Proceedings of the ASME 2020
International Design Engineering Technical Conferences
and Computers and Information in Engineering Conference
IDETC/CIE2020
August 17-19, 2020, Virtual, Online

DETC2020-22730

PAIRED COMPUTER-AIDED DESIGN: THE EFFECT OF COLLABORATION MODE ON DIFFERENCES IN MODEL QUALITY

Hamza Arshad
Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada

Vrushank Phadnis
Mechanical Engineering
MIT
Cambridge, USA

Alison Olechowski
Mechanical & Industrial Engineering
University of Toronto
Toronto, Canada

ABSTRACT

We present the results of an experiment investigating two different modes of collaboration on a series of computer-aided design (CAD) tasks. Inspired by the pair programming literature, we anticipate that partners working in a fully synchronous collaborative CAD environment will achieve different levels of quality in CAD models depending on their mode of collaboration – one in which the pair is free to work in parallel, and another where the pair must coordinate to share one control.

We found that a shared CAD control led to significantly better overall CAD quality than parallel CAD control. In addition, the shared control mode led to more complete and consistent CAD models, as well as the tendency for participants to follow instructions to correctly replicate features for the design task. As is predicted in the literature, a trade-off relationship (albeit weak) between quality and speed via the parallel collaboration was found. In contrast, the shared control mode shows no clear relationship between speed and quality.

Collaborative CAD is increasingly seen as an appealing tool for modern product design teams. This study suggests that the benefits of this tool are not solely the effect of the tool itself, but result from the collaboration style of the designers using the tool.

Keywords: Computer-Aided Design; collaboration; quality.

1. INTRODUCTION

Design engineers use computer-aided design (CAD) tools to model products, allowing the design team to display their cumulative design choices without costly prototyping. The advancement of CAD technology since the late 20th century has had a large impact on design engineers and allowed for more innovation in design choices [1], [2]. Further innovation in design can be attributed to an increase in collaboration between designers working on the same product; however, since these files are stored in static files on personal computers, problems arise relating to effective CAD modelling, version updating, and even communication [3].

A growing development is the use of virtual CAD environments where users collaborate with each over an online connection, allowing designers to be connected even if they do not share the same physical space. How this type of collaboration affects the CAD models, and the users' abilities to CAD, are questions being addressed by a developing body of research. One area of particular interest is how CAD quality is affected.

In this paper, we explore designer collaboration in a virtual, synchronously collaborative CAD environment. We present the results of an experiment to explore the effect of two different modes of paired CAD collaboration on the quality of the resulting CAD models.

2. BACKGROUND

2.1 Collaboration in Engineering

Collaboration is a fundamental part of engineering product design and thus the emergence of modern virtual collaboration tools has been accompanied by studies of the capabilities and potential effects of these tools at the organization level [4]–[6]. These studies build on the foundational literature on concurrent engineering in which the parallelization of engineering tasks is achieved by an integrated multidisciplinary team with the aim of reducing time to market [7], [8]. Much of the existing literature is focused on organizational-level tools and antecedents to collaboration. We seek understanding on the pair-level effect of collaboration in engineering work.

One source of relevant literature on performance effects of collaboration is that of software engineers collaborating via pair programming. Pair programming is when two programmers together produce one artifact (one piece of code) [9]. The traditional format of pair programming is for two partners to sit at the same workstation, with one partner – the driver – controlling the computer inputs, while the other partner actively observes the driver's work, "watching for defects, thinking of alternatives, looking up resources, and considering strategic implications" [9]. The role of driver is changed periodically between the partners.

An alternative to traditional pair programming is distributed pair programming, where the partners synchronously collaborate on the same artifact from different locations. They both see a copy of the artifact and at least one of the pair has control of the screen [10]. Distributed pair programming has been found to maintain many of the advantages of traditional pair programming [10].

2.2 Collaboration in Computer-Aided Design

Modern, commercially-available CAD tools support real-time, synchronous collaboration, but have been previously described as lacking in fulfilling the needs of collaborative design (leading to coordination, rework, and versioning issues) [11], [12]. Yet little research exists to establish our understanding of undertaking engineering work, or CAD work, with this new style of collaboration at the designer-level.

Generating a series of directions for future work, Eves et al. conducted an experimental study of collaborative CAD software which suggests that, compared to single-user CAD, multi-user CAD increases awareness of teammates' activities and increases communication between team members [13]. This exploratory study does not draw conclusions about performance effects of collaborative CAD.

2.3 Collaboration and Quality

Studies examining the effect of increased collaboration via virtual teams on quality have shown mixed results, with some evidence of collaboration leading to higher quality but slower decision making [14], [15]. In fact, quality is often argued to be a variable with a trade-off with speed in product design [16]–[18], since a push to design quickly may lead to process short-cuts or undershoot of performance specifications.

Studies have examined the effect of partnered work on detecting errors in automated decision-making contexts [19]. This work describes how the effect of a second person is not well-understood, and could be influenced by: social facilitation, where the individual's performance is improved by the mere presence of others; drive theory, which suggests that this presence of others can have a motivation-related effect on task completion, which may result in positive or negative outputs; or, social loafing, which leads to less effort exerted by the individual and therefore a loss in per-person effectiveness.

Studies of paired programmers have found that pairs generate higher quality code with fewer defects [9], [20], [21], but these studies have not yet differentiated between variations of collaboration mode. It is argued that the higher quality of code resulting from pair programming can be explained by the fact that "four eyeballs are better than two, and a huge number of defects are prevented right from the start" [9].

2.4 Quality in CAD

Quality in CAD is implicitly considered in the way designers are trained, mostly motivated by reusability of CAD in the industry and software-specific best practice training that exist [22], [23]. Yet examinations of CAD model quality is itself under-discussed in the literature. González-Lluch et al. present a taxonomy of CAD model quality as a structure for evaluating quality assurance and testing tools [24]. In that study, feature-based modeling like that presented in our study is classified as "procedural modeling," and the authors point to a study by Company et al. who identify six quality dimensions for CAD models, along with a rubric, meant as an aid for CAD training [25]. Quality is measured as the evaluation of the model as: valid, complete, consistent, concise, simple and capturing design intent. Each of these sub-dimensions is accompanied by two to four specific characteristics.

Camba et al. describe best practices for achieving reusability of CAD models [26]. Reusability is related to quality, in that following specific strategies of parametric design will lead to internal structures that are conducive to alteration and reuse. Three formal strategies for CAD reusability are discussed, suggesting that typical CAD training focuses on specific procedural commands, rather than modeling strategies.

3. RESEARCH QUESTION

We present the findings of an experiment that compares the CAD quality output from paired collaboration strategies. The first strategy, analogous to traditional pair programming, we call Shared CAD Control (SCC), whereby the two designers share one control of the CAD interface and model. The second collaboration strategy, which we call Parallel CAD Control (PCC), is analogous to distributed pair programming, and allows both partners the freedom to design independently on the same CAD model. The collaboration types' workflow are illustrated in Figure 1.

We ask how CAD quality will be affected by these two CAD collaboration strategies. Based on the pair programming and error detection literature, we expect that Shared CAD Control will result in higher quality models on average.

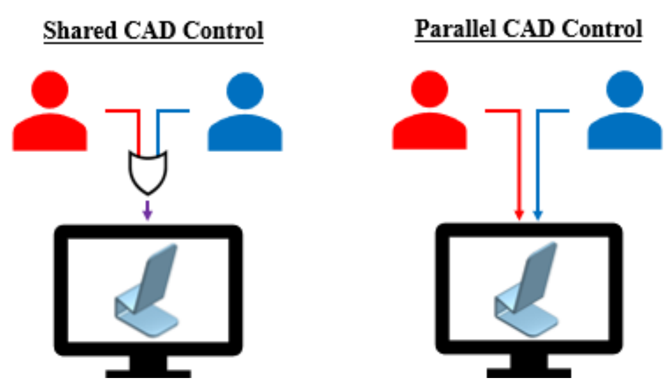

FIGURE 1: ILLUSTRATION OF THE COLLABORATION METHOD STRATEGIES

4. METHOD

4.1 Experiment overview

In order to collect CAD quality data, we recruited participants to participate in a two-hour session divided into four main phases as shown in Figure 2.

The pre-study survey was used to understand the demographics and experience of the users, while the post-study survey was used to understand user awareness and satisfaction during CAD tasks, the results of which would be used in future studies. The Training period was used to familiarize participants with the CAD software that they would be using, while Baseline I was a test of the skill of users. After Baseline I, users were paired with another participant and were randomly assigned to work in either the SCC or PCC collaboration method. Pairs were given five minutes to test this collaboration before the actual experiment was conducted, during which time they could consult with the study investigators on any collaboration questions.

The Experiment phase lasted 45 minutes. Participants were provided an initial CAD file of a phone holder, along with instructions in the form of a series of tasks required to be completed by the pair. It is important to note that Baseline I and Baseline II had tasks that were fully prescribed (i.e. exact dimensions and 3D shapes were provided; no design decisions were required). This experimental design choice was made to test participants' CAD design abilities rather than conceptual design abilities. Furthermore, participants were removed from the study if they did not reach a predetermined threshold of progress in Baseline I in order to remove weak CAD users (i.e. potential outliers). The Experiment phase also consisted of prescribed tasks with exact 3D renderings provided but specific dimensions unspecified, as participants were encouraged to decide these changes with their partners; this is discussed further in Section 4.3. An example of a design task given to participants is provided in Appendix A.

4.2 Experimental setup

Two main distinctions can be made when discussing the overall setup of this experiment – one concerns the physical setup and the other the virtual one.

4.2.1 Physical setup

The physical setup of this experiment consisted of recruiting potential participants and setting the experimental room in which they worked. Recruitment of participants was achieved through posters around the campuses of local post-secondary institutions; all of these posters were largely concentrated in the engineering or design buildings of these campuses. A small minority of posters were put up in design labs and machining shops. Other tools for recruitment were focused on University of Toronto engineering students through the use of mailed newsletters, display boards, and online engineering student groups.

Participants were screened to provide a pool of potential participants with a minimum requirement of CAD skill. These filters were the requirement of at least 12 months of 3D CAD experience, consistent with previously conducted research [27].

The data analyzed here represents 40 individual participants in the experiment, equally divided into 10 PCC and 10 SCC pairs. The demographics of the study participants are listed in Table 1. Months of CAD experience was self-reported as 29.9 (SD = 21.1) for PCC and 34.6 (SD = 36.2) for SCC. It should be noted that the random assignment of participants to workflow resulted in variations of participant demographics between the two treatments. There may be uncontrolled effects of these differences, for example, studies have reported differences in spatial reasoning by gender [28]. In our study, outcomes analyzed by gender revealed no evidence of a statistically significant difference.

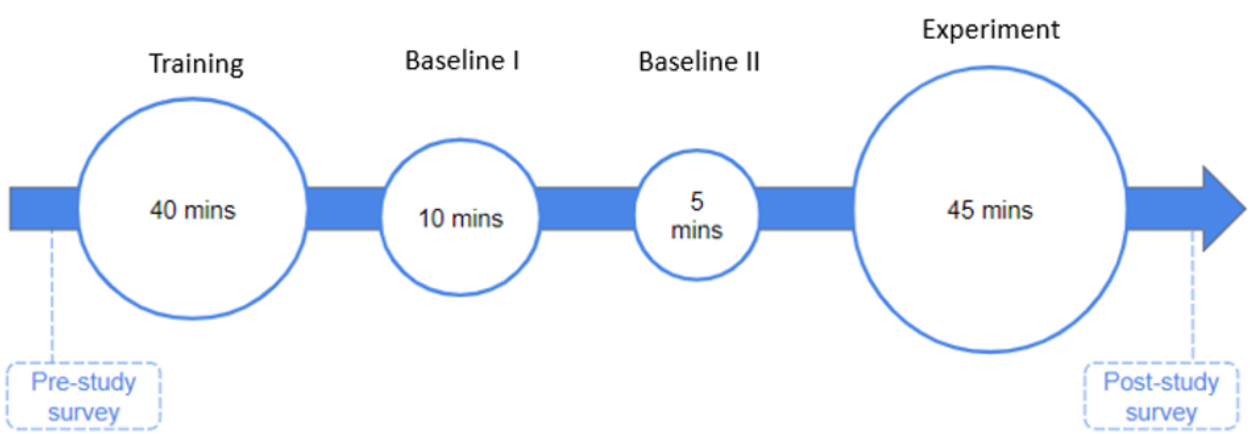

FIGURE 2: PARTICIPANTS WERE DIRECTED THROUGH FOUR MAIN PHASES OF VARYING TIME LENGTHS WITH DATA ON CAD QUALITY BEING DERIVED FROM THE EXPERIMENT (EXP) PHASE

TABLE 1: DEMOGRAPHIC MAKE-UP OF PARTICIPANTS OF THE PARALLEL CAD CONTROL (PCC) AND THE SHARED CAD CONTROL (SCC)

Variable	PCC		SCC	
	Number	%	Number	%
Male	12	60	17	85
Female	8	40	3	15
White	4	20	2	10
Asian	10	50	14	70
Latin American	1	5	0	0
Arab	2	10	1	5
Black	0	0	1	5
Mixed race	1	5	2	10
Prefer not to say	1	5	0	0
Did not respond	1	5	0	0
Total	20	100	20	100

The remainder of the physical setup involves the room setup. Four stations were set up in the four corners of the room and facing the wall, adjacent to their partner separated by a divider. This was done in order to reduce potential visual distractions participants may have from looking at the screens of other participants or from the actions of the study investigators. During the Baseline and Experiment phases, white noise played from central speakers to impede participants from hearing each other.

As seen in Figure 3, each station consisted of:

- Windows-operated computer set with the programs involved in the experiment (see 4.2.2)
- Headset to facilitate verbal communication between partners and reduce overall ambient noise during the Baseline and Experiment phases
- Webcam to capture participant face
- Reference paper of relevant commands in the CAD software

PCC utilizes Onshape (a software for shared, synchronous, cloud-based CAD) and SCC uses Use Together (a software that shares a user's screen with potential remote control) in conjunction with Onshape to allow the pairs to manipulate the same CAD file one at a time in an identical environment.

4.2.2 Virtual setup

Through the use of a "pre-phase" windows, users were forced to read instructions on the duration and expectations of the next phase; its main function was to coordinate the synchronous start of experiment participants. This provided consistency with measuring the time duration for each pair's work. The phases (other than the survey and training phases) had two main windows open when participants were working on the CAD tasks – the CAD software on the left and the CAD task list on the right. It is important to note that one task at a time was visible to participants. Participants were not allowed to go back to previous tasks.

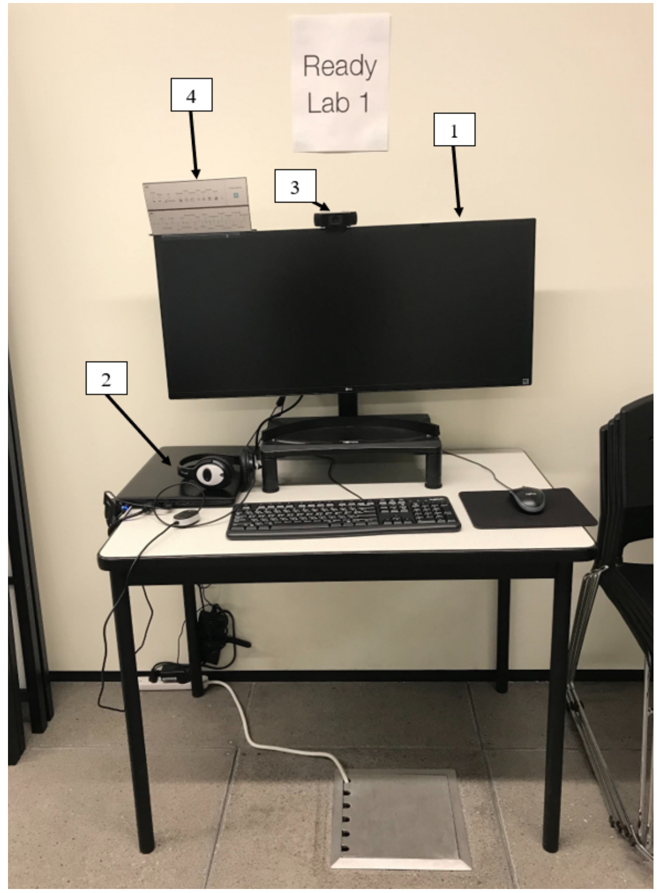

FIGURE 3: SET-UP OF A STATION IN THE EXPERIMENT WITH THE MONITOR (1), HEADSET (2), WEBCAM (3), AND REFERENCE PAPER (4)

4.3 Post-experiment data analysis

4.3.1 Scoring quality

Quality was scored via investigator evaluation with a study-specific rubric. The scoring rubric used to measure the quality of each CAD design built on work done by Company et. al [25]; their research looked into developing a method to breakdown CAD quality into definable sub-dimensions or "to convey quality criteria" in mechanical CAD (MCAD) training [25]. They divided quality into five categories - complete, concise, consistent, valid, and effective. A CAD model that was "effective" helped to convey design intent; since features in our study were fully prescribed to pairs with the exception of exact dimensions in some cases, the variable of "effective" quality was not appropriate, and thus is not considered in this study. The other categories' definitions were simplified to be applicable to our study, as shown in Table 2. Each category is subdivided into potential indicators by Company et al. For our study, not all indicators were applicable to the CAD models. Those that applied are listed in Table 2.

TABLE 2: SIMPLIFIED DEFINITIONS AND INDICATORS FOR THE SUB-DIMENSIONS CATEGORIES, ADAPTED FROM COMPANY ET AL. [25]

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

From these potential metric categories, the individual indicators (referred to as "conditions" from hereafter) for meeting the categories were derived. For each design task, each metric category was considered to determine whether they applied to the task and if so, conditions for that task were created. For example, the first task required pairs to create three holes for a charging cable and speaker ports. Specific conditions for the rubric derived for this case were:

- CAD visually resembles rendering (complete)
- 3 mm minimum width for the charging cable hole (complete)
- Mirroring used for speaker holes (consistent)
- Fully defined sketch (concise)

Two "overall" conditions were also used that required raters to look over the whole model completed by the pair. These were:

- Number of failed instances in the model tree (valid)
- Number of new parts created

These conditions were not factored into the overall quality scores for pairs, but are presented as a separate means of analyzing the two collaboration methods.

Each design task had 2-3 rubric metric categories and 6-10 rubric conditions derived from them. These were rated either a 0 (did not meet rubric condition) or a 1 (met rubric condition). A total of 81 conditions were possible for each pair in the Experiment phase to meet and 19 conditions in Baseline I; however, the experiment was designed in a manner to not allow participants to complete all of the design tasks and thus never reach the full set of conditions. This was done to allow participants to work for the complete phase time. An example of the rating is shown in Appendix B.

4.3.2 Computing quality scores

To compare the Overall Quality Score (OQS) between PCC and SCC, the average quality of the ten participants from each collaboration mode was found via equation 1.

$$(OQS)_k = \frac{\sum_{i=1}^{m}(\sum_{j=1}^{n}\frac{x_i}{n})}{m} \quad (1)$$

Where x represents either 0 (did not meet rubric point i) and 1 (met rubric point i) and is an average of n (number of rubric conditions in that design task) points. The average of the m (number of design tasks completed by participant) points was the OQS for participant k. Tasks that were left incomplete when participants ran out of time were not included in the calculation of the score – this was because the rubric is based on evaluating completed design tasks.

As stated previously, any new parts in the CAD file or any failed instances in the model tree (i.e. validity) were considered to be indicators of poor quality. These were tallied up instead of being given a score between 0 and 1 and a separate analysis was completed with them.

4.3.3 Reliability of CAD Quality Scoring

An inter-rater reliability (IRR) test was performed for the ratings of two raters using the Experiment phase's rubric, with a sample of 35% of the data. An IRR score of 0.96 was found, indicating an "almost perfect agreement" [29] between the raters. Initially, there was a major deviation among the two raters on the rubric condition of "CAD visually resembles rendering" – included to award points when the overall look of the pair's CAD matched the rendering presented to them for each design task; however, a negotiation phase between the two raters revealed that it was open to interpretation. This condition was therefore replaced with a more explicit rubric condition customized for each design task. This resulted in the IRR score of 0.96, and led to the decision to use the principal rater's score for the remainder of the study.

5. RESULTS AND DISCUSSION

5.1 Overall quality score

First, average Overall Quality Score for each collaboration mode was calculated (summarized in Figure 4A), and differences between the Paired CAD Control and Shared CAD Control were tested via a two sample t-test. The SCC configuration was found to have on average higher quality scores (0.86 ± 0.087) compared to PCC (0.71 ± 0.12). Results showed a statistically significant different in the Overall Quality Score achieved by teams in the two different collaboration modes ($t(9) = 3.2$, $p < 0.01$).

These findings agree with previous findings of pair programming research, which suggests that higher quality is achieved in paired configuration, as argued that "four eyeballs are better than two" [9]. In the SCC configuration, the designers share one control and thus there is no opportunity for parallelization of work. This means that both partners are actively focused on the model at hand, and we expect this to allow for additional quality control opportunities. This active focus by both participants allows constant feedback to be shared between the two, leading to higher chances of reiterating and improving design.

5.2 Metric Categories of quality

Next, overall quality was decomposed to its elements in order to see if the trend of difference in quality exists for each metric category. In particular we test for differences in the quality dimensions of: complete (CMP), concise (CNC), and consistent (CNS). The results of this analysis are outlined in Figure 4B and the mean results are outlined in Table 4 with the associated statistics.

TABLE 4: QUALITY SCORES BY METRIC CATEGORIES AND TEST OF DIFFERENCE

Metric	Collaboration Type	Mean	Standard Deviation	p value
Complete	PCC	0.74	0.093	0.049*
	SCC	0.84	0.12	
Concise	PCC	0.76	0.24	0.069
	SCC	0.93	0.12	
Consistent	PCC	0.58	0.32	0.038*
	SCC	0.85	0.19	

* significance < 0.05

Those pairs working in the SCC configuration had on-average higher quality in each of the metric categories of interest. The metrics complete and consistent were found to have a statistically significant difference between the two modes of collaboration at the 5% level. A complete CAD model was one that follows all dimensional and design criteria to replicate the rendering provided, while a consistent CAD model is one that is fully constrained and dimensioned with no new parts. Un-dimensioned sketches in the experimental CAD platform (Onshape) are indicated with a difference in color. This is a strong indicator of error, and again we might expect from the error-checking literature [30] that the extra attention from the non-parallelized worker in the SCC mode resulted in better quality, and also provided error-checking when replicating the rendering in the CAD model.

The fourth metric category of CAD quality, validity (occurrence of failed model instances), was also investigated. Participants were instructed not to add new parts to their models, and thus addition of new parts is considered poor quality.

PCC pairs were likely on average (2.10 ± 2.60) to have 10.5 times more failed instances in their model tree compared with SCC pairs (0.20 ± 0.42). A two sample t-test revealed that the difference between these results was statistically significant: $t(9) = 2.3$, $p < 0.05$. This may again be attributed to the decrease in monitoring in PCC compared to SCC as failed instances appear

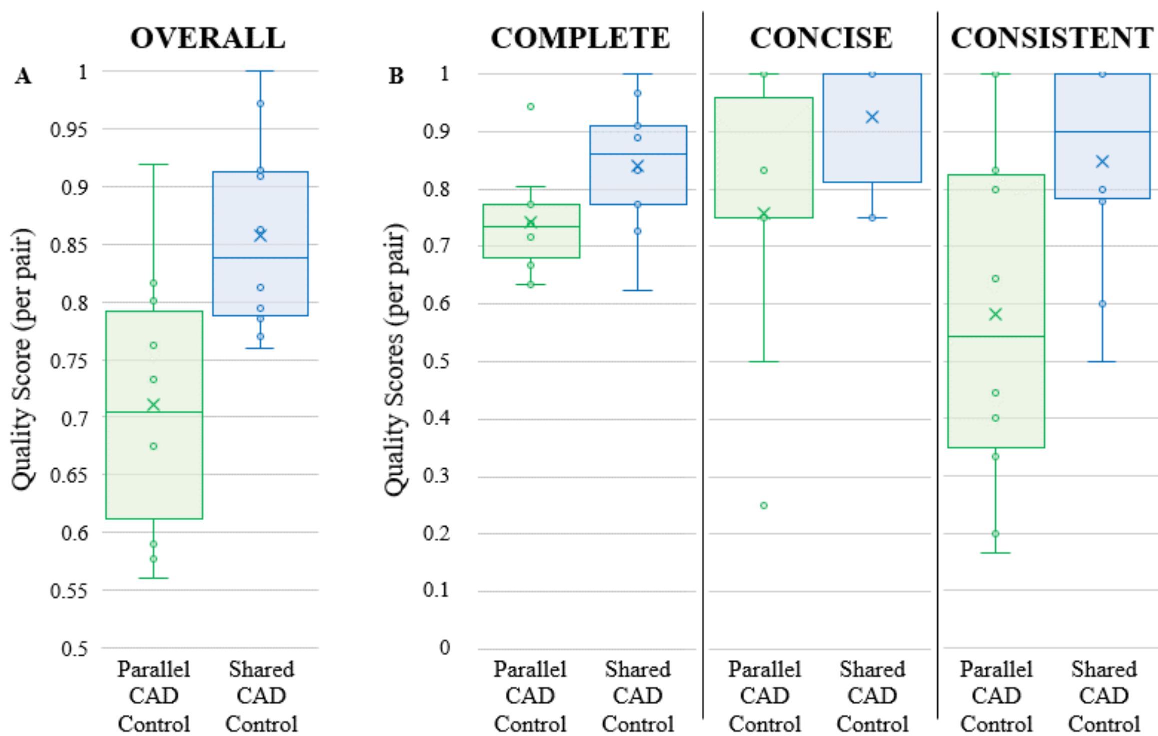

FIGURE 4: BOX-AND-WHISKER PLOTS OF **A)** OVERALL QUALITY SCORE FOR PCC (N = 10) AND SCC (N = 10), AND **B)** QUALITY SCORES FOR PCC AND SCC DIVIDED INTO THREE OF THE QUALITY METRIC CATEGORIES: COMPLETE, CONCISE, CONSISTENT. GREATER IS BETTER WITH 1 BEING MAXIMUM. SECOND AND THIRD QUARTILES ARE SHOWN.

as errors in the software platform and can be detected more readily by two individuals with focused attention.

No SCC pairs disregarded the instruction to refrain from creating new parts, whereas PCC pairs created on average 0.20 new parts in the model tree. This difference could be attributed to PCC pairs focusing less on the explicit instructions and more on the modeling, while in SCC pairs only one partner can be actively modeling at a time, and thus the other partner may pay more attention to task and experiment instructions.

5.3 Quality by task number

This section will analyze the quality scores of each design task for PCC and SCC pairs to identify whether there are particular elements of the CAD work required by a particular design task that may lead to the difference in observed quality by mode. The task quality score was created by finding the mean of each task per individual and finding the mean for each collaboration type. The results are listed in Table 5.

TABLE 5: QUALITY SCORES BY TASK NUMBER

Task Number	PCC n	PCC Mean	SCC n	SCC Mean	p value
1	10	0.64	10	0.73	0.37
2	10	0.80	10	0.90	0.29
3	10	0.80	9	0.93	0.07
4	10	0.90	9	0.96	0.11
5	9	0.73	7	0.80	0.54
6	6	0.42	3	0.89	0.03*
7	5	0.53	2	0.92	0.07
8	3	0.33			
9	2	0.67			
10	1	0.4			

* significance < 0.05

Task six resulted in the most significant difference in quality which is of high interest as the task was considered to have the most features to be implemented. The design task was to create a pen holder and incorporated the use of extrude, draft, proper sketching techniques, proper filleting techniques, and dimensional specifications to follow. SCC could have a larger quality score here because of the increase in these number of features, as reiteration and improved design choices are more readily applied when more CAD features are provided.

5.4 Speed analysis

In addition to the scored quality of the experimental models, we further examine the speed at which the participants achieved the models. Speed was calculated as "time per task" where a low time per task indicates a high speed. Speed is analogous to person-hours of work; time per task is calculated as the sum of time spent on a task by both partners; in the SCC case, the time per task is doubled to indicate that two people are working on any given task.

The experimentally derived relationship between time per task and quality is illustrated in Figure 5. The results showed that PCC has a weak, positive trend between time per task and quality (therefore a negative trend between speed and quality) while SCC has a very weak trend in the opposite direction. The fact that PCC involves more reiteration might result in this negative correlation. The SCC positive correlation is unexpected and requires further research to explain.

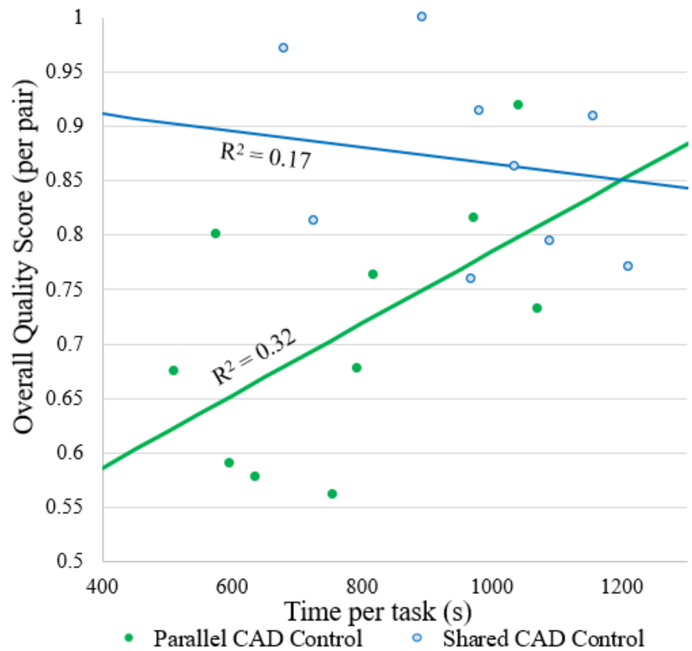

FIGURE 5: OVERALL QUALITY SCORES WITH RESPECT TO TIME PER TASK (SLOWER TEAMS TO THE RIGHT ON THE X-AXIS) FOR PCC AND SCC, WITH THEIR TRENDLINES

The low R-squared values resulting from a linear analysis of the data are typical of complex experimental analysis. Nevertheless this data can be interpreted as motivation to further explore the relationship between quality, speed and collaboration style in a larger data.

6. CONCLUSIONS AND FUTURE WORK

This study revealed evidence that the two collaboration methods discussed – PCC and SCC – lead to differing levels of quality in CAD models. This study reports on these different quality measurements and expands into the specific sub-dimensions that constitute quality as a whole. It was further shown that at a statistically significant amount:

- Overall, SCC pairs create higher quality designs than PCC
- SCC pairs more frequently fully dimension, constrain, and correctly replicate their CAD models than PCC pairs
- SCC pairs, when faced with a high number of features for one task, create higher quality models than PCC

In some aspects, the study confirms the trend observed in other contexts, that having a second pair of eyes and constant feedback on work prevents errors from being introduced in the design process.

As future work, it would be worthwhile to normalize the quality scores with different factors that could have had an effect – substantial or not. These include normalizing based on experience, in which case scores from the Baseline I rubric will be essential to gauge the "real" experience of the participants. Another approach of interest would be to break down design tasks to a fundamental level that allows the study investigators to create new design tasks with similar difficulty and thus normalize scores based on each design task.

Tasks presented in this study were a mixture of design tasks and prescribed modeling, since an approach that created pure design tasks would have made the analysis of different pairs challenging – this is because of the introduction of different approaches to creative design thinking and the difficulty to quantify these in a repeatable rubric. Next steps would be to find an optimal mix to favour design choices to a degree that still allows for study investigators to employ a fair rubric.

Lastly, further investigation of the relationship, or lack thereof, between the quality scores and speed would allow for a more thorough analysis of this fundamental design trade-off in PCC and SCC collaborations.

ACKNOWLEDGEMENTS
This study would not be possible without the support of Ready Lab members, especially Meaghan Vella and Jennifer Wu, for helping us troubleshoot the experimental procedure and conducting experiments. This research was generously supported by the Dean's Spark Fellowship, and the MITACS Globalink Research Award.

REFERENCES
[1] K. T. Ulrich and S. D. Eppinger, *Product Design and Development*, 5th ed. New York, New York, USA: McGraw Hill, 2011.
[2] "60 Years of CAD: These milestones shape the success story of Computer Aided Design since 1957," *CADENAS*, 2018. .
[3] S. S.-F. Smith, "An Evaluation of Internet-Based CAD Collaboration Tools," *J. Technol. Stud.*, p. 7, 2004.
[4] T. Marion and S. Fixson, "The Influence of Collaborative Information Technology Tool Usage on NPD," *Proc. Des. Soc. Int. Conf. Eng. Des.*, vol. 1, no. 1, pp. 219–228, 2019.
[5] D. Wu, J. Terpenny, and D. Schaefer, "A Survey of Cloud-Based Design and Engineering Analysis Software Tools," in *Proceedings of the ASME 2016 International Design Engineering Technical Conferences & Computers and Information in Engineering Conference IDETC/CIE 2016*, 2016, pp. 1–9.
[6] D. Zissis, D. Lekkas, P. Azariadis, P. Papanikos, and E. Xidias, "Collaborative CAD/CAE as a cloud service," *Int. J. Syst. Sci. Oper. Logist.*, vol. 4, no. 4, pp. 339–355, 2017.
[7] W. F. Bronsvoort and F. W. Jansen, "Feature modelling and conversion-Key concepts to concurrent engineering," *Comput. Ind.*, vol. 21, no. 1, pp. 61–86, 1993.
[8] C. H. Loch and C. Terwiesch, "Communication and Uncertainty in Concurrent Engineering," *Manage. Sci.*, vol. 44, no. 8, pp. 1032–1048, Aug. 1998.
[9] L. Williams, R. R. Kessler, W. Cunningham, and R. Jeffries, "Strengthening the case for pair programming," *IEEE Softw.*, vol. 17, no. 4, pp. 19–25, 2000.
[10] D. Stotts, L. Williams, N. Nagappan, P. Baheti, D. Jen, and A. Jackson, "Virtual teaming: Experiments and experiences with distributed pair programming," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2753, pp. 129–141, 2003.
[11] W. Shen, Q. Hao, and W. Li, "Computer supported collaborative design: Retrospective and perspective," *Comput. Ind.*, vol. 59, no. 9, pp. 855–862, 2008.
[12] F. He and S. Han, "A method and tool for human-human interaction and instant collaboration in CSCW-based CAD," *Comput. Ind.*, vol. 57, no. 8–9, pp. 740–751, 2006.
[13] K. Eves, J. Salmon, J. Olsen, and F. Fagergren, "A comparative analysis of computer-aided design team performance with collaboration software," *Comput. Aided. Des. Appl.*, vol. 4360, pp. 1–12, 2018.
[14] R. K. Purvanova, "Face-to-face versus virtual teams: What have we really learned?," *Psychol. J.*, vol. 17, no. 1, pp. 2–29, 2014.
[15] L. L. Gilson, M. T. Maynard, N. C. Jones Young, M. Vartiainen, and M. Hakonen, "Virtual Teams Research: 10 Years, 10 Themes, and 10 Opportunities," *J. Manage.*, vol. 41, no. 5, pp. 1313–1337, 2015.
[16] M. A. Cohen, J. Eliashberg, and T. H. Ho, "New product development: The performance and time-to-market tradeoff," *Manage. Sci.*, vol. 42, no. 2, pp. 173–186, 1996.
[17] L. O. Morgan, R. M. Morgan, and W. L. Moore, "Quality and Time-to-Market Trade-offs when There Are Multiple Product Generations," *Manuf. Serv. Oper. Manag.*, vol. 3, no. 2, pp. 89–104, 2001.
[18] P. Cankurtaran, F. Langerak, and A. Griffin, "Consequences of new product development speed: A meta-analysis," *J. Prod. Innov. Manag.*, vol. 30, no. 3, pp. 465–486, 2013.
[19] L. J. Skitka, K. L. Mosier, M. Burdick, and B. Rosenblatt, "Automation Bias and Errors: Are Crews

Better Than Individuals?," *Int. J. Aviat. Psychol.*, vol. 10, no. 1, pp. 85–97, 2000.
[20] V. Balijepally, R. Mahapatra, S. Nerur, and K. H. Price, "Are Two Heads Better than One for Software Development? The Productivity Paradox of Pair Programming," *MIS Q.*, vol. 33, no. 1, pp. 91–118, 2009.
[21] T. Dybå, E. Arisholm, D. I. K. Sjøberg, J. E. Hannay, and F. Shull, "Are Two Heads Better than One? On the Effectiveness of Pair Programming," *IEEE Softw.*, vol. November/D, pp. 12–15, 2007.
[22] C. Moreno, "Developing CAD Standards: A Complete Guide," *Autodesk University*, 2017. [Online]. Available: https://medium.com/autodesk-university/developing-cad-standards-a-complete-guide-8e6933ad17ac. [Accessed: 23-Feb-2020].
[23] Onshape Learning Paths, "Introduction to CAD," *Onshape*, 2020. [Online]. Available: https://learn.onshape.com/learning-paths/introduction-to-cad. [Accessed: 23-Feb-2020].
[24] C. González-Lluch, P. Company, M. Contero, J. D. Camba, and R. Plumed, "A survey on 3D CAD model quality assurance and testing tools," *CAD Comput. Aided Des.*, vol. 83, pp. 64–79, 2017.
[25] P. Company, M. Contero, J. Otey, and R. Plumed, "Approach for developing coordinated rubrics to convey quality criteria in MCAD training," *CAD Comput. Aided Des.*, vol. 63, pp. 101–117, 2015.
[26] J. D. Camba, M. Contero, and P. Company, "Parametric CAD modeling: An analysis of strategies for design reusability," *CAD Comput. Aided Des.*, vol. 74, pp. 18–31, 2016.
[27] J. Zhou, V. Phadnis, and A. Olechowski, "Analysis of Designer Emotions in Collaborative and Traditional Computer-Aided Design," in *ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conferences*, 2019.
[28] S. Hsi, M. C. Linn, and J. E. Bell, "Role of spatial reasoning in engineering and the design of spatial instruction," *J. Eng. Educ.*, vol. 86, no. 2, pp. 151–158, 1997.
[29] M. L. McHugh, "Interrater reliability: The kappa statistic," *Biochem. Medica*, 2012.
[30] L. J. Skitka, K. L. Mosier, M. Burdick, and B. Rosenblatt, "Automation bias and errors: Are crews better than individuals?," *Int. J. Aviat. Psychol.*, vol. 10, no. 1, pp. 85–97, 2000.

APPENDICES

Appendix A

An example of a design task provided to participants in EXP is shown.

Renderings:

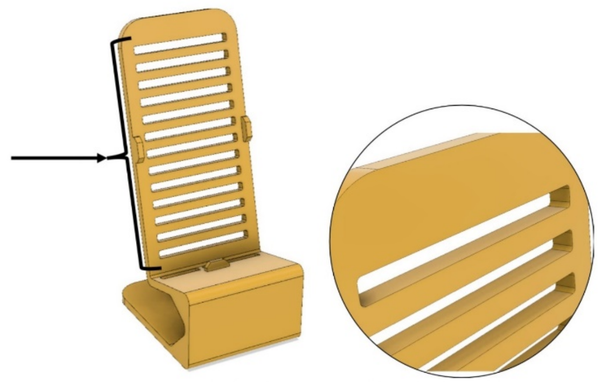

Figure 8. Phone holder with slots for air circulation (left), Closeup (right)

Phone drawing:

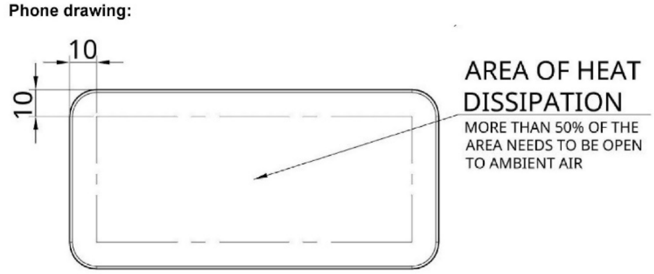

Figure 9. Area of heat dissipation on the back panel of the phone [in mm]

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

Appendix B

An attempt by an SCC and PCC pair to do the Appendix A design task is shown. PCC's filleting was not correctly applied as demonstrated in the rendering shown to pairs. Hence, PCC had a score of 0 for the rubric condition of "Fillets applied to all holes" as opposed to 1 for SCC.