

What Can We Learn from Augmented Reality (AR)?

Benefits and Drawbacks of AR for Inquiry-based Learning of Physics

Iulian Radu[†]

Graduate School of Education
Harvard University
Cambridge, MA, USA
iulian_radu@gse.harvard.edu

Bertrand Schneider

Graduate School of Education
Harvard University
Cambridge, MA, USA
bertrand_schneider@gse.harvard.edu

ABSTRACT

Emerging technologies such as Augmented Reality (AR), have the potential to radically transform education by making challenging concepts visible and accessible to novices. In this project, we have designed a Hololens-based system in which collaborators are exposed to an unstructured learning activity in which they learned about the invisible physics involved in audio speakers. They learned topics ranging from spatial knowledge, such as shape of magnetic fields, to abstract conceptual knowledge, such as relationships between electricity and magnetism. We compared participants' learning, attitudes and collaboration with a tangible interface through multiple experimental conditions containing varying layers of AR information. We found that educational AR representations were beneficial for learning specific knowledge and increasing participants' self-efficacy (i.e., their ability to learn concepts in physics). However, we also found that participants in conditions that did not contain AR educational content, learned some concepts better than other groups and became more curious about physics. We discuss learning and collaboration differences, as well as benefits and detriments of implementing augmented reality for unstructured learning activities.

KEYWORDS

Augmented Reality; Physics Education; Collaborative Learning

ACM Reference Format

Iulian Radu and Bertrand Schneider. 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, Glasgow, Scotland, UK. ACM, New York, NY, USA. 12 pages.
<https://doi.org/10.1145/3290605.3300774>

1 INTRODUCTION

Augmented reality (AR) is becoming affordable and popular and its increased adoption is generating a growing interest for educational use - both in formal and informal environments. In the formal space, teachers are increasingly using technology-enhanced hands-on learning activities to foster deep conceptual understanding, such as interactive simulations or activities involving sensors. In the informal space, we are currently witnessing the birth of the “maker” cultural movement where everyday people collaboratively tinker with physical and digital materials, in order to explore, modify or create physical artifacts. In such environments, people engage in self-driven inquiry-based learning, and are indirectly exposed to a variety of STEM concepts. We believe augmented reality has the potential to radically transform formal and informal education by making challenging concepts visible to novices.

In this research we critically investigate the benefits and drawbacks of augmented reality for inquiry-based learning. We focus on a collaborative activity that explores electromagnetism concepts, because electromagnetism is a topic that is often encountered in both maker spaces and traditional physics classrooms; it is one of the most difficult topics to master for students of all ages [1,2,5]. An activity typically taught in electromagnetism curriculums, and pursued in makerspaces, is the construction of speakers. Audio speakers involve different physical phenomena - such as flow of electric current, amplification and alternation of electricity, generation of magnetic fields from electricity, production of forces acting to vibrate membranes, audio waves, etc. These phenomena interact

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
CHI 2019, May 4–9, 2019, Glasgow, Scotland UK
© 2019 Association for Computing Machinery.
ACM ISBN 978-1-4503-5970-2/19/05...\$15.00
<https://doi.org/10.1145/3290605.3300774>

with each other in complex ways while being invisible to the naked eye, thus making the concept difficult to understand. Yet these phenomena are critically important for understanding the physics of electromagnetism.

We believe that emerging technologies, such as Augmented Reality, have the potential to address this issue and radically transform STEM learning by making challenging concepts accessible to students. Augmented Reality headsets, such as the Microsoft HoloLens, allow students to see virtual “holograms” in the physical world. It is therefore possible to design activities where learners can visualize and interact with dynamic representations of hidden forces (e.g. visualizing electrons, magnetic fields, light or radio waves).

In this research, we take a critical perspective on the educational benefits of augmented reality and explore its use for learning electromagnetism. This paper fills an important gap in the literature: prior research has explored how delivering educational content through augmented reality is beneficial in comparison to traditional media (such as printed materials, videos, or PC-based simulations). In such comparisons, it is unclear whether significant effects are due to the differences between the mediums of presentation, for example in terms of dynamic vs. static nature of educational representations, ease of user interactivity, ability to visualize representations in 2D or 3D, or differences in the informational content presented to users.

This project contributes to our understanding of AR uses in education by rigorously observing the positive and negative effect of AR technology. We expose participants to an activity that involves interacting with an interactive physical model of an audio-producing speaker. We provide dynamic visual representations of electromagnetism concepts that are aligned to the physical interactive system and we investigate how the presence or absence of such representations influences collaborative learning while keeping our experimental conditions as similar as possible. Furthermore, we investigate how much of the learning effects are due to the novelty of AR technology, by comparing a condition involving just physical interaction with the system without AR visualizations and the same physical system with simple AR visualizations (with no educational content). In all conditions we measure participant learning, collaboration and attitudes. Through this approach we contribute to a much needed understanding of the benefits and detriments of AR technology in educational settings.

2 RELATED WORK

2.1 Augmented Reality as Educational Medium

In the CHI community, there has been a long tradition of combining physical and digital representations using Augmented Reality to support learning (for example with Tangible User Interfaces; [31]). AR is a technology allowing superimposition of digital content on physical spaces and objects using a projector or a headset [3]. AR applications are beginning to be used in classrooms, on different platforms such as handheld applications [36], hands-free low-interactivity Google Cardboard experiences [37], and more interactive high-fidelity headsets such as Microsoft HoloLens [38]. Each type of AR platform provides different affordances for learning [13]. In this study we used the Microsoft HoloLens because of its ability to simulate complex phenomena and accurately align virtual content on physical objects, while allowing students to use their hands to interact with the learning content. Studies have explored how AR educational applications compare to traditional educational approaches such as textbooks [6], instructional videos [10], or PC-based instruction [5]. Research has shown that AR improves student understanding of structures that are either spatially complex or invisible to the naked eye [15, 16, 10]. Benefits come from visualization of otherwise inaccessible representations, and from permitting students to perform low-cost and low-risk experiments [7], providing information in the most educationally-relevant context [4], allowing physical practice for learning of tacit knowledge [11], providing motivation for students to engage with content and peers [5, 23], encouraging collaboration by equalizing access to information [17], and facilitating reciprocal teaching and authentic cooperative inquiry [16]. From an educational perspective, learning scientists have formulated theories about the benefits of Multiple External Representations (MERs) [32]. MERs support conceptual understanding by encouraging the use of multiple strategies, offering different perspectives on a problem and taking advantage of users’ familiarity with one representation to help transition toward a more complicated representation [32]. AR has the potential to provide synergies between physical and digital representations, which has been studied in [33,34]. Such cross-media investigations are critical for understanding the educational potential of a new medium. Our understanding of the factors that positively impact learning is still preliminary, however: learning can benefit from novelty effects of being exposed to a new technology, the added physicality of interacting with physical items, the ability to see

information in 3D instead of 2D, the ability to interact rather than watch, and ability to collaborate with colleagues in exploring a common domain [12, 13, 14].

As AR rapidly continues to broaden its popularity, there is an urgent need to explore the nuances that make this technology effective or ineffective. We need to critically understand how to best design this medium to foster new kinds of learning experiences; without this understanding it is unlikely that AR will be used to its full potential, and its adoption may suffer due to unmet expectations. In this project we study the benefits and drawbacks of AR technology by comparing between different versions of an interface designed for physics education. We start with a physical setup and use augmented reality to add dynamic 3D representations. Compared to traditional media such as videos or PC applications, a tangible physical interface by its very nature provides educational benefits due to its ability to physically embody educational concepts [20], evoke gestural interactions [18], and provide ease of exploration [24]. We compare the physical interface with the same interface augmented with simple AR content. Additionally, we compare the learning effects between different kinds of AR representations.

2.2 Augmented Reality for Electromagnetism Education

Electromagnetism comprises the set of concepts relating the properties of electricity to magnetic fields. It is a topic that students of all ages struggle with [1, 2, 5]. Students must understand and internalize abstract knowledge that is invisible to the naked eye (such as the shape of magnetic fields and flow of electric currents) and which has no simple real life referent (such as what voltage is, or how magnetic fields are generated from the flow of electricity). Existing studies have explored the effect of adding educational representations to physical objects in order to teach electrical and electromagnetism concepts. AR representations of electricity flowing through real circuits have been researched (e.g. [19, 21, 22]), along with AR visualizations of magnetic fields [25, 26], and electromagnetism concepts [5, 6, 8]. The AR applications compared to non-AR show improvement in student abilities to visualize structural phenomena [5, 6], reduced cognitive load [22], improvements in motivation and self confidence [19, 6]. Understanding of theoretical knowledge has mixed results, with some research showing improved understanding [19], while others did not [5, 22]. Some research shows that non-AR representations may be more valuable for understanding some concepts of magnetism [25]. In this project we are interested to understand which

specific topics are best suitable for teaching through AR representations, and to understand how the presence of tangibility influences these results. We are not aware of existing studies that compare student learning electromagnetism through a tangible interface vs. its augmented counterpart. In this study we present this approach and contribute a nuanced understanding of the benefits and drawbacks AR representations for physics education.

3 SYSTEM DESIGN

We have designed a system guided by three perspectives:

Involving multiple phenomena: Traditionally, computer-based educational simulations have focused on single phenomena that can be the center of learner attention; but we wanted to explore the power of augmented reality in unstructured learning environments that involve multiple interrelated physical phenomena, because they are more representative of real-world situations and projects that students typically tackle in informal learning environments (e.g., makerspaces).

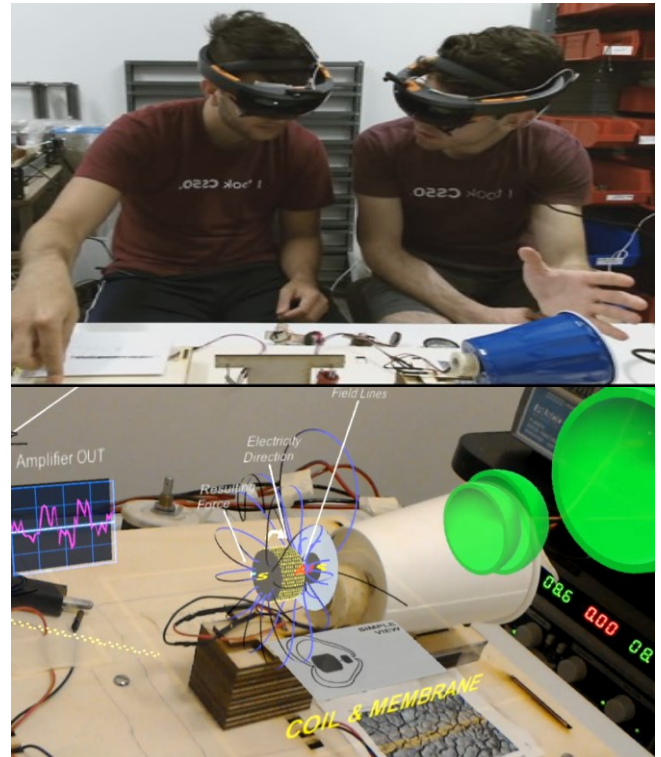


Figure 1. The Augmented Reality system developed for this project (top image: two users interacting with the speaker activity; bottom image: the magnetic fields around the coil and the magnet that are generating the sound waves).

Informed by student difficulties and domain experts:

We wished to create learning activities that students and teachers would find valuable. Through our design process, we visited physics classrooms and participated in conversations with a physics teacher and domain experts to identify a learning activity that is problematic (but also engaging) for a wide range of students.

Constrained by technology: The AR headset is limited in its capability of detecting and generating visualizations of what is occurring in the real system. Therefore, user interactions are limited such that the system could produce valid representations through the device, while still allowing a degree of agency and learner-driven exploration.

For our final design, we focused on the activity of building sound-producing speakers. This activity is typically taught in high-school and university-level physics, and exposes students to a system that converts energy from electric current, to magnetic fields, and finally to sound waves. The physics teacher advisor (mentioned in point 2 above) indicated that, even after a few classroom sessions, students can build a working speaker but have trouble explaining how it works. This activity is especially suited for AR technology because it can provide contextualized visualizations of the underlying phenomena in physics.

object. (Note, AR visualization misalignments are due to the photo camera, and not visible to participants)

Our final system consists of an interactive hardware system that replicates an audio speaker (Fig. 1). The system is composed of multiple HoloLens devices networked together. Physical electronic modules allow groups of learners to collaborate while observing 3D visualizations of invisible phenomena occurring in the physical space. Sound is produced by a diaphragm membrane with an attached magnet. The diaphragm is located next to a coil of wires, which receives amplified electric signals from a control board. Participants can push buttons on the control board to play music from a smartphone or send constant forward or backward current through the system. Participants can also control the placement of the diaphragm membrane, change the type of coil used, and adjust the amplification.

The augmented reality features of the activity provide interactive visual representations of physics phenomena. Interactions with the hardware activates AR visualizations (Fig. 2) of electric current (yellow electrons moving along the physical wire, charts showing voltage strength), magnetic fields (curved lines around the coiled wires and magnets, and coaxial planar rings around straight wires), and sound waves (green semispheres).

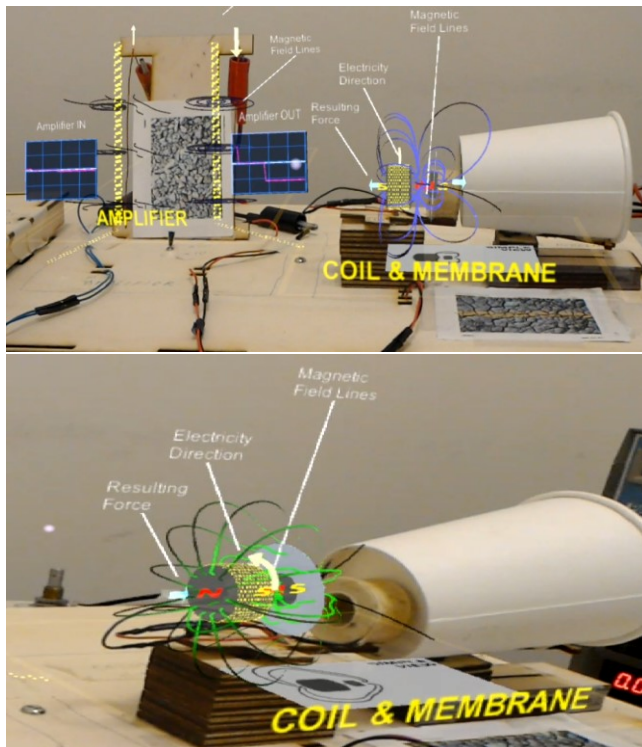


Figure 2. Changing AR visualizations of electricity, amplification, and magnetic fields overlaid on the physical

4 RESEARCH QUESTIONS AND STUDY DESIGN

We conducted a study to investigate how participant attitudes, learning outcomes and collaboration are affected by AR technology. Specifically, we measured these variables in relation to the presence or absence of AR representations of electromagnetism. Furthermore, we wanted to know how much of the effects came from the novelty of the technology. Our research questions were:

RQ1: Are participant attitudes influenced by the presence of educational AR representations ?

RQ2: Is the understanding of learning content influenced by the presence of educational AR representations ?

RQ3: Is group collaboration influenced by the presence of educational AR representations ?

RQ4: Does the mere presence of AR technology (without any educational content) affect participant experience ?

Participants were randomly assigned to four experimental conditions (Table 1). The “Non HoloLens” condition involved the same activity without wearing a HoloLens device. Participants in the “HoloLens Simple” condition wore the HoloLens device and saw limited AR visualizations (which only included outlines of the major system components and visualization of sound waves being

emitted from the speaker. Participants in the “AR Scaffold” condition wore the Hololens device, but for the first 10 minutes of the activity they only saw visualizations similar to the “Hololens Simple” group; after 10 minutes they saw the AR layer of magnetic fields; after 15 minutes they also saw the AR layer of electric current; and after 20 minutes they saw information from the poster added into the AR experience. Participants in the “AR Full” group experienced all the AR layers from the start.

These experimental conditions were chosen in order to control for the effect of exposure to physicality and information exposure. We had Non-Hololens and Hololens-Simple groups in order to test the effect of novelty and excitement that may come with experiencing even basic AR technology. We also had two types of AR educational groups, AR Scaffold vs AR Full, because learning theories [35] suggest that presenting increasingly complex representations facilitates learning.

5 METHODS

5.1 Participants

Participants were recruited from the study pool of a laboratory at a university in the northeastern United States. Participants who signed up for a study session were required 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. Pairs of participants were randomly assigned to one of the four experimental groups. We recruited 15 participant pairs for each condition (N=120). For the analysis, we removed sessions in which technical issues were encountered, and removed outlying participants whose pre-test score was beyond 2 standard deviations from the mean, which resulted in 14 pairs in each condition (N=112).

5.2 Dependent Variables

Attitude Metrics: Participant attitudes towards the **user experience** were measured using the survey instrument in [9]. The survey contains 5-point Likert scale questions about users’ perception of aesthetics, endurability, focus, novelty, involvement and usability. We also measured participant changes in attitudes towards their abilities to engage in physics / **physics self-efficacy** [28] through a 5-point Likert scale question before and after the experience. Attitude metrics are listed in Table 2 shown in the Results section.

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 [30]. The learning test contained multiple-choice and open-ended questions measuring several aspects of conceptual knowledge. For coding open-ended questions, a coding scheme was created for each question. The coding scheme was a simple decision tree which required researchers to look for specific concepts (For example: “If answer mentions magnetic field, does it mention that it is created by electricity? Y/N” or “If answer mentions membrane, does it mention what moves the membrane? (m) Magnetic field; and/or (e) Electricity”). Two researchers coded separate test questions; each question was graded by only one researcher. All learning metrics are listed in Table 3 of the Results section. Illustrative questions are provided in Figure 3.

Table 1. Information representations presented to each condition. (X = information present at all times. D = information presented after specific delay; “repr.” = “representation”)

Experimental Conditions	No-EdAR		EdAR	
	Non Hololens	Hololens Simple	AR Scaffold	AR Full
Magnetic field repr. (AR)			D	X
Electricity repr. (AR)			D	X
Electromagnetism poster (AR)			D	X
Electromagnetism poster (printed)	X	X	X	X
Sound visual repr. (AR)		X	X	X
Label & outlines (AR)		X	X	X
Labels & outlines (printed)	X	X	X	X
Hololens device		X	X	X
System interactivity	X	X	X	X

Participant understanding of **magnetic field shapes** was measured through multiple-choice questions and open-ended drawing questions (Fig 3). Two **transfer questions** measured participants’ ability to apply knowledge to other situations: “*Is it possible to build a motor that is moved through electric signals? If yes, explain how.*” and “*One day while you are hiking through nature, you accidentally drop your iron keys into a hole in the ground. Your keys*

are made of iron, and iron is attracted to magnets. In your backpack, you have a soft long wire and a square battery. Could you retrieve your keys using only these materials?"

1. How is electrical energy turned into sound inside a speaker?

Please answer using 2-3 full sentences.

Signal is the input, it goes through an amplifier, then go through a coil wire to generate magnetic field, the field attracts/repels the magnet attached to the "ap" which then generate different sound.

4. A speaker produces sound from the vibration of a membrane. Why does the membrane vibrate? Please explain how the membrane is being moved.

Please answer using 2-3 full sentences.

THE MAGNET ATTACHED TO THE MEMBRANE MOVES ACCORDING TO THE CURRENT IN THE NEARBY COIL, CAUSING THE MEMBRANE TO VIBRATE BASED ON CHANGING CURRENT

3. If the direction of electric current is suddenly inverted, the magnetic field:

☐ (a) Does not change ☒ (c) Inverts

☐ (b) Magnifies ☐ (d) Weakens

5. If the magnetic field is suddenly inverted, the speaker membrane:

☐ (a) Is pulled closer ☐ (c) Is pushed away

☐ (b) Does not move ☒ (d) Moves, but the direction cannot be determined

6. If the direction of electric current is suddenly inverted, the speaker membrane:

☐ (a) Is pulled closer ☐ (c) Is pushed away

☐ (b) Does not move ☒ (d) Moves, but the direction cannot be determined

13. When the magnet is strongly attracted to the coil, the field shape is:

☒ (a) ☐ (b) ☐ (c) ☐ (d) ☐ (e)

☐ (f) None of the above

19. Draw the magnetic field around the three following components while they carry current

A single wire

A wire coiled 5 times

A wire coiled 50 times

6. For the case when Music is playing...

a) Draw the magnetic field to the wire before and after the amplifier.

Looser lines

more tighter denser lines (strong mag field)

b) Assume music is playing. In the two charts below, plot the current in the wire before and after the amplifier.

Current before the amplifier

Current after the amplifier

Figure 3. Examples of questions from the learning test.

Understanding of the **amplifier's effect on electricity and magnetic field** was measured through one specific question and through coded open-ended questions (Fig 3). Multiple-choice questions and coded open-ended questions

measured understanding of the relationship between **magnetic field and movement** (e.g. Figure 3 question 5), **electricity and movement** (e.g. Figure 3 question 6), **electricity and magnetic field** (e.g. Figure 3 question 3)

The concept of sequential reasoning indicates the style in which participants answered the open-ended question of "How is electrical energy turned into sound inside the speaker?" A large number of responses included a narrative which explained the connection between different components as a sequence (Figure 3 q1 top) rather than directly explaining the core physics phenomena driving the speaker. Sequential reasoning is a misconception that leads to student difficulties in understanding electronic circuits and can indicate shallow understanding of content [29].

Electromagnetism: Electric current flowing in a wire always creates a magnetic field. The direction and strength of current influences the magnetic field.

Coiled Wire: Combining multiple wires in a coil makes the magnetic field stronger.

Figure 4. Information shown on the physical poster in front of the participants.

Collaboration Metrics: Collaboration metrics were qualitatively coded for each pair of participants across several dimensions using a validated rating scheme described in [27]. 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. Two researchers coded the study session videos (the unit of analysis was the video), overlapping on 20% of video sessions. The coding scheme included definitions and examples for each level of each dimension in order to guide raters (ex: *sustained mutual understanding* example of low scores: "partners ignoring each others' insights; not

listening to each other; talking over each other without common ground”, example of high scores: “partners make sure partners understand; if any advanced vocabulary is used, explain meaning when communication suffers; actively trying to see if other person is on same page”). Further examples are provided in Table 4.

5.3 Experimental Protocol

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 (Table 1). 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.

5.4 Statistical Analysis

Our experimental conditions represent a hierarchical design with two nested factors (Table 1). The main factor, Presence of AR Educational representations contains two levels: present (EdAR) and not present (No-EdAR). The nested factor “Condition” contains two levels within each of the main factor levels: Non-Hololens, Hololens Simple (both under No-EdAR), and AR Scaffold, AR Full (both under EdAR). Statistical testing for relative learning gains was performed using ANOVA nested models, and followed by nonparametric independent two-sample Wilcoxon rank-sum tests when ANOVA assumptions were not met. Statistical testing for attitudes and collaboration metrics was performed using independent two-sample Wilcoxon rank-sum.

6 RESULTS

In this section we address our four research questions:

6.1 RQ1: Are participant attitudes influenced by the presence of AR educational representations ?

Participant post-study **attitudes towards the user experience** were significantly higher than neutral under both EdAR (overall $V=1596$, $p<0.01$) and No-EdAR conditions ($V=1653$, $p<0.01$), indicating that both types of experiences had a positive effect on participants. Comparing differences between EdAR and No-EdAR groups, we found that participant attitudes were significantly higher in the EdAR groups, on measures of Aesthetics ($W=2078$, $p=0.005$) and Involvement ($W=1944$, $p=0.045$). Other metrics of user experience attitudes were not statistically different between EdAR and No-EdAR groups.

We tested for significant differences between the experimental nested sub-conditions (ie: between AR Scaffold vs. AR Full; and between Non-Hololens vs. Hololens-Simple). We found no statistical difference between the two EdAR groups (AR Scaffold vs. AR Full). However, between the No-EdAR groups (Non-Hololens and Hololens Simple) groups, the Non-Hololens group scored significantly lower on participant ratings overall ($W=497$, $p=0.004$), specifically on topics of Aesthetics ($W=453$, $p<0.01$), Curiosity ($W=459$, $p<0.01$), Endurability ($W=435$, $p<0.01$), Focus ($W=492$, $p=0.020$), Interest ($W=458$, $p=0.002$), Involvement ($W=464$, $p=0.005$).

Table 2. Measured collaboration dimensions; significant differences ($p<0.05$) from neutral ratings, and between AR and Non-AR groups (* = statistical differences found within the AR or Non-AR subconditions)

Attitude Metrics and Examples	Difference from neutral	Difference EdAR vs. No-EdAR
Aesthetics <i>I liked the graphics and images in this activity</i>	Both	AR > *
Curiosity <i>I continued the activity out of curiosity</i>	Both	non sig *
Endurability <i>I was really drawn into this activity</i>	Both	non sig *
Focus <i>I was absorbed in my task</i>	Both	non sig *
Interest <i>I felt interested in this activity</i>	Both	non sig
Involvement <i>Doing this activity was worthwhile</i>	Both	AR >
Usability <i>I found this activity easy to use</i>	Both	non sig
(Post-Pre) Change in curiosity <i>I am curious to learn more about how electronics work</i>	NonAR only	non sig
(Post-Pre) Change in physics self efficacy <i>I easily learn physics topics</i>	AR only	AR >

Participant **changes in attitudes towards physics self-efficacy** was not significantly different than zero in No-EdAR groups, but was significantly higher than zero in the EdAR groups ($V=593$, $p<0.01$). Comparing between the EdAR and No-EdAR groups, we found that physics self-efficacy was significantly higher in EdAR groups than No-EdAR groups ($W=2078$, $p=0.04$), indicating that EdAR has a stronger influence on changing participant attitudes towards their own learning. No statistical differences was detected between the individual sub-conditions associated

with each EdAR and No-EdAR groups, indicating that AR educational representations or presence of AR technology may not have a significant effect.

Participant **changes in attitudes in curiosity towards the physics content** followed a reverse trend: change in curiosity was not significantly different than zero in EdAR groups, but was significantly higher than zero in the Non-AR groups ($V=399$, $p=0.011$), possibly indicating that Non-AR group participants are left more curious. However, no difference was detected when comparing between AR and Non-AR groups, or when testing the individual sub-conditions.

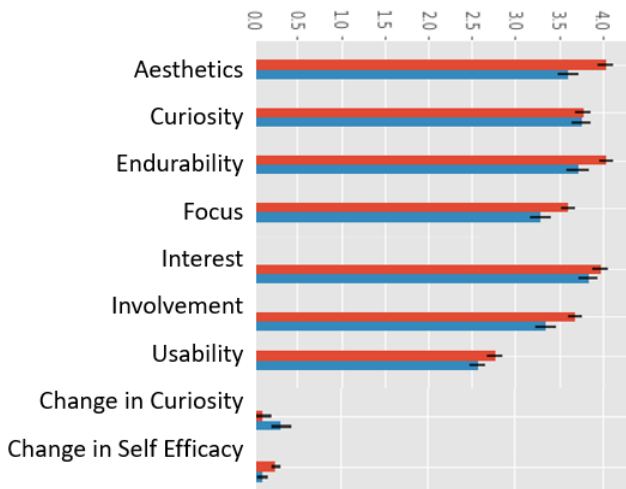


Figure 5. Group differences in attitudes (Range 0-5; Red = EdAR group, Blue = NoEdAR group; Bars = standard error)

6.2 RQ2: Is learning influenced by the presence of AR educational representations ?

Participant relative learning gains were significantly different than zero in EdAR condition (overall $t=7.55$, $p<0.01$) and No-EdAR conditions (overall $t=4.48$, $p<0.01$), indicating that the activities encouraged some understanding of the learning topics.

Participants in EdAR groups had statistically higher relative learning gains than compared to No-EdAR groups on their ability to identify and draw magnetic field shapes ($W=1942$, $p=0.047$), their understanding the relationship between electricity and magnetic fields ($W=1946$, $p=0.044$), and on answering the transfer question about constructing an electromagnet (“One day while hiking... Could you retrieve your keys using only these materials?”) ($W=1923$, $p=0.01$). No statistical differences were detected between the individual sub-conditions associated with each AR and Non-AR groups.

On the other hand, participants in No-EdAR groups had statistically higher relative learning gains than compared to EdAR groups on their ability to understand the relationship between magnetic fields and movement ($W=1139$, $p=0.006$), and were more likely to exhibit sequential reasoning ($W=1288$, $p=0.03$). No statistical differences were detected between the individual sub-conditions associated with each AREdARE and No-EdAR groups

Table 3. Measured conceptual knowledge; significant differences ($p<0.05$) from zero relative learning gains, and differences between AR vs NonAR groups. (** = significant differences at individual questions level)

Conceptual Knowledge	Difference from zero	Difference EdAR vs. No-EdAR
Shapes of Magnetic Field	Both	AR >
Transfer Questions	Both	non sig **
Sequential Reasoning	Both	NonAR >
Amplifier effect on Electricity and Mag Field	Both	non sig
Relationship between Electricity and Mag Field	Both	AR >
Relationship between Movement and Electricity	Both	non sig
Relationship between Movement and Mag Field	Both	NonAR >

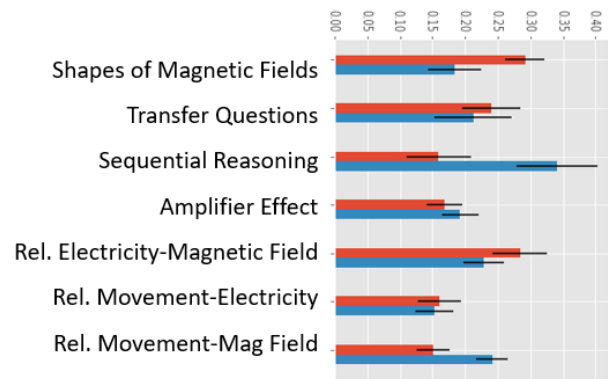


Figure 6. Group differences in relative learning gains (Range 0-1; Red = EdAR group, Blue = NoEdAR group; Bars = standard error)

6.3 RQ3: Is group collaboration influenced by the presence of AR representations ?

Two researchers double-coded 20% of the video recordings and achieved interrater reliability of Cohen Kappa 0.67, indicating “good” agreement.

Ratings of collaboration were significantly different than zero in EdAR groups and No-EdAR groups, on all metrics except Time Management. All the collaboration metrics

were higher than zero, indicating that productive group collaboration occurred in the experimental conditions. Comparing between EdAR and No-EdAR groups, participants in EdAR groups had statistically higher ratings of Time Management compared to No-EdAR groups ($W=446$, $p=0.049$), indicating that in No-EdAR groups participants were more likely to run out of time.

Within No-EdAR group sub-conditions, participants in the Non-Hololens condition had significantly higher scores of Technical Coordination compared to Hololens-Simple groups ($W=48$, $p=0.01$). These scores were higher because participants were more likely to use the physical compass to measure magnetic fields. No statistical differences were detected on other collaboration metrics within the No-EdAR groups, and no statistical differences were detected on collaboration metrics between sub-conditions associated with EdAR groups.

Table 4. Measured collaboration dimensions; significant differences ($p<0.05$) from neutral ratings, and between AR and Non-AR groups (* = statistical differences found within the AR or Non-AR subconditions)

Collaboration Metrics Examples of high collaboration	Difference from neutral	Difference EdAR vs. No-EdAR
Sustained mutual understanding <i>Ensure partners understand, are on same page</i>	Both	non sig
Dialogue management <i>Smooth flow of communication, little overlaps</i>	Both	non sig
Information pooling <i>Ask questions to seek each other's perspective</i>	Both	non sig
Reaching consensus <i>Coming to shared understanding / agreement</i>	Both	non sig
Task division <i>Task division balanced and explicit</i>	Both	non sig
Time management <i>Deadline met, detailed time planning</i>	non sig	AR >
Technical coordination <i>All tools used, including physical compass</i>	Both	non sig *
Reciprocal interaction <i>Partners hold equal status, leadership balanced</i>	Both	non sig

6.4 RQ4: Does the mere presence of AR technology (without any educational content) affect participants' learning experience?

We address this question by comparing subgroups within the No-EdAR condition (i.e. differences between Non-

Hololens groups which lacked AR technology, and Hololens-Simple groups which used AR technology).

The group lacking AR technology (Non-Hololens group) scored significantly lower on participant **attitude ratings** of Aesthetics ($W=630$, $p<0.01$), Curiosity ($W=626$, $p<0.01$), Endurability ($W=616$, $p<0.01$), Focus ($W=552$, $p=0.02$), Interest ($W=598$, $p<0.01$), Involvement ($W=580$, $p=0.01$), indicating that the mere presence of AR technology created significant gains in motivation.

In **relative learning** gains, there were no statistically significant differences between Non-Hololens and Hololens-Simple groups, as was expected since the learning content was intended to be the same between the two conditions.

There were statistical differences in **collaboration**, specifically for the score of Technical Coordination, whereby Non-Hololens groups scored higher than Hololens-Simple ($W=48$, $p=0.01$). This score is high when participants used the physical compass to measure magnetic fields.

7 DISCUSSION

This study highlights some benefits and drawbacks of using augmented reality in education. In this section we discuss our findings and offer preliminary design guidelines.

7.1 Learning with augmented reality

Participants who saw AR representations of electromagnetism were significantly more effective in developing understanding of the invisible structures of magnetic fields, understanding the connection between electrical currents and magnetic fields, transferring knowledge on how to construct electromagnets, and finishing the task on time (compared to participants who did not see them). All experimental groups had access to the same basic information from printed materials showing static visualizations of electromagnetic fields. However, the AR groups differed on several factors, including: the availability of additional AR-based representations, the alignment of the physical components to their virtual representations, the dynamic nature of virtual representation and the aesthetically engaging nature of the visualizations. Users could concurrently observe the direction of electricity while watching magnetic field shapes, thus experiencing concurrent exposure to two learning concepts (magnetic field shapes and their relationships to electricity and magnetism), which would

explain significant learning gain differences in those topics. Providing such dynamic representations aligned in a physical context through AR, allows learners to easily keep track of relevant information while exploring the dynamic nature of relationships between important variables.

7.2 When no AR is just as helpful

Results indicate that participants in all experimental conditions, regardless of the presence or absence of AR, did experience significant learning gains, positive attitudes, and positive collaboration. This indicates that the non augmented-reality experience is beneficial in itself as an effective learning activity for teaching electromagnetism concepts. Adding AR educational visualizations to an already effective experience may not always be valuable for learning. Our analysis did not find that AR representations were valuable for multiple metrics of collaborative learning. Conditions not involving AR representations were just as effective at motivating students (as measured by the engagement survey); just as effective at fostering collaboration (on all measured dimensions, except for time management); just as effective, or even more effective, at learning concepts such as the effect of amplifiers, and relationships between physical movement and magnetic fields / electricity. Developing AR applications is a costly endeavor requiring the creation of interactive 3D experiences through use of specialized engineering skills and expensive technologies (in the case of the Microsoft HoloLens). This study provides a first step toward critically investigating which educational topics this technology is valuable for, and for which topics other lower-cost approaches may be more suitable.

7.3 When AR prevents kinesthetic learning

Having AR educational representations was detrimental for some factors. The groups that had AR educational content scored lower on understanding the relationship between magnetic field and physical movement. Participants who wore the HoloLens device (even those groups which never saw AR educational representations) tended to gloss over the physical compass or poster. The non-use of compass is partly explainable by the availability of the magnetic field representations, which prevented users from measuring magnetic fields. Overall, our findings indicate that HoloLens participants focused less on physical materials and sensations (i.e. the feeling of movement caused by magnetic field forces). This was likely caused by highly stimulating AR visualizations, which may have prevented learners from focusing on more kinesthetic information, and indicates that, even when multiple representations are

presented together, the salience and ease-of-use influences how well participants integrate the representations.

7.4 New technology effects on engagement

Student engagement is a critical component of learning, and augmented reality has been shown to be highly motivational for users. It is unclear whether high engagement is due to the novelty of the technology or the nature of immersive representations. In our study we assessed different dimensions of engagement to identify which aspects are affected by AR technology. We found that multiple dimensions of engagement are impacted by simple AR representations. All experimental conditions involving HoloLens were not statistically different from each other (regardless of the presence or absence of AR educational representations), and the condition lacking AR technology was significantly lower on multiple dimensions of engagement: aesthetics, curiosity, endurance, focus, interest, involvement. This indicates that AR visualizations - regardless of any educational content - have an effect on user engagement. Furthermore, learning gain were not significantly different across conditions on several concepts, even though participant engagement was high. This observation brings a critical perspective on research findings of AR engagement, highlighting that engagement may be simply due to the exposure to new technology or new visualization techniques, irrespective of the presence of educational content. Engagement was significantly higher in AR conditions involving educational representations, but not across all metrics: when participants used the system with AR educational representations, they only felt a deeper sense of aesthetics and involvement. Overall, this suggests that augmented reality generally has the power to increase motivation, but AR educational content only add to motivation in specific dimensions.

7.5 Impacting STEM attitudes

The conditions involving AR representations of electromagnetism were significantly more effective at changing student self efficacy towards physics, as measured by pre and post self-ratings on items such as “I easily learn physics topics” and “I am the type of student who does well in physics”. Student attitudes towards their own abilities play a significant role in driving student perseverance in learning difficult topics and in guiding their future careers. The short exposure to AR educational content in our study significantly changed student attitudes towards their own learning, indicating that repeated exposure to such experiences might have the power to

change the types of fields people choose to engage in - which is crucial in a time where STEM fields are showing a lack of participation from underrepresented groups.

7.6 Representational Misconceptions

Through interviews and qualitative observations, we found that some participants had trouble understanding the AR representations. This may occur when participants lack some basic background knowledge of the concepts taught: for example one issue we observed was that users did not know how to make sense of the magnetic field since they had no prior exposure, to this this type of representation. This led to problems such as interpreting field strength based on the size of the magnetic field lines rather their density. In a classroom setting, teachers should make sure that more explanations are provided about interpreting the AR visualizations so that students do not develop misconceptions. We are planning to further investigate this effect in future work.

7.7 Sequential layering of information

We tested two methods of layering educational representations: one in which all representations were presented from the start (AR Full) and another in which layers of electricity, magnetism, and poster information, were presented sequentially (AR Scaffold). Our analysis did not find statistically significant differences between these two experimental conditions. Informal observations suggest that AR-Full participants seemed to be overwhelmed by the amount of information and had difficulty noticing important events in the system (e.g., changes in the magnetic field). In future analysis, we plan to more deeply analyze the differences between these conditions and the benefits of layering information, especially as it relates to different kinds of learners. Existing research has shown that progressively revealing more information is beneficial for increasing student curiosity, lowering cognitive load, and increasing learning gains [35]. The layering of representations may be helpful for novices to imagine and understand invisible phenomena, but there are concerns of increased student reliance on educational representations. Students may leave the experience feeling excited and having a high perception of their own knowledge, but not being able to apply the knowledge in situations where AR scaffolds are not available. Future research should investigate the benefits of layered presentation through AR, focusing on sequential addition of educational representations, as well as sequential removal of representations.

7.8 Limitations Imposed by AR Technology

The technology used in this study constrained the depth of interactivity in the educational experience. The AR system could not easily track the movement or states of physical objects, thus restricted opportunities to create simulations that accompany more open-ended inquiry, such as participants connecting wires in different ways, or exploring effects of moving the membrane in 3D space. With advances in AR tracking technology we expect that more interactive experiences can be created for inquiry-based learning. Furthermore, a limitation of AR headsets is that they cover participants' faces, thus reduce ability to make eye contact or communicate using nonverbal emotional expressions. Although our study did not detect differences in overall collaboration caused by AR headsets, future research should investigate these differences in other contexts where nonverbal communication may contribute to successful collaboration.

8 ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1748093. We would like to thank all our study participants, the anonymous reviewers, the Harvard Decision Science Lab, and all the dedicated research assistants who helped in the research data collection, including Adam Petty, Alice Huang, Iva Markicevic, Francisca Astudillo and Yanru Wang.

REFERENCES

- [1] Belcher, J. W., & Besette, R. M. (2001). MIT educators share success. *SIGGRAPH Computer Graphics*, 35(1), 18–21.
- [2] 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(7 Suppl. 1), S12–S23
- [3] Milgram, P., Takemura, H., Utsumi, A. and Kishino, F., 1995, December. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies* (Vol. 2351, pp. 282-293). International Society for Optics and Photonics.
- [4] Dunleavy, M. and Dede, C., 2014. Augmented reality teaching and learning. In *Handbook of research on educational communications and technology* (pp. 735-745). Springer, New York, NY.
- [5] Ibáñez, M.B., Di Serio, Á., Villarán, D. and Kloos, C.D., 2014. Experimenting with electromagnetism using augmented reality: Impact on flow student experience and educational effectiveness. *Computers & Education*, 71, pp.1-13.
- [6] Dünser, A., Walker, L., Horner, H. and Bentall, D., 2012, November. Creating interactive physics education books with augmented reality. In *Proceedings of the 24th Australian computer-human interaction conference* (pp. 107-114). ACM.
- [7] Kaufmann H, Du’nser A (2007) Summary of usability evaluations of an educational augmented reality application. In: *Proceedings of the 2nd international conference on virtual reality*. Springer, pp 660–669
- [8] Martín-Gutiérrez, J., Fabiani, P., Benesova, W., Meneses, M.D. and Mora, C.E., 2015. Augmented reality to promote collaborative and autonomous learning in higher education. *Computers in Human Behavior*, 51, pp.752-761.

- [9] O'Brien, H. L., E. G. Toms, E. K. Kelloway, and E. Kelley, "Developing and evaluating a reliable measure of user engagement," *Proceedings of the American Society for Information Science and Technology*, vol. 45, no. 1, pp. 1–10, Jan. 2008.
- [10] Valimont RB, Vincenzi DA, Gangadharan SN, Majoros AE (2002) The effectiveness of augmented reality as a facilitator of information acquisition. In: *Digital avionics systems conference*, vol 2, Irvine, CA, USA, pp 7C5-1–7C5-9
- [11] Tang A, Owen C, Biocca F, Mou W (2003) Comparative effectiveness of augmented reality in object assembly. In: *Proceedings of the conference on human factors in computing systems—CHI'03*, p 73
- [12] Bujak, K.R., Radu, I., Catrambone, R., Macintyre, B., Zheng, R. and Golubski, G., 2013. A psychological perspective on augmented reality in the mathematics classroom. *Computers & Education*, 68, pp.536-544.
- [13] Radu, I., 2014. Augmented reality in education: a meta-review and cross-media analysis. *Personal and Ubiquitous Computing*, 18(6), pp.1533-1543.
- [14] Cuendet, S., Bonnard, Q., Do-Lenh, S. and Dillenbourg, P., 2013. Designing augmented reality for the classroom. *Computers & Education*, 68, pp.557-569.
- [15] Wijdenes, P., Borkenhagen, D., Babione, J., Ma, I. and Hallihan, G., 2018, April. Leveraging Augmented Reality Training Tool for Medical Education: a Case Study in Central Venous Catheterization. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (p. CS11). ACM.
- [16] M. Dunleavy and C. Dede, "Augmented reality teaching and learning," in *Handbook of research on educational communications and technology*, Springer, 2014, pp. 735–745.
- [17] A. Morrison et al., "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*, 2009, pp. 1889–1898.
- [18] Roth, W.M., 2001. Gestures: Their role in teaching and learning. *Review of educational research*, 71(3), pp.365-392.
- [19] Beheshti, E., Kim, D., Ecanow, G. and Horn, M.S., 2017, May. Looking inside the wires: Understanding museum visitor learning with an augmented circuit exhibit. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 1583-1594). ACM.
- [20] O'Malley, C. and Fraser, D.S., 2004. Literature Review in Learning with Tangible Technologies. 2004. NESTA Futurelab.
- [21] Chan, J., Pondicherry, T. and Blikstein, P., 2013, June. LightUp: an augmented, learning platform for electronics. In *Proceedings of the 12th International Conference on Interaction Design and Children* (pp. 491-494). ACM.
- [22] Bellucci, A., Ruiz, A., Díaz, P. and Aedo, I., 2018, May. Investigating augmented reality support for novice users in circuit prototyping. In *Proceedings of the 2018 International Conference on Advanced Visual Interfaces* (p. 35). ACM.
- [23] D. N. E. Phon, M. B. Ali, and N. D. A. Halim, "Collaborative augmented reality in education: A review," in *Teaching and Learning in Computing and Engineering (LaTiCE)*, 2014 International Conference on, 2014, pp. 78–83.
- [24] Antle, A.N., Droumeva, M. and Ha, D., 2009, June. Hands on what?: comparing children's mouse-based and tangible-based interaction. In *Proceedings of the 8th International Conference on Interaction Design and Children* (pp. 80-88). ACM.
- [25] Cai, S., Chiang, F.K., Sun, Y., Lin, C. and Lee, J.J., 2017. Applications of augmented reality-based natural interactive learning in magnetic field instruction. *Interactive Learning Environments*, 25(6), pp.778-791.
- [26] Mannuß, F., Rubel, J., Wagner, C., Bingel, F. and Hinkenjann, A., 2011, October. Augmenting magnetic field lines for school experiments. In *Mixed and Augmented Reality (ISMAR)*, 2011 10th IEEE International Symposium on (pp. 263-264). IEEE.
- [27] 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(1), 63–86.
- [28] Unfried A, Faber M, Wiebe E. Gender and student attitudes toward science, technology, engineering, and mathematics. The Friday Institute for Educational Innovation at North Carolina State University. 2014.
- [29] DesPortes, K., Anupam, A., Pathak, N. and DiSalvo, B., 2016, October. Circuit diagrams vs. physical circuits: The effect of representational forms during assessment. In *Frontiers in Education Conference (FIE)*, 2016 IEEE (pp. 1-9). IEEE.
- [30] Cuendet, S., Bumbacher, E., & Dillenbourg, P. (2012, October). 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
- [31] Ishii, H. and Ullmer, B., 1997, March. Tangible bits: towards seamless interfaces between people, bits and atoms. In *Proceedings of the ACM SIGCHI Conference on Human factors in computing systems* (pp. 234-241). ACM.
- [32] Ainsworth, S. 2006. DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction* 16, 3: 183–198.
- [33] Schneider, B., Strait, M., Muller, L., Elfenbein, S. Shaer, O., and Shen, C. 2012. Phylo-Genie: engaging students in collaborative "tree-thinking" through tabletop techniques. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems (CHI '12)*, 3071–3080.
- [34] Antle, A.N., Wise, A.F. and Nielsen, K., 2011, June. Towards Utopia: designing tangibles for learning. In *Proceedings of the 10th International Conference on Interaction Design and Children* (pp. 11-20). ACM.
- [35] Fyfe, E.R., McNeil, N.M., Son, J.Y. and Goldstone, R.L., 2014. Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review*, 26(1), pp.9-25.
- [36] I. Radu, E. Doherty, K. DiQuollo, B. McCarthy, and M. Tiu, "Cyberchase shape quest: pushing geometry education boundaries with augmented reality," in *Proceedings of the 14th International Conference on Interaction Design and Children*, 2015, pp. 430–433.
- [37] S. H. Kidd and H. Crompton, "Augmented learning with augmented reality," in *Mobile Learning Design*, Springer, 2016, pp. 97–108.
- [38] M. G. Hanna, I. Ahmed, J. Nine, S. Prajapati, and L. Pantanowitz, "Augmented Reality Technology Using Microsoft HoloLens in Anatomic Pathology," *Archives of pathology & laboratory medicine*, vol. 142, no. 5, pp. 638–644, 2018.