Impacts of Augmented Reality on Collaborative Physics Learning, Leadership, and Knowledge Imbalance

Iulian Radu, Harvard University Graduate School of Education, iulian_radu@gse.harvard.edu
Bertrand Schneider, Harvard University Graduate School of Education, bertrand_schneider@g.harvard.edu

Abstract: Emerging technologies such as Augmented Reality (AR), have the potential to radically transform science education by making challenging concepts visible and accessible to novices. In this project, we designed a Hololens-based system in which participants learned about the physics involved in audio speakers (Radu & Schneider, 2019). We analyzed participant dyad where educational AR representations were present or not, focusing on the relationships between collaboration, learning, leadership, and knowledge imbalance. We find that, overall, AR representations improved time management, learning of structural concepts but reduced learning of physical concepts. The effects of leadership were mediated by the presence of AR: the presence of leaders in AR was linked to higher learning gains and better collaboration, whereas these effects were not present without AR. Finally, for groups with imbalanced knowledge, AR seemed to benefit participants in configurations where less-knowledgeable participants took the lead in discussions. These results indicate that AR can be beneficial for equalizing the effects of imbalanced collaboration. We discuss the implications of those results for the design of CSCL learning activities using augmented reality.

Introduction

In this research we investigate the benefits and drawbacks of augmented reality for inquiry-based learning, specifically in relation to participant agency and leadership. We focus on a collaborative activity that allows pairs of students to explore concepts in electromagnetism. Electromagnetism is a topic that is often encountered in both maker spaces and traditional physics classrooms, and it is one of the most difficult topic to master because it combines understanding of physical objects (ex: magnets, wires) and abstract concepts (ex: magnetic field shapes, voltage, electricity) (Belcher and Bessette 2001; Ibáñez et al. 2014; Maloney et al. 2001).

We focus on how AR technology in this educational context relates to leadership and individual agency. A prevailing issue in classroom group collaborations is the effect of unequal participation and unequal knowledge among group members. Some people are naturally more dominant while others more passive, and this may be enhanced by the amount of each person's domain expertise (Salomon & Globerson, 1989). This imbalance can create negative effects whereby the more passive or less-knowledgeable students do not contribute as much as their peers, potentially leading to less effective problem solving, collaboration breakdowns, and "free rider" effects. Augmented reality environments have the potential to mediate the effects of participant leadership and knowledge imbalances. On one hand, AR environments can provide a plethora of holographic visualizations, which can be beneficial for participants by allowing passive or lower-knowledge participants to easily follow the communications of their peers; they allow participants to more easily ask for clarification or interrupt by pointing at representations; they allow more knowledgeable participants to explain by use of the available representations; and they can allow more equitable interaction with the experience, thus providing agency to participants who may naturally be more passive. On the other hand, AR environments require the use of specialized hardware, including the use of head-mounted devices which may cover the participants' eyes, thus making communication difficult by masking nonverbal cues such as facial expressions.

In this study we investigate these topics by studying dyads using head-mounted Microsoft Hololens AR devices, in conditions where educational AR representations are either present or not. We predict that the presence of AR information will balance collaboration because participants have shared access to information, thus one person will be less likely to dominate the experience; additionally, we expect the availability of shared representations to potentially increase the participation of more passive group members. For this reason, we predict that educational AR visualizations will have positive effects on collaborative processes. Finally, we observe how groups behave when more active participants (drivers) have high or low knowledge, and expect that groups with more knowledge imbalance will be negatively impacted in the condition where groups do not have access to educational AR information because a "free rider" effect (Salomon & Globerson, 1989).

Related work

Augmented reality systems have been developed for a wide range of educational uses, such as for learning geometry (Radu et al. 2015), chemistry (Yu-Chien 2006), and history (Chang et al. 2015). Specifically for physics education, systems have been built for visualizing electricity (Belcher and Bessette 2001) and magnetism (Maloney et al. 2001). Augmented reality systems can support student learning of spatial structures, improve performance on physical tasks, and have lasting effects on participant memory (for a comparative media review see Radu, 2014). While the effects of AR experiences on learning have been studied, there is relatively less research focused on understanding the impact of AR on the dynamics of co-located collaboration. Previous research projects have focused on designing AR infrastructures that support social interactions (Billinghurst, Clark, and Lee 2015), for example by allowing a remote expert to inhabit a physical space, allowing multiple users to annotate physical environments, and allowing users to have different views of the same environment. Research has also found that student groups using augmented reality can benefit from increased collaboration because they are more motivated to engage with the experience (Phon, Ali, & Halim, 2014) and because the experiences simulate real-world professional collaborations (Dunleavy & Dede, 2014). Furthermore, it has been argued that AR experiences can balance leadership behaviors of participants, since group members can have access to shared visualizations and one person is less likely to control the group resources (Morrison et al., 2009).

We contribute to this research agenda by specifically studying how collaboration aspects of leadership and knowledge imbalance are impacted by the presence of educational AR, presented as holograms on physical artifacts. A prevailing issue in group collaborations is the effect of unequal participation among group members. Unequal participation is caused by multiple factors including unequal knowledge relevant to the activity, unequal ability to control the activity, or unequal personal interest and initiative. In situations where team members do not (or cannot) contribute equally, this typically results in lower learning gains (Chen 2006). Previous research indicates that when a resource is limited among team members, this encourages one person to dominate the interaction and creates imbalanced participation (Church, Hazlewood & Rogers 2006). In such contexts, participants may simply follow along, leading to decreased learning gains and poor collaboration (Salomon and Globerson 1989). Shared interfaces such as tangible objects have the potential to balance participation as each user has access to the learning content (Church, Hazlewood & Rogers 2006), especially when such interfaces allow participants to have shared control and awareness of information available to the group (Yuill and Rogers 2012).

Augmented reality system and research questions

In this section, we describe how prior work informed the design of our system, the study we conducted, our research questions and the methodology we used to answer them.

We have designed a tangible model of a sound-producing speaker, augmented with interactive educational holograms visible to two users through the Microsoft Hololens augmented reality headset. This educational system mixes physical content (physical wires, magnets, sounds, magnetic forces, movement) and virtual content (visualizations of electricity, magnetic fields, force directions). The participants can interact with the system by playing music or constant electric current in different directions; modify amplification; switch between different electromagnet coils; and slide the vibrating membrane. They also can use a compass to measure magnetic fields manually. The system is shown in Figure 1.

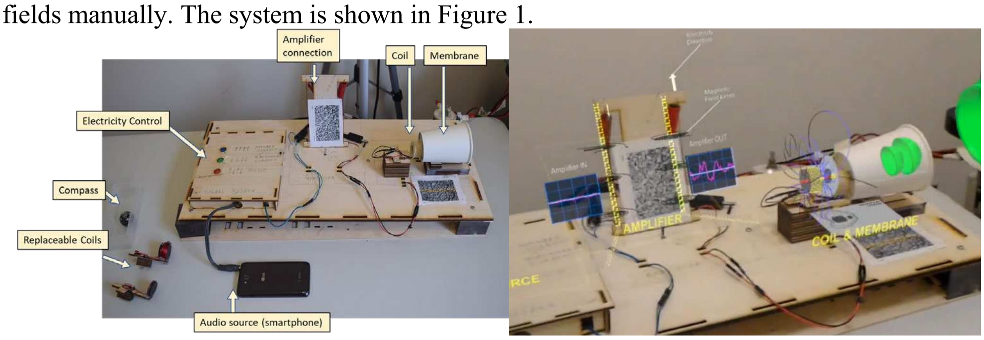

Figure 1. View of the system without educational AR representations (left) with components labeled; view of the system with educational AR showing magnetic fields, electricity and related representations (right – all holograms and labels are shown as visible in AR).

In (Radu & Schneider, 2019), we have investigated how participants' learning is influenced by augmented reality representations at the individual level (results are summarized in Fig. 3). Participants were randomly assigned to four experimental conditions which comprised a nested design with two factors (primary factor: presence of AR educational representations, sub-factor: type of technology features). All conditions had access to the physical system and were able interact by changing electricity, generate sound, move the speaker membrane, change the speaker coils, and change amplification; furthermore, all conditions had access to a physical poster that explained electromagnetism concepts, had labels showing the function of pieces of the physical system. The groups which did not have educational AR representations (noEdAR) were split into 2 subgroups – dyads which did not wear the Hololens AR devices, and dyads who did but only saw minimal AR representations composed of holographic labels of major system components, and holograms of sound waves which represented the sound when the system played music. The participants who saw educational AR representations (EdAR) had access to the same information, but could also see interactive visualizations of magnetic fields, electron flows, and electricity graphs. These EdAR groups were split into 2 subgroups in which the presentation of AR visualizations was either presented all at once, or sequenced by a timer. All conditions performed the same study activities.

Our analysis of individual learning gains indicates that participants who used AR technology had significantly higher learning gains on specific concepts such as understanding spatial structures (ex: shapes of magnetic fields through questions, i.e. *"Draw the magnetic field around a single wire"*), higher ability to transfer knowledge to different situations (ex: transfer questions, i.e. *"Is it possible to build a motor that is moved through electric signals? If yes, explain how."*), but were significantly worse at understanding physical effects (ex: questions about relationship between physical movement and magnetic fields, i.e. *"If the magnetic field is suddenly inverted, the speaker membrane: a) is pulled closer; b) is pushed away; c) does not move"*).

Figure 2. A dyad pair interacting with the system while seeing AR through Microsoft Hololens headsets.

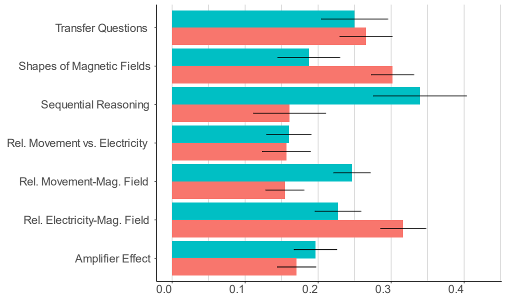

Figure 3. Individual differences in relative learning gains (Green: without educational AR; Red: with educational AR; whiskers: standard error).

In the current research we focus on the effects of AR on collaboration and participant leadership and agency, and analyze the data aggregated at the group level. We, compare learning and collaborative behaviors between two conditions: participants who used the system and saw educational AR visualizations; and participants who used the system but did not see educational visualizations. We predict that, due to the increased availability of educational visualization, participants who see educational AR will exhibit less affected by imbalanced leadership and imbalanced knowledge in the group. Our research questions are as follows:

RQ1: *Is overall collaboration and learning impacted by the presence of educational AR?*

RQ2a: *How does participant leadership imbalance impact learning, collaboration, and interaction?*

RQ2b: *How does leadership differ with the presence of AR?*

RQ3a: *How does leadership imbalance impact learning, collaboration, and interaction?*

RQ3b: *How does the effect of driver-follower imbalance differ with the presence of AR?*

Method

Participants were recruited from the study pool of a laboratory at a university in the northeastern United States. Participation required subjects to not know each other, have no significant prior physics knowledge, be born on/after 1976, speak English fluently, have at least a bachelor's degree, and wear no bifocal glasses. All participants first completed a pre-test, followed by a short written introduction to relevant physics concepts. Participants then worked on the speaker activity for 30 minutes under different experimental conditions (see Figure 3). During this period, all participants worked on a worksheet and saw a poster of printed physics knowledge on the wall. The study ended with a post-test and debriefing. The variables of interest are as follows:

Independent variables

We looked at three different independent variables in our analyses: presence / absence of AR, leadership and knowledge imbalance. We provide our operationalization of those constructs below:

Presence of Educational AR: Groups were randomly assigned to conditions in which educational AR was present (EdAR) or not present (NoEdAR), as described above.

Leadership: Groups were categorized under two conditions depending on participant leadership: dyads where a leader was present, and dyads where no leader was present. Leadership was considered to be present when both partners in a group do not initiate actions equally (i.e. when the maximum score was not recorded for the qualitatively-observed "reciprocal interaction" dimension, defined in Table 2 below).

Driver Follower Knowledge: We rated the participation of each person in a group as either being a "driver" (the participant who typically initiated actions overall) or "follower" (the participant who was more passive overall). We accounted for differences pretest knowledge, resulting in four group configurations: LL, LH, HL, HH where the first letter refers to the driver and the second letter refers to the passenger. For example, HL indicates a high-knowledge driver, and low-knowledge passenger; LH indicates low-knowledge driver and high-knowledge follower.

Dependent variables

Learning Metrics: We measured participant learning through pre- and post-tests. Participants' learning was compared using relative learning gains, a measure of the relative improvement that occurred between pre-post test scores (Cuendet et. al. 2012). The learning test contained multiple-choice and open-ended questions measuring several aspects of conceptual knowledge. All learning metrics are listed in Figure 3 and described in detail in (Radu & Schneider, 2019).

Collaboration Metrics: Collaboration metrics were qualitatively coded for each pair of participants across several dimensions using a validated rating scheme described in Meier, Spada & Rummel (2007). The scale evaluates collaboratives processes through a 5-point scale on the following dimensions: sustained mutual understanding, dialogue management, information pooling, reaching consensus, task division, technical coordination, and reciprocal interaction. Examples are provided in Table 1.

Interaction Metrics: The backend of the tangible interface system recorded how much users interacted with different system components, such as changes in electricity direction, movement of speaker membrane, changes of speaker coils, and changes in amplification.

Table 1: The measured dimensions of collaboration, based on Meier, Spada & Rummel (2007)

Collaboration Metrics	**Example**
Sustained mutual understanding	Ensure partners are on same page. Speakers make their contributions understandable for their collaboration partner rather than ignoring each other's insight.
Dialogue management	Smooth flow of communication with little overlaps. Smooth volley of conversation with little interruptions; partners make sure to have each other's' attention before transitioning to other topics
Information pooling	Ask questions to seek each other's perspective. Partners contribute their insight effectively or ask useful questions to seek opinions.
Reaching consensus	Coming to shared understanding / agreement. Both partners come to a shared conclusion; if there is disagreement, they resolve it through critical consideration
Task division	Task division is balanced, and tasks are explicitly distributed between partners through discussion.
Time management	Deadline met, detailed time planning. Partners monitor the time throughout their cooperation and make sure to finish the current subtask or topic with enough time to complete the remaining subtasks.
Technical coordination	All tools used, including physical compass. Partners use all technology at their disposal, including features such as physical magnetic compass, using different coils, referring to the physical poster, etc.
Reciprocal interaction	Partners hold equal status in working relationship; both take agency in leading the discussion instead of one partner dominating the working relationship.
Overall Scores	Overall Communication (avg. dimensions 1,2), Overall Joint Information Processing (avg. dimensions 3,4), Overall Coordination (avg. dimensions 5,6,7), Overall Collaboration (avg. dimensions 1-8)

Results

In the original study, we recruited 30 participant pairs in the two primary conditions (N=60) and removed sessions where technical issues were encountered or sessions that contained outliers with exemplary pretest knowledge (2.5 standard deviations beyond the pretest mean score), resulting in N=56 sessions (N=28 in EdAR, N=28 in noEdAR).

For qualitative measurement of the degrees of collaboration and leadership, two researchers double-coded 20% of the video recordings and achieved interrater reliability of Cohen Kappa 0.67, indicating "good" agreement. For statistical tests, normality assumptions were not met for continuous dependent variables, thus we performed Wilcoxon Rank-Sum nonparametric tests for differences between conditions of Presence of Educational AR, and for conditions of Leadership; and we performed Kruskal-Wallis H nonparametric tests followed by post-hoc tests for differences between conditions of Driver-Follower Knowledge.

RQ1: Are overall collaboration and learning impacted by the presence of educational AR?

We tested for collaboration and learning differences between EdAR and noEdAR conditions. The EdAR groups were significantly better at **time management** (W=446, p=0.050), and significantly better at answering **near transfer questions** (W=568, p=0.003), but significantly worse at understanding relationship between **magnetic fields and movement** (W=248.5, p=0.019). These learning effects mirror our results of analysis at the individual level presented in (Radu & Schneider, 2019), where we found these and additional effects detectable in the larger sample when participants are considered individually (Figure 3). We found no other statistically significant effects when analyzing differences at the dyad level.

RQ2a: How does participant leadership imbalance impact learning, collaboration, and interaction?

Dyads were categorized into two groups: dyads where a leader was present (N=22 overall; N=11 in EdAR, N=11 in noEdAR), and where no leader was present (N=34 overall; N=17 in EdAR, N=17 in noEdAR).

We analyzed whether there is a significant effect of leadership on dependent variables across all groups. We find that dyads with leadership generally had higher relative learning gains on **all transfer questions** (W=253.5, p=0.044); but had weaker collaboration in **sustained mutual understanding** (W=511, p=0.010), **reaching consensus** (W=584, p<0.001), **task division** (W=511, p=0.004), **overall joint information processing** (W=562, p=0.001), **overall communication** (W=492, p=0.033), and **overall collaboration** (W=584.5, p<0.001). This indicates that, although leadership causes some learning gains, it also produces visible negative effects on collaboration.

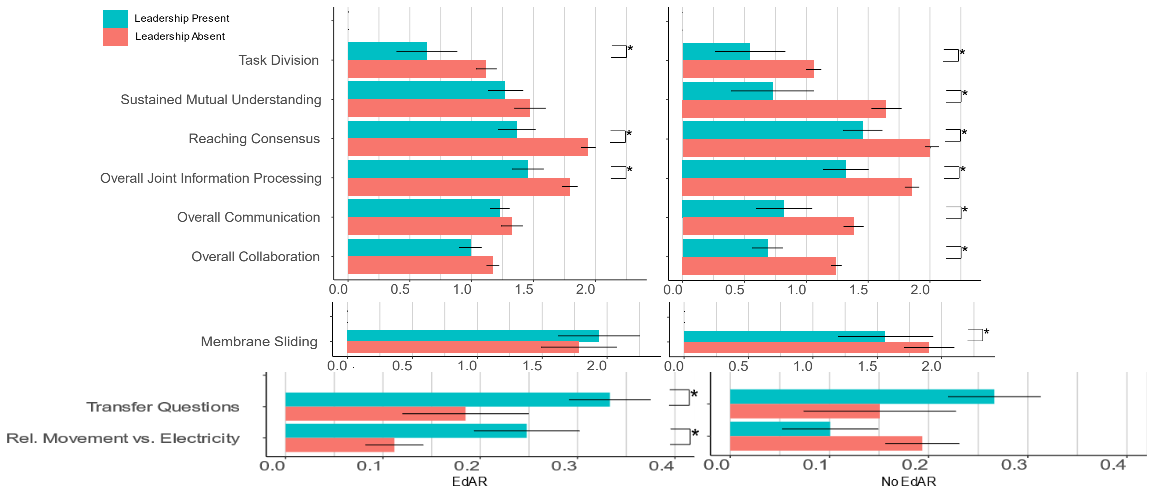

Figure 4. Significant effects of driver-follower configurations in EdAR (left) and noEdAR (right). *=sig. difference between leadership groups. Whiskers=standard error.

RQ2b: How does leadership differ with the presence of AR?

Since AR seems to have an influence on group leadership, we investigated the effects of leadership in the presence and absence of educational AR. Within EdAR groups, we found that groups with leadership had higher gains on **transfer questions** (W=49.5, p=0.040) and understanding the relationship between **electricity vs. movement** (W=47.5, p=0.031); also, groups with leadership had lower scores on **reaching consensus** (W=147.5, p=0.001), **task division** (W=126, p=0.047), and **overall joint information processing** (W=137, p=0.027), compared to groups without leadership. In contrast, in the noEdAR condition, the presence of leadership had no statistical impact on learning gains, and had worse impact on collaboration and system interaction. For groups in the noEdAR condition, similar to EdAR condition, groups with leadership were worse on **reaching consensus** (W=144.5, p=0.001), **task division** (W=130, p=0.040), **joint information processing** (W=144, p=0.009), but had additional negative effects through lower scores on **interaction** with speaker membrane (W=82, p=0.039),

sustained mutual understanding (W=140.5, p=0.016), **overall communication** (W=138, p=0.029), and **overall collaboration** (W=162, p=0.001). These results indicate that within the EdAR experience, leadership imbalance has a stronger effect of emphasizing learning gains, and has less negative effects on collaboration and system interaction.

To understand the differences between these conditions in relation to leadership effects, we sampled video recordings from EdAR and noEdAR groups. One theme observed is that in situations where EdAR was present, the visual representations helped the more passive participant follow the more dominant participant's explanations: it allowed the passive participants to interrupt and ask for clarification by referencing to the AR representation. Table 2 illustrates an example in EdAR (left) where one participant is teaching the other by using magnetic field polarity representations. In contrast, in the Non-AR example (right) one participant is trying to teach but ends up dominating the discussion while the other participant is unable to keep up with the explanations due to the lack of shared representations.

Table 2: Quotes from participants in moments of teaching

EdAR Group	**Non EdAR Group**
P1: "The direction is, like, the south pole, like N S N S is the backward one." P2: "Uh huh." P1: "Can you see that?" <pointing at AR> P2: "Yeah." P1: "Starting from the left, N S -" P2: "Yeah." P1: "-N S. And, for forward current, it's -" P2: "Oh." P1: "-S N N S." P2: "S N. Ok. So, this is...so backward was [picks up pen]" P1 [confirms, as P2 draws]: "Backward, N S !"	P2: "So alright here's an interesting thing, when you play music, that's making little currents going forward and backwards, but if no forward current is allowed, it does this <pushes button> this just overpowers everything, and only gonna pull, and when you switch it to forward, only pushes, and when you switch it to music it does little tiny currents like really fast, and then that's making it vibrate, and as it vibrates it's shaking this cup, and as that vibrates it's shaking the air, and that air is shaking our ears, and shaking our brains." P1: "Really cool" P2: "What do you think?" P1: "Good story" [nervous laughter] P2: "You like my story? Ok, so how does that help us get to the question"

RQ3a: How does group leadership imbalance impact learning, collaboration, and interaction?

Each group was categorized into one of the four driver-follower conditions. LL (EdAR N=9, noEdAR N=9), LH (EdAR N=6, noEdAR N =6), HL (EdAR N=9, noEdAR N=10), HH (EdAR N=4, noEdAR N=3). We analyzed the effects of different types of driver-follower configurations across all groups. No statistical differences were found for relative learning gains and interactions. There was a significant effect of driver-follower knowledge on the dimension of **overall coordination** (X^2=8.9, p=0.031), and **overall collaboration** (X^2=11.9, p=0.008); in all cases, descriptive LH groups scores higher than other groups. This indicates that overall, groups had better collaboration when novices were guiding the interaction.

RQ3b: How does the effect of driver-following knowledge differ with the presence of AR?

Since the AR medium may influence how driver-follower configurations behave, we investigated the effects of this variable when educational AR was present or absent. Within the EdAR groups, there were significant effects of driver-follower configurations on the dimensions of **dialogue management** (X^2=12.4, p=0.006), and **overall collaboration** (X^2=10.3, p=0.016); in all cases, descriptive statistics show that the LH groups scored higher than other groups. Within the No-EdAR groups, there were no significant effect of driver-follower knowledge distribution. These results suggest that AR educational representations are beneficial to participants in the LH configurations (Fig. 5):

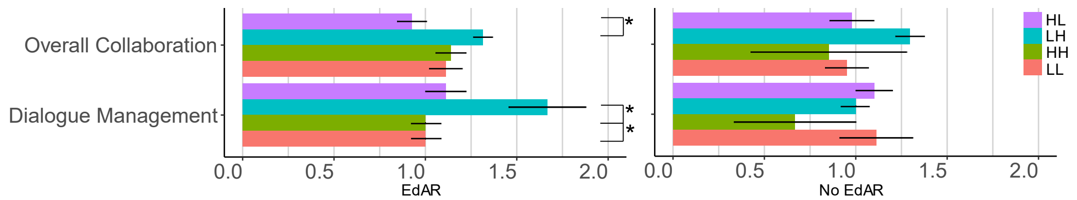

Figure 5. Significant effects of driver-follower configurations in EdAR (left) and noEdAR (right). *=sig. difference between driver-follower configuration conditions. Whiskers=standard error.

By qualitatively analyzing EdAR and noEdAR groups, we saw examples of participants with less knowledge taking initiative in guiding group discussion (Table 3). In groups where AR representations were present, participants with less knowledge could easily ask questions by pointing at hologram representations of the system, and participants with more knowledge had an easier time communicating their knowledge by referring to the visual representations. As an example, Table 3 illustrates how AR representations were useful for guiding participant discussion (left) and lack of such representations in Non-AR groups contributed to difficulties in understanding. In both examples, the participant with less knowledge is asking for clarification from the other participant; in the AR scenario the knowledge is provided by using the AR representations; in the Non-AR scenario, the person with more knowledge is unable to provide a clear explanation. These behaviors appeared to encourage communication within AR groups and may be a reason for increased benefits to dyads in the LH driver-follower configuration.

Table 3. Quotes from participants in the AR and non-AR conditions to clarify ambiguous interactions.

EdAR Group	**Non EdAR Group**
P2: "Is it because it's - is it hitting it? Is the -" P1: "I can see the vibration, or I can see how it's moving in this AR [moving hand side to side]" P2: "In the AR?" P1: "AR - the augmented reality part, like it's just - [gesturing around area surrounding cup and coiled wire with hand]" P2: "Yeah" P1: "The angle of the base is like [pointing to cup]" P2: "Oh! Oh! Oh! Yeah, yeah, yeah, yeah, yeah, yeah! I like the - this thing [moving finger back and forth in front of area between cup and coiled wire]. Okay, so the alternating currents is pulling the magnet closer or further - further" P1: "Mm-hm" P2: "Okay"	P1: "So, when it, when it's a backward current [illustrating shape of magnetic field using hands], it looks like it's pulling this way" P2 [pointing toward cup]: "On this, this thing, right?" P1: "Yeah, so I wonder - I wonder if it's almost like, when it's backward, it's like that [illustrating on paper with fingers]." P2 [confused, handing pen over]: "Are you good at drawing?" P1 [starting to draw on worksheet]: "Well, I would, um..."

Discussion

Our original hypotheses predicted that the presence of educational AR representations would have an effect on overall collaboration and learning. For collaboration, we found differences in time management, with lower values for groups which did not have educational AR representations and tended to run out of time. This indicates that the AR representations were useful for completing the task, and groups which lacked these had a harder time understanding the system and communicating since they had to generate their own representations (for example by drawing). For learning gains, analyses at dyad level did not show the same effect as the individual analysis presented in (Radu & Schneider, 2019). This is likely due to the effect of averaging the learning between a participant pair, which reduced the sample size and decreased our statistical power.

Augmented reality was also helpful for groups where participants had knowledge imbalance, specifically when the participants with high knowledge took a back seat and allowed the low-knowledge participants to drive the interaction. When this occurred, metrics of group collaboration was higher in AR groups. This indicates that AR can potentially improve collaboration in groups where participants have unequal knowledge. This might be due to the availability of information to both participants, which can equalize the effect of knowledge imbalance by allowing less-knowledgeable participants to more easily communicate points of confusion by referring to the existing AR representation. Referring to the present representations might also allow the more knowledgeable participants to teach his/her peer, which can help passive participants be more engaged in the collaborative activity.

These results point at potential future work investigating the detrimental effects of augmented reality on learning concepts focused on physicality (ex: relationship between movement and magnetic fields), that possibly may be caused by learners being hyper-focused by highly visual nature of the AR medium and paying less attention to physical sensations. On the other hand, these results indicate that AR can be beneficial for equalizing the effects of imbalanced collaboration, and future work could investigate how this can benefit situations where participants tend to be imbalanced in agency (ex: in student team projects where some participants are more high-achieving than others) or in knowledge (ex: in teams of varied expertise levels, where participants may need to take on the role of teachers).

Conclusions

In this study we analyzed collaboration, learning and interactions of dyad pairs as they experienced an AR system for learning about electromagnetism. We found that, in this context, augmented reality was generally beneficial for both learning and collaboration. Overall, AR representations improved time management, learning of structural concepts but reduced learning of physical concepts. The effects of leadership were mediated by the presence of AR: the presence of leaders in AR was linked to higher learning gains and better collaboration, whereas these effects were not present without AR. Finally, for groups with imbalanced knowledge, AR seemed to benefit participants in configurations where less-knowledgeable participants took the lead in discussions.

References

Belcher, J. W., & Bessette, R. M. (2001). MIT educators share success-Using 3D animation in teaching introductory electromagnetism. COMPUTER GRAPHICS-US, 35(1), 18–21.

Billinghurst, M., Clark, A., & Lee, G. (2015). A survey of augmented reality. Foundations and Trends® in Human–Computer Interaction, 8(2–3), 73–272.

Chang, Y.-L., Hou, H.-T., Pan, C.-Y., Sung, Y.-T., & Chang, K.-E. (2015). Apply an augmented reality in a mobile guidance to increase sense of place for heritage places. Journal of Educational Technology & Society, 18(2).

Chen, Y.-C. (2006). A study of comparing the use of augmented reality and physical models in chemistry education. In Proceedings of the 2006 ACM international conference on Virtual reality continuum and its applications (pp. 369–372). ACM.

Church, T., Hazlewood, W. R., & Rogers, Y. (2006). Around the table: Studies in co-located collaboration. Austrian Computer Society.

Cuendet, S., Bumbacher, E., & Dillenbourg, P. (2012). Tangible vs. virtual representations: when tangibles benefit the training of spatial skills. In Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design (pp. 99–108). ACM.

Dunleavy, M., & Dede, C. (2014). Augmented reality teaching and learning. In Handbook of research on educational communications and technology (pp. 735–745). Springer.

Ibáñez, M. B., Di Serio, Á., Villarán, D., & Kloos, C. D. (2014). Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. Computers & Education, 71, 1–13.

Maloney, D. P., O'Kuma, T. L., Hieggelke, C. J., & Van Heuvelen, A. (2001). Surveying students' conceptual knowledge of electricity and magnetism. American Journal of Physics, 69(S1), S12–S23.

Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. International Journal of Computer-Supported Collaborative Learning, 2, 63–86.

Morrison, A., Oulasvirta, A., Peltonen, P., Lemmela, S., Jacucci, G., Reitmayr, G., … Juustila, A. (2009). Like bees around the hive: a comparative study of a mobile augmented reality map. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1889–1898). ACM.

Phon, D. N. E., Ali, M. B., & Halim, N. D. A. (2014). Collaborative augmented reality in education: A review. In Teaching and Learning in Computing and Engineering (LaTiCE), 2014 International Conference on (pp. 78–83). IEEE.

Radu, I., Doherty, E., DiQuollo, K., McCarthy, B., & Tiu, M. (2015). Cyberchase shape quest: pushing geometry education boundaries with augmented reality. In Proceedings of the 14th International Conference on Interaction Design and Children (pp. 430–433). ACM.

Radu, I. & Schneider, B. (2019). What Can We Learn from Augmented Reality (AR)? Benefits and Drawbacks of AR for Inquiry-based Learning of Physics. In 2019 CHI Conference on Human Factors in Computing Systems Proceedings (CHI 2019), May 4–9, 2019, Glagsow, Scotland, UK. ACM.

Radu, I. (2014). Augmented reality in education: a meta-review and cross-media analysis. Personal and Ubiquitous Computing, 18(6), 1533–1543.

Salomon, G., & Globerson, T. (1989). When teams do not function the way they ought to. International Journal of Educational Research, 13(1), 89–99.

Yuill, N., & Rogers, Y. (2012). Mechanisms for collaboration: A design and evaluation framework for multi-user interfaces. ACM Transactions on Computer-Human Interaction (TOCHI), 19(1), 1.

Acknowledgements

This material is based on work supported by the National Science Foundation, Grant No. 1748093. We thank all our participants, reviewers, the Harvard Decision Science Lab, and assistants who helped with data collection.