The Role of Augmented Reality in Multi-Device Small Group Learning Ecosystems

Taehyun Kim, James Planey, and Robb Lindgren
thk4@illinois.edu, planey@illinois.edu, robblind@illinois.edu
University of Illinois at Urbana-Champaign

Abstract: Augmented reality (AR) has shown benefits for individual learning with single devices and platforms. However, there has been sparse research on the learning effects of more complex technology ecosystems where multiple devices allow for different interactions with the same simulation. This study describes one such multi-device ecosystem where undergraduate students in an astronomy course were presented with a small group problem-solving task where both AR headsets and tablet computers could be used to explore a night sky simulation. We investigated whether students learned the targeted astronomy concepts, and whether the ways the devices were used affected students' learning and their perceptions of the activity. Results indicate an overall learning gain from pre to post, and while participants who used AR did not score higher than students who did not use AR, there was an association between how each device was used and one of the key learning outcomes.

Introduction

Augmented reality (AR) describes a genre of technologies that blend the virtual world with the real world, often through dynamic overlays of "context sensitive virtual information" (Klopfer & Squire, 2008, p. 205). A number of devices have been used to achieve AR, including camera-based applications on mobile devices and immersive headsets. Studies of the efficacy of AR technologies in education have shown benefits on assessments of individual learning outcomes in a range of content areas (Chen et al., 2017; Garzón & Acevedo, 2019; Ibáñez & Delgado-Kloos, 2017; Scavarelli et al., 2021). Many researchers have cited the affordances AR has for enhancing situated action—providing context-responsive visualizations that focus attention and alter a person's perspective in ways that promote learning (Dunleavy et al., 2009; Wu et al., 2013). Radu (2014) notes the potential of AR to leverage people's natural capacity for working with spatial information, and several others have pointed to the ability to create a more personalized learning experience given what the system knows about the individual learner and what data can be collected about their learning behaviors (Dunleavy & Dede, 2014; Huang et al., 2016).

Although a range of AR applications have shown benefits for individual learning with single devices and platforms, we know that learning in the future will increasingly take place in more complex spaces where there are multiple people, multiple devices, and multiple modalities with which to interact. The types of learning challenges that students will encounter in these spaces will require them to work with and reconcile multiple representations (White & Pea, 2011). There has been some research on face-to-face multi-user technologies such as large multi-touch display surfaces that have been shown to be effective for facilitating collaborative learning (e.g., Mercier & Higgins, 2013). Likewise, there have been a handful of studies on multi-user "mixed reality" environments (e.g., Kim et al., 2021; Lui & Slotta, 2014). To our knowledge, however, there has been little if any research on the learning effects of technology ecosystems where multiple devices allow for different perspectives and different interactions with the same phenomenon or simulation.

There are important and unresolved design questions concerning how to coordinate multiple devices to both support synergistic interactions between groups of students and ensure that individuals in those groups come away with robust understandings and the ability to flexibly apply their knowledge to new contexts. In particular, there are questions specifically about immersive technologies such as AR and the role that it can and should play in these layered ecosystems. In these more complex configurations, the question is more than whether or not someone uses AR, but rather how they are using it, for what purpose, and with whom? This opens up additional questions about how to assess the impact of AR integration into the learning process. These implementations must take a multidimensional assessment approach, examining learning in the context of users' perceptions of the technologies, tasks, and group interactions. While there has been some early work attempting to explore these dynamics (Bork et al., 2021; Radu & Schneider, 2019), the mixed results from these studies motivates the need for additional research.

The current study explores the issue of learning effects in a multi-device ecosystem in the context of undergraduate astronomy courses. Astronomy is a well-suited content area for this investigation because it is a domain where students are frequently asked to reconcile multiple spatial perspectives to explain complex phenomena such as the causes of seasons (Plummer & Maynard, 2014). The specific research questions for this study are:

RQ1. To what extent did students learn the target astronomy skills and concepts while working in a multi-device (AR and tablet computers) small group ecosystem?

RQ2. How are learning outcomes and perceptions of/attitudes toward the learning activity different between students who used the AR device and those who did not?

RQ3. How did the ways that individuals used the devices affect learning outcomes and their perceptions of/attitudes toward the learning activity?

Methods

Participants and Study Design

The participants were 77 students recruited from an introductory astronomy course at a university located in the mid-western United States. All participants in the study consented to have their classroom activity data collected as part of this research. The study took place during participants' 50 mins weekly lab session for a consecutive three-week period during the Fall 2021 semester. In Week 1, the research team introduced the simulation software and let students explore it using both AR and tablet computers that were provided to them. This exploration was guided with a task sheet that highlighted the functionality that participants would be asked to use in the upcoming weeks. In Week 2, students worked on a seasons lab which was part of the students' regular course material. Students also had the opportunity to use our software and devices while they were working on the lab tasks in Week 2 of the study. The purpose of these first two weeks was to let students become familiar with our software and devices; no data from these weeks were used in the analysis described in this paper. In Week 3, the students engaged in a multi-device small group learning task designed by the research team called "Lost at Sea." During this week, students were randomly assigned to groups of three to four such that they could work on the task collaboratively.

There were a total of 8 lab sessions and each session had three to four groups (22 groups in total). Groups had the entire 50 mins of the session to work on the task together. One AR device (the Microsoft HoloLens 2) and two tablet computers were provided to each group. The research team encouraged the groups to use both devices while they explored the simulation in the first two weeks and in the introduction to the Lost at Sea task, but once the students started working on the task independently, they were able to make their own choices about which devices to use and when. All student interactions on the Lost at Sea task were video recorded using a tripod-mounted camera, and screen captures were taken from all the devices including AR. Students completed pre and post learning assessments as well as an individual post questionnaire administered at the end of the session.

Multi-Device Ecosystem Design

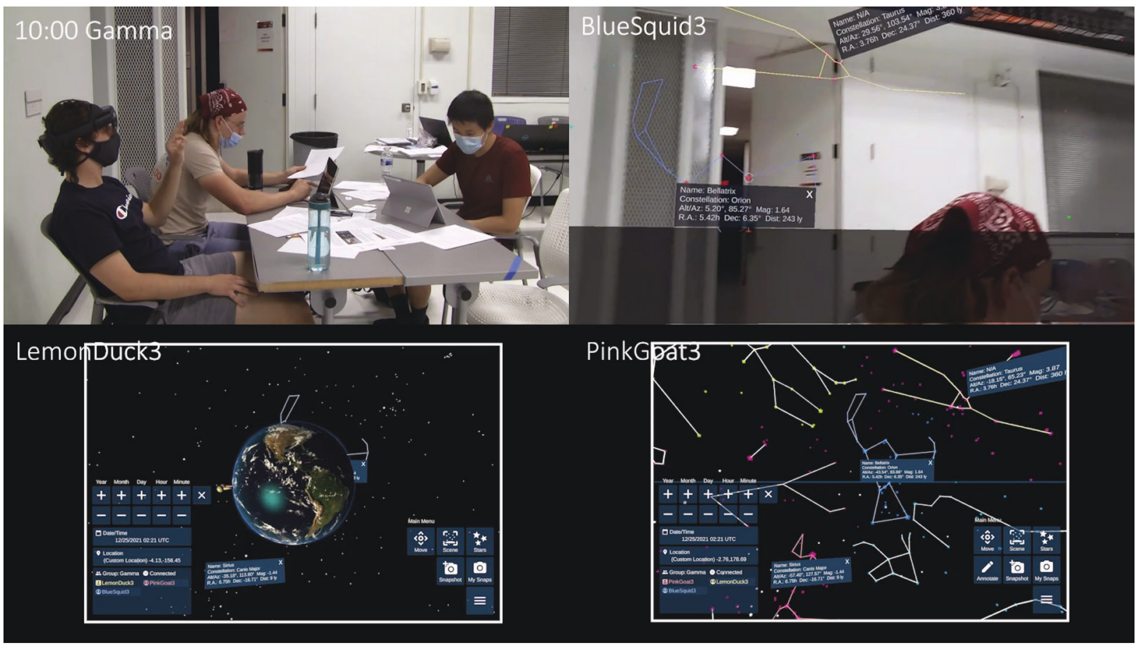
The digital sky simulation software utilized in this analysis was the product of multiple co-design cycles between the research team and collaborating astronomy educators. The system provides users with a stellar catalogue that can be explored across both space (location on Earth) and time. In addition, the system is networked between tablet and AR users, allowing all group members to annotate the night sky (e.g., sketching lines), highlight specific constellations of interest, share location information, and share their night sky perspectives collaboratively as they work. Interaction on the tablets is touch-based, while interacting in AR leverages hand and gesture detection to allow the user to tap interface items, pinch holograms to reposition them, and tap their index and thumb together to select distant objects such as stars.

Figure 1 shows a group accessing the three simulation representations available in the system, all accessible in both the tablet and AR platforms. The system was designed so that the interface was as consistent as possible across AR and tablets to reduce the need to "re-learn" the software when entering AR (see blue input and info panels in Figure 1). "Horizon view" (top-right) provides the user with a first-person view of the sky as if standing on Earth. "Star view" (bottom-right) removes the horizon limitations from the perspective, giving the user access to the full celestial sphere. "Earth view" (bottom-left) places the user in orbit above the Earth with the ability to place location pins with which to change their position on the Earth or confer with other users about potential sky viewing locations. All group software interactions (interface manipulations, view changes, annotations) were written to log files for later analysis.

Task Design

A multi-part problem solving narrative was created for students to address in the "Lost at Sea" task. The goal was to promote authentic problem solving and in-depth utilization of the simulation. The narrative describes a crewed space capsule that has splashed down at night at an unknown location. The simulation contains a selection of specially bookmarked locations with no associated latitude or longitude data, and students are tasked with using their night sky knowledge to determine the latitude and longitude of their capsule splash-down site. There are three parts to the task: Subtask 1 has the group locating stars or reference constellations to determine what hemisphere they are located in; Subtask 2 then asks the group to identify constellations that can serve as reference points for North, South, East, and West; finally, Subtask 3 asks them to formally calculate their latitude and longitude. The ability to use night sky data to determine the latitude and longitude is the overarching learning goal for this activity. Supporting task documents outlining the task goals, as well as supplemental information on the process of latitude and longitude calculation, are accessible to the group. What makes longitude calculations especially challenging is that the supporting instructions reference the most common process of leveraging the position of the sun, but groups must transform that knowledge to calculate longitude using the night sky instead.

Figure 1
Composite screenshot of the multi-device learning ecosystem showing the three views of the simulation: Top right (AR in Horizon view), bottom right (Tablet in Star view), and bottom left (Tablet in Earth view).



Measures

We conducted two paper-based assessments: pre and post learning assessments and post questionnaires about perceptions and attitudes. Pre and post learning assessments were designed to measure whether students' knowledge of how to calculate latitude and longitude changed based on their interactions with the multi-device learning ecosystem described above. Both learning outcomes entail a multi-step process, and a short open response prompt allowed the research team to evaluate the completeness and accuracy of students' understanding of this process. The prompts in both the pre and post learning assessments were the same: "Write as much as you know about the steps for calculating latitude/longitude based on the stars visible in a given location." To assess students' responses, two researchers (1st and 2nd author) looked through all students' answers together and scored them based on a rubric developed during pilot testing of the multi-device learning ecosystem (Kim et al., 2021). If students' answers did not contain any of the key astronomy concepts or they did not respond, we scored them as 0. If students' answers only contained basic concepts, then we scored them as 1 (e.g., we need to find a certain star's height to calculate latitude; calculating longitude is related to the time and location). If students' answers contained all desired astronomy concepts, then we scored them as 2 (e.g., we need to find Polaris and measure its height from the horizon, and it is the latitude for Northern hemisphere; we need to find when a certain star is at the highest point in both Greenwich and crash site to find the time difference then multiply it by 15 to calculate longitude). The post questionnaire consisted of five-point Likert scale items asking how they felt about their perceptions of and attitudes toward the multi-device group collaborative learning activity: "While working with your group, how often did you rely on your group members for help on a task or concept?", "While working in your group, how often did your group members rely on you for help with a task or concept?", "While working in your group, did collaborations occur frequently between students using AR and students using tablets?".

In addition, open-ended questionnaires asked participants what tasks they completed while they were using each device. The first and second authors also classified these open-ended post-questionnaire items based on the coding scheme detailed in Table 1. The coding scheme was applied for both types of device-use (tablet and AR), but for tablets only level 2 and 3 were employed, because all students used the tablets at least once for exploring or measuring.

Table 1
Classifying how students used technology

Level		Description	Example
0	Not used	Student did not use the device at all.	I didn't use AR this week.
1	Adjust to the technology	A student used the device but adjusted to the technology itself rather than using it for problem-solving.	Mostly focused on understanding AR.
2	Exploring	A student used the device to explore the night sky and clicked on some stars and constellations.	I selected a few constellations in the AR.
3	Measuring	A student used the device to calculate or measure for the task completion (e.g., location change, tracking time, etc.).	I used the tablet to change locations, set the time, and find constellations.

Results

Our first research question asks whether students learned the targeted astronomy content (calculating longitude and latitude) while they worked on the multi-device small group activity. Wilcoxon signed rank test (paired) was conducted on the students' pre and post learning assessment scores. The result shows that the pre assessment median score is significantly different from the post assessment median score for both latitude ($p < 0.001$, effect size $r = 0.56$) and longitude ($p < 0.001$, effect size $r = 0.54$). Table 2 shows that for both latitude and longitude, the number of students with 0 points decreased and the number of students with 1 and 2 points increased on the post-assessment. Since longitude is a more complex calculation, it is not surprising that scores for this measure were lower than for latitude.

Table 2
Pre-assessment and post-assessment score distribution

	Latitude		Longitude	
Score	Pre	Post	Pre	Post
0	59	31	74	48
1	12	17	3	21
2	6	29	0	8
Total Participants	77	77	77	77

Our second research question asks whether there is a difference between students who used AR devices and students who did not in terms of learning outcomes and perceptions of/attitudes toward the learning activity. First, we conducted the Mann-Whitney U test to measure the learning outcomes. The result shows that whether or not students use AR during the lab session did not affect either of the learning outcomes (e.g., latitude: $p = 0.82$, longitude: $p = 0.32$). Combining this with Wilcoxon signed rank test result (research question 1), we can conclude that students learned desired content regardless of their use of the AR devices. We did not conduct the same measurement for the tablet since all the students used tablets. In addition, we conducted a One-way ANOVA to measure perceptions of the learning activity. The result also revealed that whether students used AR devices or not does not impact these perceptions (e.g., Individual reliance on group: $p = 0.43$, Group reliance on individual: $p = 0.39$, Frequency of collaboration: $p = 0.74$).

Our last research question asks how the way in which individuals used each device affected learning outcomes and their perceptions of the learning activity. First, we conducted an Ordinal Logistic Regression to measure the relationship between how students used each device and learning outcomes. The estimated model shows that:

For latitude

$$\text{Logit}\ (P(Y \leqq 1)) = 2.67 - 0.29 * \text{AR usage} - 0.95 * \text{Tablet usage}$$
$$\text{Logit}\ (P(Y \leqq 2)) = 3.65 - 0.29 * \text{AR usage} - 0.95 * \text{Tablet usage}$$

For longitude

$$\text{Logit}\ (P(Y \leqq 1)) = 0.14 - 0.12 * \text{AR usage} - (-0.21) * \text{Tablet usage}$$
$$\text{Logit}\ (P(Y \leqq 2)) = 3.65 - 0.12 * \text{AR usage} - (-0.21) * \text{Tablet usage}$$

Based on the model, the odds ratios are presented in Table 3. For *latitude*, for every one unit increase in students' AR usage (i.e., Adjust to the technology → exploring), the odds of being more likely to get a higher score (i.e., 1 or 2 versus 0) is multiplied 1.33 times (i.e., increases 33%), holding constant all other variables. For students who used tablets for measuring, the odds of being more likely to get a higher score (i.e., 1 or 2 versus 0) is 2.60 times that of students who used tablets for exploring, holding constant all other variables. For *longitude*, for every one unit increase in students' AR usage, the odds of being more likely to get a higher score is multiplied 1.13 times, holding constant all other variables. For students who used tablets for measuring, the odds of getting a higher score (i.e., 1 or 2 versus 0) is 19% lower [i.e., (1 - 0.81) * 100%] than the students who used tablets for exploring, holding constant all other variables.

Table 3
Odds ratios for each learning outcome

	Latitude		
	OR	2.5%	97.5%
AR usage	1.33	0.88	2.03
Tablet usage	2.60	0.89	8.10
	Longitude		
	OR	2.5%	97.5%
AR usage	1.13	0.73	1.78
Tablet usage	0.81	0.28	2.47

Furthermore, we conducted a Two-way ANOVA to measure the relationship between how students used each device and their perceptions of the activity (see Table 4). Results revealed that there is a significant difference in the extent to which they felt that they relied on their group members based on how the students used the tablet computers ($F = 3.92$, $p = 0.05$). How they used AR also has a marginally significant effect on students' perceived group reliability ($F = 2.23$, $p = 0.08$), so we decided to look more deeply using the Tukey post hoc test to see if there are any significant pairwise differences. There is a marginally significant pairwise difference between 3 (measuring) and 1 (adjusting to the tech), ($p = 0.06$). These findings indicate that the way students used tablets influenced students' perceived dependency towards the group, and there is a pairwise difference when students used AR as a measuring technology compared to when students used AR simply to acclimate to using the technology itself on the students' perceived dependency towards the group. This means that the students who used AR for measuring felt less dependency towards the group compared to the students who used AR solely to learn how to use the technology. There were no significant differences for other perception questions yielded by different types of technology usage (see Table 4).

Table 4
How different usage of each technology affects students' perceptions of the learning activity
(* indicates $p < .05$)

	Two-way ANOVA				
Individual reliance on group	df	Sum sq	Mean sq	F	p
AR usage	3	4.57	1.52	2.23	0.08
Tablet usage	1	2.64	2.64	3.94	0.05*
Residuals	72	48.32	0.67		
Group reliance on individual	df	Sum sq	Mean sq	F	p
AR usage	3	0.46	0.15	0.24	0.87
Tablet usage	1	0.00	0.00	0.00	0.98
Residuals	72	45.34	0.63		
Frequency of collaboration	df	Sum sq	Mean sq	F	p
AR usage	3	3.00	1.00	0.93	0.43
Tablet usage	1	0.47	0.47	0.44	0.51

Residuals	72	77.77	1.08	
Tukey's Honestly-Significant-Difference (TukeyHSD) Test for AR Usage and Individual Reliance on Group				
AR usage	diff	lwr	upr	p adj
1 - 0	0.74	-0.17	1.64	0.15
2 - 0	0.11	-0.51	0.74	0.97
3 - 0	-0.15	-0.86	0.55	0.94
2 - 1	-0.63	-1.48	0.23	0.22
3 - 1	-0.89	-1.80	0.03	0.06
3 - 2	-0.26	-0.90	0.37	0.70

Discussion

The significant pre to post learning gains measured here suggest that productive learning can occur when students engage with problem-solving tasks within a multi-device small group ecosystem. In and of itself, this is a meaningful result, showing that the combination of AR and tablet technology integrated through a multi-part problem solving task did not detract from, and may have benefitted, the core instructional goals of the implementation. Interestingly, the learning gains experienced as a result of engaging with the multi-device ecosystem was not affected by whether or not a participant used the AR headset. One possible interpretation is that the learning effects of this intervention stem from the overall ecosystem and not from a specific device; a participant does not necessarily need to put on an AR headset themselves to receive the benefit of someone in the group having and sharing that perspective. Likewise, participants' perceptions of whether they relied on the other members of the group and whether the group relied on them were not affected by AR usage. This again suggests that participants felt that they were interacting with a dynamic system of people and technology as opposed to a single device.

In general, it was clear from this analysis that simply accounting for whether students used the AR device or not does not provide a complete picture of how the technology influences learning and perceptions of the activity. When we examined *how* the technology was being used, we began to see a more nuanced relationship between AR usage and performance on the problem-solving tasks. Specifically, the categorization of usage into exploring and measuring behaviors proved to be a useful lens for analysis. We discuss these results in terms of our two primary learning outcomes: calculating *latitude* and calculating *longitude*. We are keeping our discussion of these outcomes separate because the results were slightly different, and because the tasks themselves present different levels of challenges, as described below.

When examining student learning performance around the calculation of *latitude*, while the AR usage levels showed some benefit to the likelihood of an increased assessment score, tablet users showed substantially higher likelihood when engaging in tasks related to measuring as opposed to exploring. These relationships are potentially due to the affordances of the AR and tablet technologies, highlighting differences in the user interface actions required for exploration or measurement. Exploration interactions (such as panning the sky to locate a star) in the tablet required the use of arrow key touch buttons, while within AR a user needs no interface interactions (simply rotating their head and body) to modify their view. Therefore, it is much more natural and easier to explore and find constellations in the night sky with the AR environment. In contrast, measuring interactions (changing time, drawing annotations) use similar combinations of buttons (touch or holographic) and selection actions to complete. The condensed layout and faster response of the tablet interface (as opposed to "pressing" a holographic button) may have been a contributing factor in facilitating the likelihood of higher knowledge scores for the tablet users.

With the calculation of *longitude*, we see again that moving up levels of AR usage still shows a positive relationship with higher level post assessment scores although it slightly decreased compared to the latitude learning outcome. The result for tablet usage here is reversed, with tablet users engaging in exploration showing increased likelihood of higher assessment scores when compared to those who engaged in measuring. As we mentioned in the Task Design section of the paper, the longitude task was intentionally designed to be more difficult to solve since it requires students to do a higher-level calculation and transfer knowledge from a different scenario. Here the complexity of the longitude task potentially plays a large factor. To calculate longitude, a group needs to engage in a series of alternating exploration and measurement/manipulation behaviors, reducing the likelihood of one device type dominating the other due to an interface affordance. Although our RQ1 finding verified there is a pre to post increase on the longitude measure, it is also true that many of the students still remained at 0 or 1. This motivates us to investigate students' higher-level/complex learning outcomes by analyzing more diverse data (e.g., students' group discourse, log data, etc.) beyond their self-reporting.

When examining student perceptions of the activity, we see a significant relationship between their ratings of how often they felt they relied on the group based on their level of tablet usage. In other words, if the students used tablets at a higher level (e.g., measuring), they perceived they were more reliant on their group members. This may mean that students tried to solve problems by interacting with group members rather than solving everything by themselves when trying to calculate or measure something. For example, in order to calculate longitude, information on two particular locations is required at the same time. To do this, a student using a tablet at a higher level may have sent other students to each of those two locations to find information. In addition, the notable pairwise difference for perceived reliance on the group in AR between level 3 (measuring) and level 1 (simply adjusting to the AR technology) may indicate that if students used AR for measuring, they did not think they relied on the group compared to when students use AR simply to get familiar with it. This makes sense because if students are using AR to get familiar with it, they are likely to ask group members some questions about how to use the device, while if students are using AR for measurement, they are likely more focused on the task and less in need of relying on their group.

Taken together, these results contribute to an increased understanding of how groups can begin to engage in a complex problem solving task with the support of AR technologies. The analysis performed here provides insight into the role of user perceptions of technology and group interactions and the relationship of these perceptions with learning outcomes. Moving forward, it is clear that effective designs that incorporate AR into small group learning environments alongside other technologies must be cognizant of the affordances attributed to each platform, and the role they can play in facilitating specific patterns of interaction and measurements of learning. In the current study, we primarily focused on the students' self-reports of their usage of the technology. Furthermore, this study was conducted during existing university class sessions, not in a controlled experimental environment. As to keep a classroom atmosphere, we wanted to give students autonomy over which device to use and how much to use, rather than force them to use all devices evenly. Therefore, this paper mostly focuses on the effectiveness of a multi-device learning ecosystem as a whole, and how the ecosystem should be designed based on the students' usage of each device. To have a more detailed understanding of how the students' different device usage affected their learning and what are the educational effectiveness of individual device usage, our future research will include discourse analysis (e.g., topic modeling and association rule mining) along with log data analysis. We will investigate how the students' discourse topics vary when they are using different devices, how those different discourse topics are associated, and if there are any relationships between the topics and certain types of student-software interactions.

References

Bork, F., Lehner, A., Eck, U., Navab, N., Waschke, J., & Kugelmann, D. (2021). The effectiveness of collaborative augmented reality in gross anatomy teaching: a quantitative and qualitative pilot study. *Anatomical Sciences Education, 14*(5), 590–604.

Chen, P., Liu, X., Cheng, W., & Huang, R. (2017). A review of using Augmented Reality in Education from 2011 to 2016. *Innovations in Smart Learning*, 13–18.

Dunleavy, M., & Dede, C. (2014). Augmented reality teaching and learning. *Handbook of Research on Educational Communications and Technology*, 735–745.

Dunleavy, M., Dede, C., & Mitchell, R. (2009). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *Journal of Science Education and Technology, 18*(1), 7–22.

Garzón, J., & Acevedo, J. (2019). Meta-analysis of the impact of Augmented Reality on students' learning gains. *Educational Research Review, 27*, 244–260.

Huang, T.-C., Chen, C.-C., & Chou, Y.-W. (2016). Animating eco-education: To see, feel, and discover in an augmented reality-based experiential learning environment. *Computers & Education, 96*, 72–82.

Ibáñez, M. B., & Delgado-Kloos, C. (2018). Augmented reality for STEM learning: A systematic review. *Computers & Education, 123*, 109-123.

Kim, T., Planey, J., Lindgren, R., & Kang, J. (2021). Integrating Immersive Technology into Small Group Learning Environments. In de Vries, E., Hod, Y., & Ahn, J. (Eds.), *Proceedings of the 15th International Conference of the Learning Sciences - ICLS 2021*. (pp. 1137-1138). Bochum, Germany: International Society of the Learning Sciences.

Klopfer, E., & Squire, K. (2008). Environmental Detectives—the development of an augmented reality platform for environmental simulations. *Educational Technology Research and Development, 56*(2), 203–228.

Lui, M., & Slotta, J. D. (2014). Immersive simulations for smart classrooms: exploring evolutionary concepts in secondary science. *Technology, Pedagogy and Education, 23*(1), 57–80.

Mercier, E. M., & Higgins, S. E. (2013). Collaborative learning with multi-touch technology: Developing adaptive expertise. *Learning and Instruction, 25*, 13–23.

Plummer, J. D., & Maynard, L. (2014). Building a learning progression for celestial motion: An exploration of students' reasoning about the seasons. *Journal of Research in Science Teaching, 51*(7), 902-929.
Radu, I. (2014). Augmented reality in education: a meta-review and cross-media analysis. *Personal and Ubiquitous Computing, 18*(6), 1533–1543.
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 *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1–12). Association for Computing Machinery.
Scavarelli, A., Arya, A., & Teather, R. J. (2021). Virtual reality and augmented reality in social learning spaces: A literature review. *Virtual Reality, 25*, 257–277.
Wu, H.-K., Lee, S. W.-Y., Chang, H.-Y., & Liang, J.-C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education, 62*, 41–49.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No.#1822796. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation. We would like to thank the entire CEASAR staff for making this study possible, particularly Nathan Kimball and his team at Concord Consortium.

