



Augmented Reality Technology Used for Developing Topographic Map-Reading Skills in an Earth Science Course and its Potential Implications in Broader Learning Venues

Sarah Baumann¹ · Leilani A. Arthurs¹

Accepted: 16 November 2022

© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract

Topographic map-reading skills are critical for certain professions but can be difficult to learn. The purpose of this pilot study is to provide insight on the role augmented reality technology can play in the development of topographic map-reading skills. Using a situated cognition theoretical framework, this study tracks the development of students' skills in three different instructional approaches using the Topographic Mapping Assessment (TMA), instructor observations, and student feedback. Using a quasi-experimental research design, 85 college-level students in eight sections of an introductory undergraduate geoscience laboratory course were assigned to a control group ($n = 19$) that was instructed using the standard curriculum (paper-and-pencil lab exercises and field trips), a 2-D group ($n = 14$) that completed six activities using 2-D maps, or an augmented reality sandbox (ARS) group ($n = 52$) that completed six activities requiring both 2-D maps and augmented reality technology. Results from multi-level analyses of covariance suggest no significant difference in overall post-instruction scores, except female students in the ARS groups ($n = 17$) tended to score higher than students in the control group ($n = 11$), potentially indicating this method can increase outcomes for females in STEM. Other identified instructional benefits of using the ARS include increased collaboration between students, greater visibility to the instructor of student difficulties and challenges, and improved ability for the instructor to provide real-time feedback and guidance.

Keywords Topographic map-reading instruction · Map skill development · Augmented reality sandbox · Situated cognition theory

Introduction

Topographic maps are important tools for many professions, including geoscientists, geographers, landscape designers, and emergency responders (Newcombe et al., 2015). To effectively use a topographic map, the reader must both understand the symbolic features of the map and use spatial skills to mentally convert a 2-D map into a 3-D landscape (Atit et al., 2016). Novice readers struggle with both steps (Rapp et al., 2007). While teaching symbology is relatively easy, teaching spatial skills can be challenging. During more than 50 years of research from diverse disciplines, including cartography, psychology, and geology (Lobben, 2004), scholars studied how

experts read topographic maps (Gilhooly et al., 1988; Ishikawa & Kastens, 2005; Pick et al., 1995) and how they differ from novices. Despite this, gaps remain in our understanding of how instruction can more effectively facilitate the development of expert-like topographic map-reading skills among novices, especially regarding new technologies.

While spatial skills are thought to be a good predictor of success in science, technology, engineering and mathematics (STEM) (Wai et al., 2009), existing research also suggests students' spatial skills can improve with training (Newcombe & Shipley, 2015; Uttal et al., 2013). However, training that develops students' topographic map-reading and other spatial skills are rarely taught in primary or secondary school (Kastens & Ishikawa, 2006). Thus, incoming college students have a range of spatial abilities and topographic map experience (Gold et al., 2018; Ormand et al., 2014). There is an indication that students with more experience or exposure may score higher on topographic map assessments (Atit et al., 2016; Giorgis et al., 2017; Murakoshi & Higashi, 2015). Downs and Liben (1991)

✉ Leilani A. Arthurs
leilani.arthurs@colorado.edu

¹ Department of Geological Sciences, University of Colorado
Boulder, Boulder, CO, USA

hypothesized that without adequate guidance, only incoming students with stronger abilities will develop more expert-like spatial skills and better succeed in advanced STEM courses and careers.

In addition, research findings are mixed in their conclusions about the relationship between gender, topographic map-reading skills, and spatial skills. Some studies show no difference in topographic map-reading skills between genders (e.g., Moore, 2018; Murakoshi & Higashi, 2015), while others show one gender scores higher than the other (Carbonell-Carrera et al., 2017; Giorgis et al., 2017; Jackson et al., 2019; Newcombe et al., 2015). Research indicates these differences may be due to previous exposure to games and toys that develop spatial skills, (e.g., construction-based toys, video games; Gold et al., 2018). Regardless of the reason, further research can help identify how instruction can assist groups of students with a range of incoming experiences and skills to develop topographic map-reading.

This study investigates one method for facilitating the development of introductory-level undergraduate students' topographic map-reading skills. The method involves using an augmented reality sandbox (ARS) and related learning environment as a cognitive scaffold between a 2-D paper map and a corresponding 3-D landscape constructed in a sandbox (Holton & Clarke, 2006). The ARS concept, developed by UC Davis' W.M. Keck Center for Active Visualization in the Earth Sciences (KeckCAVES), is a physical sandbox with an Xbox that dynamically projects topographic contour lines according to the elevation of the sand. The ARS allows for multi-modal instruction where users can interact tactilely with the sand, observe how projected contour lines change as they reshape the landscape, and instantaneously test out questions and predictions they might have about how the geometry of contour lines changes as they shape and reshape the landscape.

Although the geoscience education community is excited by the potential of the ARS, studies to date have yielded inconclusive findings about its association with map-reading skills (Giorgis et al., 2017; Evans et al., 2018; Richardson et al., 2018; Jackson et al., 2019; McNeal et al., 2020). Studies that investigated learning gains after a single topographic map activity with an ARS show positive associations with affective outcomes but suggest no significant learning gains over traditional 2-D activities (Hale et al., 2019; Jackson et al., 2019; Soltis et al., 2020; Woods et al., 2016). Many have suggested that increasing time or more appropriate activities may be key to effective instruction with an ARS (Giorgis et al., 2017; Jackson et al., 2019; McNeal et al., 2020; Soltis et al., 2020).

This pilot study investigates the association between using an ARS multiple times and the development of students' topographic map-reading skills. The study takes place over a 16-week period in the naturalistic setting of

introductory-level geology labs. The research questions guiding this study are:

Q1: To what extent does utilizing augmented reality technology facilitate the development of topographic map-reading skills?

Q2: Is instruction with augmented reality technology associated with different levels of map-reading skill development for different genders?

Theoretical Framework

The theoretical framework for the present study is situated cognition theory, as described by Wilson and Myers (2000). Situated cognition theory asserts that learning and knowledge are indelibly linked to the context or environment in which they are acquired. Attention to the learning environment, including “language, activities of individuals and groups, cultural meanings and differences, tools (including computer tools and environments), and the interaction of all of these together” are crucial components of the learning system (Wilson and Meyers, 2000, p. 70).

Of particular interest to this study are the interactions between individuals and tools. Tools help to build “the mediums, forms, or worlds through which cognition takes place” (Wilson and Meyers, 2000, p. 71). This study views the augmented reality sandbox (ARS) as an instructional tool that can be used to facilitate students learning about topographic maps and develop their topographic map-reading skills. In this study, the interactions between individuals, including student-student and student-instructor interactions are key components of teaching and learning. As such, this study examines the context in which learning occurs with an ARS.

Research Design and Methodology

To capture the role of tools and classroom interactions, this study uses a quasi-experimental research design and mixed-method research methodology. In contrast with true experimental design, quasi-experimental designs do not randomly assign students to intervention groups. This design choice can limit the ability to infer causality; however, it is an appropriate choice depending on the research question and available resources (including student access). For this study, the quasi-experimental design allows for research to take place within the naturalistic settings of classrooms and the field (i.e., geology field trips) while ensuring equal treatment of students in the same course or section. In this study, students' assignments to the intervention groups were not random because section enrollment determined group assignment and students self-selected to enroll in their chosen section.

Additionally, this study utilizes a mixed-method methodology in order to best answer the research questions. A mixed-method methodology gathers and analyzes both quantitative and qualitative data. This allows researchers to develop a more complete understanding of a situation and benefit from the strengths of both types of data, including breadth and replicability provided via quantitative methods and narrative depth through qualitative observations. In this study, quantitative results come from surveys and assessments, while qualitative instructor observations and student feedback provide a richer understanding of what occurred in the classroom between quantitative measures.

Methods

Setting

This study took place in an introductory physical geology laboratory course at a large public university (undergraduate population 27,000). The one-credit course fulfills a general science education requirement and is popular among non-geology majors. This course can be taken independent of an introductory geology lecture course. All sections are taught by graduate student instructors, and each section has a maximum enrollment of 20 students. The course consists of a 3-hour in-class exercise or a field trip once per week for 16 weeks.

In addition to their in-class work, students are expected to complete a preparatory pre-lab activity prior to each class meeting. Eighty-five students in eight different lab sections, taught by six different graduate instructors, participated in this study.

Student Population

Eighty-five students, out of the 119 enrolled (71% participation rate), consented to participate in this study. Of these students, five percent were geology majors and 27% were STEM majors. The majority of students (85%, $n = 72$) took the course to fulfill the general education requirement described above. Table 1 displays the demographics of students in each intervention group. All demographic fields were self-reported by students. For gender, students were asked to pick between “Female,” “Male,” “Non-binary/third gender,” “Prefer not to say,” or to self-describe another gender.

Three Instructional Approaches

The study tracked the developmental progress of three groups. All three groups received standard instruction consisting of lab manual geoscience exercises and field trips during the 16-week semester. Standard instruction for each weekly 3-hour class meeting typically includes (1) a pre-lab

Table 1 Demographics of students in each intervention group

	ARS $n = 52$	2-D $n = 14$	Control $n = 19$	Total $n = 85$	Institution $n = 26,643$
Undeclared ($n = 2$)	2%	0%	5%	2%	
Non-STEM major ($n = 56$)	67%	71%	58%	66%	
STEM major other than geology ($n = 23$)	27%	21%	32%	27%	
Geology major ($n = 4$)	4%	7%	5%	5%	
Male ($n = 48$)	65%	43%	42%	56%	56%
Female ($n = 35$)	33%	50%	58%	41%	44%
Other ($n = 2$)	2%	7%	0%	2%	NA
Asian ($n = 6$)	10%	0%	5%	7%	8%
Black or African American ($n = 3$)	4%	0%	5%	3%	3%
Hispanic or Latino ($n = 8$)	10%	7%	11%	9%	12%
White ($n = 58$)	62%	86%	74%	67%	67%
More than One Race or Ethnicity ($n = 5$)	8%	7%	0%	6%	NA
Other Races and Ethnicities ($n = 3$)	8%	0%	0%	4%	8%
Prefer not to say ($n = 2$)	2%	0%	5%	2%	NA
Freshman ($n = 19$)	21%	14%	32%	22%	19%
Sophomore ($n = 34$)	38%	64%	26%	40%	24%
Junior ($n = 16$)	19%	7%	26%	19%	22%
Senior ($n = 14$)	17%	14%	16%	16%	26%
5 + year Senior ($n = 2$)	4%	0%	0%	2%	9%
First-Generation Student ($n = 8$)	13%	0%	5%	9%	17%
Not First-Generation Student ($n = 77$)	87%	100%	95%	91%	83%

homework assignment completed at home, (2) 20–30 min of in-class lecture, and (3) an in-class paper-and-pencil lab exercise. Depending on the class meeting, the in-class lab exercise is either an in-the-classroom or a field-trip exercise.

The control group only received standard instruction and completed the topographic map lab exercise described in the lab manual. The intervention groups received modified instruction (2-D group) or modified instruction with additional instruction (ARS group). The 2-D group received standard instruction in which one lab exercise and five pre-lab homework assignments were adapted to further incorporate 2-D topographic map-reading exercises. The ARS group received standard instruction in which one lab exercise and five pre-lab homework assignments were adapted to incorporate 2-D topographic maps and the ARS. Figure 1 outlines the instructional schedule for each group in the study. The primary instructor for each section independently determined how they incorporated participation and completion of activities into students' grades. Half of the instructors gave credit for completion, while the other half assigned scores based on the accuracy of responses.

The only topographic-map lab exercise in the standard instruction was used to develop the modified in-class lab exercise for the 2-D and ARS groups. All three groups completed one of three versions of a topographic-map lab exercise early in the semester. The 2-D and ARS groups, during the second half of the semester, also completed five modified pre-lab homework assignments before class meetings with a field trip. Students in the 2-D group completed the pre-lab homework assignment at home. Students in the ARS group completed the pre-lab homework assignment during special ARS office hours in a classroom/lab setting.

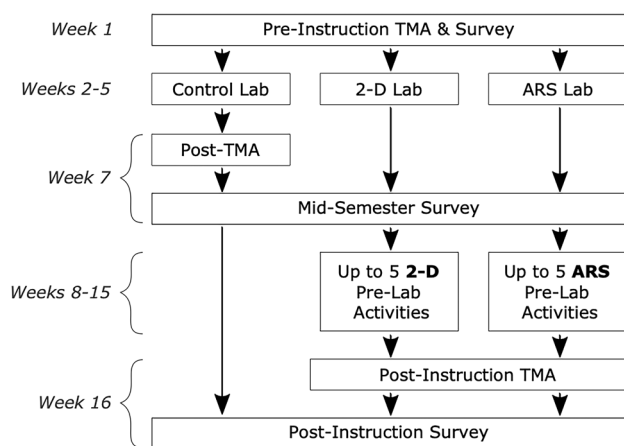


Fig. 1 General order of activities. Week numbers represent the range in which a given stage was completed. Not all sections completed an activity every week

The modified pre-lab homework assignments for 2-D and ARS groups were developed using the standard in-class lab exercises for each field trip as a starting point. Each field site was used to design a pre-lab homework assignment aimed at highlighting certain topographic features in that area and/or topographic map conventions. Furthermore, the modified pre-lab homework assignments for the 2-D and ARS groups were designed with analogous questions between groups. To complete the modified pre-lab homework assignments for the ARS group, students started by using 2-D maps, then used an ARS, and often ended with a 2-D map intended to aid the transfer of students' map-reading skills to a more traditional medium. Each ARS-based pre-lab homework assignment took between 10 and 20 min. Students completed these assignments with instructor and peer assistance. The same instructor held all ARS office hours. Additional details describing the individual activities can be found in the supplementary online resources.

Data Collection

Topographic Map Assessment

The topographic mapping assessment (TMA) was used to evaluate topographic map-reading skills. Jacovina and colleagues (2014) designed the TMA to evaluate novice map readers' ability to complete a range of topographic map-reading skills including identifying rivers, selecting appropriate topographic profiles, and determining what is visible from a given point on a map. Additionally, Newcombe and colleagues (2015) assessed the TMA for reliability and suitability for a range of novice female topographic map readers using item response theory. For the final data analysis in this study, the TMA and its accompanying scoring rubric were modified. Three TMA items were removed because of concerns about question clarity and consistency in grading. The scoring rubric for seven TMA items were revised to clarify the range of acceptable responses, resulting in a maximum TMA score of 20 points. A more detailed discussion of the revisions made to the TMA and accompanying scoring rubric can be found in the supplementary online resources.

The TMA was administered on the first day of class, prior to any instruction, and again during the midterm (control group) or the final exam (2-D and ARS groups). Although the pre-instruction TMA scores were not factored into students' grades, each instructor handled the post-instruction TMA differently. Four sections completed the post-instruction TMA for credit (both 2-D sections and two ARS sections), one was not for course credit (one ARS section), and three were given during a test but not part of their test grade (two control sections and one ARS section).

Instructor Observations

Throughout the semester, the ARS instructor compiled notes about (1) her and other instructor's observations of students' progress and (2) her interactions with instructors, other students, and course materials. After facilitating each ARS-based lab exercise or pre-lab homework assignment, the instructor reflected on the strengths and weaknesses of the activity, students' comments and attitudes, and the peer-to-peer and instructor-interactions that occurred in the classroom. At the end of the semester, these observations were compiled and analyzed for common trends and changes over time.

Student Feedback

Students completed different surveys at the beginning, middle, and end of the semester, which asked about their interests, hobbies, demographic information, previous experience, and opinions of the course. These surveys were adapted from surveys used in earlier studies by Thorndyke and Goldin (1981), Atit and colleagues (2016), and Gold and colleagues (2018). In addition to these surveys, students were also given the opportunity to provide feedback on ARS activities after their last pre-lab homework assignment with the ARS.

Data Analysis

Given some students added the course after the first week of the semester when the pre-instruction TMA was administered and some students were absent from class on the day the post-instruction TMA was administered, matched pre/post-instruction TMA was collected for a total of 80 students. Similarly, not all students completed every pre-lab homework assignment or lab exercise. Thus, the number of comparable activities completed by students in the 2-D and ARS groups varied widely, from one to six activities, although the majority completed three or more (84%).

A single grader scored all matched pre/post TMA responses. A second grader independently graded 12% of the multiple-choice responses ($n = 20$) and 100% of free-response questions on the TMA. Initial inter-grader agreement was low due to difficulty judging free-response questions. To code these items with greater reliability, a transparency overlay rubric was developed. Each overlay contains regions suitable for correct and partially correct responses. With the overlays to aid in scoring, inter-grader agreement was 92%. Any discrepancies in grades were discussed and reconciled, resulting in 100% inter-grader agreement.

Once coded, descriptive statistics of the TMA scores and sub-scores were calculated for students in each intervention

group and demographic group. To assess group differences in pre-instruction TMA scores, both one-way and factorial analyses of variance (ANOVA) were used. For the association between each instructional approach and TMA scores, this study utilized a multi-level model to account for the nested nature of the quasi-experimental design. Preliminary examination determined pre-instruction TMA score does not interact significantly with any other predictor variables. Additionally, the random effect of gender was not significant. The ultimate model for the main multi-level ANCOVA was

$$\begin{aligned} postTMA_{ij} = & \alpha_{00} + \mu_{0j} + \alpha_{01}Group1_j + \alpha_{02}Group2_j \\ & + (\alpha_{10} + \mu_{1j})preTMA_{cij} + \alpha_{20}S_{ij} \\ & + \alpha_{21}S_{ij}Group1_j + \alpha_{22}S_{ij}Group2_j + \epsilon_{ij}. \end{aligned}$$

The dependent variable was post-instruction TMA score for each student (i) in each section (j). The fixed effects for this model were mean-centered pre-instruction TMA score ($preTMA_c$), contrast-coded gender (S) of each student, two orthogonal intervention group predictors ($Group1$, $Group2$), and the interaction between gender and intervention group. Additionally, there were two random effects for section (μ_{0j}) and pre-instruction TMA score (μ_{1j}). Different codes were used to find the simple effects of gender and intervention group. The lme4 package (Bates et al., 2014) in R (version 3.6.3) was used to estimate these models. This method uses the restricted maximum likelihood approach to determine model fit and Satterthwaite's method to calculate significance tests and degrees of freedom.

Results

Student Participation

All 85 students completed the post-instruction survey and TMA, but only a subset completed the other assessments. This resulted in 80 students with both pre- and post-instruction TMA scores. Initial descriptive statistics (Table 2) and significance tests indicated there was no difference in pre-instruction TMA scores of students in each section of the course ($t_{(72)} = 1.17$, $p = 0.233$) or each group in the study ($t_{(77)} = 0.46$, $p = 0.806$). A subset of students completed each survey; 63 students completed the beginning-of-the-semester survey, 65 completed the mid-semester survey, and 13 provided additional feedback about ARS activities at the end of the semester.

Trends in TMA Scores

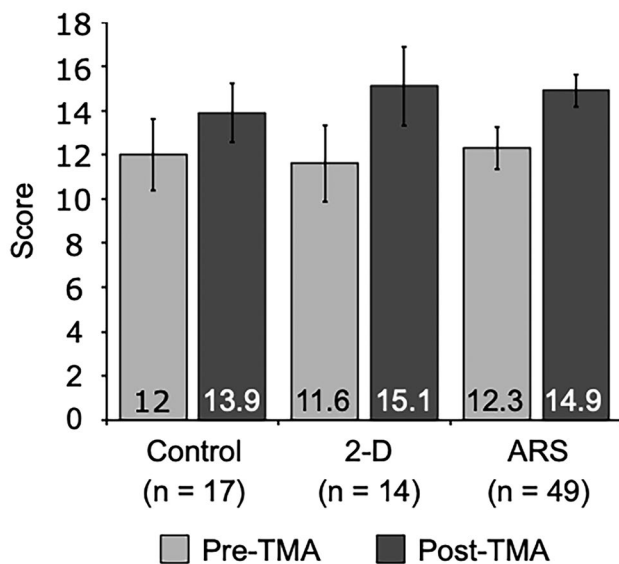
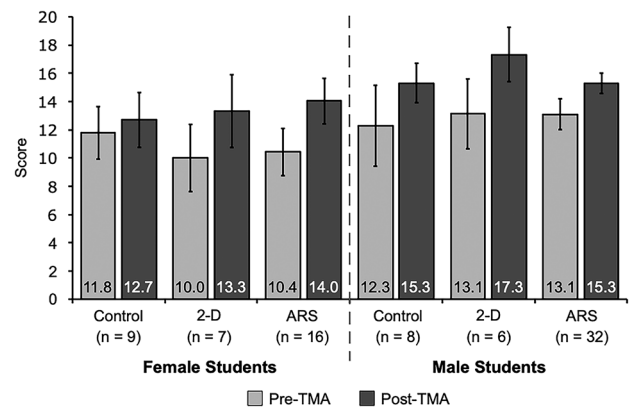
Students in the ARS group and the 2-D group had higher average post-instruction TMA scores than the control group (Table 2 and Fig. 2). Results from single degree-of-freedom

Table 2 Descriptive statistics for sections and intervention groups. Values are average and standard deviation (in parentheses) for each group

Section	Intervention	<i>n</i>	Pre-TMA	Post-TMA	Activities
1	Control	7	11.0 (3.2)	14.0 (2.3)	1.0 (NA)
2	Control	10	12.8 (3.5)	13.9 (3.3)	1.0 (NA)
	Control	17	12.0 (3.4)	13.9 (2.8)	1.0 (NA)
3	2-D	7	10.0 (3.3)	13.0 (3.2)	5.0 (1.9)
4	2-D	7	13.1 (2.7)	17.2 (2.3)	4.8 (1.2)
	2-D	14	11.6 (3.3)	15.1 (3.4)	4.9 (1.5)
5	ARS	14	12.2 (3.7)	15.3 (2.7)	5.2 (0.8)
6	ARS	10	11.1 (4.5)	15.0 (2.7)	2.3 (1.6)
7	ARS	10	14.2 (2.3)	15.3 (1.9)	4.0 (1.5)
8	ARS	15	11.8 (2.7)	14.1 (2.8)	5.5 (1.1)
	ARS	49	12.3 (3.4)	14.9 (2.6)	4.4 (1.7)

tests suggest no differences between intervention groups when the pre-instruction TMA score was controlled. The 2-D group had the highest adjusted post-instruction TMA score, but it was not significantly higher than the ARS group ($M_{\text{diff}}=1.08$, $t_{(14.2)}=-1.15$, $p=0.269$) nor the control group ($M_{\text{diff}}=1.71$, $t_{(15.2)}=-1.75$, $p=0.100$). The adjusted post-instruction average TMA score for the ARS group was larger than the control group's ($M_{\text{diff}}=0.63$, $t_{(12.7)}=0.96$, $p=0.357$), but the difference was not statistically significant.

As previous research has been inconclusive on the differences between the topographic map-reading skills of male and female students, this study compared the scores of different genders in each intervention group (Fig. 3). As a whole, female students scored significantly lower than

**Fig. 2** Average pre- and post-instruction TMA scores for each of the three intervention groups. Error bars represent margin of error**Fig. 3** Average pre- and post-instruction TMA scores for male and female students in each of the three intervention groups. Error bars represent margin of error

male students on the pre-instruction TMA ($M_{\text{diff}}=2.09$, $t_{(74)}=2.41$, $p=0.019$), and the post-instruction TMA, ($M_{\text{diff}}=2.62$, $t_{(74)}=3.79$, $p<0.001$). Even after accounting for pre-instruction TMA score, this statistically significant difference remained ($M_{\text{diff}}=1.54$, $t_{(70.9)}=2.58$, $p=0.012$).

Results from multi-level ANCOVA indicate an interaction between intervention group and gender. The difference between the adjusted post-instruction TMA score of male and female students in the ARS group ($M_{\text{diff}}=0.22$) is less than that of the 2-D group ($M_{\text{diff}}=2.10$, $t_{(67.6)}=-1.41$, $p=0.163$) and the control group ($M_{\text{diff}}=2.31$, $t_{(72.4)}=-1.71$, $p=0.092$). Female students in the ARS group scored higher than female students in the control group ($M_{\text{diff}}=1.68$, $t_{(24.4)}=1.80$, $p=0.084$), a marginally significant difference, but slightly lower than female students in the 2-D group ($M_{\text{diff}}=0.14$, $t_{(24.7)}=0.11$, $p=0.918$). In contrast, male students in the ARS group scored lower than male students in the control group ($M_{\text{diff}}=0.41$, $t_{(37.0)}=-0.46$, $p=0.645$) and male students in the 2-D group ($M_{\text{diff}}=2.01$, $t_{(33.6)}=-1.71$, $p=0.100$) but these differences were not statistically significant.

Instructor Observations and Student Feedback

Beyond TMA scores, instructor observations and student feedback provide insight on the role of augmented reality in a classroom and shifts in teaching and learning over time. The ARS instructor and regular instructors for students in the ARS or 2-D groups observed their students seemed more adept at topographic mapping tasks. While on field trips, regular instructors noted self-location tasks went smoother and faster than they had in previous semesters in the absence of the modified instruction students in the ARS and 2-D groups received. In addition to their performance in the field, the ARS instructor noticed progress in students' topographic map-reading skills from the beginning to the end of

the semester. On early topographic map activities, students struggled with basic concepts; however, they were more fluid in their responses and had a better grasp of relevant terminology by the end of the semester. For example, one of the first ARS activities asked students to draw a simple topographic map based on what they saw on the ARS. This request confused some students, who attempted to draw both the 3-D sand landscape and the contour lines. By the end of the semester, students better understood how to draw a simple topographic map, indicating growth in understanding the relationship between topographic maps, contour lines, and the real-world landscape. These observed changes, however, did not translate to significant gains over the other two groups in the study, as measured using the TMA.

In addition to practicing topographic map-reading skills, instructor observations and student comments indicate the ARS generally had a positive effect on the classroom environment, although this may wane with time. The first time students used the ARS, many were very excited. Several students asked if they would be able to use it again and three students (four percent of respondents), without any specific prompting, commented that working with the ARS was helpful during the mid-semester survey. Over time, this trend shifted. The end-of-the semester survey revealed students generally had positive feedback about the interactive nature of the course (mentioned by 19%), but three students said they were tired of the ARS activities and working on them outside of class time. When asked explicitly what they thought of the ARS activities, most respondents ($n = 11$, 85%) responded positively—they enjoyed the activities or found them helpful. One commented the ARS was the “least boring way to learn topographic maps.” But some said they found the activities repetitive or that they did not want to come outside of class to do these activities ($n = 5$, 39%).

Finally, the ARS instructor observed the role the ARS can play in instruction. While working with students using the ARS, the instructor found the ARS a useful tool for receiving feedback on student understanding, fostering conversations, and testing hypotheses. The activities and the small-group nature of instruction with the ARS allowed the instructor to supervise the progress students were making and discuss any misunderstandings or misconceptions. Curiosity about the ARS and the maps made using it sparked student questions, allowing for conversations that may not have occurred with a paper map. When students had a question or misunderstanding, the ARS proved a useful tool to test scenarios out. Early in the semester, students treated the ARS as a toy or game but quickly learned to use it as a tool for completing a task. For example, when students were practicing “line-of-sight” tasks, some students built different hills to see how the visibility between two points changed relative to slope and distance. Less instructor intervention and assistance were needed over time as students relied more on their peers or

the ARS as a tool to answer questions before approaching the instructor.

Discussion

Intervention Group Pre/Post TMA Scores

In terms of developing students’ topographic map-reading skills, this study found no significant differences in TMA scores among students who were in the two intervention (2-D and ARS) groups. Students in the 2-D group in this study scored slightly higher than those in the ARS group. This finding is similar to the findings of Moore (2018), Evans and colleagues (2018), and Jackson and colleagues (2019). Moore (2018) proposed their 2-D participants may have performed better because they got more experience practicing visualization, while Jackson and colleagues (2019) suggested the novelty of the ARS may have distracted their students. While this study did not directly assess either of these hypotheses, both could have played a role in the results of this study.

Initial instructor observations support Jackson and colleagues’ (2019) hypothesis that the novelty of the ARS may play a negative role in student learning gains. Students were excited to use and play with the ARS to the extent that it was difficult to focus on the activity. The ARS instructor observed the ARS was no longer as novel after several activities and students started to use the ARS more effectively as a tool. Researchers have suggested that the positive student affect (e.g., higher learning motivation, student engagement) from instruction with an ARS or other forms of augmented reality may improve learning (Bacca et al., 2014; Diegmann et al., 2015; Soltis et al., 2020; Vaughan et al., 2017), but the results of this study indicate this interplay is more complex. Instruction with an ARS has the ability to initially engage and motivate students beyond traditional activities (Jackson et al., 2019; Soltis et al., 2020), but the impact of this on student learning over multiple instructional sessions is an area for future research.

In addition to the role of novelty, the level of scaffolding may play a role in this study’s results. It has been suggested that tools like the ARS lighten the cognitive load for students, allowing them to reach a deeper understanding of the concept in question (Cohen & Hegarty, 2014; Hale et al., 2019; McNeal et al., 2020; Richardson et al., 2018; Taylor et al., 2004). Moore (2018) proposed that this reduction in cognitive load also reduces the visualization practice of ARS participants relative to 2-D participants. Other research on topographic map-reading aids lends support to this hypothesis. Carbonell-Carrera and colleagues (2018) found that students using augmented reality topographic maps rated activities easier and quicker to do than those using 2-D maps. This is also consistent with research that find topographic map-reading aids (e.g., ARS, colored

contour lines, stereo effects) correspond with higher topographic map-reading skill if they are present for the final assessment (Phillips et al., 1975; Potash et al., 1978; Rapp et al., 2007; Richardson et al., 2018), but research that removes the map-reading aid yields insignificant differences (Giorgis et al., 2017; Jackson et al., 2019; Richardson et al., 2018). Similarly, this study found that students progressed considerably during ARS sessions, but the 2-D group ultimately scored higher than the ARS group on the final 2-D assessment. Moore's hypothesis is an area for future research considering several studies, including the present study, have found higher learning gains in 2-D groups (Evans et al., 2018; Moore, 2018; Jackson et al., 2019).

Intervention vs. Control TMA Scores

This study also found no significant difference between the intervention groups (ARS and 2-D) and the control group. Three possible implications of this finding include: (1) the amount of instructional time in this study was long enough for significant learning benefits, (2) the modified instructional materials do facilitate the development of topographic map-reading skills, and (3) the TMA was not able to detect the learning gains in topographic map-reading skills for this population of students. The first two implications have been suggested before (Evans et al., 2018; McNeal et al., 2020). In order to teach a complex skill like topographic map-reading it may take more time or very carefully designed instruction to succeed in a short period of time. However, course instructors for lab sections with students in the intervention groups noticed their students were more adept at topographic map-reading field-tasks than students in courses they taught in previous semesters (i.e., followed the same standard instruction as the control group). The ARS instructor also observed students were more proficient in their use of terminology and problem-solving strategies toward the end of the semester relative to earlier in the semester. The student progress observed by instructors seem to contradict these first two implications.

The final possible implication is an area for future research. Clark and colleagues (2008) discovered that novice map readers are capable of inferences that lead to high scores on topographic mapping assessments without sophisticated topographic map-reading skills. Thus far, every study assessing learning gains with an ARS has used a written test and most used the TMA (Jackson et al., 2019; McNeal et al., 2020; Moore, 2018; Pollack, 2019; Richardson et al., 2018). None of these studies has found a significant difference between ARS and non-ARS interventions. While it is possible that researchers may have not yet found an optimal method for using the ARS to develop students' topographic map-reading skills, it is also possible that instruments used to assess topographic map-reading skills are not

valid and reliable for wider use beyond the study population in which it was developed. For example, Jackson and colleagues (2019) found poor psychometric reliability for the subset of TMA items used in their study and Moore (2018) identified possible ceiling effects in his study. Thus, future studies could be done to (1) further develop the TMA for use with populations beyond the small group of female psychology majors with which it was originally developed and (2) develop additional measures of topographic map-reading skills that, for example, could include field-based tests.

Gender and Skill Development

This study found significant differences in topographic map-reading skills relative to gender. Female students in this study had lower pre- and post-instruction TMA scores than male students, consistent with previous studies (Giorgis et al., 2017; Jackson et al., 2019; Newcombe et al., 2015). This study found female students in the ARS group improved more than female students in the control group, while male students improved the least in the ARS group. Although some researchers are concerned that augmented reality may hinder female students and other students with less experience with computerized displays (Habig, 2019; Niedomysl et al., 2013), our findings suggest female students benefitted from using the ARS. The reasons for this are unclear, signaling an avenue for possible future research.

Instructional Benefits

Observations of the classroom environment shed light on other potential benefits to instruction with an ARS. Diegmann and colleagues (2015) conducted a systematic literature review on augmented reality in education and identified commonly cited benefits, which include supporting student-centered learning, aiding collaborative learning, and increasing interactivity. Observations from instructors and students' comments (as described in the results section) indicate instruction with an ARS also has these benefits.

One major benefit of the ARS as an instructional tool is its role in sparking questions and discussion about topographic maps. Instruction with a single ARS limits the number of students who can use it at the same time due to the size of the ARS. The ARS instructor found this a valuable time to check with students and assess their conceptual understanding. Students' fascination with the technology itself drove many students to experiment, which led them to make observations about the resulting contour lines or ask questions. Overall, incorporating the ARS as an instructional tool contributed to a learning environment that encouraged inquiry and conversation.

A second instructional benefit to the ARS is curriculum design. During this study, many new topographic

map-reading activities were developed for the 2-D and the ARS groups. To develop these activities, the ARS was used as a tool to conceive potential activities, before translating the activities for 2-D map use. The authors of this study found the ARS very helpful in the creation of new exercises and constructing new ways of thinking about topographic-map instruction.

Limitations

Two possible limitations to our study include the small sample sizes and the instrument used to assess students' topographic map-reading skills. The small sample sizes in this study mean the results are not universally generalizable. However, these small sample sizes are inherent to the study of ARS use in a naturalistic setting. Although the quantitative data obtained using the TMA and the qualitative data obtained from instructors and students' comments are derived from small sizes, they directly address the two research questions that this study aimed to answer. Our findings signal avenues for possible future research both in terms of (1) further developing the ARS as a teaching and learning tool and (2) investigating how the ARS can be used to facilitate not only student learning but also curriculum design.

Although the TMA is the only commonly used measure of topographic map-reading skills, it has identified limitations. Like Jackson and colleagues (2019) and Moore (2018), we identified limitations in using it. However, the fact that it is commonly used opens the possibility of future meta-analyses of studies using this instrument. Given the TMAs identified limitations and its popularity as the only available measure, we suggest further development and of the TMA be conducted to increase its utility with more diverse subjects.

Conclusion

The primary goal of this study was to provide insight on the association between multiple instructional periods with an ARS on the development of college students' topographic map-reading skills. Overall, students' TMA scores were consistent with the findings of previous research showing no significant learning gains between ARS and non-ARS use (Giorgis et al., 2017; Jackson et al., 2019; McNeal et al., 2020; Richardson et al., 2018). This is despite exposing students to more than a single session using an ARS. Female students in the ARS group improved their topographic map-reading skills relative to female students in the control group, but male students in the ARS group improved the least compared to all three groups. Finally, this study found

the ARS can be employed as a useful tool for providing feedback to instructors, facilitating conversations and collaborations, and designing curriculum.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10956-022-10011-2>.

Acknowledgements Thanks are due to Chelsie Kowalski, Daniel Fulton, and Laura Lipsky for their assistance in coding and transcribing data and to Dr. Mimi Engel for her guidance on statistical analysis. Additionally, we are grateful to Sid Gustafson and Jonathan Allured for their help in constructing the Augmented Reality Sandbox used in this study. Finally, we thank all of the instructors and students who provided feedback on and participated in this project.

Data Availability The data that support the findings of this study are available on request from corresponding author LAA. The data are not publicly available due to them containing information that could compromise research participant privacy/consent.

Declarations

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional review board at the authors' institution (reference number: 18–0746) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants in the study.

Conflict of Interest The authors declare no competing interests.

References

- Atit, K., Weisberg, S. M., Newcombe, N. S., & Shipley, T. F. (2016). Learning to interpret topographic maps: Understanding layered spatial information. *Cognitive Research: Principles and Implications*, 1(1), 1–18.
- Bacca, J., Baldiris, S., Fabregat, R., & Graf, S. (2014). Augmented reality trends in education: A systematic review of research and applications. *Educational Technology & Society*, 17(4), 133–149.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2014). Fitting linear mixed-effects models using lme4. arXiv preprint <https://arxiv.org/abs/1406.5823>
- Carbonell-Carrera, C., Avarvarei, B. V., Chelariu, E. L., Draghia, L., & Avarvarei, S. C. (2017). Map-Reading Skill Development with 3D Technologies. *Journal of Geography*, 116(5), 197–205.
- Carbonell-Carrera, C., Jaeger, A. J., & Shipley, T. F. (2018). 2D Cartography Training: Has the Time Come for a Paradigm Shift? *ISPRS International Journal of Geo-Information*, 7(5), 197.
- Clark, D., Reynolds, S., Lemanowski, V., Stiles, T., Yasar, S., Proctor, S., Lewis, E., Stromfors, C., & Corkins, J. (2008). University students' conceptualization and interpretation of topographic maps. *International Journal of Science Education*, 30(3), 377–408.
- Cohen, C. A., & Hegarty, M. (2014). Visualizing cross sections: Training spatial thinking using interactive animations and virtual objects. *Learning and Individual Differences*, 33, 63–71.
- Diegmann, P., Schmidt-Kraepelin, M., Eynden, S., & Basten, D. (2015). Benefits of augmented reality in educational environments-a systematic literature review. In Thomas, O. & Teutenberg, F. (Eds.) *Wirtschaftsinformatik Proceedings*, 1542–1556.
- Downs, R. M., & Liben, L. S. (1991). The development of expertise in geography: A cognitive-developmental approach to

- geographic education. *Annals of the Association of American Geographers*, 81(2), 304–327.
- Evans, M., Fleming, B., & Drennan, G. (2018). Can the augmented reality sandbox help learners overcome difficulties with 3-D visualisation? *Terrae Didactica*, 14(4), 389–394.
- Gilhooly, K. J., Wood, M., Kinneer, P. R., & Green, C. (1988). Skill in map reading and memory for maps. *The Quarterly Journal of Experimental Psychology Section A*, 40(1), 87–107.
- Giorgis, S., Mahlen, N., & Anne, K. (2017). Instructor-Led Approach to Integrating an Augmented Reality Sandbox into a Large-Enrollment Introductory Geoscience Course for Nonmajors Produces No Gains. *Journal of Geoscience Education*, 65(3), 283–291.
- Gold, A. U., Pendergast, P. M., Ormand, C. J., Budd, D. A., Stempien, J. A., Mueller, K. J., & Kravitz, K. A. (2018). Spatial skills in undergraduate students—Influence of gender, motivation, academic training, and childhood play. *Geosphere*, 14(2), 668–683.
- Griffin, T. L. C., & Lock, B. F. (1979). The perceptual problem in contour interpretation. *The Cartographic Journal*, 16(2), 61–71.
- Habig, S. (2019). Who can benefit from augmented reality in chemistry? Sex differences in solving stereochemistry problems using augmented reality. *British Journal of Educational Technology*, 51(3), 629–644.
- Hale, K. S., Campbell, G., Riley, J., Boyce, M., & Amburn, C. (2019). Augmented Reality Sandtable (ARES) Impacts on Learning. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63(1), 2149–2153.
- Hod, Y., & Twersky, D. (2020). Distributed spatial Sensemaking on the augmented reality sandbox. *International Journal of Computer-Supported Collaborative Learning*, 15(1), 115–141.
- Holton, D., & Clarke, D. (2006). Scaffolding and metacognition. *International Journal of Mathematical Education in Science and Technology*, 37(2), 127–143.
- Ishikawa, T., & Kastens, K. A. (2005). Why some students have trouble with maps and other spatial representations. *Journal of Geoscience Education*, 53(2), 184–197.
- Jackson, D., Kaveh, H., Victoria, J., Walker, A., & Bursztyn, N. (2019). Integrating an augmented reality sandbox challenge activity into a large-enrollment introductory geoscience lab for nonmajors produces no learning gains. *Journal of Geoscience Education*, 67(3), 1–12.
- Jacovina, M., Ormand, C., Shipley, T., & Weisberg, S. (2014). Topographic Map Assessment. Retrieved from <https://www.silc.northwestern.edu/topographic-map-assessment-tma/>
- Kastens, K. A., & Ishikawa, T. (2006). Spatial thinking in the geosciences and cognitive sciences: A cross-disciplinary look at the intersection of the two fields. *Special Papers-Geological Society of America*, 413, 53–74.
- Lobben, A. K. (2004). Tasks, strategies, and cognitive processes associated with navigational map reading: A review perspective. *The Professional Geographer*, 56(2), 270–281.
- McNeal, K. S., Ryker, K., Whitmeyer, S., Giorgis, S., Atkins, R., LaDue, N., & Pingel, T. (2020). A multi-institutional study of inquiry-based lab activities using the Augmented Reality Sandbox: Impacts on undergraduate student learning. *Journal of Geography in Higher Education*, 44(1), 85–107.
- Moore, J. W. (2018). Spatial feedback as a mechanism to improve topographic map literacy using the augmented reality sandbox (Masters Thesis, Northern Illinois University, DeKalb, IL). Retrieved from <https://commons.lib.niu.edu/handle/10843/21583>
- Murakoshi, S., & Higashi, H. (2015). Cognitive characteristics of navigational map use by mountaineers. *International Journal of Cartography*, 1(2), 210–231.
- Newcombe, N. S., & Shipley, T. F. (2015). Thinking about spatial thinking: New typology, new assessments. In *Studying visual and spatial reasoning for design creativity* (pp. 179–192). Springer, Dordrecht.
- Newcombe, N. S., Weisberg, S. M., Atit, K., Jacovina, M. E., Ormand, C. J., & Shipley, T. F. (2015). The lay of the land: Sensing and representing topography. *Baltic International Yearbook of Cognition, Logic and Communication*, 10(1), 1–57.
- Niedomysl, T., Elldér, E., Larsson, A., Thelin, M., & Jansund, B. (2013). Learning benefits of using 2D versus 3D maps: Evidence from a randomized controlled experiment. *Journal of Geography*, 112(3), 87–96.
- Ormand, C. J., Manduca, C., Shipley, T. F., Tikoff, B., Harwood, C. L., Atit, K., & Boone, A. P. (2014). Evaluating geoscience students' spatial thinking skills in a multi-institutional classroom study. *Journal of Geoscience Education*, 62(1), 146–154.
- Phillips, R. J., Lucia, A., & Skelton, N. (1975). Some objective tests of the legibility of relief maps. *The Cartographic Journal*, 12(1), 39–46.
- Pick, H. L., Heinrichs, M. R., Montello, D. R., Smith, K., Sullivan, C. N., & Thompson, W. B. (1995). Topographic map reading. *Local Applications of the Ecological Approach to Human-Machine Systems*, 2, 255–284.
- Pollack, C. F. (2019). Investigating the Augmented Reality Sandbox: An Exploration of the Development and Implementation of a Reproducible STEM Resource in Secondary Education Geoscience (Masters Thesis, George Mason University, Fairfax, VA).
- Potash, L. M., Farrell, J. P., & Jeffrey, T. S. (1978). A technique for assessing map relief legibility. *The Cartographic Journal*, 15(1), 28–35.
- Rapp, D. N., Culpepper, S. A., Kirkby, K., & Morin, P. (2007). Fostering students' comprehension of topographic maps. *Journal of Geoscience Education*, 55(1), 5–16.
- Richardson, R., Sammons, D., & Delparte, D. (2018). Augmented Affordances Support Learning: Comparing the Instructional Effects of the Augmented Reality Sandbox and Conventional Maps to Teach Topographic Map-reading skills. *Journal of Interactive Learning Research*, 29(2), 231–248.
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, 2(6), 110–114.
- Soltis, N. A., McNeal, K. S., Atkins, R. M., & Maudlin, L. C. (2020). A novel approach to measuring student engagement while using an augmented reality sandbox. *Journal of Geography in Higher Education*, 1–20.
- Taylor, H. A., Renshaw, C. E., & Choi, E. J. (2004). The effect of multiple formats on understanding complex visual displays. *Journal of Geoscience Education*, 52(2), 115–121.
- Thorndyke, P. W. & Goldin, S. E. (1981). Ability differences and cognitive mapping skill (No. RAND/N-1667-ARMY). Rand Corp Santa Monica CA.
- Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139(2), 352–402.
- Vaughan, K. L., Vaughan, R. E., & Seeley, J. M. (2017). Experiential learning in soil science: Use of an augmented reality sandbox. *Natural Sciences Education*, 46(1), 1–5.
- Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101(4), 817–835.
- Wilson, B. G. & Myers, K. M. (2000). Situated cognition in theoretical and practical context. *Theoretical foundations of learning environments*, 57–88.
- Woods, T. L., Reed, S., Hsi, S., Woods, J. A., & Woods, M. R. (2016). Pilot study using the augmented reality sandbox to teach topographic maps and surficial processes in introductory geology labs. *Journal of Geoscience Education*, 64(3), 199–214.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.