





A Multimodal Experimental Approach to Study CAD Collaboration

Vrushank S. Phadnis¹ , David R. Wallace² and Alison Olechowski³ 

¹Massachusetts Institute of Technology, vphadnis@mit.edu

²Massachusetts Institute of Technology, drwallac@mit.edu

³University of Toronto, olechowski@mie.utoronto.ca

Corresponding author: David Wallace, drwallac@mit.edu

Abstract. Computer-Aided Design (CAD) collaboration has been studied since the early days of CAD research, but standardized metrics and experimental procedures for synchronous, cloud-based CAD research are lacking in the literature. In this work, we lay out an empirical approach to investigate collaboration in CAD. Our work is unique in its relevance to the nuances of synchronous CAD collaboration. We first define metrics of interest: speed, quality, communication, satisfaction, and UI (user interaction). We then introduce an experimental toolkit that leverages automated and manual data capture methods. Lastly, we deploy our toolkit in a pilot study setting to reveal preliminary insights and validate the workings of our method. Although preliminary, our findings suggest that pairs were slower than single CAD users because of coordination overheads involving communication and model-tree-scanning.

Keywords: multimodal experiments, CAD collaboration, protocol analysis, emotion tracking, cursor tracking.

DOI: <https://doi.org/10.14733/cadaps.2021.328-342>

1 INTRODUCTION

Only recently, fully synchronous CAD platforms have been deployed at commercial scale [27]. The shift to synchronous collaboration was enabled by modern cloud architecture, which allows for multiple designers to manipulate the same CAD geometry. This style of working is new to the design community and its implications are under-studied in prior research. Synchronous CAD promises to be an important topic of research as more design teams are collaborating globally [17].

Having multiple people working on the same CAD file has been shown to reduce modeling time [23]. It is bound to also affect other aspects of design work like quality, creativity, communication, and decision making. The complexity of executing design projects with teams working in synchronous CAD warrants novel design research methods [19].

Experiments provide a unique advantage by allowing researchers a higher specificity in defining research questions [2]. 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 this paper, we present an experimental method to answer exploratory research questions (RQ) about collaborative CAD.

In particular, we are interested in the following:

- RQ1: What metrics are of interest in studying the performance of CAD users?
- RQ2: What CAD collaboration insights can we derive from using our experimental method?

2 LITERATURE REVIEW

Research tools have evolved over the past few decades. However, the overarching types of research strategies have not changed. Figure 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 our 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.

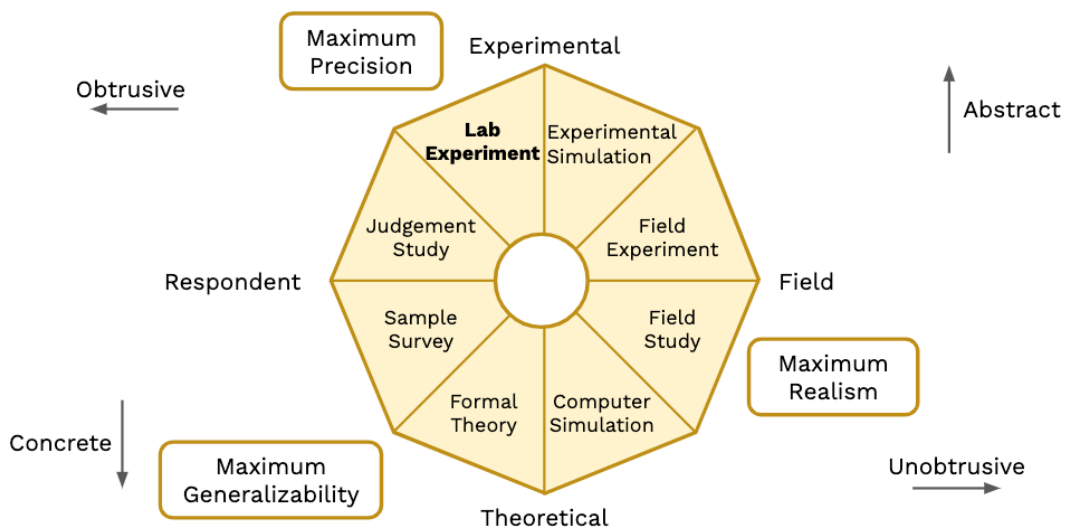


Figure 1: Research strategies. Adapted from [20].

2.1 Experiments in Synchronous CAD Research

Experimental methods have been used by other researchers to study synchronous CAD. The nascent state of this research area meant that these studies used CAD software that was still in beta testing phase and under development. Particularly, multi-user CAD (MUCAD) developed as an add-on to Siemens NX brings synchronous capability to a traditional CAD software [5, 23]. This tool has been part of multiple synchronous CAD research studies. Stone *et al.* proposed a method to predict the optimal size for a team working in MUCAD [24]. 26 students were assigned to teams of one to four to model three CAD models of varying difficulties. The authors found completion-time-prediction accuracy was correlated with team sizes, i.e. completion times for larger teams was predicted with higher accuracy. The prediction function was created using a regression model with the number of CAD features and their dependencies taken as input. Learning effects played an important role in this work as teammates complimented each other's skills. It is also noteworthy that simpler parts were modeled faster by single users vs. MUCAD teams. The quality of models was better in multi

user teams. This work was limited by the use of a pre-production version of CAD software and its participant pool size.

Another study by Stone *et al.* investigates MUCAD team dynamics [23]. This study consisted of 63 students divided into 21 teams of three. Overall, MUCAD teams performed faster than individuals tackling the same task with traditional CAD. A second finding of this study suggests that high performing teams communicated less but more effectively. And last, a higher mismatch in CAD skill levels correlated to low performance in MUCAD teams. Final scores were modified to compensate for inconsistencies in participants' experience from CAD software limitations.

Eves *et al.* conducted a study comparing MUCAD teams to teams using single user CAD systems connected by an email server [10]. In this work, eight teams of four worked on a multi-day CAD project. Team performance was rated by expert judges who evaluated screenshots of the participants' CAD files. This study was limited by a bias from mismatch in skill in using MUCAD and use of pre-production CAD software. These limitations were compensated for by modifying the judges scores. The final results did not show any statistically significant difference in performance between singles vs. MUCAD teams. However, MUCAD teams had higher awareness of team activities and communicated more in comparison to single user teams.

These early insights from research in synchronous CAD are promising and help us direct our work. Particularly, we take inspiration from literature to define our metrics of interest. Speed, quality, and communication were adopted as metrics in our work based on Stone *et al.* and Eves *et al.*'s work.

2.2 Data Capture Techniques

Advent of new sensing techniques and faster computing capabilities have enabled researchers to extract a wide variety of data types. Sivanathan *et al.* presented a ubiquitous data capture method for CAD work [22]. Their toolkit aggregates real-time logging of multiple data types like CAD metadata, eye tracking, mouse logging, and electroencephalography (EEG). The upside of the framework presented by Sivanathan *et al.* was its applicability outside of research settings. This toolkit was proposed to be used by practicing engineers to monitor CAD user behavior in real world. However, this setup requires a complicated sensor suite and is obtrusive to designers' natural way of working.

Another toolkit presented by Liu *et al.* used EEG, galvanic skin resistance (GSR) and electrocardiography (ECG) to capture psycho-physiological data. This dataset was used to compute designers' emotions and compared to logs of CAD system UI [14]. A fuzzy model was developed to derive emotions and results were then validated through 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 that can perform design protocol analysis methods by microstate analysis [15]. In this work, a model to map EEG signals to design events of interest is developed. Then a case study is analyzed using the. The authors present a detailed comparison between the automated protocol analysis results and ratings by expert human coders. This work concluded with recommendations on the choice of algorithms in automating design protocols using EEG. Like Sivanathan *et al.*, a limitation of Nyugen *et al.*'s methodology is its reliance on body sensors leading to an obtrusive experimental setup.

Rahman *et al.* present a software-based approach to capture design behavioral data. A CAD based platform, ENERGY3D was developed to study solar energy systems [18]. This work is non-intrusive to the participants but is limited in its variety of variables captured. Outside of the CAD context, Ostergaard *et al.* present an experimental method to study collaboration patterns in design review meetings [16]. In this work, they compare groups vs. individuals working through a design review session. Data collection is done by evaluating participant worksheets post-experiment. In this work, the simplicity of the experimental setup is less obtrusive but limits the amount and variety of data available.

In our toolkit, we capture multiple data streams based on video post-processing. We leverage the benefits of automated and manual techniques to collect experimental data. Such multimodal analysis has been successfully used to investigate human social behavior [1]. This approach requires no body mounted sensors and provides a more natural setting for participants.

Manual coding methods like protocol analysis have a long precedence in design research and will be discussed next [12]. Researchers have investigated the effects of virtual vs. co-located design processes using protocol analysis methods [25]. 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 [12]. Verbalization techniques are common in protocol analysis. For example, Anwar *et al.* modeled cognitive behavior in conceptual design work using such techniques [13]. Verbalization or ‘think aloud’ techniques have limitations like data validity, steep learning curve, and some tasks are not suitable for verbalization [7]. Another popular protocol analysis method involves participants recollecting their experience post-experiment. Such retrospective protocol analysis is 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 accounted for all possible events of interest beforehand [9].

3 METHODS OVERVIEW

3.1 CAD Working Styles

To narrow down the scope of our toolkit and the pilot study, we chose three distinct CAD working styles: single person working in traditional CAD, single person working in synchronous CAD, and lastly pairs working in synchronous CAD. The three working styles help us test all possible modes in which our toolkit might be applied in a future study. We chose Onshape (OS) as our collaborative CAD software. We also added a traditional CAD working style with Solidworks (SW) as a benchmark.

To comply with the ethics board requirements and bring structure to our discussion, we have adopted the following name coding: SWS1, SWS2 and SWS3 for single SolidWorks participants, OSS1, OSS2, and OSS3 for single Onshape participants, OSP1_1 and OSP1_2 for participants working in Onshape pair 1, up to OSPn_1 and OSPn_2 for participants working in Onshape pair n. See Figure 2 for a pictorial representation of this naming scheme.

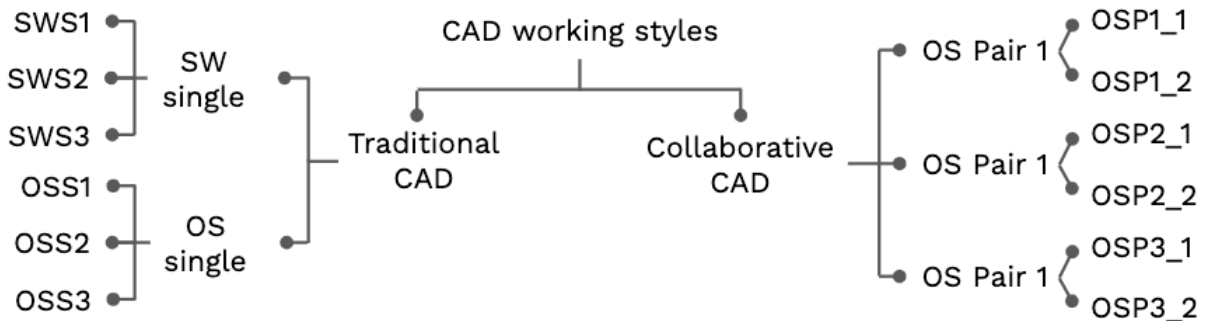


Figure 2: Naming scheme of different CAD working styles.

3.2 Metrics of Interest

Understanding synchronous CAD collaboration warrants analysis of both CAD and non-CAD activities. As pointed out in the literature review, we chose speed and quality as metrics of CAD work. In addition, communication, satisfaction, and UI statistics were documented as non-CAD related metrics. We refined our metrics based on observations from the pilot study data. For example, all tracks in the manual coding Section 3.3.1 were identified by studying multiple participant videos. We were able to add metrics posteriori because of the post-processing approach of our toolkit. A

summary of all metrics of interest and their associated measurement tools are shown in Figure 3. Metrics were operationalized using three tools/methods: protocol analysis, cursor tracking and emotion tracking. These will be further discussed in Section 3.3.

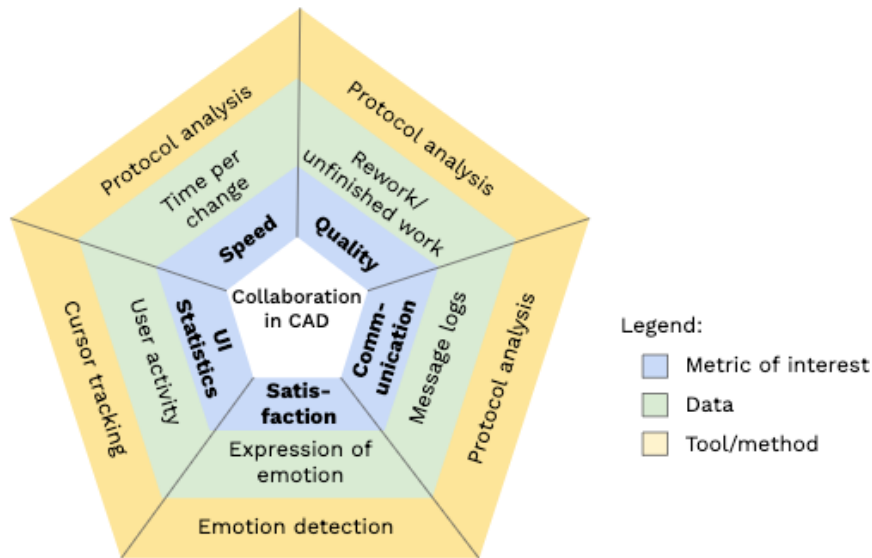


Figure 3: Toolkit architecture showing mapping of metrics of interest to data.

3.3 Data Generation Methods

This section describes the software specifications and post processing methods used to implement our multi-modal approach. All data streams captured during the pilot study were routed using the architecture shown in Figure 4. As can be seen, we used a variety of data traces and storage options to connect all stations in our experiment setup. A description of each data type follows. Note that some metrics are redundant/overlapping and were used to triangulate/cross-reference results.

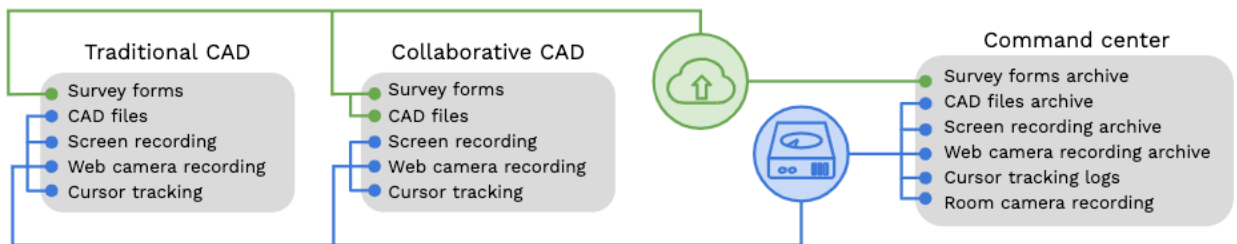


Figure 4: Data trace interconnections.

3.3.1 Manual coding of screen recording

Manual 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-behavior-structure (FBS) are popular in the design research community [11]. 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 [21].

The five tracks from Figure 5, used in our protocol analysis are: CAD track (change number), communication, geometry rotation, roll back, and help menu access. These were derived from

watching multiple screen recording videos from our pilot run. All videos were annotated using a video coding software, V-Note, shown in Figure 5. A detailed description of each track in our coding schema follows.

The CAD track helped us record time taken to complete CAD tasks and derive speed and quality metrics. This is consistent with metrics used by other design researchers studying synchronous CAD [23]. We included three subcategories into this track: active/productive work, rework, and incomplete/unfinished work. The statistics generated from this track were used to assess the participants performance in our pilot study.

Communication messages were time stamped and tagged 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.

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

Rolling back the model tree is a common strategy used in CAD work. It becomes particularly relevant in dealing with coupling effects of CAD changes. Each rollback was documented with the time for which the model stayed in this rolled back state.

Our participants use the software's help menu primarily for information on feature menu options. Time spent on referring to the help menu was catalogued in the help access track.

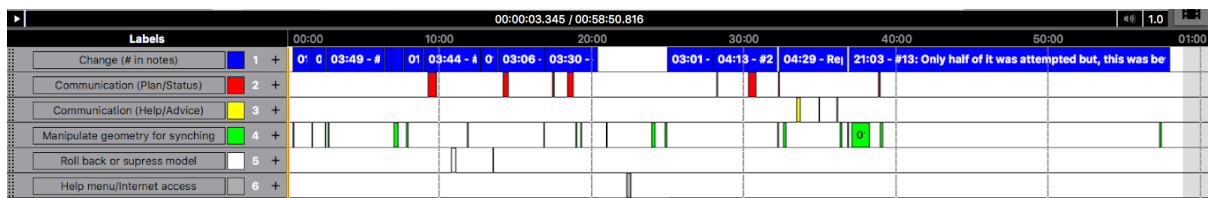


Figure 5: Manual coding session in progress.

3.3.2 Cursor tracking

We built a post processing object detection algorithm based on template matching using OpenCV. This helped us capture cursor locations in our study. The script scanned pixels of an input frame and correlated it to multiple input template images to find a match. This method relies on the screen recording video and obviates the need to record more footage or maintain additional software during the pilot study.

Cursor tracking was implemented on video footage of OS pairs and OS single working styles only, as SWS users were working on a different operating system. The change in graphics from two different operating systems posed challenges 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, which is important to future work.

High thresholds were set for the cursor-template matches, and multiple calibration tests were performed on still frames and video samples. This was important to mitigate false detection and ensured repeatability of our cursor tracking results. To further counter false detections, data points from the same location for more than 10 seconds were flagged. With our current parameter settings, the algorithm was able to recover more than 85% of cursor locations. An even higher recovery rate can be achieved by running cursor tracking natively, a simple extension of our experimental set-up requiring additional CPU resources. Cursor location data was categorized in regions of interest as seen in Figure 6. As can also be seen, the point locations from the software were mapped to heat maps overlaid by the regions of interest.



Figure 6: (Left) Screen recording showing regions of interest. (Right) Heat map of cursor locations. Regions of interest: 1. Feature Selection, 2. Communication, 3. Model Tree, 4. CAD Graphic Area.

3.3.3 Emotion tracking using web camera

Affectiva is a commercial emotion detection solution that has been successfully used in other design research projects. An interface using their software developer kit (SDK) was built with the ability to upload videos and generate emotion statistics [28].

The software works by tracking facial cues and mapping them to seven basic emotions. An example frame being processed in the software and sample output is shown in Figure 7. Facial expressions were tracked and processed to calculate emotion metrics for each participant.

The SDK output was in the form of a rating from 0-100, signifying the level of expression of a given emotion (see Figure 7). In other words, a rating of 0 meant the emotion was not detected and 100 meant the emotion was fully expressed. Our code sampled frames at 2 frames per second (fps) to match sampling frequency of other software in our toolkit. Overall, we had a 81% detection rate for facial videos processed from the pilot study.



Figure 7: (Left) Emotion recognition algorithm showing points of significance. (Right) Data stream from sample video footage [28].

4 PILOT STUDY

The primary goals for our pilot study 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.

4.1 Lab Setup

The experimental setup was built in a pre-existing facility in the MIT Behavioral Research Lab (BRL). Figure 8 shows our experimental setup at the MIT BRL. This space was used to host the experiments and training sessions.



Figure 8: (Left) Typical participant work station. (Right) An experiment in progress.

Our focus was to be as non-intrusive as possible and emulate a design office setup. Participants were asked not to use personal electronic devices during the pilot study and adequate storage space was provided to accommodate participants' belongings. It was important for us that participants were in physical and audio isolation from each other to simulate remote virtual collaboration. To enable this, the physical space inside our lab was divided using dividers and slide out partitions. We also added white noise in the environment to address any residual audio.

4.2 Design Task

The study was set up such that participants roleplayed toy designers working on an early stage concept of a toy car. After reviewing the initial CAD file, participants were tasked with a list of 42 changes that represented feedback gathered from customer review sessions. Examples from the change list are shown in Table 1.

Sr. No.	Change request	Points
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

Table 1: Examples from participant change list.

We purposely designed our change list to stay focused on the user-software interaction and required our participants to perform drafting specific changes only. The prescriptive nature of our change list minimized variability in interpretation of the changes 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. Every participant was given a demo of unique features in synchronous CAD including a tutorial on the collaborative features of the software.

4.3 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. 16 people participated in the study but we qualified only 12 participants data points 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.

4.4 Pilot Study Results

Our pilot study led to 9 CAD files, as shown in Figure 9. Single working style participant CAD files are not as elaborate as their paired counterparts. This is due to fewer changes implemented by single working style participants. This difference is captured more quantitatively in the following sections.

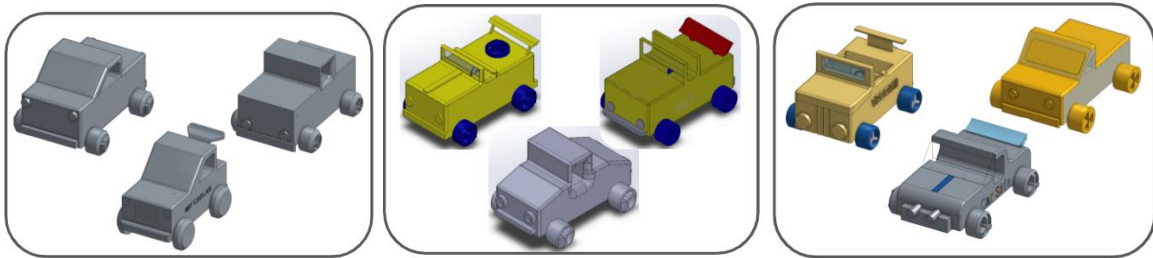


Figure 9: CAD files from (left) OS Singles, (center) SW Singles, and (right) OS Pairs.

4.4.1 Manual coding

Cohen's Kappa is a measure of inter-rater reliability (IRR) and is often used to check the reproducibility of a coding schema [6]. To check the validity of our coding schema, a second coder rated 99 video samples. Video samples were distributed to have at least one representation of our coding schema from each track per participant of our pilot study. These ratings were cross referenced with their original categories 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 [26].

Participant success in our pilot study was compared by calculating the number of changes completed in 60 minutes. As can be seen from the first two rows of Table 2, paired participants were slower at implementing changes in comparison to individuals working by themselves. It is noteworthy that the average time taken to complete a change is not directly indicative of the total number of changes implemented in our pilot study. This distinction can be explained by further analysis of the non-CAD related components of the protocol analysis.

All ratios and OME values are tabulated in Table 2. Our definition of overall modeling efficiency (OME) helps compare relative performance of participants. We adapted this approach from manufacturing literature, where overall equipment efficiency (OEE) is a popular aggregate metric to compare production efficiency of machines [8]. We defined three ratios to compute OME number: availability ratio, speed ratio, and quality ratio. Availability ratio captures the total time available to participants to perform CAD 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 at which participants executed changes. This ratio was calculated by baselining 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 incorrect implementation

and leaving changes incomplete. OME was calculated as product of all ratios, as seen in equation below.

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

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 changes

Q'= Total number of changes attempted by participant

Working styles	SW Single	OS Single	OS Pair
Average number of changes	13	11	8
Speed loss ratio	0.87	0.76	0.51
Average non-CAD time (secs)	555	168	867
Availability ratio	0.85	0.95	0.76
Average number of quality losses	3	2	1.5
Quality loss ratio	0.81	0.87	0.84
Average OME	0.6	0.6	0.3

Table 2: Summary of OME calculations.

4.4.2 Cursor tracking

Modelling the cursor tracking data proved challenging, given its size and complexity. To simplify our analysis, we chose to aggregate data for each working style. Table 3 shows the summary of cursor locations detected in each region of interest from Figure 6. This information augments the qualitative insights shown by heat maps.

	OS Single (%)	OS Pair (%)	Difference: OS Pair – OS Single
Feature Selection	11.6	9.7	-1.8
Model Tree	6.9	10.3	3.4
Communication	0.0	5.4	5.4
CAD Graphics Area	81.5	74.5	-7.0

Table 3: Table showing cursor activity summary for OS Singles and OS Pairs.

4.4.3 Emotion tracking

Similar to the cursor tracking data, emotion tracking data was computed at a high frequency and then down sampled to be consistent with cursor tracking data. Emotion metrics were generated by filtering our data with a threshold value. All instances of sustained emotion were catalogued in time intervals. Table 4 shows the aggregate values of the seven basic emotions split between OS single and OS pairs participants. As can be seen from the last column in Table 4, paired participants

expressed more emotions than individuals. A more elaborate analysis of this emotion data can be found in Zhou *et al.*'s work [28].

	OS Single	OS Pair	Difference: OS Pair – OS Single
Joy	23.3	177.3	154
Sadness	7.3	12.0	4.7
Disgust	444.0	251.3	-192.7
Contempt	328.7	587.7	259
Anger	9.3	37.7	28.4
Fear	0.0	8.0	8
Surprise	296.7	511.7	215

Table 4: Summary of emotions captured during pilot study [28].

4.4.4 Aggregating all data streams

Analyzing multiple data streams together can prove challenging and makes it difficult to compare the three working styles in a comprehensive manner. In an effort to review all data qualitatively, we created dotted plots as seen in Figure 10. These plots show the manual protocol analysis data alongside the cursor tracking dataset.

As the two datasets were not mutually exhaustive, it was difficult to set up a quantitative analysis. Dotted plots are a good visual indicator to study participant behavior at a high level. For example, it is seen that OSP3_1 rolled back the CAD geometry more often compared to OSP3_1. Laying out multiple streams side by side helped us triangulate information. The dotted plots augment information presented by the protocol analysis and cursor activity by adding a time dimension. For example, the cursor activity in Table 3 give us an aggregate view of the participants movement within the CAD environment, but the dotted plots help us understand the distribution of time spent over the timeline of our pilot study, *i.e.* if the transitions are happening regularly or in concentrated spurts.

5 DISCUSSION

From looking at the raw data generated through the manual coding method, we observed that pairs were slower than individuals. This finding was also supported in the performance analysis; wherein overall performance of pairs judged using OME was lower. The speed loss ratios contributed the most to lowering OME scores. As a result, we did not observe any explicit assembly bonus effects *i.e.* paired CAD work was not summative of individual CAD work.

The differences in user behavior of pairs *vs.* individuals is depicted in Table 3. The additional activity of communication is applicable only to pairs and is an extra overhead. Secondly, the cursor activity for pairs was higher in the model tree region. We suspect model-tree-scanning was used by pairs to seek awareness of their counterpart's work. This is peculiar to collaborative CAD and is a coordination overhead in addition to communication.

Emotion analysis output suggests that pairs emoted more than individuals. This can potentially be related to user satisfaction and we note that pairs expressed more positive emotions. All data presented in the pilot study results was generated using archival video footage. Post-processing of data alleviated impetus on the research staff to analyze data during the pilot study.

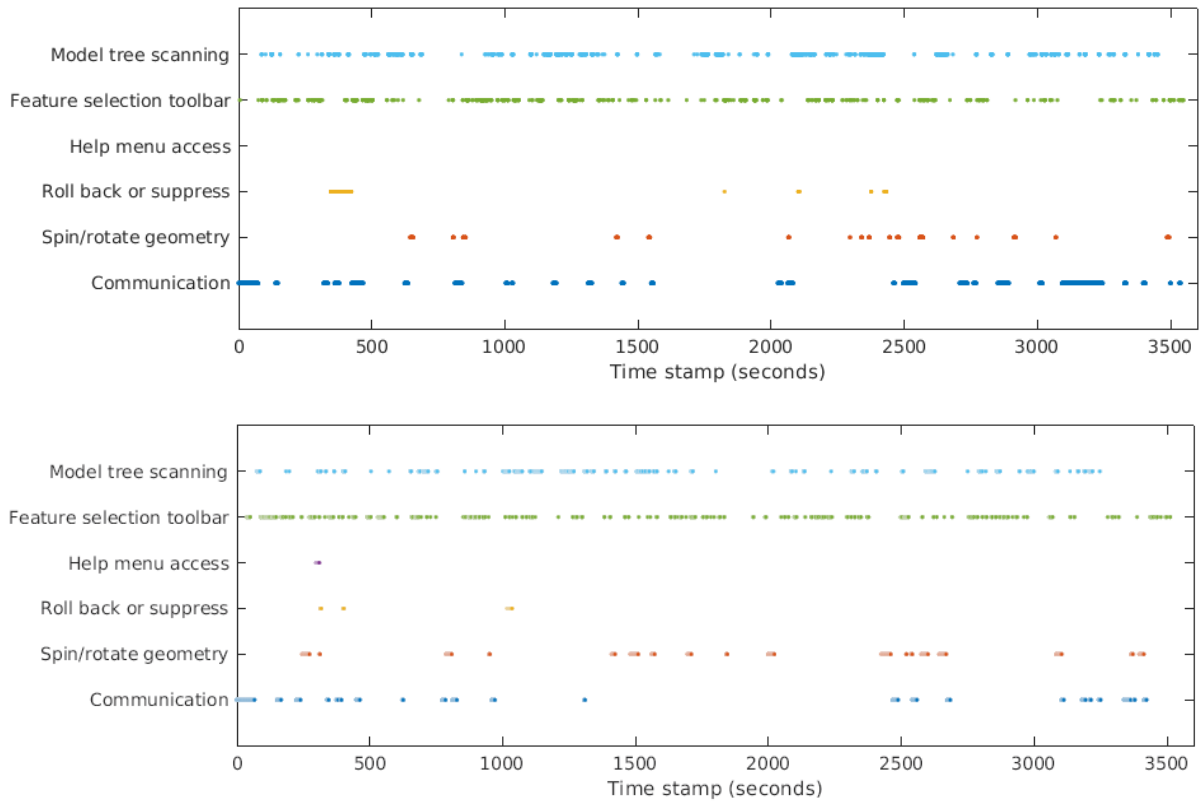


Figure 10: (Top to Bottom) Dotted plots for OSP 3, showing OSP 3_1 on top and OSP3_2 on bottom.

5.1 Limitation

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 generalizability to the real world. Rigorous training of student participants can help minimize this disconnect and will be implemented in future. Finally, our design task was kept drafting focused to bring clarity to participants and ease measurement of data. However, purely drafting focused exercises limit the richness of our results. Lastly, our post processing approach lent convenience in running the experiment, but at the cost of data loss from video post processing algorithms.

5.2 Future Work

We relied on the number of years of CAD experience as a recruitment qualification for participants' incoming skill in our pilot study. In future runs, we would like to build a calibration mechanism within our study to augment self-reporting. This can be done using standardized tests like the Purdue Spatial Visualization Test (PSVT) [4]. Richer communication mediums like audio and video conferencing are now becoming common in design collaboration. Restricting participants to only written communication was limiting and we would like to implement audio communication in the future. Lastly, our work would benefit from local installations of cursor tracking and emotion recognition tools that run natively in background.

6 CONCLUSIONS

The multimodal experimental toolkit provided consistency and standardization in executing our experiments. We posed two research questions in Section 1, pertaining to our work. We report our findings as below:

- RQ1: Figure 3 captures our metrics of interest; namely speed, quality, communication, satisfaction, and UI statistics. A successful application of these metrics was enabled by the toolkit in the pilot study.
- RQ2: Preliminary collaborative CAD insights from our pilot study are outlined below:
 - Single participants were faster than pairs. This is consistent with literature on computer supported collaborative work (CSCW). [3].
 - Communication and an increase in model-tree-scanning were identified as potential coordination overheads and are unique to pair CAD.
 - Pairs emoted more than single CAD participants.

ACKNOWLEDGEMENT

The authors would like to thank Janice Zhou, Kevin Leonardo, participants of our study and staff of MIT BRL. We would also like to thank lab members of MIT CADLab and UofT Ready Lab for their invaluable support and insights.

Vrushank S. Phadnis, <https://orcid.org/0000-0001-6526-0848>

Alison Olechowski, <https://orcid.org/0000-0001-5557-654X>

REFERENCES

- [1] Alameda-Pineda, X.; Ricci, E.; Sebe, N.: Multimodal Behavior Analysis in the Wild: Advances and Challenges, Academic Press, 2018. <https://doi.org/10.1016/C2017-0-01387-3>
- [2] Babbie, E. R.: The Practice of Social Research, Wadsworth Publishing Company, 2001. <https://www.cengage.com/c/the-practice-of-social-research-14e-babbie/9781305104945/>
- [3] Balijepally, V.; Mahapatra, R.; Nerur, S.; Price, K.: Are Two Heads Better than One for Software Development? The Productivity Paradox of Pair Programming, MIS Quarterly 33, 2009, 91. <https://doi.org/10.2307%2F20650280>
- [4] Bodner, G. M.; Guay, R. B.: The Purdue Visualization of Rotations Test, The Chemical Educator 2, 1997, 1–17. <https://doi.org/10.1007/s00897970138a>
- [5] Briggs, J.C.; Hepworth, A.I.; Stone, B.R.; Coburn, J.Q.; Greg Jensen, C; Red, E.: Integrated, Synchronous Multi-User Design and Analysis, Journal of Computing and Information Science in Engineering, 2015. <https://doi.org/10.1115/1.4029801>
- [6] Chickarello, D.; Righter, J.; Patel, A.; Summers, J. D.: Establishing a Protocol to Observe Leadership Behaviors Within Engineering Design Teams, Volume 7: 30th International Conference on Design Theory and Methodology, 2018. <https://doi.org/10.1115/DETC2018-85869>
- [7] Chiu, I.; Shu, L. H.: 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. <https://doi.org/10.1115/DETC2010-28675>
- [8] deRon, A. J.; Rooda, J. E.: Equipment Effectiveness: OEE Revisited, IEEE Transactions on Semiconductor Manufacturing 18, 2005, 190–196. <https://doi.org/10.1109/TSM.2004.836657>
- [9] Ennis, C. W.; Gyeszly, S. W.: Protocol analysis of the engineering systems design process, Research in Engineering Design 3, 1991, 15–22. <https://doi.org/10.1007/BF01580065>

- [10] Eves, K.; Salmon, J.; Olsen, J.; Fagergren, F.: A comparative analysis of computer-aided design team performance with collaboration software, *Computer-Aided Design and Applications*, 2018. <https://doi.org/10.1080/16864360.2017.1419649>
- [11] Gero, J. S.; Kannengiesser, U.: A function-behaviour-structure ontology of processes, *Design Computing and Cognition*, 2006, 407–422. https://doi.org/10.1007/978-1-4020-5131-9_21
- [12] Jiang, H.; Yen, C.-C.: Protocol Analysis in Design Research: a review, in *International Association of Societies of Design Research*, (Korean Society of Design Science, 2009, pp. 147–156. <http://www.iasdr2009.or.kr/Papers/Orally%20Presented%20Papers/Design%20Method/Protocol%20Analysis%20in%20Design%20Research%20-%20a%20Review.pdf>
- [13] Khaidzir, K. A. M.; Lawson, B.: The cognitive construct of design conversation, *Research in Engineering Design* 24, 2013, 331–347. <https://doi.org/10.1007/s00163-012-0147-8>
- [14] Liu, Y.; Ritchie, J. M.; Lim, T.; Kosmadoudi, Z.; Sivanathan, A.; Sung, R. C. W.: A fuzzy psycho-physiological approach to enable the understanding of an engineer's affect status during CAD activities, *Computer-Aided Design* 54, 2014, 19–38. <https://doi.org/10.1016/j.cad.2013.10.007>
- [15] Nguyen, P.; Nguyen, T. A.; Zeng, Y.: Segmentation of design protocol using EEG, *Artificial Intelligence for Engineering Design, Analysis and Manufacturing* 33, 2019, 11–23. <https://doi.org/10.1017/S0890060417000622>
- [16] Ostergaard, K. J.; Wetmore, W. R.; Divekar, A.; Vitali, H.; Summers, J. D.: An experimental methodology for investigating communication in collaborative design review meetings, *CoDesign* 1, 2005, 169–185. <https://doi.org/10.1080/15710880500298520>
- [17] Ozer, M.; Cebeci, U.: The Role of Globalization in New Product Development, *IEEE Transactions on Engineering Management* 57, 2010, 168–180. <https://doi.org/10.1109/TEM.2009.2025492>
- [18] Rahman, M. H.; Schimpf, C.; Xie, C.; Sha, Z.: A Computer-Aided Design Based Research Platform for Design Thinking Studies, *Journal of Mechanical Design* 141, 2019. <https://doi.org/10.1115/1.4044395>
- [19] Red, E.; French, D.; Jensen, G.; Walker, S. S.; Madsen, P.: Emerging Design Methods and Tools in Collaborative Product Development, *Journal of Computing and Information Science in Engineering* 13, 2013. <https://doi.org/10.1115/1.4023917>
- [20] Runkel, P. J.; McGrath, J. E. *Research on Human Behavior: A Systematic Guide to Method*, New York, 1972.
- [21] Shah, D.; Kames, E.; Clark, M.; Morkos, B.: 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. <https://doi.org/10.1115/DETC2019-98423>
- [22] Sivanathan, A.; Lim, T.; Ritchie, J.; Sung, R.; Kosmadoudi, Z.; Liu, Y.: The application of ubiquitous multimodal synchronous data capture in CAD, *Computer-Aided Design* 59, 2015, 176–191. <https://doi.org/10.1016/j.cad.2013.10.001>
- [23] Stone, B.; Salmon, J.; Eves, K.; Killian, M.; Wright, L.; Oldroyd, J., et al.: A Multi-User Computer-Aided Design Competition: Experimental Findings and Analysis of Team-Member Dynamics, *Journal of Computing and Information Science in Engineering* 17, 2017. <https://doi.org/10.1115/1.4035674>
- [24] Stone, B.; Salmon, J.; Hepworth, A.; Red, E.; Killian, M.: Methods for Determining the Optimal Number of Simultaneous Contributors for Multi-user CAD Parts, *Proceedings of CAD'16*, 2016. <https://doi.org/10.1080/16864360.2016.1273578>
- [25] Tang, H. H.; Lee, Y. Y.; Gero, J. S.: 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* 32, 2011, 1–29. <https://doi.org/10.1016/j.destud.2010.06.004>
- [26] Viera, A. J.; Garrett, J. M.: Understanding interobserver agreement: the kappa statistic, *Fam. Med.* 37, 2005, 360–363. <https://pubmed.ncbi.nlm.nih.gov/15883903/>

- [27] Wu, D.; Rosen, D. W.; Lihui, W.; Schaefer, D.: Cloud-based design and manufacturing: A new paradigm in digital manufacturing and design innovation, *Computer-Aided Design*, 2015, 59–1-14. <https://doi.org/10.1016/j.cad.2014.07.006>
- [28] Zhou, J. (janice); Phadnis, V.; Olechowski, A.: Analysis of Designer Emotions in Collaborative and Traditional Computer-Aided Design, Volume 7: 31st International Conference on Design Theory and Methodology, 2019. <https://doi.org/10.1115/DETC2019-98516>