Teacher-Guided Educational VR : Assessment of Live and Prerecorded Teachers Guiding Virtual Field Trips

Christoph W. Borst   Nicholas G. Lipari   Jason W. Woodworth

University of Louisiana at Lafayette

Abstract

We present a VR field trip framework, Kvasir-VR, and assess its two approaches to teacher-guided content. In one approach, networked student groups are guided by a live teacher captured as live-streamed depth camera imagery. The second approach is a standalone (non-networked) version allowing students to individually experience the field trip based on depth camera recordings of the same teacher. Both approaches were tested at two high schools using a VR environment that teaches students about solar energy production via tours of a solar plant. We show that our live networked approach can produce promising test score gains and very high ratings of co-presence, affective attraction, overall opinion, etc. Results show a benefit of live networked VR, as the standalone approach had lower performance in terms of gains and most ratings, although its ratings were still positive. We further consider possible differences of school environment (dedicated vs. integrated classroom), and we conclude with tradeoffs and implications to benefit future design of educational VR.

Keywords: Collaborative VR, education, avatars, Kinect.

Index Terms: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality; K.3.0 [Computers and Education]: General

1 Introduction

We investigate two educational VR approaches with a virtual field trip given at high schools. Real field trips have promising effects such as long-term information recall [10] and a first-person view that communicates size and spatial relationships of objects. VR field trips may provide such benefits without requiring students to leave the home or classroom. VR also allows environments to be augmented with additional educational content, such as embedded educational animations and simulations of rare events.

Our approaches focus on incorporating a teacher into VR using depth-camera-based teacher imagery from a Kinect device (Figure 1). One approach uses a live networked teacher and the other approximates the experience with prerecorded teacher segments, to provide a broadly-deployable substitute. Contributions include:

- We present a VR field trip framework, Kvasir-VR, for placing teachers in VR with a novel combination of depth camera imagery, heterogeneous displays, and virtual mirror techniques.
- We assess learning and student experience in real high school classrooms. The study of a live depth-camera-based teacher extends knowledge about such representations (Section 2.3), showing effectiveness in a new and real-world application.
- Results show benefits of live-guided educational VR over a non-networked substitute, although both were rated positively. Compared to prior work comparing in-person and prerecorded teachers, we study the topic in the new context of live and prerecorded video of teachers in VR. This shows a level of tradeoff expected for the convenience of prerecorded content.
- Results show a likely difference in learning between schools, and we relate it to distraction level of classroom types.
- We discuss tradeoffs between the approaches and implications for the long-term design of virtual field trips.

Figure 1: A student wearing a head-mounted display stands at a virtual tower overviewing a virtual solar plant (the background duplicates the student's view for illustration). A teacher (seen to the left) provides an introduction to the plant and is represented by a mesh captured by a Kinect depth camera.

2 Related Work

2.1 Educational VR

VR has long been suggested as a way to enhance education [33]. VR can produce experiences that are vividly remembered, along with other effects that seem to hinge on immersive or embodied experiences [5]. Motivation and engagement are fundamental to effective instruction, and they may be increased with good VR experiences [2,24]. Additional benefits can theoretically be achieved with multi-user, interactive, collaborative VR [14,15].

Based on several studies in a meta-analysis, games and VR can be more effective for education than traditional instruction [19]. For example, Virvou and Katsionis [30] reported that a VR game can be highly motivating and achieve better educational effects than traditional software. More recent examples have considered VR for a wide range of topics including computational thinking [21], aviation safety training [8], and firing range training [4].

2.2 Computer-based Tours of a Solar Energy Plant

VR models of energy facilities have been used to teach students about alternative energy production, e.g., wind farms [1] and solar plants [6]. Ritter [26] described our preliminary work to teach solar

This is an author-formatted version. Original publication: C. W. Borst, N. G. Lipari, and J. W. Woodworth, "Teacher-Guided Educational VR: Assessment of Live and Prerecorded Teachers Guiding Virtual Field Trips", IEEE Conference on Virtual Reality and 3D User Interfaces 2018; 18-22 March, Reutlingen, Germany.

thermal power plant concepts using three forms of an audio-visual tour including a 2D video version, a desktop-based 3D game, and an HMD-based VR version of the game. VR produced 11% and 13% average normalized test score gains in two high school studies (7% and 6% raw gain). Suggested factors limiting gain included distraction in a multi-activity classroom for one study and internal distractions in the VR itself related to novelty and intensity of the experience. Higher gain (43% normalized, 14% raw) was reported for university students, in a controlled laboratory setting, with a revised prototype and preliminary networking aspects. The highest gain for any high school study was 42% (21% raw) for a desktop game in a dedicated room. Statistical tests were not conducted.

Our new VR field trip is based on teaching and testing the same topics as before [26], using the tour-like approach, with evaluation in the same classrooms. Our new work differs fundamentally by focusing on embedded teachers to guide the tour and by including questions to students. We use different 3D models, although they model the same real power plant. We assess live and prerecorded teachers, rather than different presentation media, so we include measures such as social presence and affective attraction to assess students' experiences of the teacher.

2.3 Related Techniques

Technological aspects are not the focus of this paper, but we briefly mention prior work on depth-camera-based telepresence, heterogeneous displays, and virtual mirrors. We combine such techniques to allow a non-immersed teacher to guide students.

Several researchers considered livestreamed 3D meshes, based on depth camera data, as promising 3D video-like avatars for telepresence, e.g., [3,18]. Such mesh representations may provide richer and more immersive communication than conventional avatars [3]. Recent work uses multiple depth cameras for more complete user capture [3]. In our work, a student faces a teacher and single-camera capture of the teacher's front provides a reasonable representation that is easily captured. We have not found prior studies in which such depth-camera-based 3D avatars were deployed and assessed for education in classrooms.

Heterogeneous displays or views, also called asymmetric interfaces, may lend themselves to leader-follower exercises such as factory planning walkthroughs [22]. Combining different views could make collaborative tasks easier on users [23]. The asymmetric interface in our educational VR approach includes a virtual mirror teacher interface (Figure 2), somewhat like early video-based mirror worlds [17,29].

3 Virtual Field Trip Framework and Application

Our approach, named Kvasir-VR, uses rendering, networking, and interface techniques further detailed by Ekong [9] and Woodworth [32]. The studied application is based on a real field trip at a real power plant [7], but it adds interactable objects and educational content such as animations and see-through to device internals.

3.1 Student Interface

Students are immersed in the environment using standard consumer hardware including a head-mounted display (HMD) and tracked hand-held controller (Figure 1). For safety and practical reasons related to classroom use, students were seated in our experiments.

Students visit various educational stations and use ray-based pointing to trigger embedded educational elements such as animations. To provide a simple interface requiring minimal explanation and practice, only one button is used for all interactions (all programmable buttons simply perform in the same way).

Students move between stations by selecting teleportation targets (viewing platforms at educational stations) with the ray. A multi-step teleport motion was chosen to minimize motion sickness risks for seated students while still providing some sense of the path and of the environment's spatial layout. The teleport motion first steps the student to the target position through several intermediate positions with short fades between them, and then rotates the student through several steps about their vertical axis, to keep seated students facing a consistent real direction while they are virtually oriented to face educational content. During teleportation, a student also sees a leading indicator (an arrow) that moves smoothly just ahead of the student as both an in-progress indicator and to help users anticipate and understand motion.

Figure 2: The teacher guides students from a large TV interface with a Kinect mounted at its bottom. This interface leaves the teacher's face clear of obstructions. The TV visuals include: 1) a mirror view of the environment, 2) duplicated student views at the TV's lower left, 3) webcam views of students, in ovals hovering at positions for teacher eye gaze towards the students, and 4) pointing cues to help the teacher point correctly in 3D.

3.2 Teacher Interface

A teacher guides immersed students from a projection-based or TV-based teacher interface (Figure 2 shows a TV typically used). The teacher provides verbal descriptions of the environment and educational concepts, asks questions to check understanding and attention, directs students to click on interactable objects or teleport targets, and monitors students overall. Besides verbal descriptions, pointing is the main aspect of communication. Pointing is a key mechanism for guiding others in collaborative VR [20].

For a display with a narrow field of view (FOV), such as a TV, the teacher is presented with a "virtual mirror" view with an exaggerated field of view. Thus, the teacher sees themselves as mirrored in this view, somewhat like the mirrored view in video conferencing tools like Skype. This view supports the teacher pointing at objects behind the teacher (virtually) while facing the students. We expect that the mirror view limits immersion and naturalness for the teacher, but we prioritize student experience over teacher experience. The mirror view has substantial practical advantages. It is more practical for classroom deployment where large-FOV projection systems are infeasible. Compared to an HMD view, it maintains a clear view of the teacher's face and eyes, preserves oversight of the teacher's real environment, and may be more comfortable for extended multi-session use.

The teacher is aided by additional visuals (Figure 2). Students' views of the environment are reproduced as inset images and may be checked by the teacher to ensure a student sees what is intended or to resolve communication problems. Pointing cues, visible only to the teacher, allow the teacher to point to correct depths with a mirror view [32]. Live webcam videos of students are overlayed in semitransparent ovals. The ovals are positioned for correct teacher gaze towards a student, using perspective projection that considers teacher head tracking and TV geometry. The teacher's view also includes minimal representations of student interaction rays.

3.3 Depth Camera Mesh and Networking

The teacher is captured by a Kinect sensor for representation as a mesh, creating a video-like, but 3D, avatar. Color and depth images from the Kinect are streamed from the teacher to student computers using an FFmpeg-based Unity plugin [9]. The plugin also allows recording and playback of Kinect data for creating standalone field trips based on prerecorded teacher clips. Custom Unity shaders render the teacher mesh by: 1) setting vertex positions and texture coordinates based on Kinect data received as textures, 2) removing background or invalid vertices, and 3) filtering depth and edges for improved appearance. Audio for voice is added using Teamspeak (networked) or standard recordings (standalone).

Figure 3: Virtual Solar Plant. *Top*: A teacher describes the (animated) flow of heated water into the boiler. *Bottom*: Quiz response board in the standalone version.

3.4 Energy Plant Field Trip

3.4.1 Motivation

We applied the Kvasir-VR framework to teach solar energy concepts to high school students in science and engineering programs. We developed a virtual solar power plant modeled after a real pilot-scale energy plant [7], for a virtual field trip activity that initially resembled real field trips at the facility. This provided schools who could not visit the real facility with an alternative way of experiencing the tour. The VR tour was additionally extended with educational content beyond that visible in real tours, and interaction was added to increase student involvement. The networked aspect of the live teacher supports both local and remote teachers. Remote operation allows students to meet remote teachers who have specialized expertise about the subjects being taught.

3.4.2 Activity Overview

The virtual field trip for our studies is summarized by the following main steps. In between the listed steps, the teacher asks a student to teleport to the next station (a single person activates group teleport).

1) Students meet the teacher at a tower overlooking the facility. The teacher checks communication mechanisms (audio/video), shows how interaction involves pointing and pushing a button, and gives an overview description of the plant.

2) At solar collectors, the teacher describes the heating and flow of water in a pipe. A student is asked to trigger an animation showing sun rays heating the pipe, and a student is asked to flip a power switch that turns on a solar tracker to rotate the collectors toward the sun. Students are asked to estimate the size of the collectors, to recall the main purpose of the solar collectors, to say how they think more water could be heated, and to state where the heated water is traveling.

3) Near an evaporator/boiler and related pipes (Figure 3), the teacher explains fluid flow and heat exchange through this part of the plant. A student is asked to open a valve to show animated water flow through external pipes, and a student is asked to trigger see-through rendering of the evaporator to show additional fluid flows, including refrigerant flow. Students are asked to identify two of the fluids flowing in and out of the evaporator (hot water and vaporized refrigerant).

4) Near a turbine and generator, the teacher explains the transfer of energy from high-pressure vaporized refrigerant to turbine rotation to electricity from the connected generator. A student selects the turbine to reveal see-through rendering of internal flow and rotation, and a student selects the generator to link it to the turbine for rotation. Students are asked to state the main purpose of each of these two devices.

5) Near a condenser and cooling tower, the teacher explains how vaporized refrigerant is condensed back into a liquid, and how this involves water cooled at the cooling tower. A student selects the condenser for see-through rendering showing remaining flows of water and refrigerant through the exchanger and through associated pipes. Students are asked to identify some of the fluids (cold water entering, cooled refrigerant leaving) and to determine where the cooled refrigerant is going. Finally, in a brief wrap-up, the teacher reviews the main path of energy from the sun to generated electricity and discusses the capacity and motivation for solar energy plants.

3.4.3 Differences between live and prerecorded teachers (i.e., networked and standalone approaches)

We implemented both a live networked and a non-networked standalone version of the teacher-guided field trip. In the networked version, a live teacher guides small groups of students from the teacher interface described in Section 3.2. The students are each locally placed at the same virtual viewing location as each other (on a teleport platform), with other student representations appearing offset to the left or right according to seating order in a real classroom. Students see minimal representations of other students, consisting of generic head avatars, which are typically not in view, and minimal ray representations. Students and the teacher all hear each other through a networked voice server. Environment responses to student actions (teleport, animation) are reflected at all networked computers. In our studies, the teacher selected a specific student for each such interaction, rotating through the students to ensure all were involved. Questions requiring verbal response were first asked to the entire group (e.g., "can anyone tell me…"), but, when no student responded, the teacher asked specific students.

Students may benefit from a live teacher being able to correct misunderstandings and respond more meaningfully to a wide range of student responses. This two-way communication tends to occur after the teacher asks questions to check student knowledge. The apparent presence of a real person itself may also improve student interest or motivation to perform well. Overall, two-way communication may enhance learning and better resembles a real field trip, but the communication takes extra student time.

The standalone version was developed using a prerecorded teacher to allow students to learn the material independently, at reduced deployment cost and complexity, and with each student experiencing all interactions. It is intended to closely resemble the live version, but with some difference resulting from limitations of a prerecorded teacher. In our studies, the prerecorded teacher was a Kinect-based recording of the same teacher, following the same guiding script, and recorded at the same teacher interface (the teacher also wore the same clothing and hair style in all deployments). The main difference is that a prerecorded teacher cannot respond with the understanding of a real teacher when a student freely answers questions about the environment. Thus, the standalone version uses an in-world multiple-choice quiz to allow constrained responses (Figure 3). Students select an answer with ray-based interaction. In case of an incorrect response, the quiz pauses briefly to provide feedback and then repeats the question.

4 Experiment Methods

4.1 Overview and Independent Variables

We assessed and compared the teacher-guided VR approaches using a between-subjects study design. The main independent variable was the type of approach (networked vs. standalone), with differences described in Section 3.4.3, and with random assignment of subjects to approaches. The field trip teacher was unfamiliar to the students and was located in a separate room from students to avoid effects of the teacher's real physical presence and to resemble remotely-guided educational VR (teacher and student rooms were about 11 meters apart and networked through a local switch. Virtually, the teacher appeared 2 meters or less in front of students).

Secondarily, we include classroom type as another independent variable (dedicated vs. integrated). Prior work [26] speculated that an integrated classroom, with multiple student groups working on different activities, created a distraction for VR tours, and a quiet dedicated VR space was recommended. In our study, classroom types reflect the facilities available at two different schools, i.e., they reflect uncontrolled school characteristics rather than random assignment of students to conditions. However, we believe classroom type is the clearest difference between the schools, and we did not find significant differences in pre-test or immersive tendencies scores (Sections 4.3 and 5.1).

Both approaches (networked and standalone) were tested in both schools, for an overall 2 X 2 between-subjects design.

4.2 Dependent Measures and Related Hypotheses

4.2.1 Test Score Gain

We assessed learning with a multiple-choice test of knowledge about solar plant devices and their purposes, fluids, and energy conversion steps. The same test was given before (pre) and after (post) the VR field trip. Test score gain was computed, especially normalized gain [11], per student. This normalized gain is the increase in score, from the pre-test to post-test, divided by the maximum possible increase (the increase that would have resulted from a perfect post-test score). In contrast to raw gain, normalized gain is less dependent on, or less correlated to, prior knowledge.

We avoided forming a specific (directional) hypothesis about the gains of different VR approaches because the approaches have tradeoffs that may favor either one (Section 3.4.3). We wanted to assess their educational effectiveness with a standard measure and compare them. Regarding classroom type, we expected the dedicated classroom to produce best results.

4.2.2 Questionnaire Items

Subjective aspects of the VR experience were assessed using a mix of well-established questions and questions more tailored to our educational VR. To ensure the total questionnaire size was manageable to subjects, we abbreviated questionnaires from prior work. Except where stated otherwise, question responses were ratings from 1 to 7. Semantic anchors were placed below values 1, 4, and 7. The questions can be summarized as follows:

1) Social Presence: Eight questions from Networked Minds Social Presence [12], with two each from subscales Co-presence, Attentional Allocation (distraction, reverse coded), Perceived Message Understanding, and Perceived Behavioral Interdependence. Within each subscale, one question asked about the subject's impression of the teacher and the other asked about the teacher's impression of the subject (even though all questions were answered by the subject).

2) Attraction: Four questions following Herbst et al. [13], asking about the teacher: one rating teacher competence (Cognitive Attraction), two on positivity and friendliness (Affective Attraction), and one on Rated Threat.

3) SUS Presence: Four questions resembling the first four presence questions from Usoh et al. [28], to ask about the subject's sense of presence in the solar plant.

4) Engagement: Two questions asking if subjects were "motivated to complete the application" and if they "found the experience captivating and could easily focus". We have named this scale Engagement, although it may differ from some definitions.

5) Usability: Two questions asking subjects if the system was "easy to use" and if its functions were "well integrated".

6) SSQ (simulator sickness): Five simulator sickness questions based on SSQ [16], asking about general discomfort, fatigue, headache, eye strain, and difficulty focusing or concentrating. Ratings were from 1 (none) to 4 (extreme).

7) Other: Additional questions ranged from the extent to which the experience was "liked" overall to asking students to judge if the teacher was real or prerecorded.

We hypothesized that the networked approach could produce higher scores on items such as Social Presence, e.g., due to the two-way communication with a live teacher rather than prerecorded segments. However, we did not expect these results to extend to aspects such as SUS Presence, which do not focus on the teacher.

4.3 Subjects

We report results from 88 high school students. An additional 26 students tried a VR field trip but were excluded from study because 5 of them reported previous exposure to the VR field trip and 21 experienced failures such as a VR device not displaying images.

The students were from science and engineering classes such as chemistry, calculus, robotics, and engineering. Teachers advertised the activity to students. Table 1 summarizes demographics per condition. These reflect a smaller available class at the dedicated school and low female class enrollment in these programs. Students were slightly younger (lower grade level) at the integrated classroom, but immersive tendency and pre-test scores (Section 5.1) were not found to differ between any conditions statistically.

Table 1: Some subject demographics per condition: mean score on an immersive tendencies questionnaire (IT score), number of subjects (male, female, unspecified), and mean age in years.

Approach	Classroom	IT score	M, F, U	Age
Networked	*Dedicated*	5.04	10, 3, 1	16.6
	Integrated	4.90	31, 3, 1	14.9
Standalone	*Dedicated*	5.05	10, 0, 2	15.9
	Integrated	4.97	21, 1, 5	14.4

4.4 Procedure and Apparatus

Subjects and their guardians provided assent and consent forms prior to the study. Subjects entered the experiment area in groups of 4 students, except in a few cases where only fewer students were available. Subject assignment to groups was random per classroom, and groups were alternatingly assigned the two different VR approaches (networked, standalone). Subjects first met the field trip teacher in VR (virtually) and only met physically after the experiment. Subjects were seated near each other, in a single row, and remained seated for the entire experiment.

Immediately before experiencing VR, subjects provided basic demographic information, filled out an abbreviated immersive tendencies questionnaire [31], and took the pre-test, all with pencil and paper. Subjects were then given an equipment overview. Student and teacher stations used laptops with GeForce 1080 graphics, internally (MSI GT73VR Titan Pro-003) or with an external graphics extender (for Alienware 15 R2 laptops). Subjects donned the VR displays (Oculus Rift CV1) and fit them by adjusting straps and an inter-pupil distance dial. Subjects with eyeglasses were told they could wear their glasses with the Rift by expanding the straps and taking care when putting on the display.

The proctor checked that subjects reported being comfortable and able to see an environment clearly, and then handed an Oculus Touch controller to each subject. The proctor next placed headphones with voice microphones on each student (HyperX Cloud Gaming headsets; built-in Oculus headphones were removed from the displays). Once all students were ready, the VR field trip began. In the networked condition, the VR teacher appeared virtually in front of subjects and began by asking if he could be seen and heard. In the standalone condition, a sign told students to point and click with the controller to begin the prerecorded teacher, and students used a ray-based quiz to indicate if they could see and hear the teacher. The field trip then continued as in Section 3.4.2. Any technical failure with equipment that was not immediately correctable resulted in subject omission from the study, but such subjects could experience the VR in later non-experiment sessions.

Immediately after the VR field trip, subjects took the post-test and then filled out written questionnaires in the following order: 1) a questionnaire about the teacher (Social Presence, Attraction), 2) a questionnaire about other aspects (SUS Presence, Usability, other questions), and 3) a questionnaire about problems (SSQ).

5 Results and Discussions

5.1 Test Score Gains

Educational effectiveness is summarized by normalized gains in Figure 4 and Table 2. A 2 X 2 between-subjects ANOVA detected a significant difference between approaches [$F(1,84)=4.62, p=.035, \eta_p^2=.052$]. Given the low power of a between-subjects design and a small number of dedicated-classroom subjects, it is also notable that there was a near-significant difference between classroom types [$F(1,84)=3.87, p=.053, \eta_p^2=.044$] and a near-significant interaction [$F(1,84)=2.83, p=.096, \eta_p^2=.033$]. Readers may prefer to call these "trends" or 0.1-level significance. The likely interaction is explained by direct comparisons: the networked approach produces better results in the dedicated classroom than in the integrated classroom [$t(47)=2.40, p=.020, d=.758$], but the standalone approach was not detected better in the dedicated room than in the integrated room [$t(37)=0.236, p=.814, d=.082$].

The main result is that the live networked field trip produced better results than the standalone version. Secondarily, although weaker demonstrated overall, the dedicated classroom gave better results than the integrated classroom, particularly with the networked teacher. The live networked teacher in a dedicated classroom produced the highest mean gain by far, of 66.0%, compared to other means between 23.9% and 28.6%.

A lack of substantial difference between the two standalone cases suggests that students in the dedicated class (room) were not consistently better learners than those in the integrated class. This, as well as similar pre-test scores and immersive tendencies, aligns with our belief that the critical aspect of classrooms (i.e., schools) was distraction level and not student qualities. A possible explanation for the overall gain pattern (Figure 4) is that neither the distracting classroom nor the standalone approach kept the students engaged with educational VR content. Both an effective approach and a low-distraction environment were needed for good learning.

Mean gain for all conditions was higher than gains reported for the prior VR solar plant field trips in high schools, and the networked approach in a dedicated classroom produced higher gain than any prior reported approach (Section 2.2). Furthermore, we consider a gain of 66.0% to be a good result in light of Hake's choice of 70% as a "high gain" [11], which is usually not achieved.

Table 2: Mean scores: pre-test, post-test, and normalized gain. (The gain is not computable directly from the listed scores because it is "average-of-gains", not "gain-of-averages"). Mean time spent (minutes) by subjects in each condition is also given.

Approach	Classroom	Pre	Post	Gain	Time
Networked	*Dedicated*	53.9%	82.5%	66.0%	15.9
	Integrated	52.5%	67.0%	28.6%	17.4
Standalone	*Dedicated*	53.0%	67.4%	26.8%	9.36
	Integrated	54.2%	65.0%	23.9%	8.96

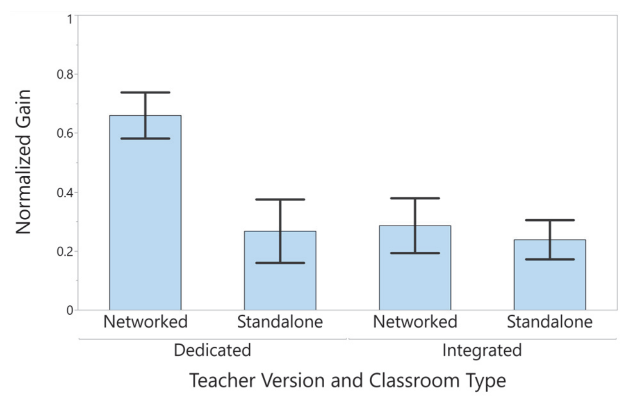

Figure 4: Educational effectiveness of the four conditions: Means and standard errors (+/- 1 SE) of normalized gain in test score.

5.2 Questionnaire Results

5.2.1 Overview

We compare networked and standalone approaches in terms of questionnaire responses. We do not detail classroom types here because associated tests, performed for all items, did not reveal significant classroom differences. However, there was a near-significant result for Social Presence [$Z=1.65, p=.099$] related to the Attentional Allocation subscale [$Z=1.89, p=.058$]. Although weak evidence by itself, this is consistent with the pattern and expectation of distractions in the integrated classroom.

Multi-item scale results are summarized in Figure 5. Scale scores were computed as means of the contributing items. Within Social Presence, Attentional Allocation item ratings were first reversed (subtracted from 8) because the questions were phrased as ratings of Distraction.

Scale scores were positive throughout (Figure 5) for both networked and standalone approaches, considering that 4 was labeled as a neutral answer in our questionnaires, and mean and median scores range from about 5 to 6.5 throughout (except SSQ, which measures motion sickness, and its low ratings are positive).

Table 3 summarizes statistical comparison between networked and standalone approaches (Wilcoxon Rank Sum tests). Where significant effects were found in these scales, we include subscales. We do not apply global corrections such as Bonferroni correction, and note conservative readers can be cautious about any borderline significant p-values.

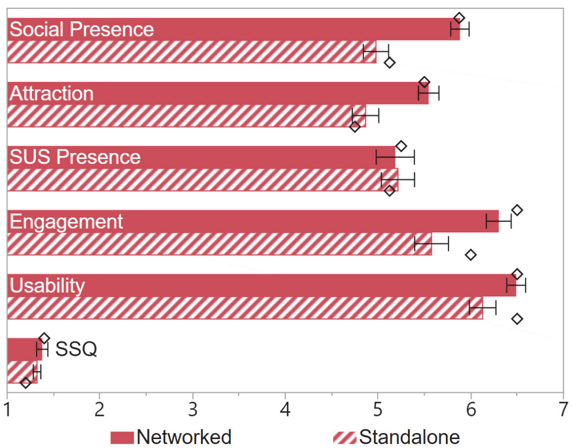

Figure 5: Overall scales: Bars show means and standard errors. Diamond markers show medians. Note that SSQ (simulator sickness) scores could only range from 1 to 4.

Table 3: Networked and standalone approaches compared in terms of questionnaire responses; Wilcoxon Rank Sum tests.

Scale	Z	*p*	*r*
Social Presence	-4.71	**< .001**	.502
Co-Presence	-3.99	**< .001**	.425
Distraction	-1.04	.298	.111
Message Understanding	-5.44	**< .001**	.583
Behavioral Interdependence	-3.67	**< .001**	.391
Attraction	-3.42	**< .001**	.369
Cognitive	-4.14	**< .001**	.452
Affective	-2.39	**.017**	.258
Threat	1.23	.220	.134
SUS Presence	-.230	.818	.025
Engagement	-3.24	**.001**	.347
Usability	-1.85	.064	.198
SSQ sickness	.102	.919	.011
Extra Questions			
Liked the experience	-1.75	.079	-.188
Too slow / Too fast	1.86	.063	.201
Understood teacher pointing	-2.25	**.025**	-.241
Understood teacher explanations	-3.17	**.002**	-.338
Teacher helped understanding	-3.54	**< .001**	-.377
Animated objects helped	-2.79	**.005**	-.299
In-game questions/quizzes helped	-.591	.554	-.063

5.2.2 Social Presence

As hypothesized, subjects reported significantly higher social presence with the networked approach that had a live teacher, although both approaches received positive ratings. A similar result was found for the subscales dealing with co-presence, message understanding, and behavioral interdependence (Figure 6). However, no significant difference was detected in reported distraction from the teacher (attentional allocation), and the plot suggests any existing difference would likely be small.

Students strongly perceive teacher presence of a depth-camera-based live teacher (networked approach), as shown by the median co-presence score of 7 (the maximum possible).

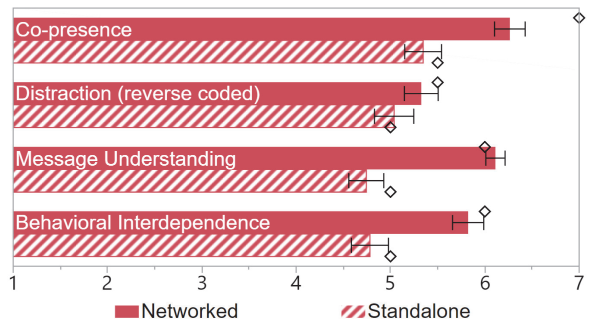

Figure 6: Social Presence subscales: Note Attentional Allocation is shown as reverse-coded Distraction. Bars show means and standard errors. Diamond markers show medians.

5.2.3 Attraction

Attraction scores overall resembled social presence results. Attraction subscales are plotted in Figure 7. Students responded very positively to the live teacher, as shown by the maximum possible median scores of 7 for affective and cognitive attraction (networked approach). The standalone approach also produced positive ratings, especially the affective attraction median of 6, yet prerecorded teacher competence was rated as two full points below the live teacher (cognitive attraction median of 5). Rated threat was low for both approaches, with no significant effect detected.

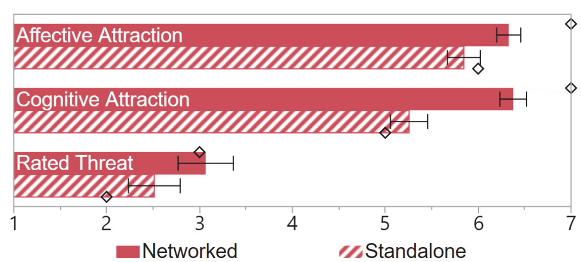

Figure 7: Attraction subscales: Bars show means and standard errors. Diamond markers show medians.

5.2.4 Presence

Students in both approaches report a positive, yet not very high, sense of their own presence in the solar plant, with overall mean and median around 5.2 (Figure 5). No significant differences were detected (Table 3). The plotted scores, and our own inspection of subitems, suggests any existing difference would be minimal.

5.2.5 Engagement

Students reported high motivation and focus for both approaches, with median scores of 6 and 6.5 (Figure 5). The networked approach scored significantly higher (Table 3). Based on this self-report, students are highly engaged, especially with a live teacher.

5.2.6 Usability

Both approaches received a very high median usability score of 6.5, with no statistically detected difference between approaches. This suggests that the single-button ray selection and teleportation interface worked well in both approaches and that increased interaction requirements in the standalone approach did not meaningfully affect perceived usability.

5.2.7 Simulator Sickness (SSQ)

There were few reports of problems such as discomfort, eye strain, or headaches, and no difference was detected between approaches. Of the 5 SSQ questions and two approaches, the only median rating above 1 (“none”) concerned the ability to focus or concentrate with the standalone approach, where the median was 2 (“slight”).

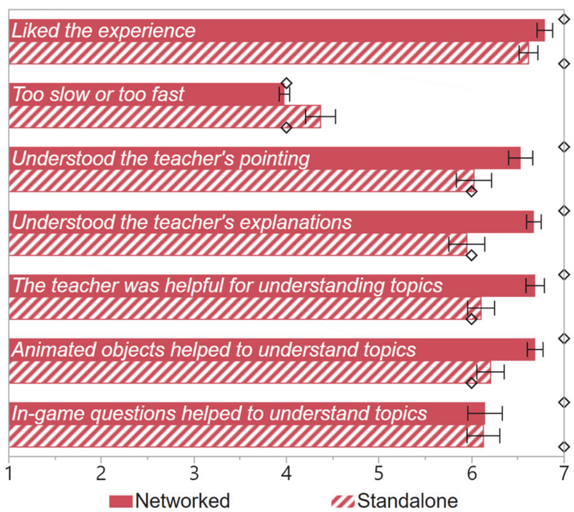

Figure 8: Additional ratings: Bars show means and standard errors. Diamond markers show medians.

5.2.8 Additional Ratings

Figure 8 summarizes additional items assessing the VR field trips.

Like: The first item asked overall opinion in terms of liking the experience. Both approaches resulted in the maximum possible median score, 7, with no statistically significant difference.

Speed: Speed was rated on a scale from “too slow” (1) to “too fast” (7). Median scores for both approaches were 4, matching the text anchor “well paced”. Statistically, there was near significance. The plotted values suggest little meaningful difference, if any.

Understood teacher: Two questions rated the understandability of the teacher’s pointing and explanations. These resembled Message Understanding, but here the intent was to consider aspects of educational content and pointing related to the teacher’s interface and teacher representation. The networked approach received the highest possible median, 7, significantly higher than standalone. Nonetheless, the standalone rating was also high (median 6).

Component helpfulness: Subjects were asked to rate the value of the teacher, animated objects, and in-game questions in terms of their helpfulness for understanding. The questions about the teacher and the animated objects revealed the same response pattern stated above for “understood teacher”. However, the third question did not result in a statistically detected difference – both approaches had similar mean and median ratings. We conducted additional statistical tests to compare “in-game question” ratings to ratings from the other two questions. Results suggest only that subjects rated animated objects as being more helpful than in-game questions [$Z=2.26$, $p=.024$], but we do not consider the apparently small difference meaningful enough to guide design.

5.2.9 Live or Prerecorded?

We additionally asked students to indicate if they believed the field trip teacher was live, prerecorded, or a mix of live and prerecorded parts. In a past demonstration, a live teacher was misjudged as an AI system. No subjects in our study reported the live networked teacher as entirely prerecorded, but 9 reported the live teacher as a mix of live and prerecorded. Of the subjects experiencing the prerecorded teacher (standalone), 2 reported believing the teacher was live and 3 reported the teacher was a mix.

5.2.10 Freeform response items

We asked subjects to specify the most positive and negative aspects of the experience, and to make any other comments they wanted to. For the question about best aspects, the most frequent answer types were: interactions with surroundings or objects (18 subjects), benefits of seeing content in 3D/VR (18), teaching effectiveness (15), and answers about presence (11). For negative aspects, one answer type dominated: graphical glitches in the teacher representation (28 subjects). The most common “other” comment was general positive feedback such as “it was really good” (16 subjects), and some subjects made suggestions such as adding more interaction (4) or making longer field trips (3).

6 Conclusion and Future Work

We presented an approach to VR field trips using a live teacher represented by depth camera imagery, and we showed its effectiveness in terms of student learning and subjective experience. We also presented and compared a standalone version that approximates the experience by using a prerecorded teacher.

Our study showed an educational benefit of live networked VR over a non-networked approximation, especially in terms of test score gains. We believe this results from a live teacher being more interesting to students and better being able to correct misunderstandings and respond to questions. This result can be seen as echoing studies where conventional live lectures produced higher gains than videos, e.g., [25]. However, live VR avatars are not the same as live in-person lectures, and effectiveness could hinge on factors such as avatar type. There are also contrasting studies, for example, without learning gains for live lectures compared to videos giving students substantial control of pausing and replay [27]. More work is needed to understand what types of control and interaction best contribute to learning in VR.

Standalone approaches remain relevant for research, considering practical advantages: deployment is relatively simple as it does not require a live expert instructor, the associated extra equipment, coordination, or extra setup time. Our standalone field trip took less student time due to reduced communication, and the difference could be used to present more content. Standalone approaches scale up readily to large numbers of students and more general audiences.

Further research is needed to identify which specific aspects of the live teacher presence or guidance (or the presence of other students in the environment) contributed to learning. For example, results could be related to student beliefs about the extent to which a real person observes them, and not just to better verbal responses by the teacher when students give incorrect answers to questions. Deeper consideration of these aspects can help guide other approaches aiming to address some of the tradeoffs between live and prerecorded teachers, for example:

1) A mixed approach, in which a live teacher is only present or active at key moments, or plays more of a monitoring role. Such an approach may be more scalable while still retaining some benefits of a live teacher.

2) We expect that, eventually, increasingly advanced autonomous agents (AI systems and avatars) may mimic a live VR teacher

well enough to reduce or eliminate the benefit of a real expert teacher. Our results may help predict the benefit achievable with a realistic VR agent over a simple sequencing approach, and further study could help guide agent development. On the other hand, if effectiveness hinges on student belief that a real teacher or real classmates are involved, then no agent known as artificial may be able to produce the same effect.

ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1451833 and by additional support from the Mozilla Gigabit Community Fund and a US Ignite Application Development Award. Models and animations were built by Adam Prejean and Jarret Delle Donne. Paul Smith and Adil Khokhar assisted with data entry. Sam Ekong, Terrence L. Chambers, and Kenneth A. Ritter III made fundamental contributions in prior work, as cited. Ekong contributed additional software development and scene editing. Chambers directs the real energy center and provided additional input on phrasing of the VR teacher's explanations. We thank Nicolette Darjean and Lisa Ranney for supporting classroom deployments at the David Thibodaux STEM Magnet Academy and at Comeaux High School.

REFERENCES

[1] P. Abichandani, W. Fligor and E. Fromm, "A cloud enabled virtual reality based pedagogical ecosystem for wind energy education," in *Proc. IEEE Frontiers in Education*, pp. 1-7, Madrid, 2014.

[2] A. G. Abulrub, A. N. Attridge and M. A. Williams, "Virtual Reality in Engineering Education: The Future of Creative Learning," in *IEEE Global Engineering Education Conference*, Amman, Jordan, 2011.

[3] S. Beck, A. Kunert, A. Kulik and B. Froehlich, "Immersive group-to-group telepresence," *IEEE Transactions on Visualization and Computer Graphics,* vol. 19, no. 4, pp. 616–625, 2013.

[4] K. K. Bhagat, W. Liou and C. Chang, "A cost-effective interactive 3D virtual reality system applied to military live firing training," *Virtual Reality,* vol. 20, no. 2, pp. 127-140, 2016.

[5] J. Blascovich and J. Bailenson, Infinite Reality: Avatars, Eternal Life, New Worlds, and the Dawn of the Virtual Revolution, New York, NY: Harper Collins, 2011.

[6] C. W. Borst, K. A. Ritter and T. L. Chambers, "Virtual energy center for teaching alternative energy technologies," in *Proc. IEEE Virtual Reality (VR)*, pp. 157-158, Greenville, SC, 2016.

[7] T. L. Chambers, J. R. Raush and B. Russo, "Installation and operation of parabolic trough organic rankine cycle solar thermal power plant in south louisiana," *Energy Procedia,* vol. 49, pp. 1107-1116, 2014.

[8] L. Chittaro and F. Buttussi, "Assessing knowledge retention of an immersive serious game vs. a traditional education method in aviation safety," *IEEE Transactions on Visualization and Computer Graphics,* vol. 21, no. 4, pp. 529-538, 2015.

[9] S. Ekong, C. W. Borst, J. W. Woodworth and T. L. Chambers, "Teacher-student vr telepresence with networked depth camera mesh and heterogeneous displays," in *Proc. International Symposium on Visual Computing*, pp. 246-258, Las Vegas, NV, 2016.

[10] J. H. Falk and L. D. Dierking, "School field trips: Assessing their long-term impact," *Curator: The Museum Journal,* vol. 40, no. 3, pp. 211-218, 1997.

[11] R. Hake, "Interactive-Engagement vs. Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses," *American Journal of Physics,* vol. 66, no. 1, pp. 64-74, 1998.

[12] C. Harms and A. Biocca, "Internal consistency and reliability of the networked minds social presence measure," in *Seventh Annual International Workshop: Presence 2004*, València, Spain.

[13] K. C. Herbst, L. Gaertner and C. A. Insko, "My head says yes but my heart says no: Cognitive and affective attraction as a function of similarity to the ideal self," *Journal of Personality and Social Psychology,* vol. 84, no. 6, pp. 1206-1219, 2003.

[14] S. M. Huang, S. S. Liaw and Y. C. Teng, "Developing a Collaborative Virtual Reality Learning System," *Journal of Information Technology and Application in Education (JITAE),* vol. 1, no. 2, pp. 74-79, 2012.

[15] R. L. Jackson and E. Fagan, "Collaboration and learning within immersive virtual reality," in *Proc. Collaborative Virtual Environments*, pp. 83-92, San Francisco, CA, 2000.

[16] R. S. Kennedy, N. E. Lane, K. S. Berbaum and M. G. Lilienthal, "Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness," *International Journal of Aviation Psychology,* vol. 3, no. 3, pp. 203-220, 1993.

[17] M. W. Krueger, T. Gionfriddo and K. Hinrichsen, "VIDEOPLACE—an artificial reality," *ACM SIGCHI Bulletin,* vol. 16, no. 14, p. 35–40, 1985.

[18] A. Maimone and H. Fuchs, "Encumbrance-free telepresence system with real-time 3d capture and display using commodity depth cameras," in *Proc. International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 137-146, Basel, 2011.

[19] Z. Merchant, E. T. Goetz, L. Cifuentes, W. Keeney-Kennicutt and T. J. Davis, "Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis," *Computers & Education,* vol. 70, no. C, pp. 29-40, 2014.

[20] T. Nguyen and T. Duval, "A survey of communication and awareness in collaborative virtual environments," in *Proc. IEEE Workshop on Collaborative Virtual Environments (3DCVE)*, Minneapolis, MN, 2014.

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

[22] S. Pick, S. Gebhardt, B. Weyers, B. Hentschel and T. Kuhlen, "A 3d collaborative virtual environment to integrate immersive virtual reality into factory planning processes," in *Proc. IEEE Workshop on Collaborative Virtual Environments (3DCVE)*, Minneapolis, MN, 2014.

[23] T. Porssut and J. R. Chardonnet, "Asymetric telecollaboration in virtual reality," in *Proc. IEEE Virtual Reality (VR)*, pp. 289-290, Los Angeles, CA, 2017.

[24] J. Psotka, "Educational Games and Virtual Reality as Disruptive Technologies," *Journal of Educational Technology & Society,* vol. 16, no. 2, pp. 69-80, 2013.

[25] S. Ramlogan, V. Raman and J. Sweet, "A Comparison of Two Forms of Teaching Instruction: Video vs. Live Lecture for Education in Clinical Periodontology," *European Journal of Dental Education,* vol. 18, no. 1, pp. 31-38, 2014.

[26] K. A. Ritter, Virtual solar energy center: A case study of the use of advanced visualization techniques for the comprehension of complex engineering device. Ph.D. Dissertation, Lafayette, LA: University of Louisiana at Lafayette, 2016.

[27] B. Schreiber, J. Fukuta and F. Gordon, "Live lecture versus video podcast in undergraduate medical education: A randomised controlled trial," *BMC Medical Education,* vol. 10, no. 68, 2010.

[28] M. Usoh, E. Catena, S. Arman and M. Slater, "Using presence questionnaires in reality," *Presence-Teleoperators and Virtual Environments,* vol. 9, no. 5, pp. 497-503, 2000.

[29] V. J. Vincent, "The mandala virtual reality system: The vivid group," in *Proc. Virtual Reality Conference and Exhibition on VR Becomes a Business*, pp. 167-170, San Jose, California, 1993.

[30] M. Virvou and G. Katsionis, "On the usability and likeability of virtual reality games for education: The case of VR-ENGAGE," *Computers & Education,* vol. 80, no. 1, pp. 154-178, 2008.

[31] B. Witmer and M. J. Singer, "Measuring Presence in Virtual Environments: A Presence Questionnaire," *Presence,* vol. 7, no. 3, p. 225–240, 1998.

[32] J. Woodworth and C. W. Borst, "Design of a practical TV interface for teacher-guided VR field trips," in *Proc. IEEE Workshop on Everyday Virtual Reality (WEVR)*, pp. 29-34, Los Angeles, CA, 2017.

[33] C. Youngblut, "Educational Uses of Virtual Reality Technology," Institute for Defence Analyses, Alexandria, VA., 1998.