Gaming4All: Reflecting on Diversity, Equity, and Inclusion for Game-Based Engineering Education

Casper Harteveld, Nithesh Javvaji, Tiago Machado
Northeastern University
{c.harteveld,javvaji.n,ti.machado}@northeastern.edu

Yevgeniya V. Zastavker
Olin College of Engineering
yevgeniya.zastavker@olin.edu

Victoria Bennett, Tarek Abdoun
Rensselaer Polytechnic Institute
{bennev,abdout}@rpi.edu

***Abstract*—Game-based learning has long been heralded as a promise to transform education into a more engaging and experiential paradigm. Its implementation in engineering education has thus far shown that promise, yet more evidence is needed for widespread adoption. Our work focuses on issues of equity, inclusivity, and diversity, thus surfacing the question of whether game-based learning benefits each and every student. To date, these issues have not received much attention in game-based learning educational research landscape, despite its poignancy, given today's global social climate and the systematic lack of equity and diversity in engineering education. With a focus on gender and race/ethnicity we reflect on the outcomes of implementing game-based learning environment in undergraduate geotechnical engineering courses (*n* = 362). We find evidence that game-based learning supports positive outcomes and provides an equitable learning environment. However, we also note several challenges that are further discussed.**

***Index Terms*—game-based learning, engineering education, inclusive design, mixed reality**

I. Introduction

Game-based learning has long been heralded as a promise to transform education into a more engaging and experiential paradigm [1], [2]. Games, specifically digital or video games, have various affordances that explain their potential for education, such as that they can (1) immerse students into virtual environments, among others, to simulate phenomena that are difficult to experience in reality and practice particular roles (e.g., being a professional engineer) [3]; (2) provide immediate formative and summative feedback to instructors and students [4] by leveraging the wealth of data games generate [5]; and (3) serve as a tool for capturing and maintaining learning motivation [6], [7]. However, a key transformative aspect of game-based learning is that it allows for a shift from the traditional educational model, in which the teacher "transmits" information and the student acts as passive recipient [8], to an active learning, student-centered educational paradigm, where students have autonomy over their own learning and the role of instructors is to guide and facilitate.

This transformation to a more student-centered, experiential model is what makes game-based learning appealing for *engineering education* in particular. Because engineering practice relies on one's ability to understand potential problems and design appropriate solutions, one of the more frequently cited needs for engineering education is that students engage in practical training and gain authentic hands-on experience [9], [10]. Providing students with such experiences is difficult, costly, and often irreproducible (e.g., natural disasters). Game-based learning can address these challenges and has thus far shown that promise [11], [12]. Yet, its use in engineering education remains limited, in part, because more evidence is needed for widespread adoption.

Similarly to other Science, Technology, Engineering and Mathematics (STEM) disciplines, engineering education has an urgent need for broadening its student population, in particular with regards to women and other underrepresented groups. According to the American Society for Engineering Education (ASEE), in 2015 about 19.9% of the Bachelors students were women and 64.9% were white [13]. These numbers are remarkably similar in other Western countries and have hardly changed over the years [14], thereby pointing to the systemic problems in engineering education. While game-based learning has been primarily studied for its educational benefits for all students in engineering education (see [11], [12]), given the lack of diversity in engineering education it is important to study the impact game-based learning may have on issues of diversity, equity, and inclusion. With a few exceptions (e.g., [15]), this has not received the attention it deserves. Leveraging these concerns and opportunities, our study asks: How and to what extent does game-based learning provide equitable learning opportunities across all learners in engineering education (i.e., Gaming4All)?

Our project introduces game-based learning into geotechnical engineering curricula through a mixed reality game, *GeoExplorer*, which to date allowed hundreds of students (*n* = 362) across four institutions to experience this educational model. In the current versions of this game, which combines traditional classroom experiences (e.g., lectures, laboratory work, field data, software models and simulations) with virtual activities, students get experience with a Cone-Penetration Testing (CPT) field testing technique. In this paper, we present the perceived impact of this game on different student populations with a focus on *gender* and *race/ethnicity*. These results serve as a launching point for a more deliberate effort to consider issues of diversity, equity, and inclusion as we introduce game-based learning into engineering education.

II. Background

Prior to introducing GeoExplorer, we elaborate in this section on the use of game-based learning in engineering education. Then we discuss the issues of diversity, equity, and inclusion in engineering education in general. Finally, we discuss the issues of gender and race/ethnicity in the context of inclusive game design.

A. Game-Based Learning in Engineering Education

Game-based learning with its various enactments has been emerging as a new pedagogical practice in engineering education. Games for teaching engineering Computer-Aided Design (CAD) tools represent a category on their own [16]. For example, GamiCAD is a gamified tutorial system for the Autodesk AutoCad software [17]. In addition to research focusing on engineering tools, recent studies have been concerned with the ways in which game-based learning allows for development of both cognitive and non-cognitive skills. For example, a study of the application of a game designed to teach introductory engineering lessons to first-year students revealed that it fosters more interactions between students and instructors and improves learning and motivation [18]. Overall, a recent review found 191 papers published since 2000 where an educational game was used in an undergraduate engineering course [11]. Of these, only 62 papers discuss learning outcomes. This suggests a dearth in empirical evidence about game-based learning. The lack of empirical evidence in game-based engineering education is cited by other authors (e.g., [19]), which may be one of the reasons this educational practice is still not widespread [12].

We believe that a *mixed reality* approach may foster adoption of game-based learning in engineering education. By "mixed reality" we mean the combination of traditional curriculum elements (e.g., analyzing data, physical lab) with virtual ones (e.g., field testing in a virtual environment). We also believe that if game-based learning addresses the urgent issues of diversity, equity, and inclusion (DEI) in engineering education, it will help to foster its adoption. However, similarly to the lack of empirical evidence on game-based learning effectiveness, there is currently little understanding of how it may affect improvements in DEI in engineering education.

B. Diversity, Equity, and Inclusion (DEI)

The DEI issues in engineering education have been widely documented in education literature (e.g., see [20]). In addition to exploratory and explanatory studies, recent literature on the subject focuses on research-based instructional designs and interventions serving the goal of equitable participation of all students in engineering education [21]–[24]. With game-based learning environments identified in the recent studies as serving to support positive learning outcomes for all students, it behooves us to unpack our understanding further to seek answers to the question: How and to what extent does game-based learning provide equitable learning opportunities across all learners? To our knowledge, this question has not been addressed in the current engineering education literature and this work serves to bridge this gap. This paper is a starting point in this work: as we will discuss below in the Design section, we made a number of inclusive (game design) efforts and will report in this paper our results on some of the emergent DEI findings leveraging GeoExplorer.

C. Gender and Inclusive Game Design

It is an obvious non-functional requirement for the development of educational games to be inclusive. Unfortunately, when considering gender equity, for a long time, games were considered to be "a boys' media" [25]–[27] and this raises research concerns [28], [29]. Recent studies demonstrate that the number of women interested in video games has increased in the last decades [30], but that there are gender differences in the kind of games women and men choose to play [31]–[33].

Although this can be seen as progress, there are still a number of underlying problems about reaching more gender equity in gaming. One of them is the gamergate controversy, which surfaced a series of mysogynistic issues in the game world (e.g., internet forums, blogs, social network services, industry), considering it to be a toxic male's club, frequently harassing women and other minorities [34]–[36]. As well, the game industry mainly consists of male CEOs and developers, and, as a consequence, the companies apply a so-called "I" methodology, with the majority of the characters portraying male stereotypes [37], [38]. Situations like these, unfortunately, help to reinforce the paradigm that STEM fields are not a place for women and other minitorized groups [39].

This suggests the critical importance of efforts to understand how educational games are perceived by women and other underrepresented in STEM groups in the classroom and that special attention is placed on inclusive game design, especially when such games are applied in areas already suffering from inequities as is the case in engineering education [14].

D. Race, Ethnicity, and Inclusive Game Design

Any effort in the development of educational tools is incomplete without strategies to include groups underrepresented in engineering, including racial and ethnic minoritized groups. In fact, increasingly equity-related efforts are made to address the issues of racial disparities and creating opportunities for more equitable learning environments [40]. This requires consideration of *white supremacy* and *systemic structures of oppression*, which serve to perpetuate inequities (e.g., a toy or an educational game designed for a "masculine" persona alone may be more appealing to boys, further contributing to the video gaming world as a boy's club [41]). These issues are further complicated by often implicit and subconscious, *implicit bias* that, without harmful intentions, has far-reaching and lasting negative effects on all groups underrepresented in STEM, including racial and ethnic minorities [42]. With more emergent studies revealing that racial and ethnic equity and diversity is critical and necessary in improving educational outcomes for all learners, particularly those from minitorized learner groups [43], [44], the equity and inclusion aspects of game-based learning environments are now a requirement for educational games [40], [45], [46]. While we are aware of efforts focused on gender equity in game-based learning in

Fig. 1: Screenshot of GeoExplorer: Driving the CPT truck.

engineering education [15], to our knowledge the issues of race/ethnicity have not been addressed in this context.

III. Design

The game GeoExplorer is part of a larger effort to transform engineering education with mixed-reality game-based learning, specifically in civil/geotechnical engineering. The vision behind this game is to provide students experience with various field testing techniques. In this section, we describe how the game works and what inclusive design efforts were made.

A. GeoExplorer

Our first step in developing GeoExplorer was to focus on the topic of Cone Penetration Testing (CPT), a field testing technique for which students traditionally get little to no hands-on exposure. It involves a common in-situ method to determine the geotechnical engineering properties of soils and delineating soil stratigraphy. It does this by pushing an instrumented cone into the ground, which registers different values that enable identification of soil types across different depths. For the CPT component of this game, we leveraged an existing game called CPT-Operator, which was developed by the research institute Deltares to inform practitioners. We adapted this game for engineering education in the USA and this was the first game version that we used in this project, *GeoExplorer–Basic*. Based on student feedback and our ambitions to make the game more suitable for higher education, we iterated this version into *GeoExplorer–Plus*. Finally, we made this more of a mixed reality game by expanding the game with a website—a version that we refer to as *GeoExplorer–Advanced*. We describe all three versions in more detail below.

1) GeoExplorer–Basic: GeoExplorer–Basic follows the same gameplay pattern of the original CPT-Operator. In this version, students are immersed into a realistic 3D environment where they control a CPT truck (see Figure 1), which is used to conduct the CPT. Furthermore, they have a mobile phone with which they can find help about the game, choose a mission, see their scores, and interact with their boss. After they pick a mission, they have to drive to the CPT location using the GPS on their phone. Upon arriving at the CPT location, they enter into the truck and prepare for the CPT by taking a number of steps such as selecting and cleaning a cone, and stabilizing the truck. Finally, students conduct the CPT by observing the values on the in-truck computer. Based on the assignment, students need to decide when to stop the CPT, which is an important decision because stopping too early will not provide the data needed and stopping too late could cause the cone to break and/or accrue unnecessary expenses. In this version, each mission ends after stopping the CPT.

2) GeoExplorer–Plus: Aside from visual and gameplay improvements, the main distinction between GeoExplorer–Basic with GeoExplorer–Plus is the inclusion of CPT data and requested analysis of this data. The game does not end after stopping the CPT. Rather, users are provided with a CSV file with the CPT data, which is stored locally on their computer. This data is used to identify the soil types at the CPT location. For the calculation, students can use any kind of software. As this calculation does not require specialized software, most students opt for Microsoft Excel. After they upload their report, they are called by their boss who asks them about their conclusion. The mission ends after this call.

3) GeoExplorer–Advanced: GeoExplorer–Advanced is more significantly different because it extends the game by including a website and the inclusion of a story and contexts for each of the missions. For both GeoExplorer–Basic and GeoExplorer–Plus the game essentially revolves around the virtual environment (VE), which we built with Unity. With GeoExplorer–Advanced, the game is expanded with a Wordpress website that represents Terra Inc., a young innovative engineering company with a focus on sustainability and that seeks interns to join their company. To play the game, students have to apply for an internship position. After acceptance, they can read about the different missions, and decide which mission to join. Each mission is contextualized around sustainability. For example, one of the CPT missions is about the use of a site for a solar panel farm.

B. Inclusive Efforts

From the start, the team made a conscious effort to design for a gender-neutral visual style [47]. In the process of developing GeoExplorer–Advanced, however, the team considered additional ways in which to make the game more inclusive. First, we made the company staff and CEO representative of diverse racial/ethnic groups. Second, we incorporated the concept of sustainability to highlight the societal good that can be achieved by engineers, a concept suggested recently as a simple solution to get "women [to] enroll in droves" [48]. As described above, the concept of sustainability is emphasized through Terra Inc.'s mission statement and each of the mission descriptions. For example, the Farmhouse mission states:

> Terra Inc. has been awarded a contract by a community farm's Board of Directors to conduct CPTs to evaluate the foundation soils for a joint farmhouse and storage facility. This community shared agriculture program involves constructing a large, root cellar-style vegetable storage facility attached to the

> vegetable processing barn and farmhouse. The aim is to apply traditional subterranean root cellars on a larger scale.
>
> Local farmers will be able to store fresh produce in the cellar and process it in the farm barn. This is a low-budget farmhouse intended to encourage and promote the production by the local farmers. The entire structure will be energy-efficient and guarantees a natural freshness of the products processed and sold to the rest of the city.

Other aspects of the game were difficult to re-envision from an inclusive design perspective as GeoExplorer realistically simulates the CPT procedure. The team did discuss whether the driving elements of the game would be more appealing to male students; however, prior research actually suggests the opposite might be true [15]. For *Racing Academy*, a game to support undergraduate students learning of mechanical engineering, the findings show that there was no gender difference in the beneficial effect of playing this game, but there is some evidence that female students found it more motivating. Moreover, driving is one of the few "playful" elements of the game that we conjectured would help get players immersed into the game. It gives the sense of being on a mission instead of simply completing a simulated task.

IV. Methods

We implemented GeoExplorer at four different schools: California State University Fullerton (CSUF), Manhattan College (MC), Rensselaer Polytechnic Institute (RPI), and Southern Methodist University (SMU). The implementations followed the same curriculum and are implemented in a similar course. All courses are undergraduate courses that provide an introduction to geotechnical engineering.

A. Participants

We implemented GeoExplorer–Basic in Spring 2016 with 29 students at RPI; in Spring 2017 GeoExplorer–Plus with 70 students at CSUF; and in Fall 2017 and Spring 2018 GeoExplorer–Advanced with 263 students at all four institutions. Unfortunately, among the GeoExplorer–Advanced population, 9 students did not complete the presurvey and 53 students did not complete the postsurvey. Across the entire student population, in terms of gender identification, 67% self-identified as male, 31% as female, and the rest preferred not to say or described themselves differently (e.g., non-binary). With regards to race and ethnicity, 4% self-identified as African American, 27% as Hispanic / Latino or Latina, 12% as Asian, 41% as White / Caucasian, and the rest preferred not to answer (11%) or described themselves differently (5%). Regarding the latter, most described themselves as Arab, Middle Eastern, or a mix. Participation was a required classroom activity and IRB was obtained at each institution to study the outcomes.

B. Materials

After completing GeoExplorer, students were asked to fill out a survey. As we changed the items across the years due to research interests, we focus here on the survey items that were the same across the implementations. In total an overlap exists in 15 items across four scales, which measure the *CPT Learning* of the intervention, its *Integration* with the rest of the classroom, how students perceived *Geotechnical Engineering Learning* from it, and how the intervention impacts potentially their *Career*. All survey items are 5-point Likert scales. Table IV provides an overview of the scales.

C. Procedure

At all schools GeoExplorer was implemented as part of a module focused on CPT. Students first completed the presurvey and then received the same lecture. The slides were prepared by the research team. After this lecture, students were instructed to play at least two scenarios (of approx. 20 min each). At some schools playing the game happened in the classroom, at other schools this was in the form of a homework assignment. In the GeoExplorer–Basic and GeoExplorer–Plus implementations, students received direct access to the virtual environment (VE); for the GeoExplorer–Advanced implementation, students first go to a Wordpress website where they make an account. After they are accepted by Terra Inc. as their new intern by email, they get access to the employee portal where they can read about the different missions and download the VE.

For GeoExplorer–Basic, CPT data is not provided to players and so no reports are requested. For the GeoExplorer–Plus version, students have to directly submit their calculations through the game. This involves pausing the gameplay and then uploading their report directly in the VE. With GeoExplorer–Advanced, on the other hand, students submit their reports on the website. Only for GeoExplorer–Advanced we included achievements: by completing at least two missions and the surveys, students become CPT master and by completing all CPT missions and the surveys they become CPT Champion. For GeoExplorer–Basic and GeoExplorer–Plus the instructor sends the surveys to the students directly; for GeoExplorer–Advanced the surveys are accessed through the employee portal. The latter may explain the lower response rate.

V. Results

We first evaluate the four scales—CPT Learning, Integration, Geotechnical Learning, and Career—and then proceed to look at the results in general, for the entire population and across the three versions (i.e., Basic, Plus, and Advanced, see Table II). Following this, we look at the results on gender and race/ethnicity. Here we look at the summative scales across all three versions (see Table III) and delve deeper into the individual items for the Advanced version (see Table IV), given that the larger student sample played this version. The population of the Advanced version is in terms of gender identical to the overall population; for race/ethnicity it is fairly similar: 5% self-identified Black of African American, 25% Latino/Latina or Hispanic, 15% Asian, 43% Caucasian or White, and the rest preferred not to answer (7%) or described themselves differently (5%).

TABLE I: The factor loadings of the initial 4-factor solution and the subsequent 3-factor solutions (with loadings $>$ .40).

	Iteration 1				Iteration 2			Iteration 3		
Item	*F1*	*F2*	*F3*	*F4*	*F1*	*F2*	*F3*	*F1*	*F2*	*F3*
1								–	–	–
2	70				.68			.75		
3	.70				.71			.71		
4	.65				.69			.68		
5		.57				.55			.55	
6		.64				.55			.60	
7		.67				.72			.70	
8								–	–	–
9					.42			.40		
10	.46				.47			.47		
11	.57				.61			.58		
12			.69				.71			.71
13			.72				.72			.73
14			.71				.70			.70
15			.76				.75			.76

A. Scale Evaluation

For the scale evaluation, we considered the reliability and validity of the survey items. For reliability, we looked at the Cronbach's Alpha. All original scales have an excellent reliability ($\alpha \geq .90$), highlighting that participants responded consistently across the items for each scale. As all items were specifically developed as part of this project, we performed an Exploratory Factor Analysis (EFA) to look into validity. As the items use a Likert scale and we anticipated that the factors would correlate, we ran an EFA on the data set without any missing values (n = 281) with polychoric correlations, oblimin as rotation method, and weighted least squares as factoring solution. A scree plot analysis suggested 3–5 factors, with a 4-factor solution the most likely optimum. Thus, for our first iteration we started with the original 4-factor solution and find that no item loads high enough ($>$ .40) on the fourth factor (F4). Moreover, three items (1, 8, and 9) do not load high on any of the factors. The results are, however, "clean" because all variables load high on one factor only (see Table I). It turns out that items associated with the Integration and Career scales all load onto the same factor; in contrast, items associated with the CPT Learning and Geotechnical Engineering Learning scales load onto a single factor.

In our second iteration, we considered a 3-factor solution. The results are the same except that this time Item 9 loads high enough on one of the factors (see Table I). As Items 1 and 8 continue to not load onto any of the factors, we removed these and ran the factor analysis again for a 3-factor solution for our third iteration. The results are similar. Thus, our EFA shows that the Integration and Career scales are robust; the two learning scales less so and can be merged into a single *Learning* scale. We note that the items associated with CPT learning load higher on this newly merged Learning factor compared to the items associated with Geotechnical Engineering Learning. It may be possible that there is another latent factor that explains the relationship between these two original scales.

TABLE II: Results on the summative scales, in *M* (*SD*), across all three GeoExplorer versions.

Scale	**Basic** *n* = 29	**Plus** *n* = 70	**Advanced** *n* = 207
CPT Learning	3.98 (0.53)	4.09 (0.56)	3.96 (0.75)
Integration	3.74 (0.46)	4.33 (0.65)	3.94 (0.92)
Geotechnical Learning	3.75 (0.53)	3.97 (0.62)	3.73 (0.80)
Career	3.07 (0.93)	3.15 (1.02)	3.30 (0.99)
Learning	3.82 (0.49)	3.97 (0.57)	3.84 (0.70)

Consistent with these outcomes, we find that the CPT Learning and Geotechnical Learning moderately correlate with each other, r = .59, $p < .001$. However, highlighting a difference between these scales, CPT Learning has a weak correlation with Integration, r = .14, p = .02; while Geotechnical Learning has a weak correlation with Career, r = .22, $p < .001$. These results make sense: the integration of the classroom activities can determine how well students learn about the specific topic of CPT, whereas what students have learned about geotechnical engineering at large may influence their stance on a career in this discipline. When considering the Learning scale, the weak relationship with Integration disappears. No other significant correlations exist. Based on these results, we consider for our further analyses both the original scales and the new Learning scale. For the calculation of the scales we summarized the individual items associated with each scale and divided the total by the number of items so each scale outcome ranges from 1 to 5.

B. General

Table II shows the descriptive statistics on the scales across the three versions. The results across all versions and populations show that students agreed with learning about CPT (M = 4.00, SD = 0.69) and that it was well integrated with other classroom activities (M = 4.00, SD = 0.86). We see lesser agreement on the more indirect measures such as learning about geotechnical engineering at large (M = 3.80, SD = 0.71) or pursuing a career in this field (M = 3.24, SD = 1.01). Overall, on all scales the Plus version scores higher than the Basic and Advanced version, except on Career, where Advanced scores higher.

However, using ANOVAs we find a significant difference between the versions on Integration, $F(2,293)$ = 7.37, $p <$ .001. Post hoc tests with Tukey's HSD indicate that the Plus implementation (M = 4.33, SD = 0.65) indeed outperforms both the Basic (M = 3.94, SD = 0.46) and Advanced (M = 3.75, SD = 0.92). For the Plus implementation the research team provided additional support, which may explain this difference. Another possibility is that when we considered the differences among schools, we only found a difference on Integration, $F(3,292)$ = 5.38, p = .001, and this difference was a result between the only institution that implemented GeoExplorer–Plus and two other institutions. We also find a significant difference on Geotechnical Learning, $F(2,252)$ = 3.32, p = .04. Post hoc tests reveal here that the Plus implementation (M = 4.05, SD = 0.61) has more of an impact than the Advanced version (M = 3.76, SD = 0.79).

C. Gender

For this analysis, we focused on those who identified themselves as female (F) or male (M) only, given that only 1% reported a different gender identity. First, we do not find any significant difference on gender for each specific version or a change within a gender between versions, except for Integration, which is a general finding and applies to both genders. Therefore, based on the summative scales no difference exists on gender. From Table III we retrieve, however, some interesting observations. On Integration female students score consistently higher across versions. Then, overall, female students score higher on CPT Learning compared to Geotechnical Learning. Finally, on Career female students score lower on the Basic version, quite higher on the Plus version, and about equal on the Advanced version.

In considering the results on the individual items for the Advanced version (see Table IV), we see that female students have a higher interquartile range (e.g., on Item 1 4–5 for female students compared to 4–4 for male) on each item that makes up the CPT Learning scale, and specifically on Item 1, which is about learning how to use the CPT equipment. For Integration the difference is most noticeable with Item 7, which is focused on putting theory to practice. Thus, it seems that on items that speak to the practical aspects female students tend to score higher. In contrast, on the Career scale we see that in pursuing a concentration (Item 13) or a graduate degree (Item 14) in geotechnical engineering female students score lower. So it seems fewer female students are interested in gaining more specialized knowledge in this field. Nevertheless, none of these observations are significant: for gender no difference exists on the individual items as well.

D. Race/Ethnicity

For this analysis, we grouped the population into Asian (A), Latino/Latina or Hispanic (L), Caucasian or White (W), and other race/ethnicity (O). The other race/ethnicity consists of students who identified themselves as Black/African American, Native American/Hawaiian, or a different race/ethnicity. This group makes up 9% of our total student sample. We excluded from our analysis students who preferred not to answer this question. We also excluded the Asian group for the Plus version because only a single student identified as such. For the Basic version, it is important to note that 18 students (67%) identified as white. The other categories include only three students each. As such, we are showing the Basic version for descriptive purposes only and do not consider it for any of our inferential analyses on race/ethnicity.

Similar to gender, we generally did not observe significant differences within and between versions for race/ethnicity on the summative scales (see Table III). However, the other race/ethnicity (O) group has a significant difference with every other group on the Integration scale for the Advanced version, $F(3,170) = 5.316$, $p < .001$. This is in contrast to the Plus version (or Basic version), where this group seems to have experienced the integration better compared to the Advanced version, $t(22) = 3.04$, $p = .006$. Thus, the difference may be due to the format of the Advanced version. In contrast, we find that for the Career scale, the L group scored higher for the Advanced version compared to the Plus version, $t(48) = -2.69$, $p = .01$. As for the Career scale, while not significant, the most notable differences are observable here.

The results on the individual items (see Table IV) confirm the results based on the summative scales: no significant differences except for the items on Integration between the other race/ethnicity (O) group and the rest. There are some interesting observations, however. The A group has a lower interquartile range (IQR) on all Integration items compared to the L and W groups and on almost all Career items compared to all other groups. On the CPT Learning items, the L group has consistently the highest, followed by the A, W, and, finally, the O group. The most striking observation is that, except for Career, the O group scores consistently the lowest as determined from the IQR scores.

VI. Discussion

In this paper, we present the results of evaluating a game called GeoExplorer on four measures: CPT Learning, Integration, Geotechnical Learning, and Career. The overall results on these measures are positive, more so for the measures directly related to the game (i.e., CPT Learning and Integration) than measures that are more indirect (i.e., Geotechnical Learning and Career). Thus, our work contributes to the growing body of evidence of positive empirical results pertaining to the use of game-based learning in engineering education [11]. However, the main aim of our paper was to consider the impact of this game on different student populations, with a focus on gender and race/ethnicity. In this section we discuss our findings.

A. Generalizability and Validity

Before discussing our findings, it is first important to reflect on our student population and measures. As for gender, with 31% identifying as female, our student population includes more women engineering students compared to the average in the USA (i.e., 19.9%). Numbers reaching approx. 30% are, however, not uncommon, especially not for engineering fields that typically attract more women, such as civil engineering [14]. Regarding race/ethnicity, our numbers are strikingly similar compared to the numbers reported by ASEE [13] across engineering education. The main difference is that our student population has more students identifying as Latina/Latino or Hispanic (30% compared to 10.7%) and less as Caucasian or White (46% compared to 64.9%). More than half (55%) of our participants came from one of the four institutions and this university is considered an Hispanic-Serving Institution, which explains this disparity.

All scale measures show strong reliability. Examining the individual responses shows that the results are indeed consistently similar across most items for each scale (see Table IV). In terms of validity we find that two measures are robust:

TABLE III: Results on the summative scales, in *M* (*SD*), across all three GeoExplorer versions on gender and race/ethnicity.

	Basic						**Plus**[*]					**Advanced**					
	Gender[a] (n = 29)		*Race/Ethnicity*[b] (n = 27)				*Gender*[a] (n = 69)		*Race/Ethnicity*[b] (n = 51)			*Gender*[a] (n = 191)		*Race/Ethnicity*[b] (n = 180)			
Scale	F	M	A	L	W	O	F	M	L	W	O	F	M	A	L	W	O
CPT Learning	4.13 (0.67)	3.84 (0.29)	3.92 (0.63)	3.92 (0.14)	4.04 (0.61)	3.88 (0.18)	4.06 (0.46)	4.12 (0.57)	4.18 (0.59)	4.10 (0.48)	4.11 (0.66)	4.01 (0.78)	3.96 (0.71)	4.06 (0.64)	4.06 (0.85)	3.93 (0.72)	3.76 (0.96)
Integration	3.86 (0.55)	3.64 (0.34)	3.78 (0.69)	3.67 (0.33)	3.69 (0.49)	4.00 (0.0)	4.45 (0.46)	4.32 (0.67)	4.27 (0.69)	4.59 (0.46)	4.43 (0.63)	4.04 (0.85)	3.92 (0.94)	3.92 (0.94)	4.14 (0.80)	4.11 (0.72)	3.28 (1.28)
Geotechnical Learning	3.75 (0.67)	3.75 (0.38)	4.00 (0.75)	3.92 (0.14)	3.64 (0.56)	3.83 (0.52)	4.14 (0.55)	3.95 (0.62)	4.01 (0.66)	4.16 (0.63)	4.00 (0.27)	3.71 (0.84)	3.80 (0.72)	3.98 (0.70)	3.76 (0.86)	3.69 (0.74)	3.61 (0.90)
Career	2.98 (1.06)	3.15 (0.83)	3.33 (1.53)	3.67 (0.58)	2.97 (0.87)	3.08 (0.38)	3.39 (1.02)	3.04 (0.95)	2.87 (0.98)	3.08 (1.21)	3.50 (0.74)	3.28 (1.16)	3.34 (0.89)	3.18 (1.08)	3.47 (0.88)	3.28 (1.03)	3.30 (0.90)
Learning	3.91 (0.62)	3.74 (0.34)	3.83 (0.76)	3.89 (0.19)	3.82 (0.56)	3.75 (0.12)	4.04 (0.40)	3.98 (0.59)	4.01 (0.59)	4.11 (0.58)	4.08 (0.47)	3.85 (0.73)	3.86 (0.66)	4.01 (0.64)	3.90 (0.81)	3.79 (0.61)	3.69 (0.92)

Gender: F = Female; M = Male. *Race/Ethnicity*: A = Asian; L = Latino/Latina or Hispanic; W = Caucasian or White; O = other race/ethnicity.
[*] We excluded the Asian category for the Plus version as only a single student identified as such.

TABLE IV: Overview of the GeoExplorer-Advanced results (n = 207) on all individual survey items, in *Mdn* (*IQR*).

		Gender[a] (n = 191)		**Race/Ethnicity**[b] (n = 180)			
Item	**Description**	F	M	A	L	W	O
CPT Learning (α = .92)							
1	Learn about the operation of CPT equipment.	4 (4 - 5)	4 (4 - 4)	4 (4 - 4.75)	4 (4 - 5)	4 (4 - 4)	4 (3 - 4)
2	Visualize subsurface stratigraphy.	4 (4 - 4.5)	4 (4 - 4)	4 (4 - 4)	4 (4 - 5)	4 (4 - 4)	4 (3 - 4)
3	Simulate field conditions	4 (4 - 4.5)	4 (4 - 4)	4 (4 - 4)	4 (4 - 5)	4 (4 - 4)	4 (3 - 4.5)
4	Conduct a geotechnical site investigation using CPT equipment.	4 (4 - 4.5)	4 (4 - 4)	4 (4 - 4.75)	4 (4 - 5)	4 (4 - 4)	4 (3 - 4.5)
Integration (α = .94)							
5	The class lectures prepared me for conducting the CPT in the mock internship.	4 (4 - 5)	4 (4 - 5)	4 (3 - 5)	4 (4 - 5)	4 (4 - 5)	4 (2.5 - 4)
6	The class lectures prepared me for analyzing the results from the CPT.	4 (4 - 5)	4 (4 - 5)	4 (3.25 - 5)	4 (4 - 5)	4 (4 - 4.5)	3 (2.5 - 4)
7	The CPT Environment was a good way to put what I learned in class into practice.	4 (4 - 5)	4 (4 - 4.5)	4 (3 - 5)	4 (4 - 5)	4 (4 - 5)	4 (3 - 4)
Geotechnical Learning (α = .90)							
8	Helped me understand geotechnical engineering better.	4 (3 - 4)	4 (4 - 4)	4 (4 - 4)	4 (3 - 4)	4 (4 - 4)	4 (3 - 4)
9	Made me want to learn more about geotechnical engineering.	4 (3 - 4)	4 (3 - 4)	4 (3 - 4)	4 (3 - 4)	4 (3 - 4)	3 (3 - 4)
10	Helped show me why geotechnical engineering is important.	4 (3 - 4)	4 (3 - 4)	4 (3 - 5)	4 (4 - 4)	4 (3 - 4)	3.5 (3 - 4)
11	Was an effective way to learn about site investigation methods that are relevant to geotechnical engineering.	4 (3.5 - 4.5)	4 (4 - 4)	4 (4 - 4)	4 (3 - 4)	4 (4 - 4)	4 (3 - 4)
Career (α = .94)							
12	Take more classes that focus on geotechnical engineering.	4 (3 - 4)	4 (3 - 4)	3 (2 - 4)	4 (3 - 4)	3.5 (3 - 4)	4 (3 - 4)
13	Pursue a concentration in geotechnical engineering.	3 (2.5 - 4)	3 (3 - 4)	3 (2 - 4)	3 (3 - 4)	3 (3 - 4)	3 (3 - 4)
14	Pursue a graduate degree in geotechnical engineering.	3 (2 - 4)	3 (3 - 4)	3 (2 - 4)	3 (3 - 4)	3 (2 - 4)	3 (3 - 4)
15	Seek employment with a geotechnical engineering firm.	3 (3 - 4)	3 (3 - 4)	3 (2 - 4)	4 (3 - 4)	3 (3 - 4)	3 (3 - 4)

Gender: F = Female; M = Male. *Race/Ethnicity*: A = Asian; L = Latino/Latina or Hispanic; W = Caucasian or White; O = other race/ethnicity.

Integration and Career. The two measures on learning—CPT Learning and Geotechnical Learning—could be merged into a single Learning scale. However, from our findings we see that different response patterns emerge that are important to distinguish (e.g., on gender), which makes sense given that CPT Learning is directly impacted by the game and Geotechnical Learning indirectly. Further work is needed to refine and validate these two scales.

B. Gaming4All?

By and large our findings do not show any significant differences within and between versions of GeoExplorer, supporting that game-based learning is an equitable learning environment. However, while it is important to be cautious as many observations are not significant, our work shows areas worthy of further attention and investigation that indicate possible differences. On gender we find that female students are more favorable than their male counterparts on learning about the topic and putting what they learned in the classroom (i.e., theory) to practice. This suggests that contextualizing the work and showing its relevance is potentially helpful for this underrepresented group. A particular area that needs attention is highlighting the usefulness of acquiring more specialized knowledge, as female students were less inclined to pursue a concentration or graduate degree in geotechnical engineering.

On race/ethnicity we find a few significant differences, which we discuss below. The most striking and alarming observation is that the other race/ethnicity group, which is composed of students identifying as Black or African American, Native American or Hawaiian, and self-described (predominantly Middle Eastern, Arab, or mix), are inclined to score lower on all measures, except for Career. This suggests that the most marginalized and underrepresented groups may need more attention. Our results suggest a clear need for assisting this population in implementing game-based learning.

C. Implementation Matters

The clearest differences in our data are observed on the Integration scale, which discusses the relationship between the class lecture and CPT virtual environment. We find a difference on the Plus version implementation compared to the Basic and Advanced version, which can be explained that this implementation received more support and that this institution generally scores higher on this measure compared to the other institutions. We also find a difference between the other race/ethnicity groups for the Advanced version and the rest. In addition, while not significant, the Asian group scores consistently lower on the items associated with this scale. Game-based learning needs to be embedded well into existing curricula, and our results suggest that care is needed from a diversity, equity, and inclusion (DEI) perspective.

D. Mixed Results for Mixed Reality and Inclusive Efforts

We pursued a mixed reality approach in our work because we believe that this could foster adoption of game-based learning in engineering education—by combining (or "mixing") traditional engineering curricula elements with virtual ones. In the versions of GeoExplorer (i.e., Basic, Plus, Advanced), we increasingly included more mixed reality aspects. Furthermore, for the Advanced version we made several inclusive efforts, by considering representation and contextualizing the work around sustainability. The current results are mixed. While other factors may have played a role, on Integration and Geotechnical Learning the Plus version outperforms the Advanced the version. The difference on Integration could be explained that the Advanced version is a more complex implementation with the addition of the website. On Geotechnical Learning the difference is surprising, given that the Advanced version added more context and relevance for the tasks. At the same time, we see that the Advanced version benefited the Latina/Latino or Hispanic group regarding their career perspectives. While positive, we targeted specifically women and not this group for our inclusive efforts. Therefore, we do not see much impact on the efforts we made, both on mixed reality and inclusiveness, and the impacts we observe require further investigation.

E. Limitations

Due to the necessity of making meaningful inferences with our data, we proceeded by excluding (e.g., students who did not identify as either female or male) or grouping certain populations (e.g., Black or African American with Native American or Hawaiian), thereby, as a result, marginalizing these groups. We note that it is important to consider such groups in the future; however, this would require a larger data set. Moreover, to be more inclusive, in our future work we are including for gender the option "prefer to self-describe" and for race/ethnicity the options "South Asian" and "Middle Eastern or North African."

Related, an inherent limitation in our work are the sheer student numbers associated with the different versions. This makes a comparison between versions difficult and even forced us to exclude certain race/ethnic groups. In order to advance inclusive (game) design practices, it is of crucial importance to continue to evaluate the impact of technological innovations in a design-based research (DBR) fashion [49]. Only by systematically and rigorously evaluating design iterations we will able to determine what has a positive outcome.

Educational interventions generally suffer from a multitude of variables that may influence outcomes [3], such as the influence of instructors and institutions. We did not consider the possible influence of instructors. We did look into institutions and found this to be only of influence on the Integration scale. Aside from this scale, the numbers are quite similar across versions, institutions, gender, and race/ethnicity, suggesting that unobserved or unaccounted for variables may not play an important role. However, it may suggest that our measures do not capture the differences well enough. In addition to further validating our scales, we plan to work with a 7-point Likert scale instead of the 5-point Likert scale presented in this paper to see if we can better capture the possible differences. We also note that our results have not been adjusted for the multiple comparisons we made (i.e., for each scale). Applying a Bonferroni correction, the most conservative adjustment, would only make the Geotechnical Learning effect between the Plus and Advanced version insignificant.

The clearest limitation of our work, however, is that all outcomes are *perceived* and thus reported by the students themselves. Such measures suffer from a reporting bias and do not shed light on actual academic performance. In addition, all measures are reported at the end of the experience and do not consider students' initial starting points. It might very well be that female students are similarly interested in pursuing a career in geotechnical engineering at the end of the activity; however, they may have been less interested before the activity. We plan to explore this in our future work.

Acknowledgment

This material is based upon work supported by the National Science Foundation (DUE-141838 and DUE-1915247). We thank everyone involved with the GeoExplorer project.

References

[1] C. Harteveld, *Triadic game design: Balancing reality, meaning and play*. Springer Science & Business Media, 2011.

[2] J. Gee, *What Video Games Have To Teach Us about Learning and Literacy*. New York, NY: Palgrave Macmillan, 2003.

[3] C. Harteveld, *Making sense of virtual risks: A quasi-experimental investigation into game-based training*. Amsterdam, the Netherlands: IOS Press, 2012.
[4] M. S. Gresalfi and J. L. Barnes, "Designing feedback in an immersive videogame: Supporting student mathematical engagement," *Educ. Technol. Res. Dev.*, p. 1–22, 12 2015.
[5] V. J. Shute, "Stealth assessment in computer-based games to support learning," *Computer Games and Instruction*, p. 503–524, 2011.
[6] B. E. Wiggins, "An overview and study on the use of games, simulations, and gamification in higher education," *International Journal of Game-Based Learning*, vol. 6, no. 1, p. 18–29, 2016.
[7] M. Papastergiou, "Digital game-based learning in high school computer science education: Impact on educational effectiveness and student motivation," *Computers & Education*, vol. 52, p. 1–12, 2009. [Online]. Available: http://dx.doi.org/10.1016/j.compedu.2008.06.004
[8] J. Muller, "The future of knowledge and skills in science and technology higher education," *Higher Education*, vol. 70, p. 409–416, 2015. [Online]. Available: http://dx.doi.org/10.1007/s10734-014-9842-x
[9] N. R. Council *et al.*, *Engineering in K-12 education: Understanding the status and improving the prospects*. National Academies Press, 2009.
[10] L. D. Feisel and A. J. Rosa, "The role of the laboratory in undergraduate engineering education," *Journal of Engineering Education*, vol. 94, no. 1, pp. 121–130, 2005.
[11] C. A. Bodnar, D. Anastasio, J. A. Enszer, and D. D. Burkey, "Engineers at play: Games as teaching tools for undergraduate engineering students," *J. Eng. Educ.*, 2016.
[12] A. A. Deshpande and S. H. Huang, "Simulation games in engineering education: A state-of-the-art review," *Comput. Appl. Eng. Educ.*, vol. 19, no. 3, p. 399–410, 9 2011.
[13] B. L. Yoder, "Engineering by the numbers," 2015.
[14] J. Bossart and N. Bharti, "Women in engineering: Insight into why some engineering departments have more success in recruiting and graduating women." *American Journal of Engineering Education*, vol. 8, no. 2, pp. 127–140, 2017.
[15] R. Joiner, J. Iacovides, M. Owen, C. Gavin, S. Clibbery, J. Darling, and B. Drew, "Digital games, gender and learning in engineering: Do females benefit as much as males?" *Journal of Science Education and Technology*, vol. 20, no. 2, pp. 178–185, 2011.
[16] Z. Kosmadoudi, T. Lim, J. Ritchie, S. Louchart, Y. Liu, and R. Sung, "Engineering design using game-enhanced cad: The potential to augment the user experience with game elements," *Computer-Aided Design*, vol. 45, no. 3, pp. 777–795, 2013.
[17] W. Li, T. Grossman, and G. Fitzmaurice, "Gamicad: a gamified tutorial system for first time autocad users," in *Proceedings of the 25th annual ACM symposium on User interface software and technology*, 2012, pp. 103–112.
[18] L. F. Braghirolli, J. L. D. Ribeiro, A. D. Weise, and M. Pizzolato, "Benefits of educational games as an introductory activity in industrial engineering education," *Computers in Human Behavior*, vol. 58, pp. 315–324, 2016.
[19] A. P. Markopoulos, A. Fragkou, P. D. Kasidiaris, and J. P. Davim, "Gamification in engineering education and professional training," *International Journal of Mechanical Engineering Education*, vol. 43, no. 2, pp. 118–131, 2015.
[20] A. L. Pawley and A. Slaton, "The power and politics of stem research design: Saving the "small n."," in *ASEE Annual Conference & Exposition, Seattle, Washington. doi*, vol. 10, 2015, p. 24901.
[21] D. R. Simmons and S. M. Lord, "Removing invisible barriers and changing mindsets to improve and diversify pathways in engineering." *Advances in Engineering Education*, 2019.
[22] E. Foster, J. Karlin, M. S. Quiles-Ramos, and D. Riley, "Seeding a strategic campaign to address root causes of inequity in engi-neering education." Collaborative Network for Engineering and Computing Diversity Conference . . . , 2019.
[23] F. Zhang, V. E. Lee, R. Jin, S. Garg, K.-K. R. Choo, M. Maasberg, L. Dong, and C. Cheng, "Privacy-aware smart city: A case study in collaborative filtering recommender systems," *Journal of Parallel and Distributed Computing*, vol. 127, pp. 145–159, 2019.
[24] L. Malcom-Piqueux and S. M. Malcom, "Engineering diversity: Fixing the educational system to promote equity," *Bridge*, vol. 43, no. 1, pp. 24–34, 2013.
[25] J. R. Dominick, "Videogames, television violence, and aggression in teenagers." *Journal of communication*, 1984.
[26] G. R. Loftus and E. F. Loftus, *Mind at play; The psychology of video games*. Basic Books, Inc., 1983.
[27] S. J. Kaplan, "The image of amusement arcades and differences in male and female video game playing," *Journal of Popular Culture*, vol. 17, no. 1, p. 93, 1983.
[28] J. B. Funk and D. D. Buchman, "Children's perceptions of gender differences in social approval for playing electronic games," *Sex Roles*, vol. 35, no. 3-4, pp. 219–231, 1996.
[29] M. A. Barnett, G. D. Vitaglione, K. K. Harper, S. W. Quackenbush, L. A. Steadman, and B. S. Valdez, "Late adolescents' experiences with and attitudes toward videogames 1," *Journal of Applied Social Psychology*, vol. 27, no. 15, pp. 1316–1334, 1997.
[30] M. D. Dickey, "Girl gamers: The controversy of girl games and the relevance of female-oriented game design for instructional design," *British journal of educational technology*, vol. 37, no. 5, pp. 785–793, 2006.
[31] D. Ghuman and M. Griffiths, "A cross-genre study of online gaming: Player demographics, motivation for play, and social interactions among players," *International Journal of Cyber Behavior, Psychology and Learning (IJCBPL)*, vol. 2, no. 1, pp. 13–29, 2012.
[32] A. Eden, E. Maloney, and N. D. Bowman, "Gender attribution in online video games," *Journal of Media Psychology*, 2010.
[33] J. Jansz and L. Martens, "Gaming at a lan event: the social context of playing video games," *New media & society*, vol. 7, no. 3, pp. 333–355, 2005.
[34] A. Massanari, "# gamergate and the fappening: How reddit's algorithm, governance, and culture support toxic technocultures," *New Media & Society*, vol. 19, no. 3, pp. 329–346, 2017.
[35] S. Chess and A. Shaw, "A conspiracy of fishes, or, how we learned to stop worrying about# gamergate and embrace hegemonic masculinity," *Journal of Broadcasting & Electronic Media*, vol. 59, no. 1, pp. 208–220, 2015.
[36] T. E. Mortensen, "Anger, fear, and games: The long event of# gamergate," *Games and Culture*, vol. 13, no. 8, pp. 787–806, 2018.
[37] R. Potanin, "Forces in play: The business and culture of videogame production," in *Proceedings of the 3rd International Conference on Fun and Games*, 2010, pp. 135–143.
[38] C. Kerr, B. Francis, K. Cross, G. C. Guide, and I. Games, "Computer science stereotypes as barriers to inclusion for women and how they extend to videogames," *Health*, 2020.
[39] S. M. Jackson, A. L. Hillard, and T. R. Schneider, "Using implicit bias training to improve attitudes toward women in stem," *Social Psychology of Education*, vol. 17, no. 3, pp. 419–438, 2014.
[40] I. F. Ogbonnaya-Ogburu, A. D. Smith, A. To, and K. Toyama, "Critical race theory for hci," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI'20). https://doi.org/10.1145/3313831.3376392*, 2020.
[41] B. Friedman and H. Nissenbaum, *Bias in Computer Systems*. USA: Center for the Study of Language and Information, 1997, p. 21–40.
[42] D. Hankerson, A. R. Marshall, J. Booker, H. El Mimouni, I. Walker, and J. A. Rode, "Does technology have race?" in *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2016, pp. 473–486.
[43] A. L. Antonio, M. J. Chang, K. Hakuta, D. A. Kenny, S. Levin, and J. F. Milem, "Effects of racial diversity on complex thinking in college students," *Psychological Science*, vol. 15, no. 8, pp. 507–510, 2004.
[44] T. L. Strayhorn, "Sense of belonging and stem students of color," in *College Students' Sense of Belonging*. Routledge, 2012, pp. 76–91.
[45] C. J. Passmore, M. V. Birk, and R. L. Mandryk, "The privilege of immersion: Racial and ethnic experiences, perceptions, and beliefs in digital gaming," in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–19.
[46] A. To, J. McDonald, J. Holmes, G. Kaufman, and J. Hammer, "Character diversity in digital and non-digital games," *Transactions of the Digital Games Research Association*, vol. 4, no. 1, 2018.
[47] G. Williams, "Are you sure your software is gender-neutral?" *interactions*, vol. 21, no. 1, pp. 36–39, 2014.
[48] L. Nilsson, "How to Attract Female Engineers," *The New York Times*, Dec. 2017. [Online]. Available: https://www.nytimes.com/2015/04/27/opinion/how-to-attract-female-engineers.html
[49] W. Sandoval, "Conjecture mapping: An approach to systematic educational design research," *Journal of the Learning Sciences*, vol. 23, no. 1, p. 18–36, 2014.