Virtual Collaboration Training for Freshman Undergraduate STEM Students

Eric Nersesian, Jessica Ross-Nersesian, Adam Spryszynski, Michael J. Lee
Department of Informatics
New Jersey Institute of Technology
Newark, NJ USA
{eric.nersesian, jessica.c.ross, as2569, mjlee}@njit.edu

***Abstract*—Higher educational institutions formalize socialization for their incoming undergraduate student populations with traditional forms of physical classroom-based learning community (LC) skill-building environments; however, recent studies have shown that virtual LC environments can offer improved results over physical LC environments. This study examines whether incoming undergraduate science, technology, engineering, and math (STEM) students gain the same benefits to their academic performance regardless of whether they receive LC training in physical or virtual reality (VR) treatment. We found that either treatment of collaboration training improve the participants' academic performance in comparison to the control treatment. In addition, we found that the VR participants gave more academic help in social settings to their peers throughout the semester than their control group counterparts. Upon interviewing the two treatment group participants, we found that virtualization of collaboration may impact perceptions on leadership roles, group functions, and thinking about the future. This research shows that virtualizing LCs has the potential to expand and supplement existing learning structures, and create new ones where they were not previously available, and aims to offer a better understanding of the strengths and limitations of introducing VR technologies in higher education.**

***Index Terms*—Learning Communities, STEM Education, Educational Technology, Computer Aided Instruction, Virtual Reality**

INTRODUCTION

A learning community (LC) is a group of people who share common academic goals and attitudes, who meet semi-regularly to collaborate on classwork [1]. These groups are documented to provide a significant positive impact on participating students' academic performance [2]. Higher educational institutions, such as our university, the New Jersey Institute of Technology (NJIT), formalize LC skill training for their incoming undergraduate student population with regulated fall semester classes called Freshman Year Seminars (FYS). FYS traditionally takes the form of physical (i.e, in-person), classroom-based LC skill-building environments; however, recent studies have shown that virtual LC environments can offer improved results over physical LC environments [3].

Research has shown that LCs are beneficial to the overall educational experience of participating students [2]. For example, a student satisfaction survey from the previous academic year conducted on FYS students found that 94.33% of LC students stayed in their originally declared degree at the end of their freshman year as opposed to 90.41% for non-LC students. However, a regular percentage of incoming freshman have scheduling conflicts that prevent them from attending physically-based LCs. Recently, our university had 700 freshmen attend LC FYS and *215 attend non-LC FYS*, meaning that 23.49% of incoming freshmen need a viable alternative to physically-based LC education.

Although there is an immediate need to provide alternatives for students, institutions often hesitate to adopt new technology solutions [4], often because they cannot afford or cannot justify spending on something that does not have an immediate recognizable impact [5]. While new technologies are available that provide communication tools for groups, the lack of information available on the implementation of virtual spaces for education hampers institutions from integrating them.

Virtualizing LCs has the potential to expand and supplement existing learning structures, and create new ones where not previously available [6]. This research will lead to a better understanding of the strengths and limitations of introducing virtual reality (VR) technologies in higher education. Presenting institutions with well-researched data about the benefits of virtualization technologies will ease the reservations that administrators and other stakeholders might have.

The cost-effectiveness of those technologies further reduces obstacles that prevent VR applications being integrated as a regular part of LCs [7]. Because of this, we propose to build a LC curriculum around the hardware, Oculus Rift, and the software, Facebook Spaces, due to their low-cost and high educational market potential. What is most necessary at this stage of adoption is to examine what roles these technologies can take in creating collaborative environments for the educational process. This research will focus on examining what factors virtualization brings into play when adapting the existing toolset into an educational environment to service LCs.

RELATED WORKS

In today's educational environment, increasing demands require educators to use new techniques to improve the quality and frequency of learning experiences [8]. LCs are one well-researched approach to using structured socialization environments to maximize the potential for informal learning opportunities [1]. Higher educational institutions such as our university have formalized LC skill training for the incoming undergraduate student population with FYS.

Socialization factors such as collaboration, networking, and organizational thinking are of particular interest to FYS LC curricula, due to their critical influence on students' ability to form LCs independently, especially for STEM students that statistically spend less time honing these socialization skill sets [9]. FYS traditionally takes the form of physical classroom-based LC skill-building environments; however, recent studies have shown that virtual LC environments may produce similar results compared to physical LC environments [10].

Instructor-driven meetings in a physical location are not the only implementation of LCs. Peer mentoring is an alternative to traditional first-year seminars that relies on student mentors in less formal settings. Such programs have been successfully implemented in STEM programs, improving student outcomes, and benefiting both mentees and mentors [11]. Online LCs are also useful in reducing the negative impact of limited communication of students in online classes. Guidelines proposed by Yuan suggest using both synchronous and asynchronous technologies to overcome issues with presence and schedule, which are not limited to online students [10]. Virtual reality and related technologies offer possible solutions to issues to many common issues in education settings [12]. There is little literature that focuses specifically on LCs in this domain.

LCs vary depending on specific desired outcomes and as such, there does not exist a single, definitive set of guidelines for implementing such communities. The literature does, however, provide best practices that can be used by educators working in this domain. Theoretical frameworks should inform learning communities and similar programs [13], [14]. Implementing appropriate scaffolds and curricular structures are also considered a good practice that provides a solid foundation for a learning community [15].

The benefits of LCs and other programs targeting first-year students extend beyond academic improvements. Improved engagement with peers and instructors has a positive impact on overall engagement [16]. In more recent literature, Settle demonstrated that learning participating in LCs showed increased retention rates [17] and improved on feelings of isolation [18] in minority student cohorts.

By leveraging the strengths of physical collaboration and socialization, LCs integrate a multifaceted approach to learning acquisition. In contrast, VR offers a novel medium for experiences that may be difficult to bring into a generalized educational setting (especially due to costs of equipment and training) [19]. To determine if these benefits can be combined, this research will perform comparative analysis across a control and variable group. In comparing non-LC FYS classes to those integrated with virtual LCs and physical LCs, we will analyze how VR adoption within traditional LCs impacts overall academic performance. The differences between VR-only and physical-only LCs will provide insight into how VR experiences compare to physical LCs and whether (or which parts of) VR can be considered as a suitable alternative for physical LCs.

VR devices continue to become cheaper and offer immersive interfaces that can change the way we provide instruction, engage learners, and teach new skills [20]. These cheaper instructional tools have the potential to reach a wider and broader audience, including minorities that may not have access to other types of instruction [21]. However, much is still unknown about how to do these things effectively. Our work informs the future direction of VR-related collaboration educational technologies. This research will provide a foundation for understanding the uses of virtual environments in LC education by actively developing, iterating, and improving our VR collaboration exercise content to cover more topics within the LC curriculum.

Method

Research Aim

The aim of our study was to examine if VR instruction and collaboration has the same benefits as in-person LC training in the classroom. In addition, we wanted to provide a proof-of-concept to demonstrate that such training can occur in VR with minimal monetary cost. To address these aims, we decided to create new curriculum to test between two treatment conditions (in-person physical interactions vs. VR interactions using the exact same curriculum), and a control condition (the current LR instruction at the university using their original curriculum). Specifically, we compared the following outcomes: academic performance, and social connectedness.

Participants

Each fall semester, NJIT offers 55 LC class sections and 5 non-LC class sections to freshman undergraduate students, with approximately 25-30 students self-enrolling into each section. We visited most class sections during the first week to explain our study and recruit participants from NJIT's College of Engineering and College of Computing (representing 75% of the total student body).

Our study had three groups: control, physical treatment, and virtual treatment (see Table I). The control group used the LC course's original curriculum. The physical treatment group hosted LC education in a physical environment and worked in randomized teams each week. The virtual treatment group hosted LC education in a virtual environment and worked in randomized teams each week.

To minimize instructor effects, we randomized the participants in each group. Each treatment group had 8 teams of 3-4 participants each, for a total of 62 participants. The control group had no teams as they only had to fill out assessments for the study, and included a total of 28 participants. We compensated each participant a minimum of $150 for the duration of the study, requiring them them to participate in

Group	Subjects	Collaboration	Assessments
Control	28	None	Academic/Social/Survey
Treatment 1	29	Physical	Academic/Social/Survey/Interview
Treatment 2	33	Virtual Reality	Academic/Social/Survey/Interview

TABLE I
TREATMENT GROUP DETAILS

the study for 2 hours each Wednesday, for 10 consecutive Wednesdays (equating to $15/hour, which is 107% more than the federal minimum wage of $7.25, and 74.4% more than our state's minimum wage of $8.60).

Implementation

Implementation began with an exploration and design phase for the VR curriculum modules to take the place of physical collaborative assignments and culminate in the introduction of those modules into the non-LC curriculum as a replacement for the missing LC curriculum. We created a corpus of unique curricula spanning several socialization and collaboration topics oriented toward learning community skill sets. The collaboration curriculum was aligned so both the physical and virtual form would be equivalent. Both curricula was structured/limited by the features available within the Facebook Spaces software, which included discussions, drawing, dice, playing cards, 3D model viewing, photo viewing around a round table. Facebook spaces allows a maximum of four participants in a session, so we limited all of our groups to 3-4 participants each. We also found that this group size was optimal as it was difficult to establish steady collaboration within smaller groups (fewer than 3) or larger groups (more than 4).

For the VR treatment condition, we used Facebook Spaces as the software platform. During the initial stages of the study, VR treatment participants came for assistance and training to interact within the virtual space using the physical controls and software.

Collaboration Curriculum

The study ran concurrent to the FYS classes, which were active for the first ten weeks of the fall semester. The two treatment groups, physical and VR, went through eight weeks of the collaboration curriculum (see Table II), which included the same activity between the two conditions for one hour each week. We outline these activities below:

Lost at Sea: The goal of this activity was to introduce participants to the type of tasks they would be doing throughout the study. We asked each participant to rank a list of items given a survival situation. The participants then discussed among the entire group to establish a group rating. In most cases, the group rating score was higher than individuals' ratings, demonstating that group work often yields better results. The measures of success for this activity included: every member actively participated, participants understood that working in groups can yield better results, and participants responded generally positively to the activity.

Academic Discussions: The goal of this activity was to encourage participants to share relevant experiences in the academic domain. Participants answered a list of questions regarding their study habits. The designated organizer led a discussion based on a script of questions, including "What is your favorite class so far?," "What is the class you think will require most work?," and "Do you think study groups are effective for classes?" The measures of success for this activity included: apparent comfort of the group, attitude towards the activity, and willingness to collaborate beyond the activity.

Ship of Theseus: The goal of this activity was to facilitate discussion of an abstract problem. The participants discuss a thought experiment questioning an identity of an object, called "the Ship of Theseus" thought experiment. Over time, the ship has every part replaced with a new part, and the old parts are used to build another ship. The participants are asked to discuss within their group which is the real ship. The measures of success for this activity included: participants come to some conclusions even if they do not reach a consensus.

Is a hot dog a sandwich?: The goal of this activity was to facilitate discussion of an abstract problem. The participants discuss a thought experiment questioning an identify of an object. They go through a list of foods that are questionable whether they should be classified as a sandwich, and the group must come to a discussion on each object. The measures of success for this activity included: participants come to some conclusions even if they do not reach a consensus.

Problem Solving Activities: The goal of this activity was to facilitate solving a series of logic puzzles with defined solutions with focus on role and task division. The logic puzzles were titled "Missionaries and Cannibals" and "Rocket Ship". They can often be solved with trial and error by keeping track of individual steps and backtracking them. The measures of success for this activity included: participants solved all problems in the time allotted.

Moon Landing: The goal of this activity was to recognize progress already made as a group. This activity came from the same source as Lost at Sea and had the same structure to the activity. Since first week served as an icebreaker and the content of the activity served largely secondary roles, this reiterated the general trend for groups to outperform individual scores. The measures of success for this activity included: the group recognized that working collaboratively had better overall results, and that they recognized they made progress since the first meeting.

Pictionary: The goal of this activity was to serve as a morale boost during periods of increased stress and anxiety affecting all members of the group. In our case, midterm exams. The researcher adjusted rules to competitive or collaborative depending on group preferences and perception of flow. The measures of success for this activity included: every member participated, and group morale remains high.

Time	Physical Treatment	Virtual Reality Treatment
Week 1	Lost At Sea	Intro to Virtual Reality
Week 2	Academic Discussions	Lost At Sea
Week 3	Ship of Theseus	Academic Discussions
Week 4	Is hotdog a sandwich?	Ship of Theseus
Week 5	Missionaries and Cannibals	Is hotdog a sandwich?
Week 6	Rocket Ship	Missionaries and Cannibals
Week 7	Moon Landing	Rocket Ship
Week 8	Pictionary	Moon Landing

TABLE II
WEEKLY COLLABORATION CURRICULUM BREAKDOWN

Academic & Social Assessments

The participants filled out provided academic and social assessments in designated forms and times during the semester. Data collection points occurred in person on a one-on-one basis starting at the beginning of the semester, and ending at the end of the semester. The participants were required to fill out the social assessment every week during the semester, and an academic metric a week before all exams and final grades were due. Additionally, we interviewed both treatment group participants at the end of the semester.

The academic metric was an form asking the participant to list any classes they had an exam or grade in within the next week, and what grade they believed they would receive from F to A, which was converted to 1-5 (only integers). The participants had on average five classes and each class had between 3-4 exams along with a final class grade, resulting in a minimum of 4 grade data points.

We found that there were no established methods for measuring students' connectedness in LC settings. Therefore, we adapted our social metric from work in Social Network Analysis (SNA), a tool for describing the underlying mechanisms for social dynamic theory. Our resulting social metric was a form asking participants to list out fellow students that they had academically-related social interactions with over the past week. These "academically-related social interaction" were defined as at least five minutes of socialization, either in-person or remotely, in any location outside of the classroom where the participant was given help or gave help in regards to a class activity (e.g., an assignment). The form had room to list out five classes (the average number of classes for a freshman) and up to 12 fellow students per class with directionality (i.e., half indicated that they *provided* help to someone else, and half indicated that they *received* help from someone else).

Interviews

At the end of the semester, we conducted 1-hour semi-structured interviews with participants from both the treatment groups. We interviewed 26 participants from the VR treatment group and 19 from the physical treatment group for a total of 45 interviews. Two researcher conducted all the interviews, one for all physical treatment participants, and the other for all VR group participants. We asked a total of 23 questions intended to better understand how the different treatments affected participants' opinions about collaboration and LCs. A sample of our questions include:

1) Which activity did you like the most?
2) Was there an activity you really disliked?
3) Do the activities help students communicate better?
4) Are you more likely to ask students for help?
5) Are you more likely to socialize with other students?

We used the three-stage coding process described by Cambell et al., for the measurement of intercoder reliability for semi-structured interviews [22]. The interview lead read through all interviews and generated a list of themes. A second researcher reviewed 10% of the transcripts and reached a minimum of 87% intercoder reliability after two trials (86% than 89%). Once this was complete, the lead interviewer coded the remaining transcripts and counted code occurrences.

After consolidation, we identified 52 codes grouped into 10 themes. We coded the occurrences of each code as a binary output, no(0)/yes(1). As the resulting data are nominal/categorical, we report on the counts of the different labels (e.g., histograms); shown as percentages of 'yes' responses per group.

Codes grouped into their themes:

1) Factors for choosing NJIT:
Family, Reputation, Finance, Location, Academics
2) Current experience at NJIT:
Academic, People, Independence, Campus
3) Motivations behind collaboration:
Sharing ideas, Social, Solving problems, Shared workload, Helping
4) Significance behind collaboration:
Social skills, Mindsets, Jobs, Comfort
5) General activities participants liked or disliked:
Sports, Games, Social, Academic, Difficulty, Public speaking
6) Perception of roles in collaborative activities:
Assigned, Natural, Leader, Follower, Multiple, Power, Ability, Interest
7) Mentors and their roles on collaboration:
Guidance, Per request, Supervision, Contamination
8) Collaboration activites participants liked:
Problem solving, Novelty, Competitive, Fun, Word-based, Different views
9) Purpose of LCs:
Guidance, Comfort, Getting to know people, Building relationships, Future
10) Effects of collaboration activities:
Friends, Being heard, Easier to work, Communicating, Differences

RESULTS

Academic Differences between Treatments and Control

We gathered the academic assessments that were filled out throughout the semester and converted letter grades of F through A to a scale of 1-5. We averaged all academic assessments per participant to get a final academic metric between 1-5, which was analyzed as ordinal variables using the non-parametric Mann-Whitney test (as our data was not normally distributed). We found no significance in comparing the treatment groups ($W(N=62) = 954.0$, $p = 0.573$), meaning there was no observable difference between the physical and VR treatment group participants' self-reported semester grades. However, we did find significant difference between the physical treatment and control ($W(N=57) = 993.0$, $p = 0.016$) and between the VR treatment and control ($W(N=61) = 726.5$, $p = 0.041$), with both the physical and VR treatment groups having a significantly higher self-reported semester grades (control median = 3.792, physical median = 4.0, VR median = 4.0) (see Figure 1).

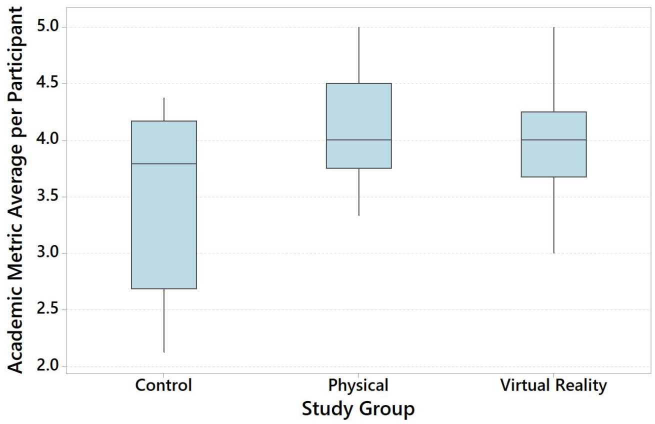

Fig. 1. Boxplot of Academic Metric Average per Participant by Study Group

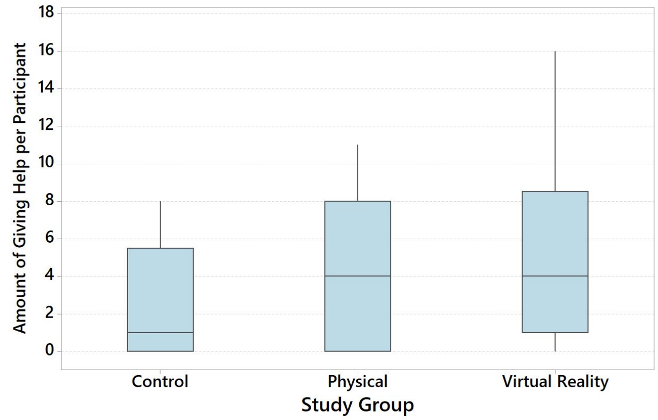

Fig. 2. Boxplot of the Social Metric "Giving Help" by Study Group

Social Assessment Difference Between VR and Control

We compared the social metrics of the three study groups by examining the total academic social interactions per semester of each participant and the directional sums of the their social interactions (giving or receiving help). We used Chi-Squared tests as the data was not normally distributed. We found a significant difference when comparing VR to the control group for giving help ($\chi^2(2, N = 57) = 5.6531, z = 2.376, p = 0.0175$), with the VR participants giving significantly more help than their control counterparts (see Figure 2).

Interview Codifications of Physical and VR Groups

For our *mentor roles on collaboration* theme, the *guidance* code appeared in 88% of the VR interviews and 32% of the physical interviews. This means the VR participants mentioned the idea of guidance 2.75 times more than the physical participants when asked about mentor roles on collaboration (see Figure 3).

For our *perception of collaboration roles* theme, the *natural* and *ability* codes appeared in 79% and 42% of VR interviews and 63% and 21% of physical interviews. This means the VR participants mentioned the ideas of natural and ability 1.25 and 2 times more than the physical participants when asked about perception of roles in collaboration. The *follower* code appeared in 58% of physical interviews and 29% of VR interviews, which means the physical participants mentioned being a follower 2 times more than the VR participants when asked about their collaboration roles (see Figure 3).

For our *motivations behind collaboration* theme, the *shared workload* and *helping* codes appeared in 58% and 38% of the VR interviews and 37% and 26% of the physical interviews. This means the VR participants mentioned the idea of shared workload 1.57 and helping 1.46 times more than the physical participants when asked about the motivation behind collaboration. The *sharing ideas* code appeared in 63% of the physical interviews and 42% of the VR interviews, which means the physical participants mentioned being a follower 1.5 times more than the VR participants when asked about their collaboration roles (see Figure 4).

For our *enjoyment factors of collaboration* theme, the *fun* and *different views* codes appeared in 68% and 32% of the physical interviews and 54% and 17% of the VR interviews. This means the physical participants mentioned the idea of fun 1.26 and different views 1.88 times more than the VR participants when asked about the enjoyment factors of collaboration. The *problem solving* code appeared in 67% of the VR interviews and 47% of the physical interviews, which means the VR participants mentioned solving problems 1.43 times more than the physical participants when asked about their enjoyment factors of collaboration (see Figure 4).

For the *significance of collaboration* theme, the *jobs* code appeared in 83% of the VR interviews and 42% of the physical interviews. This means the VR participants mentioned the idea of jobs 1.98 times more than the physical participants when asked about the significance of collaboration (see Figure 5).

For our *Purpose of Learning Communities (LCs)* theme, the *future* code appeared in 25% of the VR interviews and 0% of the physical interviews. This means the VR participants mentioned the idea of the future 25 times more than the physical participants when asked about the purpose of LCs. The *comfort* code appeared in 58% of physical interviews and 33% of VR interviews, which means the physical participants mentioned being comfortable 1.76 times more than the VR participants when asked about the purpose of LCs (see Figure 5).

Discussion

Interview Codification Interpretation

For the *mentor roles on collaboration* theme, the *guidance* code was 275% higher in the VR treatment participant interviews. This could be a result of the large amounts of technical help and training that was needed in the first few weeks to get the VR participants acclimated to the hardware.

For the *perception of participant roles in collaboration* theme, the *natural* code was 125% higher and the *ability* code was 200% higher in the VR treatment participant interviews. The *guidance* code from the previous theme may be an influential

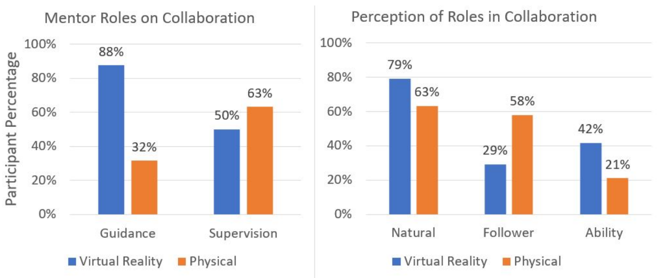

Fig. 3. Mentor Roles and Perception of Roles on Collaboration

factor since the additional guidance during the early stages of the study may have felt as a natural ability formation for the VR groups. This effect may also be seen in a 200% higher follower code for the participants in the physical treatment group, which could also be interpreted as 50% lower follower code for the participants in the VR treatment group.

For the motivations behind the *collaboration* theme, the *shared workload code* was 157% higher in the VR treatment participant interviews, while the *sharing of ideas cod*e was 150% higher in the physical treatment participant interviews. This contrast in identification of collaboration motivation between the treatment is reinforced with the additional code of *helping* being 146% higher in the VR treatment participant interviews. Since participants in the VR treatment viewed collaboration motivation involving a shared work experience, we believe they were less shy about asking a fellow group member to do a task for them. This interpretation is reinforced with the difference in the helping code between treatment. Meanwhile, the physical treatment participants had real interactions building appreciation for personal qualities such as what others were thinking and their ideas.

For the *enjoyment factors of collaboration* theme, the *problem solving* code was 143% higher in the VR treatment participants interviews, while the physical treatment participants interviews had 1.26% and 188% higher mentions of the *fun* and *different views* codes. This preference for problem solving and different views reinforces the interpretation of the different presence of the shared ideas and shared workloads codes.

For the *significance of collaboration* theme, the *jobs* code was 198% higher in the VR treatment participants interviews. This may be because VR is an emerging technology, and the introduction to VR application was set in a fictional world. These may have caused the participants to start thinking about their futures since many were engineering students. This interpretation is reinforced by the fact that the *future* code was 2500% higher in the VR treatment participants interviews from the purpose of LCs theme. We also see that the comfort code is 176% higher in the physical treatment participants interviews for this theme. This would indicate that the physical treatment participants got more comfortable with collaboration which reinforcing the interpretation from the shared ideas, and different views codes.

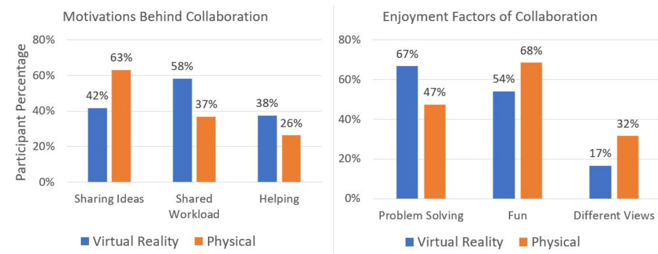

Fig. 4. Motivations and Enjoyment Factors of Collaboration

Academic & Social Assessment Interpretation

Analysis of the academic metrics revealed a statistical significance of the self-reported semester grades for the two treatment groups compared to the control group. We believe that having any type of collaboration training has direct positive impact to the incoming undergraduate STEM students' academic performance. We believe that since both the physical and VR treatment groups had statistical academic improvements over the control group that we have shown virtualization of collaboration training for FYS LCs is a viable alternative to the traditional, physical treatment.

Analysis of the social metrics revealed a statistical significance between the VR treatment and control for the amount of academic help offered by the participant to fellow students during the semester. We believe this result is explained by the interview results from the motivations behind the collaboration theme. The *shared workload* code was 157% higher and the helping code was 150% higher in the VR participant interviews. We believe the VR participants learned more about offering help and sharing work, and that was reflected in their giving help social metric.

Lessons Learned

VR Impact on Socialization Abilities: The VR treatment participants reported that VR can address the anxiety of freshman relating to their new college environment full of strangers that may carry risk of public humiliation. The participants indicated that socializing in VR helped remove the fear from public interactions with unknown fellow students. They reported that over the several weeks of the study, they felt more comfortable during the rest of their week when they were out on the campus and interacting with other students.

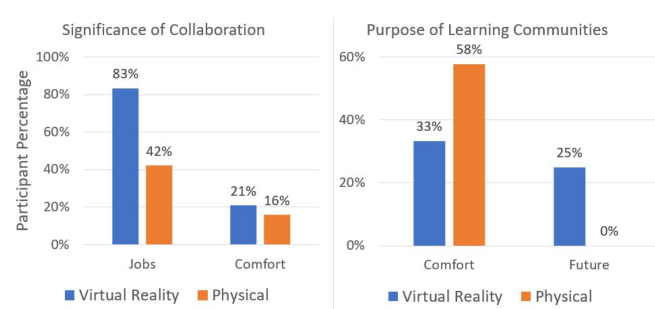

Fig. 5. Significance and Purpose of Collaboration

Preference of Cooperative vs Competitive Activities: We focused mostly on collaborative problem solving but games worked well and were usually successful in groups with past poor performance. Most groups direct their interactions around cooperation, or at least a mixture of cooperative and competitive with the competitive interactions blending into a cooperation nature, and vice versa. One group would only compete with each other and another group had very low interactions in general, but leaned toward cooperation interactions.

Open-ended vs Concrete Solutions: The participants seemed split on open-ended vs closed solution problems. Some participants pointed out that problems with defined solutions are more satisfying to solve, with the solution serving as a kind of reward. Participants that preferred the Ship of Theseus and "Is a hot dog a sandwich?" type problems, listed a variety of causes, including:

1) Open-ended problems are more challenging
2) They liked the opportunity to discuss and debate
3) These kinds of problems are different
4) Longer problems have depth to talk about

Additionally, some of the participants are inclined to lead and moderate and this could affect their preferences.

Session Breakdown: We found it useful to observe the group as they entered the room, and trying to start with activities that are at the group's emotional state and allow them to warm up. At the beginning, we stuck mostly to a single activity per session. Starting the fourth week, we slowly introduced other activities, like card games, and the week 5 activity was three different problems presented as a single activity. Some participants liked the variety of having multiple activities, others preferred to focus on one. We will continue to explore this area in future work.

Math Anxiety Breaks Collaboration: We noticed that the counting task involving math could stop collaboration as participants were afraid of making a mistake and being embarrassed in front of their group. In the future, we will avoid activities that breaks the flow of collaboration and discussion as it simply defeats the purpose of our intervention.

Three Types of Group Mentalities

Talkers: For groups preferring discussion activities, their interactions peaked around conversations. Conversations were not all collaborative, they flow between collaboration and competitive natures with the main dialogue themes exploring personal opinions on a subject. After personal explorations, they can easily come to group decisions. Usually, these groups were formed around one personality with the others feeling comfortable as a secondary personality with followup comments or confirmation or reject of direction that the group leader is going. Usually, these roles were define based on the communication abilities of each member, with the last group member having the lowest socialization ability listening, and doing something with their hands, ie drawing or playing with models, but still verbally confirming their approval of the group's direction. We am not sure if these groups only form when there is this single dominant personality.

They enjoyed all collaboration activities with personal preferences ranging on the types of activities. They usually do not need to be warmed up, or coerced into interacting with each other. They usually were very excited to start up their session and bonded quickly as a group. They came in together chatting and leaving together. They were quick with the assessments, and any logistics of the study, since they help each other. That made us think of using another assessment of checking their times to coming into the study, the speed of filling of the assessments, speed of entering into VR, and speed of finishing up activities.

Doers: For groups preferring physical activities, their interactions peaked around games, puzzles, activities that can be completed successfully with minimal conversations and more interactions involving physical manipulations with the hands. Tactile stimulation and manipulation was preferred for shared experiences. We were not sure if its because the group did not have a persona to lead the conversation situations and they preferred these more silent group activities. We were not sure if this is a hierarchy situation based on the social skill level of participants, and that all participants matured through these levels of preferred group dynamics and social interactions. For this study's purpose, we were trying to get the participants to be as comfortable conversationally as possible since the end goal is to maximize the chances that out in the real world of campus and class interactions they would ask for help when needed and offer help when capable to do so.

This style of group dynamic seemed to prefer logic puzzle activities when forced to do a conversational activity. This seemed to be because they can rely on a tactile interaction point with the group, i.e., drawing, to supplement their conversations when interacting with each other to solve the puzzle. These groups seemed to prefer collaborative tactile exercises like Pictionary and puzzle-solving as that seemed to avoid confrontations. They seemed to want to positively interact with the group but lacked the confidence or experiences to do so in full conversational situations.

Not Ripe Yet: Only two groups were not ready for any structured type of interactions. They needed unstructured playtime to pull out group interactions, and exploratory environments with opportunities for immature jokes. The simplest games of hangman and blackjack were most effective. Their collaboration fell apart under conversations and goal-oriented activities. Although both groups solved the missionaries puzzle, which was surprising, they did poorly compared to the other groups. The logic puzzle gave them something interesting to solve where they did not need to perform in front of the rest of the group. Pictionary did not work well because they were focused on making fun of each other's drawing skills which gave them performance anxiety for both the drawer and the guesser.

They needed activities with minimal structure. We gave them 3D models to play around with to get used to the presence of the other participants. We followed that up with some type of activity that gives a small amount of structure and performance but not enough to evoke sarcasm from the group and elicit performance anxiety. After that, we usually followed with a mildly structured activity when we needed to actively promote positive collaboration experiences while minimizing the potential for public humiliation. Of the two groups with this mentality, both were all males, with one all Caucasian and the other all African-American. Most girls in the study were more socially advanced, more collaboration and conversation focused, yet did seem shier and tended to not be the primary personality of conversation groups. We think that had mostly to do with the fact that they are an underrepresented minority.

Study Limitations

We were limited in developing the collaborative activities because we wanted to keep everything consistent across conditions. The software used in the VR treatment had limited tools for user interactions, so we built our activities around these limitations.

Future Work

Future studies based on this work could focus on combining collaboration research with learning science research, particularly for undergraduate computer science education where group-based problem solving and communication skills are essential for students upon entering the job market.

CONCLUSION

This study examined whether incoming undergraduate STEM students gain the same benefits to their academic performance regardless of whether they receive LC training in physical or VR treatment. We found that either treatment of collaboration training improve the participants' academic performance in comparison to the control treatment. In addition, we found that the VR participants gave more academic help in social settings to their peers throughout the semester than their control group counterparts. Upon interviewing the two treatment group participants, we found that virtualization of collaboration may impact perceptions on leadership roles, group functions, and thinking about the future.

ACKNOWLEDGMENT

We thank Oculus Education and Facebook for their support of this work. Any opinions, findings, conclusions, or recommendations are those of the authors and may not reflect the views of these other parties.

REFERENCES

[1] J. Bell, J. Falk, R. Hughes, G. Hunt, J. Parrish, M. Ruffin, K. Sacco, and G. Troxel, "Informal stem education: resources for outreach, engagement and broader impacts," *Science Education (CAISE)*, 2016.

[2] M. J. Graham, J. Frederick, A. Byars-Winston, A.-B. Hunter, and J. Handelsman, "Increasing persistence of college students in stem," *Science*, vol. 341, no. 6153, pp. 1455–1456, 2013.

[3] M. Jou and J. Wang, "Investigation of effects of virtual reality environments on learning performance of technical skills," *Computers in Human Behavior*, vol. 29, no. 2, pp. 433–438, 2013.

[4] M. Stohlmann, T. J. Moore, and G. H. Roehrig, "Considerations for teaching integrated stem education," *Journal of Pre-College Engineering Education Research (J-PEER)*, vol. 2, no. 1, p. 4, 2012.

[5] L. C. Duque, "A framework for analysing higher education performance: students' satisfaction, perceived learning outcomes, and dropout intentions," *Total quality management & business excellence*, vol. 25, no. 1-2, pp. 1–21, 2014.

[6] C. J. Boyce, C. Mishra, K. L. Halverson, and A. K. Thomas, "Getting students outside: Using technology as a way to stimulate engagement," *Journal of Science Education and Technology*, vol. 23, no. 6, pp. 815–826, 2014.

[7] M. Bower, C. Howe, N. McCredie, A. Robinson, and D. Grover, "Augmented reality in education–cases, places and potentials," *Educational Media International*, vol. 51, no. 1, pp. 1–15, 2014.

[8] S. Olson and D. G. Riordan, "Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. report to the president." *Executive Office of the President*, 2012.

[9] S. S. Carrino and W. J. Gerace, "Why stem learning communities work: The development of psychosocial learning factors through social interaction." *Learning Communities: Research & Practice*, vol. 4, no. 1, p. 3, 2016.

[10] J. Yuan and C. Kim, "Guidelines for facilitating the development of learning communities in online courses," *Journal of Computer Assisted Learning*, vol. 30, no. 3, pp. 220–232, 2014.

[11] A. M. Zaniewski and D. Reinholz, "Increasing stem success: a near-peer mentoring program in the physical sciences," *International Journal of STEM Education*, vol. 3, no. 1, pp. 1–12, 2016.

[12] L. Freina and M. Ott, "A literature review on immersive virtual reality in education: state of the art and perspectives," in *The International Scientific Conference eLearning and Software for Education*, vol. 1, no. 133, 2015, pp. 10–1007.

[13] R. Hall and Z. Jaugietis, "Developing peer mentoring through evaluation," *Innovative Higher Education*, vol. 36, no. 1, pp. 41–52, 2011.

[14] C.-M. Zhao and G. D. Kuh, "Adding value: Learning communities and student engagement," *Research in higher education*, vol. 45, no. 2, pp. 115–138, 2004.

[15] V. Tinto, "Learning better together," *Transitioning Students in Higher Education: Philosophy, Pedagogy and Practice*, p. 2, 2019.

[16] K. Krause, "Understanding and promoting student engagement in university learning communities," *Paper presented as keynote address: Engaged, Inert or Otherwise Occupied*, pp. 21–22, 2005.

[17] A. Settle and T. Steinbach, "Retention rates for the first three years of a linked-courses learning community," in *Proceedings of the 19th Annual SIG Conference on Information Technology Education*, 2018, pp. 166–171.

[18] A. Settle, J. Doyle, and T. Steinbach, "Participating in a computer science linked-courses learning community reduces isolation," *arXiv preprint arXiv:1704.07898*, 2017.

[19] D. Parmar, J. Isaac, S. V. Babu, N. D'Souza, A. E. Leonard, S. Jörg, K. Gundersen, and S. B. Daily, "Programming moves: Design and evaluation of applying embodied interaction in virtual environments to enhance computational thinking in middle school students," in *2016 IEEE Virtual Reality (VR)*. IEEE, 2016, pp. 131–140.

[20] M. C. Johnson-Glenberg and C. Megowan-Romanowicz, "Embodied science and mixed reality: How gesture and motion capture affect physics education," *Cognitive Research: Principles and Implications*, vol. 2, no. 1, p. 24, 2017.

[21] G. L. Cohen, J. Garcia, N. Apfel, and A. Master, "Reducing the racial achievement gap: A social-psychological intervention," *science*, vol. 313, no. 5791, pp. 1307–1310, 2006.

[22] J. L. Campbell, C. Quincy, J. Osserman, and O. K. Pedersen, "Coding in-depth semistructured interviews: Problems of unitization and intercoder reliability and agreement," *Sociological Methods & Research*, vol. 42, no. 3, pp. 294–320, 2013.