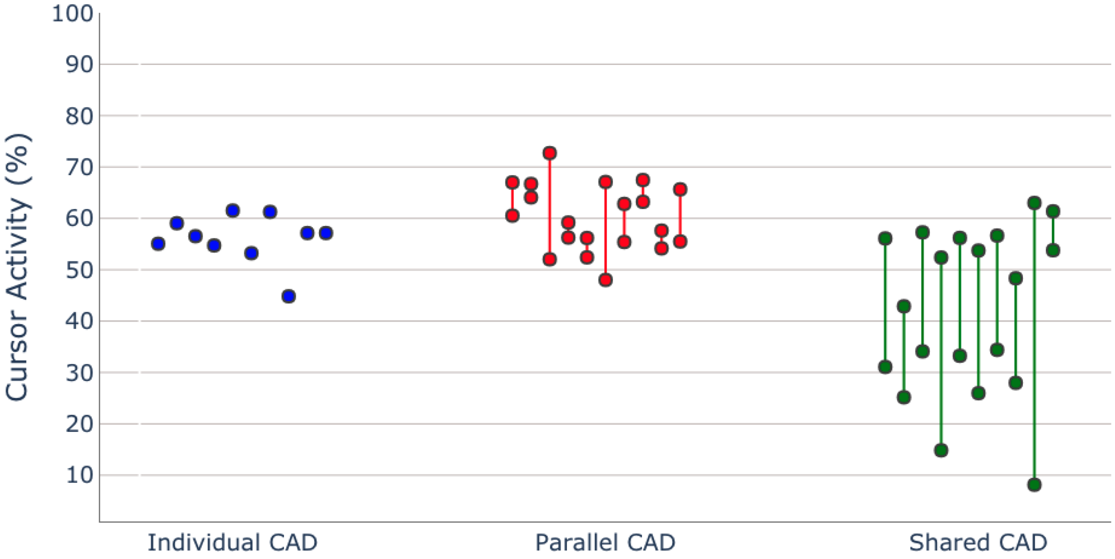


Figure 4-17: Cursor activity *vs.* working style.

To further delve into the UX activity differences, we divided the participant dashboard into regions of interest. See Figure 4-18. The cursor activity for each working style is shown below the name of each region. As can be seen, cursor activity in each region is fairly similar except in *Task instructions* where Shared CAD shows substantially higher activity.

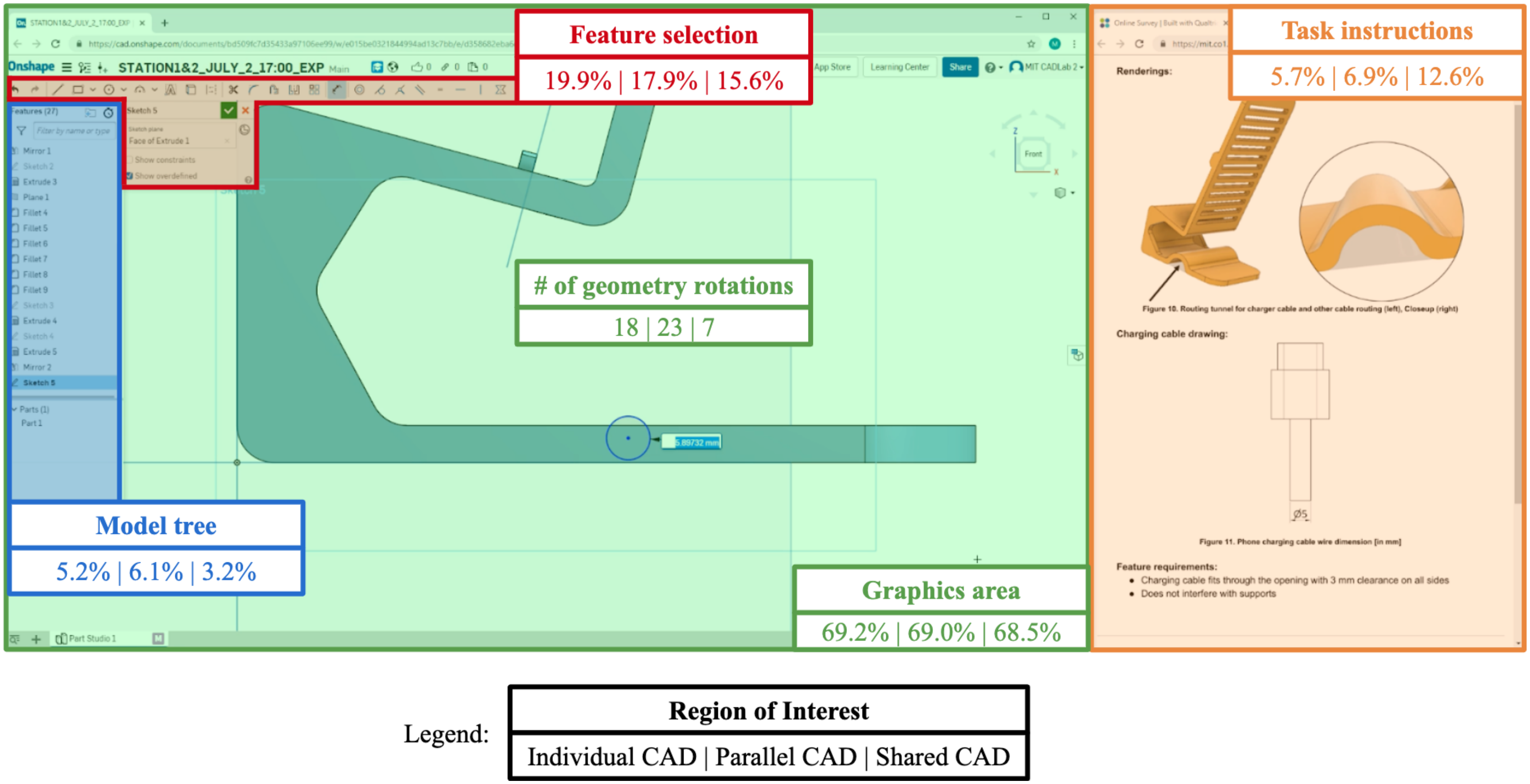


Figure 10. Routing tunnel for charger cable and other cable routing (left), Closeup (right)

Figure 11. Phone charging cable wire dimension [in mm]

Feature requirements:
- Charging cable fits through the opening with 3 mm clearance on all sides
- Does not interfere with supports

Figure 4-18: Cursor activity in regions of interest for each working style.

The heatmaps shown in Appendix Figure C-2 visually shows the distinction between cursor activity of each working style. The yellow line marks the split between *Graphics area* and *Task instructions* sections of the participant dashboard. There is a visible difference in the right half of each plot. Shared CAD participants show the most cursor activity in the design task region, followed by Parallel CAD and then Individual CAD. Thus, further validating our finding from Figure 4-18.

Satisfaction

Satisfaction scores were calculated based on a Likert scale questions from the post-study surveys. Participant responses were received as a rating of agreement: Strongly disagree to Strongly agree. The responses were then converted to a scale of 1-5 and the same is plotted in Figure 4-19. Shared CAD participants had the lowest score but also the most variability. Parallel CAD and Individual CAD participants reported similar levels of satisfaction.

In addition to the above question, we asked Shared CAD and Parallel CAD participants if they will use their method of collaboration in future. Shared CAD participants reported a converted Likert scale score of 3.0 and Parallel CAD reported 3.4.

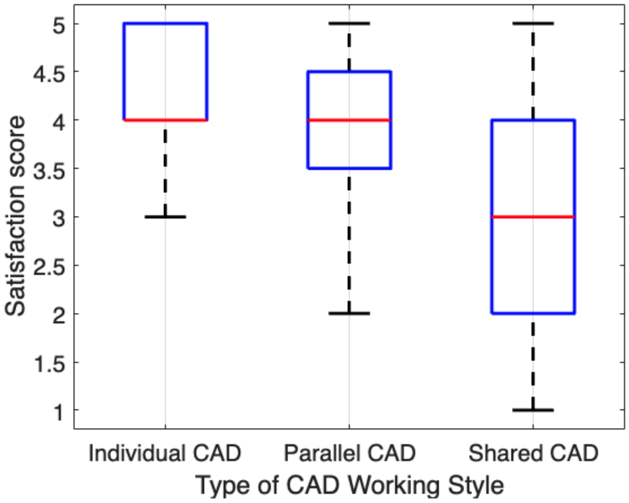


Figure 4-19: Satisfaction score *vs.* working style.

We planned to use the emotion detection toolkit to triangulate the information from self-reported satisfaction scores. This approach was consistent with the emotion detection work presented in Chapter 3. We used Affectiva's facial detection software to process participants face videos and derive emotion metrics.

However, during the implementation we ran into unanticipated issues. In some of our runs, participant faces were not entirely in the frame and sometimes support for the Toronto baseball team got in the way. Two such examples are shown in Figure 4-20.

Figure 4-20: Examples of unusable web camera captures.

The limited data from emotion tracking is shown in Table 4.10. In-spite of the low coverage numbers, we see some interesting trends. From the pilot study, we had observed that Parallel CAD participants expressed more emotions than individuals [3]. This result is supported in Table 4.10. The only exception being that Individual CAD participants expressed the most contempt.

Within pair CAD, Shared CAD participants were less emotive than Parallel CAD, except in expressing joy. In general, both pair CAD participants expressed more joy than Individual CAD. Given the small number of data points, further analysis of the emotion data will not be pursued in this thesis and will be treated separately.

4.6.3 Aggregating Quantitative Findings

In this section, we evaluate all relevant variables from previous sections together. We will consider only the pair CAD working styles in this analysis to focus our results. In particular, we will compare pair CAD performance based on speed and quality metrics, alongside differences in Phase II speed, differences in cursor (UI) activity

Table 4.10: Summary of emotions from experiment data.

Emotion (in secs.)	Individual CAD	Parallel CAD	Shared CAD	Ratio
Joy	1	35	65	1 : 35 : 65
Sadness	3	7	1	1 : 4.3 : 0.3
Disgust	28	24	21	1 : 0.8 : 0.7
Contempt	103	56	34	1 : 0.5 : 0.3
Anger	9	14	2	1 : 1.6 : 0.2
Fear	3	1	2	1 : 0.3 : 0.7
Surprise	26	43	15	1 : 1.6 : 0.5
# of files with >50% coverage	6/12	16/22	10/19	

and a pairs' combined audio activity.

For the comparative analysis we will use Pearson correlation coefficients to measure any linear relationships between variables. All combinations of Pearson's r are as shown in Table 4.11. Values of significance are marked with a '*'. These values are significant because they are higher than 0.5 in magnitude and pertain to a p-value of less than 0.05.

For Parallel CAD, we see that difference in Phase II speed negatively correlated with Phase IV speed and audio communication. Also, a higher difference in cursor activity correlated with high audio communication. In Shared CAD, we saw that Phase IV speed negatively correlated with audio communication. Although not significant, note the relationship between Phase IV speed and quality. It is negative for Parallel CAD and positive for Shared CAD. This result is in agreement with the regression model results from Section 4.6.1.

Finally, note that outside of Table 4.11, we carried out a correlation analysis of participants Phase II skill level *vs.* cursor activity. We found no statistically significant relationship in any working style.

Table 4.11: Correlation between experimental factors.

Pearson coefficient PCC | SCC	Difference in Phase II speed	Phase IV speed	Phase IV quality	Difference in cursor activity	Audio activity
Difference in Phase II speed	1	-	-	-	-
Phase IV speed	-0.7* | -0.5	1	-	-	-
Phase IV quality	0.3 | 0.04	-0.5 | 0.5	1	-	-
Difference in cursor activity	-0.3 | -0.4	-0.07 | 0.4	-0.4 | -0.3	1	-
Audio activity	-0.6* | 0.4	0.3 | -0.6*	-0.1 | -0.4	0.6* | -0.4	1

4.6.4 Open-Response Survey Questions

In addition to Likert scale questions, we asked our participants open-ended questions to gauge their frustrations. This exercise gave the research team a qualitative understanding of the participants experience.

Some excerpts that are illustrative of the common themes are shown below:

Question 1: What were some of the frustrations you experienced during the study?

Individual CAD participants:

“Most of the time spent was used to decide how to dimension parts rather than draw them”

“Not familiar with the text-editing, spend lots of time finding setting for size of font”

Parallel CAD participants:

"At times, my partner and I had deleted parts that the other was in the middle of working on, but we were able to overcome this by undoing the last command."

"Lack of cursor/pointer sharing. Often I did not understand what line/angle my partner was referring to."

Shared CAD participants:

"Only one active mouse at a time. Other than that, just learning pains of a new interface."

"Being unable to view the model from perspectives other than my partner while they were working on it"

Question 2 (only for Parallel CAD and Shared CAD): Live (synchronous) collaboration in CAD is fairly new. Based on your study experience, what are ways of improving the experience of working together in CAD?

Parallel CAD participants:

"Or at least adding an identifier to show who made the sketch. That way I can avoid messing up my partner's work."

"Have a history box displaying every single change by a specific user. To keep tract of the work in progress."

Shared CAD participants:

"get to know your partner's skills and preferences beforehand"

"Allow a 'synced' view and an 'unsynced' view where the passive partner can't change the part, but can rotate and pan."

Table 4.12 shows a summary of the recurring topics from participant responses. The categorization of responses was done by two raters having moderate agreement (based on IRR) of 82.5% for Question 1 and 78.3% for Question 2. The final ratings were decided through an arbitration process.

Table 4.12: Count of some recurring themes from open response survey questions.

	Parallel CAD	Shared CAD	Individual CAD
Question 1			
Collaboration method	37%	76%	NA
CAD tool related	42%	24%	92%
Design task related	21%	0%	9%
Question 2			
Combination of Parallel/Shared CAD work	55%	76%	
More practise/training sessions	15%	12%	
Improve and enlarge CAD feature set	30%	12%	

4.7 Summary

A summary of our findings is shown in Table 4.13. The following chapter will help explain this outcome. Note that we updated research question 1 (RQ1) to better represent our findings from Section 4.6.1.

Table 4.13: Summary of all research questions.

Research questions		Result	Details
RQ1	(Updated) On a per person basis, is Individual CAD faster than Parallel CAD and Shared CAD work?	Supported ($p<0.01$)	Although Parallel CAD completed most tasks as pairs, on per-person basis they were slower than Individual CAD but faster than Shared CAD.
RQ2	Does Shared CAD lead to higher quality work compared to Parallel CAD and Individual CAD?	Supported ($p=0.03$)	Shared CAD participants indeed performed highest quality work compared to Individual CAD and Parallel CAD participants.
S1	Do Shared CAD participants communicate as much as Parallel CAD participants?	Supported ($p<0.01$)	Shared CAD participants communicated twice as much as Parallel CAD participants
S2	Do Parallel CAD participants share work more equally compared to Shared CAD participants?	Supported ($p<0.01$)	All CAD working styles showed different amounts of cursor activity. More so, Shared CAD displayed a unique pattern of activity.
S3	Are Parallel CAD and Shared CAD participants more satisfied than Individual CAD participants?	Not supported ($p=0.06$)	The average of self reported scores in our post-study survey did not result in a statistically significant difference.

Chapter 5

Discussion and Conclusions

5.1 Discussion

5.1.1 Reflections From Pilot Study

From the pilot study data, we made an early observation that on a per person basis, pairs were slower at implementing CAD changes compared to individuals. But this hypothesis could not be statistically examined because of our limited sample size. We based our comparison on overall modeling efficiency (OME) which led us to identify losses and overheads in pair work.

We used additional data traces from cursor tracking and emotion recognition to support our OME results. We noted that pairs emoted more than individuals. This could be partly because of the additional person involved in the CAD environment. The cursor tracking data and emotion recognition data helped us identify events of interest. It was found that interaction with the graphics area of the CAD UI aroused the most emotions. This is likely because of the frustrations of using pair CAD as the current architecture of CAD tools is not conducive to real-time collaboration.

Development, deployment and validation of our toolkit was the primary motivation for the pilot study runs. We implemented a non-intrusive data collection strategy that relied on post processing. This also helped alleviate additional burden on the research staff in the form note-keeping during the pilot study. All data gathered in the pilot study was generated using the toolkit and this in itself was a testament to the successful working of the toolkit.

In summary, the pilot study served as a foundation to future work in the form of establishing a usable toolkit, building relevant experience in conducting experiments, and providing early insights in real-time CAD collaboration.

5.1.2 Experiment Findings

We deployed an improved version of the toolkit for our final experiments. We also derived new speed and quality metrics to base a comparison of CAD working styles.

Overall, our results agree with the pilot study results. Pairs were slower than individuals, on a per person basis. This result is consistent with established norms in small group literature: overheads slow down collaborative work and limit team members from reaching their full individual potential [49]. In our experiment, these overheads were likely manifested from coordination efforts, communication, and awareness seeking. For example, we see that Shared CAD participants communicated with each other for 74.79% of the study time which reduced the time available for CAD modeling activities. Similarly, Parallel CAD participants had to account for additional coordination overheads given the dynamic nature of their CAD environment. However in case of Shared CAD work, lack of editing freedom turned out to be the biggest curtailment to Phase IV speed.

On the contrary, CAD quality was expected to benefit from having multiple eyes checking the CAD file. This effect turned out to be true only in the case of Shared CAD. Alluding to the fact that sharing visual artifacts play a much bigger role in communication of CAD work. In fact, sharing the CAD visual possibly also led to significantly higher level of audio communication between Shared CAD participants. Lastly, it is noteworthy that Parallel CAD work reduces awareness compared to Shared CAD. In summary, we believe the combination of higher/frequent communication and increased awareness led to better CAD quality in Shared CAD work.

Driver-navigator style pairing is common in software pair programming. In our experiment, only Shared CAD participants demonstrated such a role-setting. The cursor activity data for all participants was similar except in Shared CAD where one participant clearly dominated the cursor activity. This might mean that the other Shared CAD participant was always focused on overseeing their partner's work. Essentially, Shared CAD pairs had to establish a mutual agreement in order to progress. This two-fold validation of decisions is possibly a reason for the higher quality in Shared CAD. Conversely, Parallel CAD participants could progress in a much less structured way, without each other's consent. We see this outcome from the Phase IV speed/quality relationship for all working styles. Slower Parallel CAD participants produced higher quality work whereas the opposite was true for Shared CAD and Individual CAD.

The correlation analysis in Table 4.11 provides clues that help us understand some of the outcomes discussed so far. In Parallel CAD, pairs with a high skill mismatch communicated less and their Phase IV speed scores were lower. This make us speculate that Parallel CAD pairs found it easiest to work by themselves in case of a high skill mismatch. In the case of Shared CAD, there was no such correlation. However, it was found that better performing pairs communicated lesser. This is consistent with high performing MUCAD teams studied in Stone *et al.*'s work [26].

5.1.3 Recommendations on Pair CAD

Firstly, we demonstrate that CAD style has an effect on the design outcome. Choosing one pair CAD style or using pair CAD throughout the design process might not be the best strategy. A hybrid model which employs pair work in a focused manner is ideal. For example, during a crucial design milestone, teams might employ Shared CAD work to produce a higher quality outcome and improve designers' confidence. On the contrary, once crucial decisions are made and there is a lot of CAD drafting remaining, teams might choose Parallel CAD work in order to attain high speed of work.

An approach suggested by a participant in the open response question is to use both Shared CAD and Parallel CAD together. This would mean that pairs have the ability to move between Parallel CAD and Shared CAD on-the-go. Although ideal, this warrants an expansive screen size and elaborate software suite.

CAD is just one software in the array of tools needed for pair CAD work. Although not studied in-depth in this work, AV communication, design documentation and team dynamics are fundamental to success in pair CAD work. We recommend that designers use synchronous collaboration tools like Google Docs, Slack, and Zoom to augment pair CAD tools.

It is important to highlight the need for adequate training and investing the time to develop a strong working relationship between pair CAD teams. Otherwise, pair CAD can lead to a low morale and lack of adoption. A glimpse of these frustrations is seen in the satisfaction scores and open response survey questions.

Overall, pairing promises to bring speed and quality gains to CAD work but teams need to be patient in evolving their unique work strategies.

5.1.4 Limitations

It is hard to provide absolute conclusions based on just one experiment. The results of our experiment are constrained by the nature of the design tasks presented to our participants and the duration of the experiment. This means that our results lay the foundation in this niche research area but do not generalize to all industry settings.

We used students as our participant pool. This has been used by other researchers and is acknowledged to limit the applicability of our results [83]. However, studying industry teams using cloud-based CAD is currently difficult because of the infancy of these CAD tools.

The duration of our study was short; limiting the potential for teams to develop an established working relationship. We posit that our results could change if the pairs had more than one opportunity to work together. A study of that magnitude was considered out of scope for time and financial constraints.

On the technical front, some of our web-camera recordings were rendered unusable because of lack of standardization in camera positioning. This meant that we had to forego the emotion recognition analysis section of the final experiment.

5.1.5 Future Work

The immediate next step would be to generalize our results by validating our findings in a professional setting. Then, the application of our results need to be studied over the entire design process; as constraints are design-phase-dependent. This next step could be done as a case study in industry. In this future work, all users should've access to both pair CAD working styles and be allowed to switch on-the-fly. Using pair CAD for a sustained period of time will give participants enough time to evolve their preferred working style. Through this work, one could understand how and when do designers use the two pair CAD working styles. More importantly, this will provide a framework to help designers in making the trade-off in choosing either pair CAD working style.

The eventual success of pair CAD work will rely on developing design supports and interventions specifically meant to support pair CAD work. We hope that wider adoption of synchronous CAD tools will justify building the needed software infrastructure to scaffold pair CAD work. An example of one such design support is outlined in the next paragraph.

We noticed the importance of awareness in aiding pair CAD work. In existing real-time CAD tools the onus is on the users to check on the status of their collaborators' work. However there are techniques used in non-CAD pair working tools that suggest a better way. For example, view-cones are used in VR applications to determine the spatial location of users. In pair programming, color-coding text and cursor locations is useful in orienting users. These awareness enhancing methods will highly improve the pair CAD experience.

Lastly, the broader discussion around pair CAD work cannot be complete without acknowledging the role of a mediator. Some form of an agent that guides pair CAD work by managing communication, maintaining documentation, ensuring equity of work. This guidance is particularly important for new adopter of pair CAD work. For example, to ensure higher quality, the agent would nudge the pairs to share screens before an important design decision is made. The exact role and description of this agent is not clear but should be determined by future research. This additional guidance will greatly reduce the learning curve and frustrations in learning to use pair CAD.

The exact manifestation can happen in the form of a third person (mentor) or software. An implementation of software-support agents has precedence. From early versions like Clippy in Microsoft Word to the current artificial intelligence (AI) enabled assistants like Google assistant, Alexa and Siri. It can be expected that the role of the agent will become more passive as pairs establish a mature working relationship.

5.2 Conclusions

The learnings from our work can be summarized as below:

- Pair CAD participants did not show any assembly bonus effect. In other words, Parallel CAD and Shared CAD participants were severely slowed down by overheads from collaborative work. On a per person basis, Individual CAD produced more work than Parallel CAD and Shared CAD

- Shared CAD lead to highest CAD quality and Parallel CAD work lead to the lowest CAD quality

- Speed and quality were positively correlated for Shared CAD and Individual CAD but negatively for Parallel CAD

- Shared CAD participants communicated twice as much as Parallel CAD participants

- Shared CAD participants displayed a clear driver-navigator style role-play but Parallel CAD participants showed greater work equity

- Parallel CAD pairs with a high skill mismatch performed worse and communicated lesser than teams with a lower skill mismatch. No such effect was found in Shared CAD.

- Better performing Shared CAD teams communicated lesser than lower performing Shared CAD teams.

Finally, I would like to revisit the observation made in Chapter 1. It is undeniable that we at the cusp of a transition to more remote work. It is also noteworthy that in the recent past, the pursuit of digitisation has been already ongoing. This makes us well situated with tools and technologies to face the new normal. However, the abrupt transition leaves us with no time to fully understand the ramifications of remote work. Through this research, we hope to influence others to create more insights to inform mindful adoption of virtual work.

Appendix A

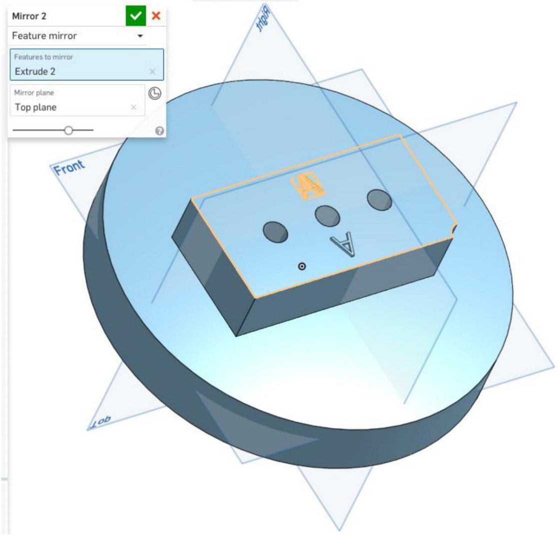
Design Tasks

The following pages show representative design tasks from Phase I, Phase II and Phase IV of the experiment.

Use features from *3D Subgroup 2* to create a sketch as shown below. You are encouraged to try all features in the subgroup and ask questions.

3D Subgroup 2:

Mirror text extrude created in last step:



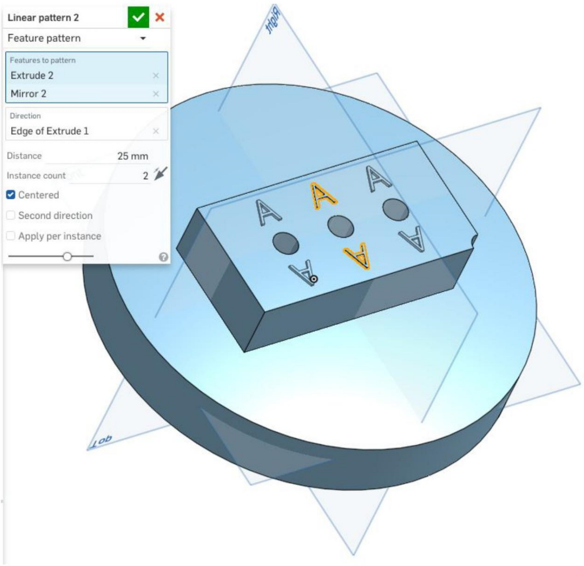
Mirror 2

Feature mirror

Features to mirror: Extrude 2

Mirror plane: Top plane

Create a linear pattern of mirrored feature from last step:



Linear pattern 2

Feature pattern

Features to pattern: Extrude 2, Mirror 2

Direction: Edge of Extrude 1

Distance 25 mm

Instance count 2

☑ Centered

☐ Second direction

☐ Apply per instance

Figure A-1: An instruction page from Phase I (training) demonstrating the use of "mirror" and "linear pattern" commands to participants.

Image of CAD feature:

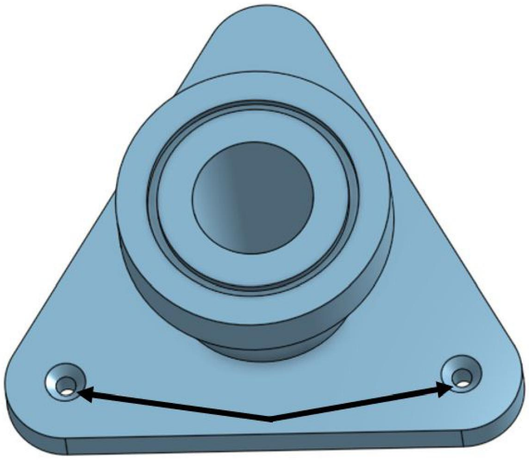


Figure 12. Mounting hole location

2D drawing:

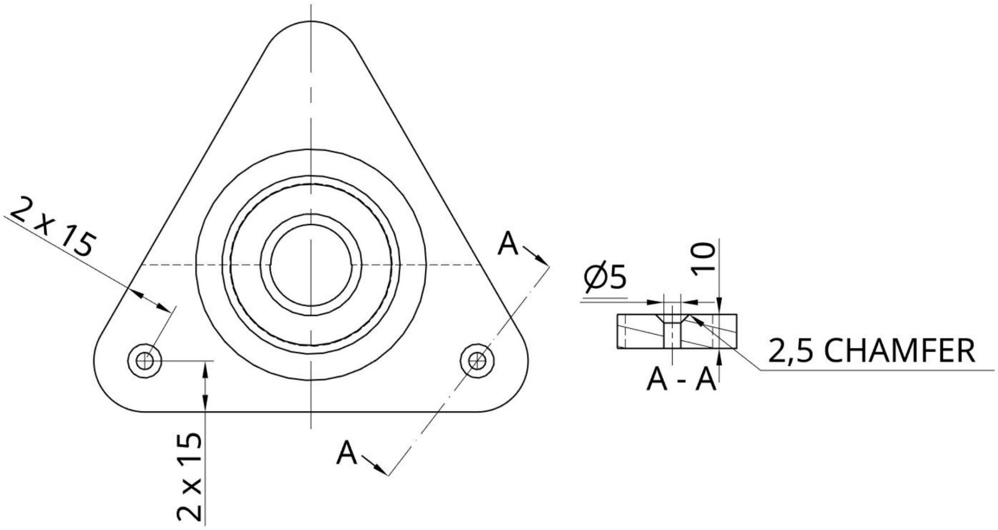


Figure 13. Dimension of mounting hole [in mm]

Feature requirements:

- Holes on thinner section of baseplate
- Holes symmetric to each other

Figure A-2: Phase II design task asking participants to add two holes to their exiting CAD file.

Renderings:

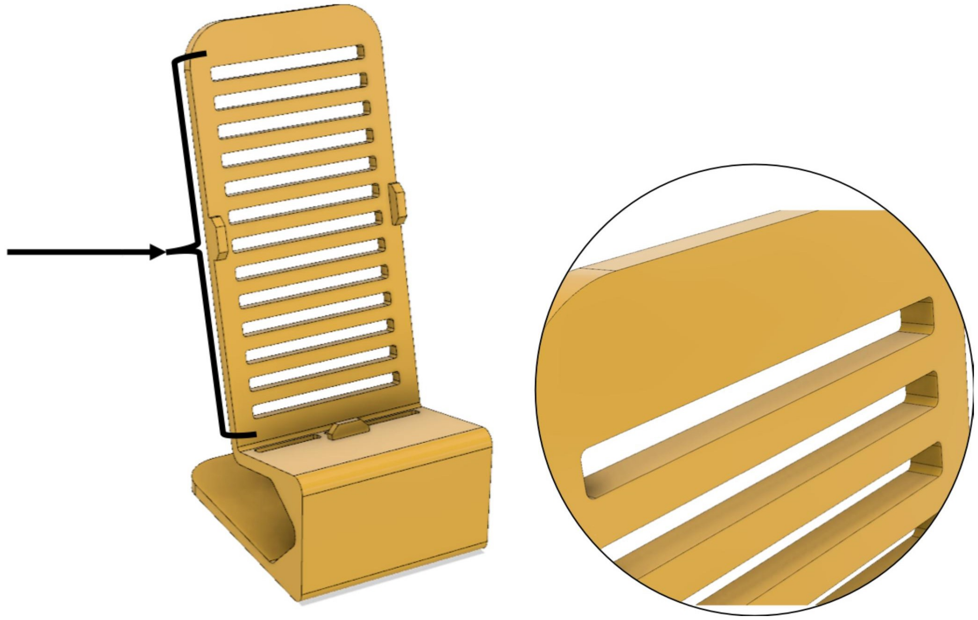


Figure 8. Phone holder with slots for air circulation (left), Closeup (right)

Phone drawing:

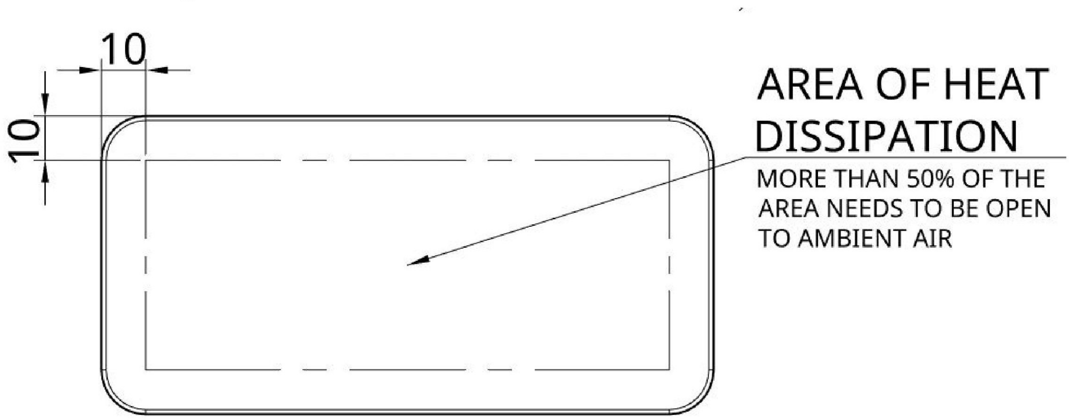


Figure 9. Area of heat dissipation on the back panel of the phone [in mm]

Feature requirements:

- More than 50% phone area is open to air in landscape orientation
- More than 50% phone area is open to air in portrait orientation

Figure A-3: Phase IV design task showing a user need; to add ventilation feature to existing CAD model.

Appendix B

Supplementary Results

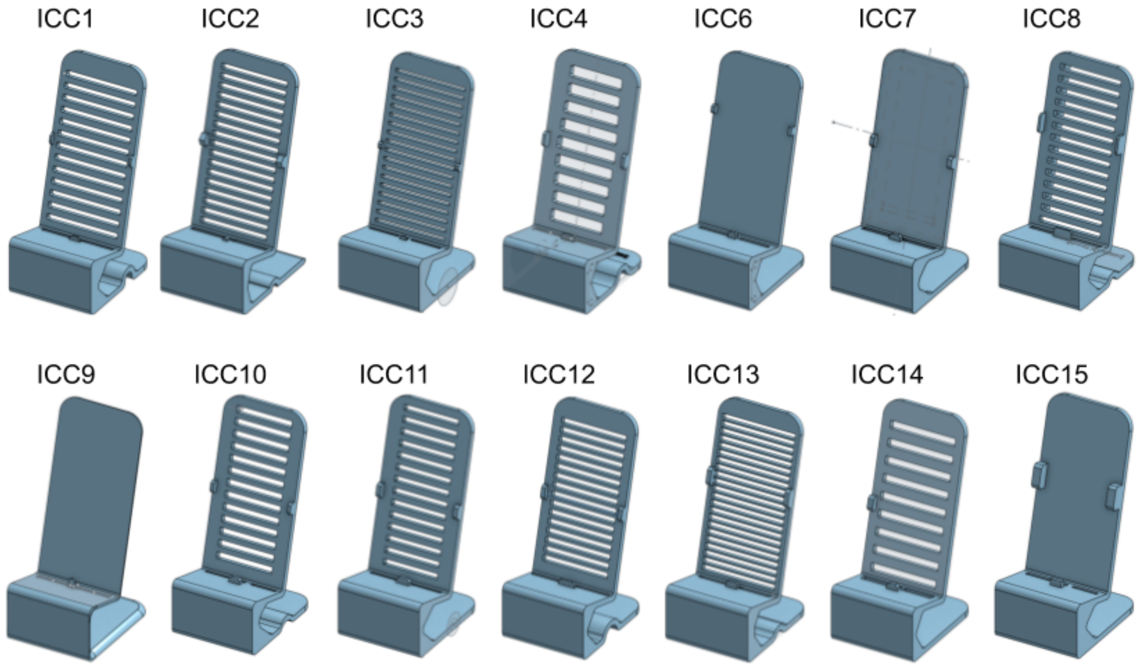


Figure B-1: CAD files by Individual CAD participants.

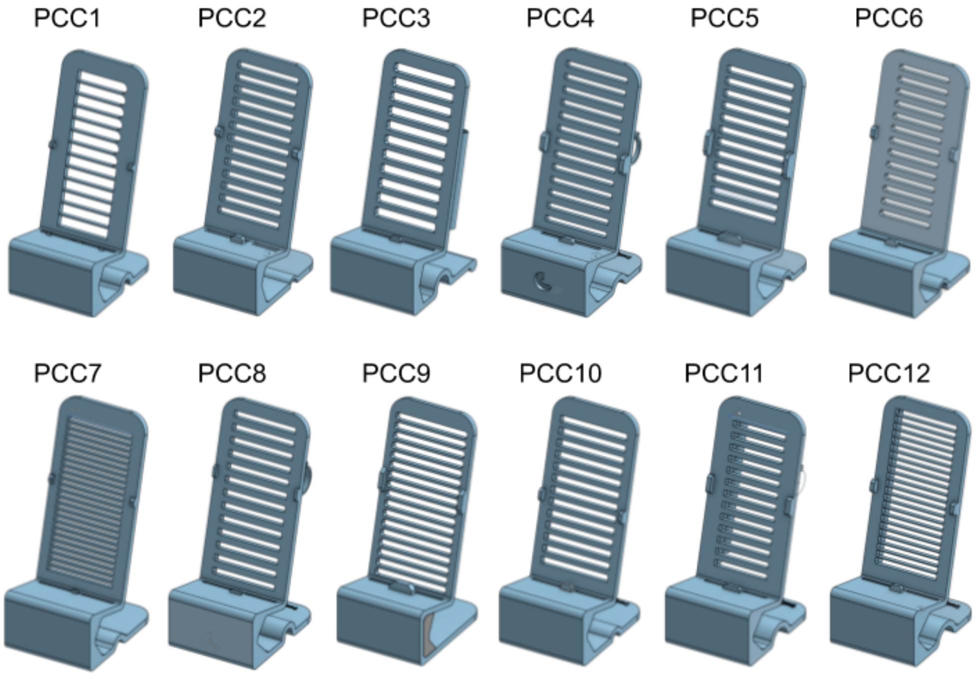


Figure B-2: CAD files by Parallel CAD participants.

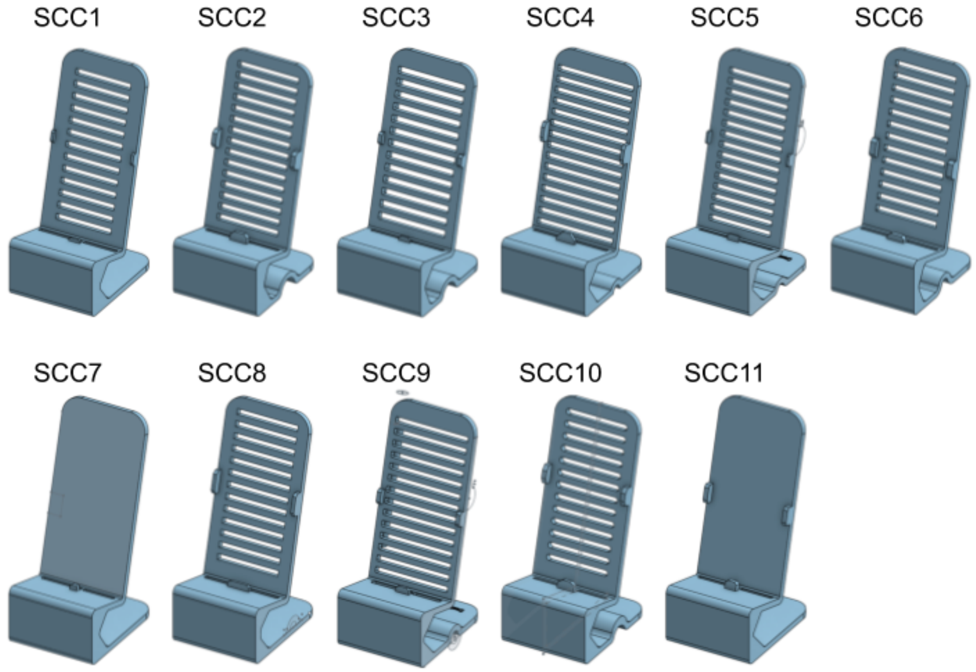


Figure B-3: CAD files by Shared CAD participants.

Legend for Table B.1, Table B.2 and Table B.3:

A = Number of tasks completed in Phase II

B = Phase II speed score

z_B = z-score equivalent of Phase II speed score

C = Phase II quality score

z_C = z-score equivalent of Phase II quality score

D = Number of tasks completed in Phase IV

E = Phase IV speed score

z_E = z-score equivalent of Phase IV speed score

F = Phase IV quality score

z_F = z-score equivalent of Phase IV quality score

Note that data in columns D, E,z_E, F and z_F for Parallel CAD and Shared CAD pairs is repeated for both participants in that pair.

Table B.1: Individual CAD performance data.

Code name	A	B	z_B	C	z_C	D	E	z_E	F	z_F
ICC 1	5	7.8	0.8	100.0	0.7	5	6.8	1.6	100.0	1.7
ICC 2	4	6.7	0.2	100.0	0.7	5	6.8	1.6	76.5	-0.1
ICC 3	4	6.7	0.2	75.0	-2.0	4	5.8	0.9	77.0	-0.1
ICC 4	1	1.0	-2.8	100.0	0.7	7	9.5	3.3	90.5	1.0
ICC 6	4	6.7	0.2	87.5	-0.6	3	4.5	0.1	74.2	-0.3
ICC 7	4	6.7	0.2	100.0	0.7	3	4.5	0.1	89.2	0.9
ICC 8	5	7.8	0.8	90.0	-0.4	6	8.5	2.6	85.8	0.6
ICC 9	4	6.7	0.2	87.5	-0.6	1	2.0	-1.6	42.9	-2.7
ICC 10	3	5.1	-0.6	100.0	0.7	5	6.8	1.6	76.2	-0.1
ICC 11	5	7.8	0.8	90.0	-0.4	4	5.8	0.9	84.1	0.5
ICC 12	4	6.7	0.2	87.5	-0.6	5	6.8	1.6	90.6	1.0
ICC 13	4	6.7	0.2	100.0	0.7	5	6.8	1.6	93.1	1.2
ICC 14	4	6.7	0.2	100.0	0.7	4	5.8	0.9	67.0	-0.8
ICC 15	3	5.1	-0.6	100.0	0.7	3	4.5	0.1	60.5	-1.4

Table B.2: Parallel CAD performance data.

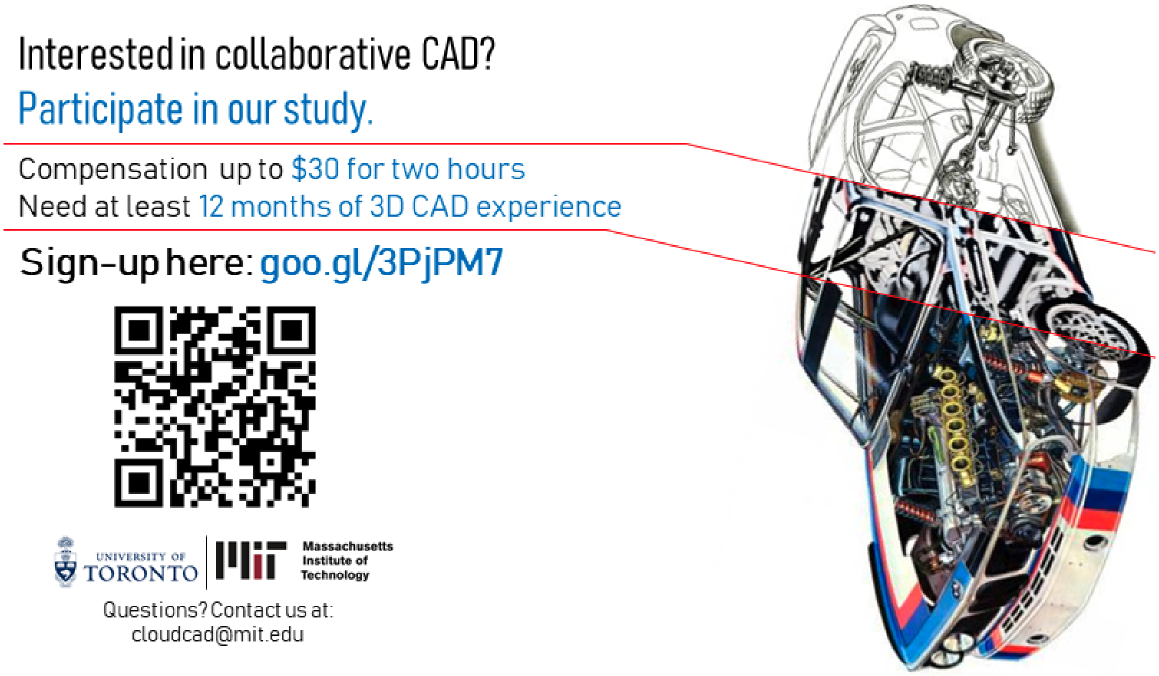
Code name	A	B	z_B	C	z_C	D	E	z_E	F	z_F
PCC 1-1	5	7.8	0.8	100.0	0.7	5	3.4	-0.6	76.2	-0.1
PCC 1-2	2	2.9	-1.8	100.0	0.7	5	3.4	-0.6	76.2	-0.1
PCC 2-1	3	5.1	-0.6	75.0	-2.0	7	4.7	0.2	59.0	-1.5
PCC 2-2	4	6.7	0.2	66.7	-2.9	7	4.7	0.2	59.0	-1.5
PCC 3-1	4	6.7	0.2	100.0	0.7	7	4.7	0.2	57.7	-1.6
PCC 3-2	4	6.7	0.2	100.0	0.7	7	4.7	0.2	57.7	-1.6
PCC 4-1	5	7.8	0.8	75.0	-2.0	8	5.4	0.6	67.5	-0.8
PCC 4-2	4	6.7	0.2	90.0	-0.4	8	5.4	0.6	67.5	-0.8
PCC 5-1	1	1.0	-2.8	100.0	0.7	5	3.4	-0.6	67.7	-0.8
PCC 5-2	4	6.7	0.2	87.5	-0.6	5	3.4	-0.6	67.7	-0.8
PCC 6-1	5	7.8	0.8	100.0	0.7	5	3.4	-0.6	73.2	-0.4
PCC 6-2	3	5.1	-0.6	70.0	-2.5	5	3.4	-0.6	73.2	-0.4
PCC 7-1	4	6.7	0.2	75.0	-2.0	5	3.4	-0.6	81.6	0.3
PCC 7-2	3	5.1	-0.6	100.0	0.7	5	3.4	-0.6	81.6	0.3
PCC 8-1	4	6.7	0.2	100.0	0.7	9	5.4	0.7	80.1	0.2
PCC 8-2	4	6.7	0.2	100.0	0.7	9	5.4	0.7	80.1	0.2
PCC 9-1	1	1.0	-2.8	100.0	0.7	4	2.9	-0.9	91.9	1.1
PCC 9-2	4	6.7	0.2	100.0	0.7	4	2.9	-0.9	91.9	1.1
PCC 10-1	2	2.9	-1.8	75.0	-2.0	7	4.7	0.2	58.9	-1.5
PCC 10-2	4	6.7	0.2	100.0	0.7	7	4.7	0.2	58.9	-1.5
PCC 11-1	5	7.8	0.8	100.0	0.7	7	4.7	0.2	79.4	0.1
PCC 11-2	4	6.7	0.2	87.5	-0.6	7	4.7	0.2	79.4	0.1
PCC 12-1	5	7.8	0.8	90.0	-0.4	7	4.7	0.2	56.0	-1.7
PCC 12-2	4	6.7	0.2	87.5	-0.6	7	4.7	0.2	56.0	-1.7

Table B.3: Shared CAD performance data.

Code name	A	B	z_B	C	z_C	D	E	z_E	F	z_F
SCC 1-1	4	6.7	0.2	100.0	0.7	5	3.4	-0.6	91.4	1.1
SCC 1-2	5	7.8	0.8	90.0	-0.4	5	3.4	-0.6	91.4	1.1
SCC 2-1	5	7.8	0.8	83.3	-1.1	5	3.4	-0.6	86.3	0.7
SCC 2-2	3	5.1	-0.6	100.0	0.7	5	3.4	-0.6	86.3	0.7
SCC 3-1	6	9.5	1.7	87.5	-0.6	6	4.2	-0.1	100.0	1.7
SCC 3-2	4	6.7	0.2	94.4	0.1	6	4.2	-0.1	100.0	1.7
SCC 4-1	5	7.8	0.8	100.0	0.7	5	3.4	-0.6	79.4	0.1
SCC 4-2	2	2.9	-1.8	100.0	0.7	5	3.4	-0.6	79.4	0.1
SCC 5-1	6	9.5	1.7	94.4	0.1	7	4.7	0.2	97.1	1.5
SCC 5-2	5	7.8	0.8	100.0	0.7	7	4.7	0.2	97.1	1.5
SCC 6-1	4	6.7	0.2	100.0	0.7	5	3.4	-0.6	75.9	-0.2
SCC 6-2	4	6.7	0.2	100.0	0.7	5	3.4	-0.6	75.9	-0.2
SCC 7-1	4	6.7	0.2	87.5	-0.6	2	1.5	-1.8	78.6	0.1
SCC 7-2	2	2.9	-1.8	100.0	0.7	2	1.5	-1.8	78.6	0.1
SCC 8-1	3	5.1	-0.6	83.3	-1.1	4	2.9	-0.9	77.0	-0.1
SCC 8-2	5	7.8	0.8	100.0	0.7	4	2.9	-0.9	77.0	-0.1
SCC 9-1	5	7.8	0.8	90.0	-0.4	7	4.7	0.2	81.3	0.3
SCC 9-2	5	7.8	0.8	100.0	0.7	7	4.7	0.2	81.3	0.3
SCC 10-1	4	6.7	0.2	100.0	0.7	5	3.4	-0.6	90.8	1.0
SCC 10-2	2	2.9	-1.8	100.0	0.7	5	3.4	-0.6	90.8	1.0
SCC 11-1	4	6.7	0.2	87.5	-0.6	3	2.3	-1.4	75.5	-0.2
SCC 11-2	3	5.1	-0.6	100.0	0.7	3	2.3	-1.4	75.5	-0.2

Appendix C

Supplementary Figures



Interested in collaborative CAD?

Participate in our study.

Compensation up to $30 for two hours

Need at least 12 months of 3D CAD experience

Sign-up here: goo.gl/3PjPM7

UNIVERSITY OF TORONTO | MIT Massachusetts Institute of Technology

Questions? Contact us at:
cloudcad@mit.edu

Figure C-1: One of the posters used for recruitment of participants.

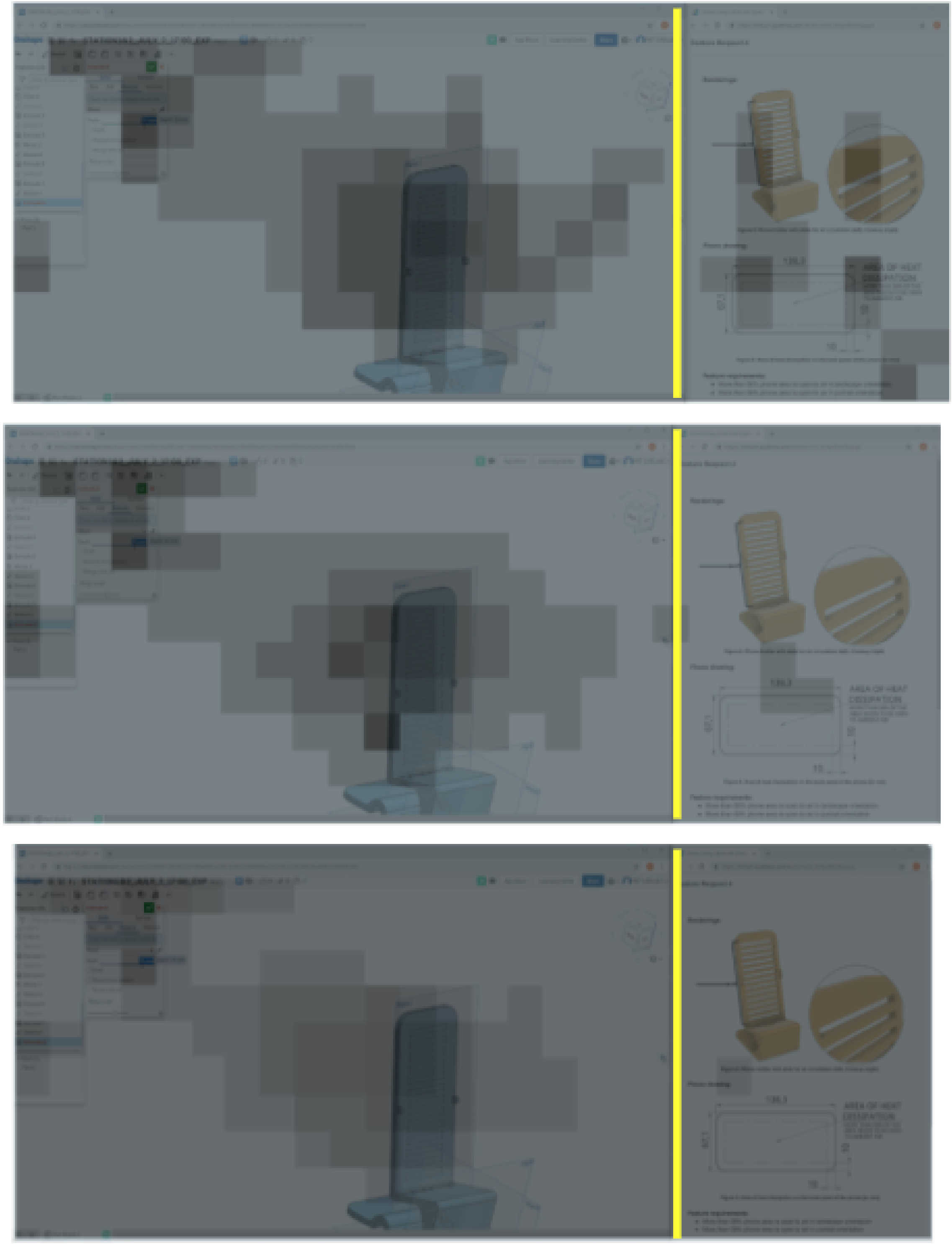


Figure C-2: Cursor activity heat-maps: (top) Shared CAD, (center) Parallel CAD and (bottom) Individual CAD. Darker color implies area with higher activity.

Bibliography

[1] “Current and future usage plans for cloud-based computer-aided design (cad) worldwide in 2017 [graph].” Jon Peddie Research, 2017. In Statista. Retrieved July 30, 2020, from https://www.statista.com/statistics/754566/worldwide-current-and-future-usage-plans-for-cad-in-the-cloud/.

[2] P. J. Runkel and J. E. McGrath, eds., *Research on Human Behavior: A Systematic Guide to Method*. New York: Holt, Rinehart and Winston, 1972.

[3] J. Zhou, V. Phadnis, and A. Olechowski, “Analysis of designer emotions in collaborative and traditional computer-aided design,” *Journal of Mechanical Design*, 2020.

[4] H. Arshad, V. Phadnis, and A. Olechowski, “Paired computer-aided design: The effect of collaboration mode on differences in model quality,” *ASME IDETC/CIE (accepted)*, 2020.

[5] J. J. Van Bavel, K. Baicker, P. S. Boggio, V. Capraro, A. Cichocka, M. Cikara, M. J. Crockett, A. J. Crum, K. M. Douglas, J. N. Druckman, and et al., “Using social and behavioural science to support covid-19 pandemic response,” *Nature human behaviour*, vol. 4, p. 460–471, May 2020.

[6] E. Brynjolfsson, J. Horton, A. Ozimek, D. Rock, G. Sharma, and H.-Y. TuYe, “Covid-19 and remote work: An early look at us data,” *NBER Working Paper No. 27344*, 2020.

[7] M. H. Olson, “Remote office work: changing work patterns in space and time,” *Communications of the ACM*, vol. 26, no. 3, p. 182–187, 1983.

[8] B. J. Bergiel, E. B. Bergiel, and P. W. Balsmeier, “Nature of virtual teams: a summary of their advantages and disadvantages,” *Management Research News*, vol. 31, no. 2, p. 99–110, 2008.

[9] S. Morrison-Smith and J. Ruiz, “Challenges and barriers in virtual teams: a literature review,” *SN Applied Sciences*, vol. 2, no. 6, 2020.

[10] P. Koutsabasis, S. Vosinakis, K. Malisova, and N. Paparounas, “On the value of virtual worlds for collaborative design,” *Design Studies*, vol. 33, no. 4, p. 357–390, 2012.

[11] M. McCrindle, *The ABC of XYZ: Understanding the Global Generations.* The ABC of XYZ, 2014.

[12] W. Zhou, E. Simpson, and D. P. Domizi, "Google docs in an out-of-class collaborative writing activity," *International Journal of Teaching and Learning in Higher Education*, vol. 24 n3, p. 359–375, 2012.

[13] V. S. Phadnis, K. A. Leonardo, D. R. Wallace, and A. L. Olechowski, "An exploratory study comparing cad tools and working styles for implementing design changes," *Proceedings of the Design Society: International Conference on Engineering Design*, vol. 1, no. 1, p. 1383–1392, 2019.

[14] J. J. Zhou, V. Phadnis, and A. Olechowski, "Analysis of designer emotions in collaborative and traditional computer-aided design," *ASME IDETC/CIE*, 2019.

[15] V. Phadnis, D. R. Wallace, and A. Olechowski, "A multimodal experimental approach to study cad collaboration," *Computer-Aided Design and Applications*, vol. 18, no. 2, 2020.

[16] V. Phadnis, D. R. Wallace, and A. Olechowski, "(in progress) are two heads better than one in cad? comparing the speed and quality of individual vs. pair work.," *Journal of Mechanical Design*, 2020.

[17] D. R. Wallace, *A Probabilistic Specification-based Design Model.* PhD thesis, 1995.

[18] K. J. Ostergaard and J. D. Summers, "Development of a systematic classification and taxonomy of collaborative design activities," *Journal of Engineering Design*, vol. 20, no. 1, p. 57–81, 2009.

[19] L. L. Martins, L. L. Gilson, and M. Travis Maynard, "Virtual teams: What do we know and where do we go from here?," *Journal of Management*, vol. 30, no. 6, p. 805–835, 2004.

[20] W. D. Li, W. F. Lu, J. Y. H. Fuh, and Y. S. Wong, "Collaborative computer-aided design—research and development status," *Computer-Aided Design*, vol. 37, no. 9, p. 931–940, 2005.

[21] G. Andreadis, G. Fourtounis, and K.-D. Bouzakis, "Collaborative design in the era of cloud computing," *Advances in Engineering Software*, vol. 81, p. 66–72, 2015.

[22] L. L. Bucciarelli and S. Kuhn, "Engineering education and engineering practice: Improving the fit," *Between Craft and Science*, p. 210–229, 1997.

[23] T. Woo, *The Ease of Agile Development in Cloud-Based CAD.* 2016.

[24] K. Eves, J. Salmon, J. Olsen, and F. Fagergren, "A comparative analysis of computer-aided design team performance with collaboration software," *Computer-Aided Design and Applications*, vol. 15, no. 4, p. 476–487, 2018.

[25] B. Stone, J. Salmon, A. Hepworth, E. Red, and M. Killian, "Methods for determining the optimal number of simultaneous contributors for multi-user cad parts," *Proceedings of CAD'16*, 2016.

[26] B. Stone, J. Salmon, K. Eves, M. Killian, L. Wright, J. Oldroyd, S. Gorrell, and M. C. Richey, "A multi-user computer-aided design competition: Experimental findings and analysis of team-member dynamics," *Journal of Computing and Information Science in Engineering*, vol. 17, no. 3, 2017.

[27] V. L. Holyoak, E. Red, and G. Jensen, "Effective collaboration through multi user cax by implementing new methods of product specification and management," *Computer-Aided Design and Applications*, vol. 11, no. 5, p. 560–567, 2014.

[28] E. R. Babbie, *The Practice of Social Research.* Wadsworth Publishing Company, 2001.

[29] L. T. M. Blessing and A. Chakrabarti, "Drm, a design research methodology," 2009.

[30] A. Sivanathan, T. Lim, J. Ritchie, R. Sung, Z. Kosmadoudi, and Y. Liu, "The application of ubiquitous multimodal synchronous data capture in cad," *Computer-Aided Design*, vol. 59, p. 176–191, 2015.

[31] Y. Liu, J. M. Ritchie, T. Lim, Z. Kosmadoudi, A. Sivanathan, and R. C. W. Sung, "A fuzzy psycho-physiological approach to enable the understanding of an engineer's affect status during cad activities," *Computer-Aided Design*, vol. 54, p. 19–38, 2014.

[32] P. Nguyen, T. A. Nguyen, and Y. Zeng, "Segmentation of design protocol using eeg," *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, vol. 33, no. 1, p. 11–23, 2019.

[33] P. Nguyen, T. A. Nguyen, and Y. Zeng, "Empirical approaches to quantifying effort, fatigue and concentration in the conceptual design process," *Research in Engineering Design*, vol. 29, no. 3, p. 393–409, 2018.

[34] M. H. Rahman, C. Schimpf, C. Xie, and Z. Sha, "A computer-aided design based research platform for design thinking studies," *Journal of Mechanical Design*, vol. 141, no. 12, 2019.

[35] K. J. Ostergaard, W. R. Wetmore, A. Divekar, H. Vitali, and J. D. Summers, "An experimental methodology for investigating communication in collaborative design review meetings," *CoDesign*, vol. 1, no. 3, p. 169–185, 2005.

[36] H. Jiang and C.-C. Yen, "Protocol analysis in design research: a review," in *International association of societies of design research*, p. 147–156, Korean Society of Design Science, Oct 2009.

[37] H. H. Tang, Y. Y. Lee, and J. S. Gero, "Comparing collaborative co-located and distributed design processes in digital and traditional sketching environments: A protocol study using the function–behaviour–structure coding scheme," *Design Studies*, vol. 32, no. 1, p. 1–29, 2011.

[38] K. A. M. Khaidzir and B. Lawson, "The cognitive construct of design conversation," *Research in Engineering Design*, vol. 24, no. 4, p. 331–347, 2013.

[39] I. Chiu and L. H. Shu, "Potential limitations of verbal protocols in design experiments," *Volume 5: 22nd International Conference on Design Theory and Methodology; Special Conference on Mechanical Vibration and Noise*, 2010.

[40] C. W. Ennis and S. W. Gyeszly, "Protocol analysis of the engineering systems design process," *Research in Engineering Design*, vol. 3, no. 1, p. 15–22, 1991.

[41] P. J. Kieslich and F. Henninger, "Mousetrap: An integrated, open-source mouse-tracking package," *Behavior research methods*, vol. 49, p. 1652–1667, Oct 2017.

[42] E. Hehman, R. M. Stolier, and J. B. Freeman, "Advanced mouse-tracking analytic techniques for enhancing psychological science," *Group Processes & Intergroup Relations*, vol. 18, no. 3, p. 384–401, 2015.

[43] P. Du and E. F. MacDonald, "Eye-tracking data predict importance of product features and saliency of size change," *Journal of Mechanical Design*, vol. 136, no. 8, 2014.

[44] E. Kwon, J. D. Ryan, A. Bazylak, and L. H. Shu, "Does visual fixation affect idea fixation?," *Journal of Mechanical Design*, vol. 142, no. 3, 2020.

[45] K. L. Phan, T. Wager, S. F. Taylor, and I. Liberzon, "Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in pet and fmri," *NeuroImage*, vol. 16, p. 331–348, Jun 2002.

[46] P. Ekman and E. L. Rosenberg, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford University Press, Apr 2005.

[47] Q. Bao, A. M. Hughes, E. Burnell, and M. C. Yang, "Investigating user emotional responses to eco-feedback designs," *ASME IDETC/CIE*, 2018.

[48] X. Alameda-Pineda, E. Ricci, and N. Sebe, *Multimodal Behavior Analysis in the Wild: Advances and Challenges*. Academic Press, Nov 2018.

[49] G. W. Hill, "Group versus individual performance: Are n+1 heads better than one?," *Psychological Bulletin*, vol. 91, no. 3, p. 517–539, 1982.

[50] K. Beck and E. Gamma, *Extreme Programming Explained: Embrace Change*. Addison-Wesley Professional, 2000.

[51] L. Williams, R. R. Kessler, W. Cunningham, and R. Jeffries, "Strengthening the case for pair programming," *IEEE Software*, vol. 17, no. 4, p. 19–25, 2000.

[52] B. J. d. S. Estácio, B. J. da Silva Estácio, and R. Prikladnicki, "Distributed pair programming: A systematic literature review," *Information and Software Technology*, vol. 63, p. 1–10, 2015.

[53] A. Begel and N. Nagappan, "Pair programming: What's in it for me? andrew," in *International Symposium on Empirical Software Engineering and Measurement*, Association for Computing Machinery, Oct 2008.

[54] T. Chau, F. Maurer, and G. Melnik, "Knowledge sharing: agile methods vs. tayloristic methods," *WET ICE 2003. Proceedings. Twelfth IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, 2003.*

[55] S. Salinger, F. Zieris, and L. Prechelt, "Liberating pair programming research from the oppressive driver/observer regime," *2013 35th International Conference on Software Engineering (ICSE)*, 2013.

[56] E. d. Bella, E. di Bella, I. Fronza, N. Phaphoom, A. Sillitti, G. Succi, and J. Vlasenko, "Pair programming and software defects–a large, industrial case study," *IEEE Transactions on Software Engineering*, vol. 39, no. 7, p. 930–953, 2013.

[57] P. Dourish and V. Bellotti, "Awareness and coordination in shared workspaces," *Proceedings of the 1992 ACM conference on Computer-supported cooperative work - CSCW '92*, 1992.

[58] M. Stefik, D. G. Bobrow, G. Foster, S. Lanning, and D. Tatar, "Wysiwis revised: early experiences with multiuser interfaces," *ACM Transactions on Information Systems (TOIS)*, vol. 5, no. 2, p. 147–167, 1987.

[59] J. Li, S. Greenberg, and E. Sharlin, "A two-sided collaborative transparent display supporting workspace awareness," *International Journal of Human-Computer Studies*, vol. 101, p. 23–44, 2017.

[60] T. Phan, W. Honig, and N. Ayanian, "Mixed reality collaboration between human-agent teams," *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 2018.

[61] S. Niu, D. Scott McCrickard, J. Nguyen, D. Haqq, L. Kotut, T. L. Stelter, and E. A. Fox, "Investigating paradigms of group territory in multiple display environments," *Proceedings of the ACM on Human-Computer Interaction*, vol. 4, no. GROUP, p. 1–28, 2020.

[62] M. Rajpal, "Effective distributed pair programming," *Proceedings of the 13th Conference on Global Software Engineering - ICGSE '18*, 2018.

[63] M. Goldman, G. Little, and R. C. Miller, "Collabode: Collaborative coding in the browser," in *Proceedings of the 4th international workshop on cooperative and human aspects of software engineering, CHASE 2011. ACM*, IEEE/ACM, May 2011.

[64] J. Schenk, L. Prechelt, and S. Salinger, "Distributed-pair programming can work well and is not just distributed pair-programming," *Companion Proceedings of the 36th International Conference on Software Engineering - ICSE Companion 2014*, 2014.

[65] D. E. Perry, H. P. Siy, and L. G. Votta, "Parallel changes in large scale software development: an observational case study," *Proceedings of the 20th International Conference on Software Engineering.*

[66] C.-W. Ho, S. Raha, E. Gehringer, and L. Williams, "Sangam," *Proceedings of the 2004 OOPSLA workshop on eclipse technology eXchange - eclipse '04*, 2004.

[67] X. Chen, S. Gao, Y. Yang, and S. Zhang, "Multi-level assembly model for top-down design of mechanical products," *Computer-Aided Design*, vol. 44, no. 10, p. 1033–1048, 2012.

[68] T. W. Maver, "Social impacts of computer-aided architectural design," *Design Studies*, vol. 7, no. 4, p. 178–184, 1986.

[69] D. H. Sonnenwald, "Communication roles that support collaboration during the design process," *Design Studies*, vol. 17, no. 3, p. 277–301, 1996.

[70] T.-J. Nam and D. Wright, "The development and evaluation of syco3d: a real-time collaborative 3d cad system," *Design Studies*, vol. 22, no. 6, p. 557–582, 2001.

[71] K. Kim and K.-P. Lee, "Collaborative product design processes of industrial design and engineering design in consumer product companies," *Design Studies*, vol. 46, p. 226–260, 2016.

[72] J. Q. Coburn, J. L. Salmon, and I. Freeman, "Effectiveness of an immersive virtual environment for collaboration with gesture support using low-cost hardware," *Journal of Mechanical Design*, vol. 140, no. 4, 2018.

[73] K. T. Ulrich, *Product Design and Development.* Tata McGraw-Hill Education, Nov 2003.

[74] L. Rising and N. Janoff, "The scrum software development process for small teams," *Software, IEEE*, vol. 17, pp. 26 – 32, 08 2000.

[75] J. S. Gero and U. Kannengiesser, "A function-behaviour-structure ontology of processes," *Design Computing and Cognition '06*, p. 407–422.

[76] D. Shah, E. Kames, M. Clark, and B. Morkos, “Development of a coding scheme for qualitative analysis of student motivation in senior capstone design,” *Volume 3: 21st International Conference on Advanced Vehicle Technologies; 16th International Conference on Design Education*, 2019.

[77] D. Chickarello, J. Righter, A. Patel, and J. D. Summers, “Establishing a protocol to observe leadership behaviors within engineering design teams,” *Volume 7: 30th International Conference on Design Theory and Methodology*, 2018.

[78] H.-E. Chen and S. R. Miller, “Can wearable sensors be used to capture engineering design team interactions?: An investigation into the reliability of sociometric badges,” *Volume 7: 29th International Conference on Design Theory and Methodology*, 2017.

[79] A. J. Viera and J. M. Garrett, “Understanding interobserver agreement: the kappa statistic,” *Family medicine*, vol. 37, p. 360–363, May 2005.

[80] P. Muchiri and L. Pintelon, “Performance measurement using overall equipment effectiveness (oee): literature review and practical application discussion,” *International Journal of Production Research*, vol. 46, no. 13, p. 3517–3535, 2008.

[81] A. J. deRon and J. E. Rooda, “Equipment effectiveness: Oee revisited,” *IEEE Transactions on Semiconductor Manufacturing*, vol. 18, no. 1, p. 190–196, 2005.

[82] R. Oechsner, M. Pfeffer, L. Pfitzner, H. Binder, E. Müller, and T. Vonderstrass, “From overall equipment efficiency (oee) to overall fab effectiveness (ofe),” *Materials Science in Semiconductor Processing*, vol. 5, no. 4-5, p. 333–339, 2002.

[83] Balijepally, Balijepally, Mahapatra, Nerur, and Price, “Are two heads better than one for software development? the productivity paradox of pair programming,” *MIS Quarterly*, vol. 33, no. 1, p. 91, 2009.

[84] M. A. Alharbi, “Exploring the potential of google doc in facilitating innovative teaching and learning practices in an efl writing course,” *Innovation in Language Learning and Teaching*, vol. 14, no. 3, p. 227–242, 2020.

[85] R. Guay, P. R. Foundation, and E. T. S. T. Collection, *Purdue Spatial Visualization Test.* 1976.

[86] M. Foschi, “Hypotheses, operationalizations, and manipulation checks,” *Laboratory Experiments in the Social Sciences*, p. 247–268, 2014.

[87] J. Faludi, F. Yiu, O. Srour, R. Kamareddine, O. Ali, and S. Mecanna, “Do student trials predict what professionals value in sustainable design practices?,” *Journal of Mechanical Design*, vol. 141, no. 10, 2019.

[88] N. Salleh, E. Mendes, and J. Grundy, “Empirical studies of pair programming for cs/se teaching in higher education: A systematic literature review,” *IEEE Transactions on Software Engineering*, vol. 37, no. 4, p. 509–525, 2011.

[89] N. Nagappan, L. Williams, M. Ferzli, E. Wiebe, K. Yang, C. Miller, and S. Balik, "Improving the cs1 experience with pair programming," *ACM SIGCSE Bulletin*, vol. 35, no. 1, p. 359, 2003.

[90] P. Company, P. Company, M. Contero, J. Otey, and R. Plumed, "Approach for developing coordinated rubrics to convey quality criteria in mcad training," *Computer-Aided Design*, vol. 63, p. 101–117, 2015.

[91] M. L. McHugh, "Interrater reliability: the kappa statistic," *Biochemia Medica*, p. 276–282, 2012.