How Immersion and Self-Avatars in VR Affect Learning Programming and Computational Thinking in Middle School Education

Dhaval Parmar*, Lorraine Lin†, Nikeetha DSouza‡, Sophie Jörg†, Alison E. Leonard†, Shaundra B. Daily§, and Sabarish V. Babu†

Abstract—We present an empirical evaluation of immersion and self-avatars as compared to desktop viewing in Virtual Reality (VR) for learning computer programming and computational thinking in middle school education using an educational VR simulation. Students were asked to programmatically choreograph dance performances for virtual characters within an educational desktop application we built earlier called Virtual Environment Interactions (VEnvI). As part of a middle school science class, 90 students from the 6th and 7th grades participated in our study. All students first visually programmed dance choreography for a virtual character they created in VEnvI on a laptop. Then, they viewed and interacted with the resulting dance performance in a between-subjects design in one of the three conditions. We compared and contrasted the benefits of embodied immersive virtual reality (EVR) viewing utilizing a head-mounted display with a body-scaled and gender-matched self-avatar, immersive virtual reality only (IVR) viewing, and desktop VR (NVR) viewing with VEnvI on pedagogical outcomes, programming performance, presence, and attitudes towards STEM and computational thinking. Results from a cognition questionnaire showed that, in the learning dimensions of Knowledge and Understanding (Bloom's taxonomy) as well as Multistructural (SOLO taxonomy), participants in EVR and IVR scored significantly higher than NVR. Also, participants in EVR scored significantly higher than IVR. We also discovered similar results in objective programming performance and presence scores in VEnvI. Furthermore, students' attitudes towards computer science, programming confidence, and impressions significantly improved to be the highest in EVR and then IVR as compared to NVR condition. Our work suggests that educators and developers of educational VR simulations, who want to enhance knowledge and understanding as well as simultaneous acquisition of multiple abstract concepts, can do so by employing immersion and self-avatars in VR learning experiences.

Index Terms—Virtual reality, computer science education, embodied cognition, self-avatars, immersion, VR in middle school education.

1 Introduction

Virtual reality (VR) has gained immense popularity in recent years. From its earliest conceptions in science fiction stories, and its beginnings with bulky, expensive hardware, VR today is ever more accessible and affordable, transitioning from research labs into everyday household use. Perhaps VR garners this appeal through its novelty, immersion, and naturalistic interactions, blurring the lines between the real and the virtual. Compelling illusion and the ability to substitute reality allows VR to simulate scenarios otherwise difficult to achieve, and provides avenues for training, therapy, and entertainment [1], [2], [3].

Virtual reality has great potential for educational applications, and research has shown numerous advantages of using virtual reality-based instruction for learning [4], [5]. Even with the increasing demand for jobs in computer science and related STEM (science, technology, engineering, and mathematics) fields, there is a lack of representation of minorities and women in western nations [6], [7]. The interest in STEM fields for girls is much lower compared to boys, and this interest is often lost during middle school [8]. In our previous work, we approached this problem by creating Virtual Environment Interactions (VEnvI) [9], a desktop VR educational application, which uses a visual programming intervention similar to Alice [10] and Scratch [11]. VEnvI uses concepts from dance and computer programming to get middle school students, especially girls, interested in computer science and teach them basics of coding. Our initial exploratory study [9] showed enthusiasm overall for immersive VR to complement learning of computer science concepts through VEnvI. This research greatly extends our initial work by specifically investigating the role of virtual reality, especially the factor of virtual self-avatars and immersion, in enhancing the learning of abstract computer science concepts, potentially facilitating embodied cognition [12]. Kilteni et al. [13] explain virtual embodiment as having three underlying subcomponents: self-location, or the "spatial experience of being inside a body"; agency, or the "sense of having global motor control"; and body ownership, or "one's self-attribution of a body". Llobera et al. [14] state that a full body ownership within VR combines the virtual and the real body into one overall entity. Such an illusion is obtained by having a co-located virtual self-avatar which moves correspondingly and synchronously with the real body. Virtual embodiment is powerful in affecting how the brain represents and interprets the body, shown by instances such as overestimation of object sizes when occupying a virtual child's body [15], affecting drumming performance

*Northeastern University
{d.parmar}@northeastern.edu
†Clemson University
{lorrain, sjoerg, aleona2, sbabu}@clemson.edu
‡Indiana University
{nfarfand}@iu.edu
§Duke University
{shani.b}@duke.edu

with different virtual bodies [16], and reduction in implicit racial bias by embodying light-skinned people in a dark-skinned virtual body [17]. Virtual embodiment also has the ability to improve speed and accuracy in distance estimation and behavioral task performance [18], [19], [20], and having a strong effect on the sense of physical presence, social presence, and self-presence [21], [22], [23]. Self-avatars are able to enhance the perception and judgment of action capabilities (affordances) in VR experiences [24]. Taken together, a large body of research has shown that embodiment in VR using self-avatars facilitates the use of a robust body-scheme and enhances perception-action coordination [25], which in-turn potentially enhances engagement and learning of tasks from a body-centric perspective [26].

Virtual self-avatars, even at a rudimentary level, provide a possibility of facilitating embodied cognition within VR. The theory of embodied cognition stems from the idea that the brain's sensory modalities, states of the body, and situated actions form the basis of cognition [27]. Having a body provides a foundation for cognition, and learning is greatly enhanced through embodied approaches [28]. Embodied cognition proposes viewing the learning process as "not a mind working on abstract problems, but a body that requires a mind to make it function" [29]. A key goal of this research is to utilize embodied processes of inquiry to help students engage in computational thinking. The idea of learning computer programming via embodied cognition builds upon the work of Papert [30], who has shown that the learning and understanding of mathematical concepts while programming is more efficient among students when their active engagement with the programming knowledge is associated with their knowledge of self, culture, and the body. An immersive, embodied experience within VEnvI afforded by VR can be used to couple dance and programming, where learners get to actually perform moves in addition to programming them and think through their virtual body, making VR a perfect platform for facilitating embodied cognition.

The idea of using dance as a medium for learning computational concepts was introduced with the Dancing Alice project [31], in which students learned basic elements of dance and then used them to program choreography for virtual characters. Throughout the study, we observed students standing up consistently in front of their computer and using their physical bodies to think through the actuation of dance moves for their characters. Furthermore, dance created new avenues for students who might not typically be interested in computing, therefore "allowing for interdisciplinary, embodied engagement, broadening their perspectives on dance and computing" [32]. The knowledge and sense of one's own body, or body-syntonicity as coined by Papert [30], has the ability to bootstrap students' intuitive knowledge in order to learn programming concepts. However, the Dancing Alice project was developed using Alice [10] and had its limitations. Students were put off by the disjointed appearance of the virtual characters, and desired "human-like" characters and movements [31]. Students also desired the ability to customize their characters by changing their gender, ethnicity, clothing, and body shape, often replicating their own identifying characteristics [33], [34]. The results of that research led us to the creation of the VEnvI software, utilizing the features of the Unity game engine and motion-captured animation to impart realism to the characters. VEnvI also introduced immersive virtual reality to students in an exploratory study to anecdotally evaluate the feasibility of VR for learning programming concepts [9], and students responded positively and with great enthusiasm towards VEnvI and an immersive viewing of their programmed moves for a virtual character.

In this research, we aim to conduct a rigorous empirical evaluation to study the effectiveness of immersive VR and self-avatar-based interaction, utilizing our VEnvI software and associated 6-week curriculum introduced in our earlier research, on teaching simple core programming concepts and computational thinking principles to middle school students. To the best of our knowledge, our contribution is one of the first studies on examining how immersion and self-avatars affect learning, telepresence, social presence, and user impressions in educational VR interactions for middle school children. The subsequent sections are organized into related work, system description, experiment design, methodology, results, discussion, and conclusions of our research and contributions.

2 Related work

2.1 Virtual reality for training

With the ability of VR to spark motivation and interest, VR has been a powerful tool for skills training. Since VEnvI aims to train students in computer programming skills using VR, it is important to look at research utilizing VR for training. For example, Seymour et al. [1] found significantly improved task performance among surgical students with the user of the Minimally Invasive Surgical Trainer (MIST-VR) system to train skills pertinent to the operating room environment. Similarly, Armstrong et al. [35] demonstrated the use of VR in military training. Their research evaluated the Virtual Reality Stroop Task (VRST), which measured reaction time in a military convoy scenario with simulated combat threats, showing the system to be on par with computerized and traditional tests of attention and executive functioning.

Research has shown that VR can improve learning. Bertrand et al. [36] showed significant improvements in learning outcomes by using immersive VR for bimanual psychomotor skills training in metrology. Similarly, Johnson et al. [37] showed that interactive virtual humans can be used effectively to educate users in interpersonal social situations, such as doctor-patient interactions and medical diagnostic procedures. Furthermore, Raij et al. [38] showed that learners potentially treat virtual interlocutors in interpersonal training simulations in a manner similar to real people, highlighting the power of VR systems with virtual humans in teaching abstract concepts in perceived real world situations, showing potential for students using VEnvI to learn abstract programming concepts.

Highly related to VEnvI is the work by Chan et al. [3], who created a dance training system using VR and 3D motion capture. Users observed a virtual trainer and mimicked the trainer's movements to learn how to dance. The users reported that the VR system increased their interest in dancing and motivated them to learn. Utilizing the potential of immersive embodied VR to empower education by exposing new knowledge, encouraging retention, and positively altering student perspectives is ripe for exploration within VR research, which VEnvI aims to do.

2.2 Virtual reality in STEM education

The use of VR to impart science, technology, engineering, and mathematics (STEM) education has been heavily explored and supported in literature. In an extensive survey on the use of 3D immersive virtual worlds in K-12 and higher education conducted by Hew et al. [4], it was found that students preferred learning in the virtual worlds because of the ability to move freely within the environment, the ability to meet and interact with virtual agents and peers, and the ability to experience the simulated 3D environment. The review of the studies suggested that virtual worlds can be helpful in stimulating social behavior among participants through the use of avatars. Similarly, Johnson-Glenberg et al. [39] conducted a review of existing research on semi-virtual environments having video-game elements and analyzed the effects on embodied STEM learning and assessment. Their analysis found significant improvements in learning of STEM concepts within the embodied, semi-virtual environments when compared to regular classroom instruction, and suggested embodiment to be a vital component in game-based learning. These surveys show the potential of immersion and embodiment within VEnvI to have a significant impact in learning computer science concepts.

In another research, Johnson-Glenberg et al. compared desktop versus VR based interaction, and interaction fidelity of mouse, controller, and watching a video of the experience on learning [40]. They found that platform did not directly affect learning, but presence, agency, and engagement facilitated learning as experienced by participants in the VR with controller condition. Dalgarno et al. [41] analyzed 3D virtual learning environments to identify the affordances impacting learning outcomes. These affordances comprised of enhanced spatial knowledge representation tasks, experiential learning opportunities, increased motivation and engagement, enhanced contextualization of learning, and richer collaborative learning possibilities as compared to 2D alternatives. Therefore, it would be impactful to study multiple platform modalities and affordances within VEnvI to formulate a clear understanding of the effects of VR on learning. VR has the power to enhance traditional modes of instruction, but it is essential to incorporate VR into the school curriculum in a way that augments the curriculum and provides additional learning benefits.

2.3 Immersive vs. non-immersive applications

Comparative studies pitching immersive virtual learning environments against traditional non-immersive training methods are crucial in understanding the impact of VR on learning. In one research, Coulter et al. [42] compared the effects of a fully immersive head-mounted display (HMD) based learning environment to a desktop-based learning environment for medical education and found that participants in the HMD condition had significantly higher knowledge gains as compared to the desktop condition. Similarly, Chittaro et al. [43] found that HMD-based immersive learning was better than traditional card-based learning for airline safety education among passengers. Participants found the immersive environment more engaging and fear-inducing than the safety cards, which possibly contributed to higher knowledge retention. In a previous research, we compared immersive head mounted display based viewing to a traditional desktop viewing metaphor on psychomotor skills learning of electrical circuitry related concepts and motor tasks in engineering students [44]. We found that immersive viewing facilitated the active learning of concepts and action of abstract concepts related to electricity much more effectively than traditional desktop simulations. These studies have implications to the current work that investigates the extent to which immersive viewing and self-avatars in active learning environments can connect motor actions associated with dance to abstract concepts in computational thinking.

Conversely, Juan et al. [45] did not find any significant differences in immersive versus desktop training in their research where students learned about the interior of the human body. However, the children enjoyed learning via either training method, suggesting that immersive VR did not have negative effects on learning. Similarly, Madden et al. [46] investigated immersive VR, desktop VR, and hands-on real world interaction on learning and attitudinal measures. Although there was no difference in learning between conditions, they found that gender and game play experience seemed to affect learning, supporting the notion that participant experience should be considered. Therefore, immersive virtual learning environments can provide learning benefits on par with, if not better than the traditional methods, and exceed in performance gains when interactivity is essential.

Immersive VR can be developed in conjunction with existing teaching methods to augment the learning experience. Research exploring the effects of immersion and presence among children within virtual learning environments is sparse, and it is important to conduct empirical studies to gain better insight into children's reactions, attitudes, and expectations of immersive embodied VR.

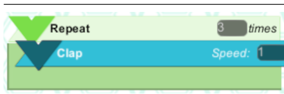
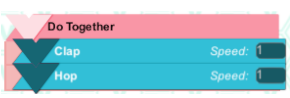
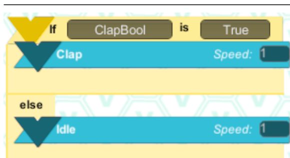
2.4 Our contributions

The VEnvI curriculum included classroom education, physical movement and exercises, desktop-based learning of computer science concepts using the VEnvI software, and immersive self-avatar-based viewing of the programmed choreographies. We conducted a novel empirical evaluation using a 3×2 multifactorial between-subjects experiment design to examine the effects of immersive embodied virtual reality via a co-located virtual self-avatar (EVR) vs. immersive viewing alone (IVR) vs. desktop viewing without VR (NVR) on presence, engagement, and cognition of programming concepts and principles in middle school students.

These effects were measured using multiple surveys, including pre- and post-cognition questionnaires based on a revised version of Bloom's taxonomy [47] and the Structure of Observed Learning Outcomes (SOLO) taxonomy [48]. The Bloom's and SOLO cognitive taxonomies are used for classification of educational learning objectives into levels representing a continuum of increasing cognitive complexity. Bloom's taxonomy classifies cognition into dimensions of remembering (identifying and retrieving knowledge), understanding (clarifying, illustrating, categorizing), application (executing and implementing), analysis (differentiating, organizing), evaluation (checking, critiquing), and creation (planning, producing). The SOLO taxonomy classifies cognition into dimensions of unistructural (identify, follow), multistructural (combine, describe, list), relational (analyze, apply, criticize), and extended abstract (create, hypothesize, reflect). A telepresence and social presence survey from the Nowak and Biocca presence questionnaire [49] that we adapted for middle school children, and a general survey to measure usability, satisfaction, and attitudes towards programming were also employed. Programming performance was also measured using a metric developed to analyze the programmed choreographies of the students. In later sections, we briefly describe the VEnvI programming tool and the EVR, IVR, and NVR conditions, outline the study design, and present the results of our empirical evaluation.

3 System Description

TABLE 1
The six programming concepts afforded by VEnvI.

VEnvI element	Concept	Functionality
	Sequence	Perform dance moves in a sequence
	Loop	Repeat a sequence of moves, a set number of times
	Parallelism	Perform an upper-body and a lower-body movement simultaneously
	Variable	Create a boolean variable and modify its value
	Conditional	Conditional branching of choreography based on a variable check
	Function	Modularize a set of moves as a reusable function

VEnvI is a visual block-programming software that allows learners to use computer programming concepts to create dance choreography for a virtual character. Students can drag-and-drop blocks of programming elements (see Table 1) and connect them together to create a program (see Figure 2). The basic elements are the blue blocks which are atomic motion-captured dance sequences that the character can perform, such as 'turn left', 'step right', or 'hop'. These sequences are divided into locomotor moves, where the character ends up in a new position at the end of the move (e.g., 'slide right'), and non-locomotor moves, where the character's position at the end of the move is the same as at the start of the move (e.g., 'clap'). These basic sequence blocks can be repeated, combined, branched, and modularized using the programming elements described in Table 1.

Students in all three conditions could view the result of their programs instantaneously by clicking on the 'Play' button and watching their character perform in the VE window on the top-left (see Figure 2) or in full-screen. Students in the immersive VR conditions (EVR and IVR) had the added ability to view their programmed choreography using an Oculus Rift head-mounted display. A Microsoft Kinect v2 motion sensor (Figure 1C) was used to track the students' bodies and to place a co-located self-avatar which mimicked their movements (see Figure 1D and Figure 1E). By nature, dance involves a great range of motion, and we did not want the students to be limited in their motion while experiencing the immersive VR. Further, our virtual embodiment design was an approximation, with the primary purpose of engaging the students in the learning experience. Therefore, to implement this low fidelity embodiment using self-avatars we opted for unobtrusive computer-vision-based tracking, for which the Kinect v2 sensor was popular and readily available to the researchers at the time of designing the intervention.

One male and one female character were available as self-avatars for the students as shown in Figure 3, and the students were assigned one or the other based on their identifying gender. The ethnicity of the avatar was not varied, and the avatar's skin texture was kept at an average tone (see Figure 3). The body type of the avatar was not matched with the students either, but the proportions of the torso, limbs, and overall height of the avatar were automatically scaled using the tracked information from the Kinect sensor. The self-avatar was co-located with the students, and they saw the avatar body in place of their own when they looked at it. The self-avatar mimicked the movements of the students. The Kinect sensor tracked 25 joints on the body, and the students could experience a good level of detail in their movements including hands, legs, feet, as well as fingers. The finger tracking was not perfect due to the distance from the tracker, and a few times the fingers on the avatar would not move with the student's fingers. For the head, only position tracking was used from the Kinect tracking data, and orientation was obtained from the Oculus Rift.

On their first interaction with the self-avatar in the EVR condition, the students acclimated themselves with their virtual body by doing simple exercises such as bringing their arms forward and to the side, lifting their legs one by one, stepping forwards, backwards, side to side, turning around, and jumping. Students spent some time observing their surroundings and looking down at their virtual body during this first encounter.

IPD was kept at a constant 60 mm, the average IPD for ages 11–13 [50], to reduce setup times between participants, so any student can use the system at any desired time during the class. The students were free to move around the classroom, but would see a semi-transparent barrier within their virtual view, indicating the end of the Kinect-trackable area. The system would lose tracking when students crossed

Fig. 1. (A) Students using VEnvI on a laptop. (B) Students performing a dance activity. (C) System setup for the immersive embodied viewing metaphor. (D) A student in VEnvI's immersive embodied interaction metaphor examining her virtual body, and (E) her embodied point of view within VEnvI showing her self-avatar.

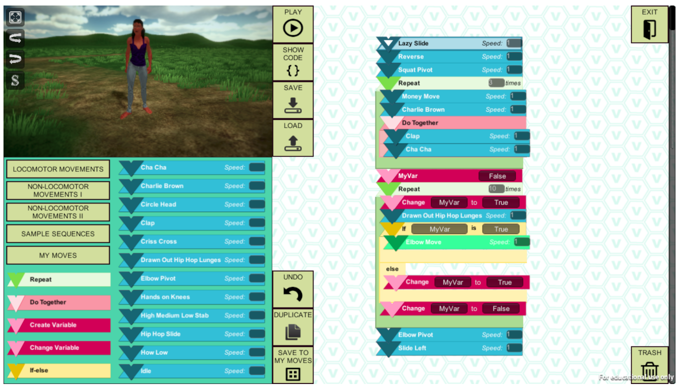


Fig. 2. The VEnvI user interface showing the VE window on the top left, the move-selection area on the bottom left, and the drag-and-drop area on the right.

Fig. 3. The female and male characters used as self-avatars.

the barrier, but the tracking would be regained as soon as the student re-entered the trackable area. The students did not express any specific issues due to lost tracking, and most students tried to stay within the tracking barrier.

The system was developed using the Unity3D game engine. The co-located self-avatar was only present in the EVR condition and was disabled in the IVR condition. The goal of the immersive visualization was to allow students to examine their characters and their programmed choreographies from a first-person perspective, letting them explore their dances in an immersive manner, and encouraging active embodied learning. At any point during the class session, students could bring their programs into the immersive VR metaphor at a separate VR workstation setup in one end of the classroom, evaluate their programs, think about modifications, dance with and even learn from their characters.

4 Comparative Evaluation

4.1 Experiment Design

In our empirical evaluation, we aim to factor out the effects of self-avatars and immersion by comparing three conditions: desktop-based non-immersive VR, a traditional technology based learning metaphor, as a negative control (NVR); HMD-based immersive VR without any virtual self-embodiment (IVR); and an immersive embodied VR condition which adds a co-located, gender-matched, body-tracked, virtual self-avatar to the IVR condition to embody the user within the virtual environment (EVR). The primary question this research aims to answer is:

RQ 1. *To what extent does the presence of immersive viewing and virtual self-avatars within an educational virtual environment facilitate learning of abstract computational concepts in middle school children?*

This research also aims to answer the following question as a secondary impact:

RQ 2. *Can immersion and virtual self-avatars positively affect presence, programming performance, and attitudes towards computer science in middle school children?*

Since research has shown improvement in affordances [51], [52], [53], perception and user engagement due to added immersion and embodiment in virtual environments [26], [54], in combination with the embodied cognition theory [27], [29], these three conditions are expected to get incrementally stronger in facilitating immersion and embodied cognition in a continuum from NVR to IVR to EVR. Our primary hypotheses are as follows:

H1. *The students in the immersive embodiment condition (EVR) will have higher cognitive scores than students in the immersive-only condition (IVR), who in turn will have higher cognitive scores than students in the desktop condition (NVR).*

H2. *The students in the EVR condition will have higher objective programming performance than students in the IVR condition, who in turn will have higher objective programming performance than students in the NVR condition.*

Furthermore, due to immersion and self-avatars each having positive effects on presence [21], [22], [23], and user perception and attitudes [16], [17], we also hypothesize the following:

H3. *The three conditions will have differential but incremental impact on presence, programming performance, and attitudes towards computer science, with the impact increasing incrementally from NVR to IVR to EVR.*

4.2 Participants

This research was conducted at a partnering middle school. 90 students (59 female, 31 male) from the 6th and 7th grades were included in this research study. Although more students initially participated, but not all students completed this research study due to time commitment and dropouts. Final number of participants who were randomly assigned to one of the three experiment conditions and completed the study were as follows: NVR (n=28), IVR (n=26), and EVR (n=36) condition. The study was advertised in the study hall sessions and elective classes such as graphics communication and aerobics. Participation was voluntary, and at each iteration only a small fraction of the class chose to participate. Informed parental consent as well as student assent for participation in the study was obtained following the guidelines approved by the institutional review board of the researchers' university. It was difficult to control and counterbalance the participants due to the low sample sizes and dropouts in each study iteration. Therefore, this study had to be conducted in four distinct iterations spanning one and a half years to bring the participant pool to a comparable size and distribution across all three conditions. Demographically, a diverse set of 59 middle school students identified as White, 12 as African American, 6 as Hispanic/Latino, 4 as Asian, 7 as Multiracial, and 2 as Other/Unspecified participated in the study. Participant age was between 11 and 14, and consisted of 31 males and 59 females.

4.3 Measures

Various measures were employed in this study to obtain quantitative and qualitative insights regarding the participants in the VEnvI program. A demographic survey was administered to gather general information about the participants. To measure knowledge gain, pre and post cognitive tests based on various levels of the revised Bloom's and SOLO taxonomies were administered to the students. These tests included questions such as:

- The first block will cause the character to clap twice. How many times will the character clap with the second set of blocks?
- What is a variable? Explain in your own words. Give an example.
- Give an example of a conditional in your life.

Presence was measured using a questionnaire adapted from the Nowak and Biocca presence inventory [49], as shown in Table 5. A pre- and post-survey was administered asking the students questions about their views on programming, to assess if the three conditions within the VEnvI program had any impact on changing their perceptions regarding programming, with questions such as:

- Do you feel like you are confident at programming?
- Do you want to learn more about programming?

Finally, a debriefing questionnaire was used to gather qualitative responses regarding students' overall experience, system usability and satisfaction.

4.3.1 Programming performance

An additional measure in this objective was to assess programming performance of the students by analyzing students' programmed choreographies within VEnvI. The scoring criterion was adapted from our previous work [55]. We define and can measure programming performance in VEnvI through elements such as number of unique and repeated move blocks, number of unique and repeated programming concept blocks (CS blocks), the highest level in the hierarchy of nested blocks, and the duration (in seconds), which were extracted from the saved VEnvI files of the final programmed performance for each student. Unique blocks within the programmed choreography represented variety. Therefore, each unique occurrence of a movement block or a CS block was awarded twice as many points as those gained when using the same blocks repeatedly. Nested blocks within the program represented complexity, and the highest level in the nesting hierarchy showed the maximum complexity reached by the student when creating the programmed choreography. Finally, the duration of the choreography was also taken into account, with longer running programs getting more points. Students were awarded 10 points for every 30 seconds of the program (calculated as 1/3 points per second). The scoring criterion for calculating the programmed performance score is shown in Table 2.

TABLE 2
Scoring criterion for programming performance within VEnvI.

Scoring Element	Points
Move block	1 per block
Unique move block	2 per block
CS block	5 per block
Unique CS block	10 per block
Highest level in hierarchy	5 per level
Duration	1/3 per second

4.4 Procedure

The weekly progression of activities for this study is shown in Table 3. All study groups (NVR, IVR, and EVR) received the same instructions and performed the same activities. Students in the EVR group visualized their programmed choreography in the immersive embodied virtual environment with a co-located self-avatar which mimicked their movements, whereas the self-avatar was absent in the IVR condition. Students in the NVR condition viewed their programmed choreography on their computer screens only. Students in both IVR and EVR conditions were introduced to the Oculus Rift and the Kinect motion sensor and were informed about the benefits and risks involved when using the VR equipment. In all conditions, students were encouraged to move and dance with their characters to think through the creation of the programmed choreography and enjoy the result.

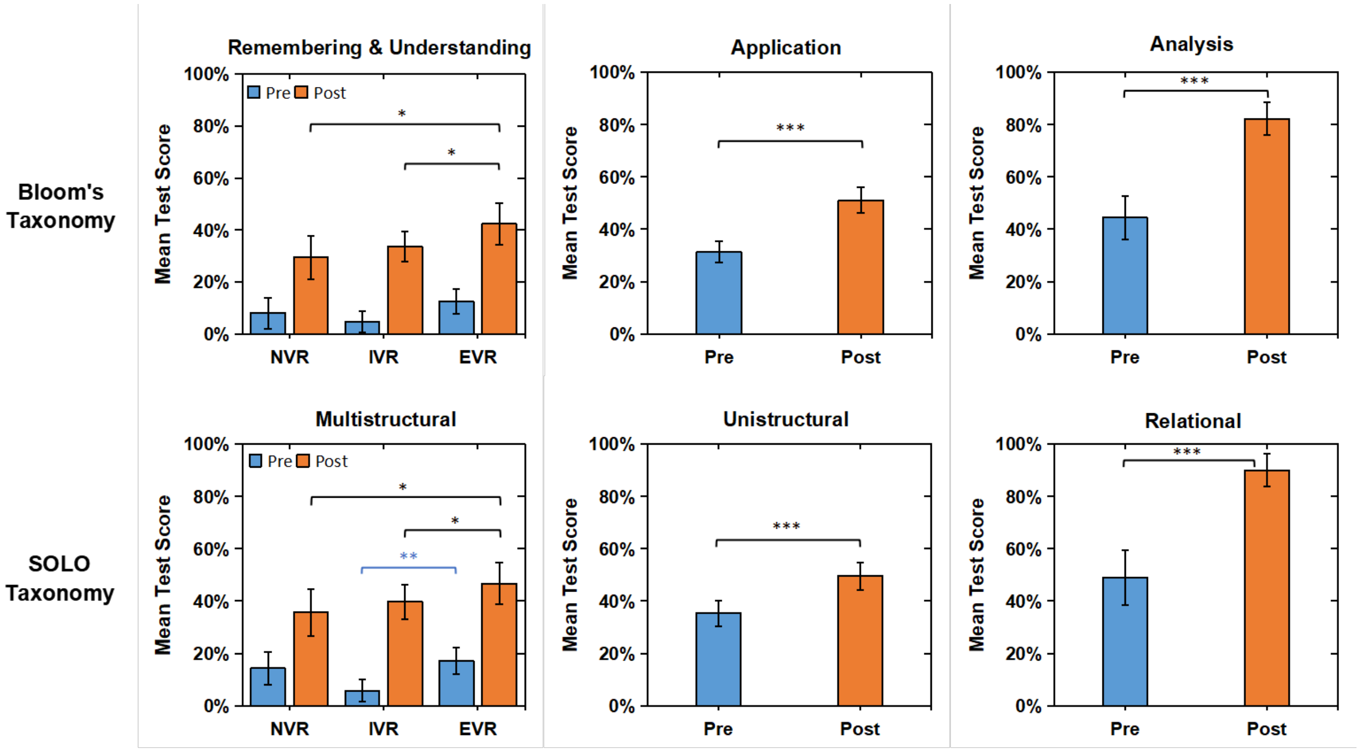


Fig. 4. Mean cognitive test scores for the remembering and understanding (top-left), application (top-center), and analysis (top-right) categories of the Bloom's taxonomy, and the multistructural (bottom-left), unistructural (bottom-center), and relational (bottom-right) categories of the SOLO taxonomy. *, **, and *** indicate a significant statistical difference with $p < 0.05$, $p < 0.01$, and $p < 0.001$ respectively. Error bars represent 95% confidence interval.

5 Results

5.1 Quantitative results: Cognition

The questions in the cognition questionnaire were analyzed using two principal cognitive taxonomies: Bloom's revised cognitive taxonomy and the Structure of Observed Learning Outcomes (SOLO) taxonomy. In analyzing the data gathered on the mean scores in each of the categories, a 2×3 mixed-model repeated measures ANOVA was employed. This was inspired and conducted in a similar manner from a methodological and statistical analysis perspective to studies on learning in VR conducted by Zanbaka et al. [56] and Suma et al. [57], which have examined the general learning effects of content presented in VR experiences. The within-subjects factors were the mean pre- and post-cognition scores in each of the dimensions of a taxonomy, indicated as 'session' in the analysis, and the between-subjects factor was the viewing metaphor (NVR vs. IVR vs. EVR), indicated as 'condition'. Parametric ANOVA analyzes were conducted on the data after carefully verifying that the underlying assumptions were met, such as the data in the samples were normally distributed and error variance between samples were equivalent. Thus, it was ensured that Box's test was not significant, Levene's test was conducted to verify homogeneity of variance, and Mauchly's test of sphericity was conducted to ensure that error variance in groups of samples are equivalent. Pairwise post-hoc tests for levels of the between-subjects variables were conducted using Tukey's HSD method, and for levels of the within-subjects variable was conducted using the Bonferroni adjusted alpha method. Table 4 shows key results from the post-hoc analysis. Figure 4 shows mean student cognitive scores for the Bloom's and SOLO taxonomies.

5.1.1 *Bloom's taxonomy*

Remembering & Understanding: The ANOVA analysis revealed a significant main effect of session, F(1,87) = 186.92, p <0.001, p. $\eta^2 = 0.70$, a main effect of condition, F(1,87) = 3.43, p = 0.04, p. $\eta^2 = 0.08$, and a session by condition interaction effect, F(1,87) = 3.07, p = 0.05, p. $\eta^2 = 0.07$.

In order to examine the condition by session interaction further, we did block analysis by comparing between conditions within the pre-test or the post-test scores. Post-hoc pairwise comparisons of the pre-test scores did not reveal any significant differences between conditions. Post-hoc pairwise comparisons using Tukey's HSD analysis of the post-test scores revealed that students in the EVR condition (M=42.36%, SD=23.77) scored significantly higher than students in the IVR condition (M=33.65%, SD=14.04), p = 0.05, 95% CI = [-3.99, 21.41], and students in the EVR condition also scored significantly higher than students in the NVR condition (M=29.46%, SD=21.57), p = 0.04, 95% CI = [0.46, 25.33]. Overall, in the post-test scores, participants in the EVR condition scored the highest and participants in the NVR condition scored the lowest, with participants in IVR scoring in the middle. Post-hoc pairwise comparisons between pre-test and post-test sessions for each condition using Bonferroni method revealed the following. In the NVR condition, post-test scores (M=29.46%, SD=21.57) were significantly higher than the pre-test scores (M=8.04%, SD=15.29), p ¡ 0.001, 95% CI = [14.71, 28.14]. In the IVR condition, post-test scores (M=33.65%, SD=14.04) were signifi-

TABLE 3
Weekly plan of activities for the VEnvI outreach program.

Week	Activity
1	• Introductions and pre-surveys. • Warm-up activities and introduction to dance.
2	• Introduction to programming. • Learning sequences and performing physical activities to demonstrate sequences. • Introduction to the VEnvI software. • Introduction to the immersive embodied metaphor (EVR and IVR condition). • Programming sequences in VEnvI.
3	• Learning loops and performing physical activities to demonstrate loops. • Learning parallelization and performing physical activities to demonstrate parallelization. • Programming loops and parallelization in VEnvI. • Students alternate between programming and viewing their programmed choreography in the immersive VR metaphor with a self-avatar (EVR condition) or without a self-avatar (IVR condition) or on a large screen or laptop display (NVR condition).
4	• Learning variables and performing physical activities to demonstrate variables. • Learning conditionals and performing physical activities to demonstrate conditionals. • Learning functions and performing physical activities to demonstrate functions. • Programming variables, conditionals, and functions in VEnvI. • Students alternate between programming and viewing their programmed choreography in the immersive VR metaphor with self-embodiment (EVR condition) or without self-embodiment (IVR condition) or on a large screen or laptop display (NVR condition).
5	• Programming for a dance challenge. • Students alternate between programming and viewing their programmed choreography in the immersive VR metaphor with a self-avatar (EVR condition) or without a self-avatar (IVR condition) or on a large screen or laptop display (NVR condition).
6	• Viewing dance challenge performances on HMD (EVR and IVR conditions) and classroom projection display (NVR condition). • Post-surveys.

cantly higher than the pre-test scores (M=4.81%, SD=10.05), p ¡ 0.001, 95% CI = [23.94, 35.78]. In the EVR condition, post-test scores (M=42.36%, SD=23.77) were significantly higher than the pre-test scores (M=12.50%, SD=14.01), p ¡ 0.001, 95% CI = [21.88, 35.81].

Application: The analysis revealed a significant main effect of session, F(1,87) = 66.05, p <0.001, p. $\eta^2 = 0.45$. Overall, a Bonferroni pairwise comparison revealed that participants scored significantly higher in the post-test session (M=51.11%, SD=23.78) as compared to the pre-test session (M=31.36%, SD=19.82), p <0.001, 95% CI = [14.82, 24.36].

A block analysis of post-hoc pairwise comparisons of cognition scores between the conditions using Tukey's HSD analysis within pre-test only or within post-test only, did not reveal any significant differences.

Analysis: The analysis revealed a significant main effect of session, F(1,87) = 59.88, p <0.001, p. $\eta^2 = 0.42$. Overall, a Bonferroni pairwise comparison revealed that participants scored significantly higher in the post-test session (M=82.22%, SD=30.27) as compared to the pre-test session (M=44.44%, SD=39.98), p <0.001, 95% CI = [14.82, 23.36].

TABLE 4
Key results among the three condition groups from the quantitative analysis of the post-test cognitive questionnaire. * indicates a significant statistical difference with p <0.05.

Condition	N	μ	SD	p
Bloom's - Remembering & Understanding				
EVR	36	42.36	23.77	0.05*
IVR	26	33.65	14.04	
EVR	36	42.36	23.77	0.04*
NVR	28	29.46	21.57	
SOLO - Multistructural				
EVR	36	46.76	23.85	0.05*
IVR	26	39.74	16.38	
EVR	36	46.76	23.85	0.05*
NVR	28	35.71	23.00	

A block analysis of post-hoc pairwise comparisons of cognition scores between the conditions using Tukey's HSD analysis did not reveal any significant differences between conditions in the post-test scores, but revealed a significant difference between conditions in the pre-test scores. Post-hoc pairwise comparisons between conditions in the pre-test session revealed that participants in EVR (M=56.94%, SD=38.08) scored significantly higher than participants in IVR (M=26.92%, SD=32.34), p <0.01, 95% CI = [6.43, 53.62].

5.1.2 SOLO Taxonomy

Unistructural: The analysis revealed a significant main effect of session, F(1,87) = 29.64, p <0.001, p. $\eta^2 = 0.27$. Overall, a Bonferroni pairwise comparison revealed that participants scored significantly higher in the post-test session (M=49.6%, SD=24.96) as compared to the pre-test session (M=35.34%, SD=23.36), p = 0.001, 95% CI = [8.56, 19.16].

A block analysis of post-hoc pairwise comparisons of cognition scores between the conditions using Tukey's HSD analysis within pre-test session only or within post-test session only, did not reveal any significant differences.

Multistructural: The analysis revealed a significant main effect of session, F(1,87) = 175.58, p <0.001, p. $\eta^2 = 0.68$, a significant main effect of condition, F(1,87) = 3.20, p = 0.05, p. $\eta^2 = 0.07$, and a session by condition interaction effect, F(1,87) = 3.31, p = 0.045, p. $\eta^2 = 0.05$.

Post-hoc Tukey's HSD pairwise comparisons revealed that students in the EVR condition (M=46.76%, SD=23.85) scored significantly higher in the post-test session than students in the IVR condition (M=39.74%, SD=16.38), p = 0.05, 95% CI = [-6.29, 20.32] and students in the EVR condition also scored significantly higher than students in the NVR condition (M=35.71%, SD=23.00), p = 0.05, 95% CI = [-1.98, 24.07]. Post-hoc pairwise comparisons using Tukey's HSD analysis of the pre-test scores revealed that participants in the EVR condition (M=17.13%, SD=14.63) scored significantly higher than participants in the IVR condition (M=5.77%, SD=10.48), p <0.01, 95% CI = [2.71, 20.02]. Post-hoc pairwise comparisons between pre-test and post-test sessions for each condition using Bonferroni method on Multistructural scores revealed the following. In the

NVR condition, post-test scores (M=35.71%, SD=23.00) were significantly higher than the pre-test scores (M=14.28%, SD=16.17), p ¡ 0.001, 95% CI = [14.17, 28.69]. In the IVR condition, post-test scores (M=39.74%, SD=16.38) were significantly higher than the pre-test scores (M=5.77%, SD=10.48), p ¡ 0.001, 95% CI = [26.44, 41.50]. In the EVR condition, post-test scores (M=46.76%, SD=23.84) were significantly higher than the pre-test scores (M=17.13%, SD=14.63), p ¡ 0.001, 95% CI = [23.23, 36.03].

Relational: The analysis revealed a significant main effect of session, F(1,87) = 48.76, $p < 0.001$, p. $\eta^2 = 0.37$. Overall, a Bonferroni pairwise comparison revealed that participants scored significantly higher in the post-test session (M=90%, SD=30.17) as compared to the pre-test session (M=48.89%, SD=50.27), $p < 0.001$, 95% CI = [31.78, 53.40].

A block analysis of post-hoc pairwise comparisons of cognition scores between the conditions using Tukey's HSD analysis within pre-test session only or within post-test session only, did not reveal any significant differences.

5.2 Quantitative results: Presence

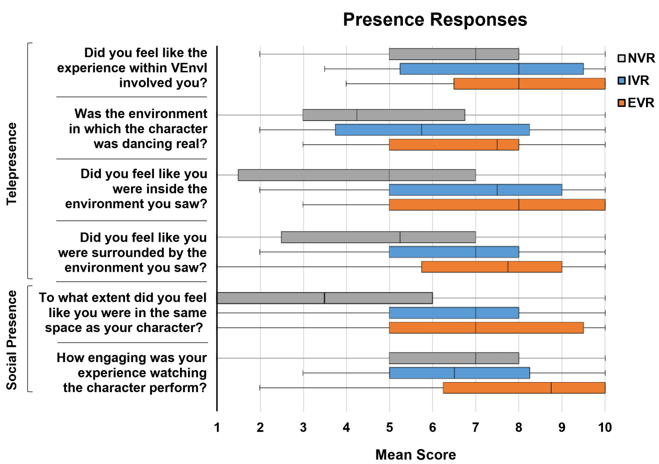


Fig. 5. Box plot for the quantitative responses for telepresence and social presence by condition.

The presence dependent variables consisted of five questions related to telepresence, three questions related to social-presence, and two questions related to identity. Each of the questions assessed a sub-dimension of the presence factors they measured and were rated by the students on a 1 (Not Intense) to 10 (Very Intense) scale. The quantitative results were treated with a univariate independent samples ANOVA on the factor of condition (NVR vs. IVR vs. EVR) on each dependent measure of presence. In each of the tests, Levene's test of equality of error variances was tested to ensure that error variances in the groups of samples were statistically equivalent before the ANOVA analysis was further conducted. Figure 5 shows the descriptive statistics of the quantitative presence responses. Table 5 shows results from the post-hoc analyses.

Regarding telepresence question **"involved"**, the ANOVA analysis revealed a significant effect of condition, F(1,89) = 3.36, p = 0.04, p. $\eta^2 = 0.07$. Overall, participants in the EVR condition rated their experience to be significantly

TABLE 5
Descriptive statistics and post-hoc results from the quantitative analysis of the presence questionnaire. *** indicates a significant statistical difference with $p < 0.001$; ** indicates a significant statistical difference with $p < 0.01$; * indicates a significant statistical difference with $p < 0.05$.

Question	Cond.	N	μ	SD	p
Did you feel like the experience within VEnvI **involved** you?	EVR NVR	36 28	8.05 6.62	1.96 2.04	0.03*
Did you feel like the environment in which the character was dancing was **real**?	EVR NVR	36 28	6.85 4.70	2.11 2.69	0.006**
	IVR NVR	26 28	6.32 4.70	2.73 2.69	0.044*
Did you feel like you were **inside** the environment you saw?	EVR NVR	36 28	7.25 4.75	2.45 3.15	0.001***
	IVR NVR	26 28	6.92 4.75	2.33 3.15	0.01**
Did you feel like you were **surrounded** by the environment you saw?	EVR NVR	36 28	7.29 5.07	2.38 2.80	0.002**
	IVR NVR	26 28	6.96 5.07	2.20 2.80	0.017*
To what extent did you feel like you were in the same **space** as your character?	EVR NVR	36 28	6.89 4.09	2.60 3.09	<0.001***
	IVR NVR	26 28	6.50 4.09	2.58 3.09	0.005**
How **engaging** was your experience watching the character perform?	EVR IVR	36 26	7.82 6.65	1.95 2.50	0.036*

more involving (M=8.05, SD=1.96) as compared to those in the NVR condition (M=6.62, SD=2.04), p = 0.03.

Regarding telepresence question **"real"**, the analysis revealed a significant effect of condition, F(1,89) = 5.11, $p < 0.01$, p. $\eta^2 = 0.11$. Overall, participants in the EVR condition rated their environment to be significantly more real (M=6.85, SD=2.11) as compared to those in the NVR condition (M=4.70, SD=2.69), $p < 0.01$. Also, participants in the IVR condition (M=6.32, SD=2.73) rated this feeling to be significantly higher as compared to those in the NVR condition (M=4.70, SD=2.69), p = 0.04.

Regarding telepresence question **"inside"**, the analysis revealed a significant effect of condition, F(1,89) = 7.75, $p < 0.001$, p. $\eta^2 = 0.15$. Overall, participants in the EVR condition (M=7.25, SD=2.45) rated their feeling of being inside the environment to be significantly higher as compared to those in the NVR condition (M=4.75, SD=3.15), $p < 0.001$. Also, participants in the IVR condition (M=6.92, SD=2.33) rated this feeling to be significantly higher as compared to those in the NVR condition (M=4.75, SD=3.15), p = 0.01.

Regarding telepresence question **"surrounded"**, the analysis revealed a significant effect of condition, F(1,89) = 7.00, $p < 0.01$, p. $\eta^2 = 0.14$. Overall, participants in the EVR condition (M=7.29, SD=2.38) rated their feeling of being immersed and surrounded by the environment significantly higher as compared to those in the NVR condition (M=5.07, SD=2.80), $p < 0.01$. Also, participants in the IVR condition (M=6.96, SD=2.20) rated this feeling to be significantly higher as compared to those in the NVR condition (M=5.07, SD=2.80), p = 0.02.

Regarding social-presence question **"space"**, the analysis revealed a significant effect of condition, F(1,89) = 8.97, p <0.001, p. η^2 = 0.17. Overall, participants in the EVR condition (M=6.89, SD=2.60) rated their feeling of being in the same space as their character significantly higher as compared to those in the NVR condition (M=4.09, SD=3.09), p <0.001. Also, participants in the IVR condition (M=6.50, SD=2.58) rated this feeling to be significantly higher as compared to those in the NVR condition (M=4.09, SD=3.09), p <0.01.

Regarding social-presence question **"engaging"**, the analysis revealed a significant effect of condition, F(1,89) = 4.60, p = 0.04, p. η^2 = 0.07. Overall, participants in the EVR condition (M=7.82, SD=1.95) rated their experience to be significantly more engaging than participants in the IVR condition (M=6.65, SD=2.50), p = 0.04. No significant difference was found between EVR and NVR conditions, likely due to the high amount of variance in the NVR scores (Figure 5). This may indicate that the participants in the NVR condition interpreted engagement in different ways, such as engagement with the character, engagement with the performance, or engagement with the technology.

5.3 Quantitative results: attitudes towards programming

The participants answered a pre- and post-experience questionnaire, which consisted of a mix of nominal variables (yes or no type responses) as well as Likert scale responses ranging from strongly disagree to strongly agree. For the nominal-type responses, we subjected the data to a non-parametric analysis. In analyzing the Likert scale responses, we found that the data was not normally distributed and the assumption of homogeneity of variance was violated. Therefore, these responses were also analyzed using a non-parametric statistical analysis, in accordance with HCI and statistics best practices. A related-samples Wilcoxon signed-rank test was conducted on responses for questions asked to understand student opinions regarding programming and dance in a pre and post fashion. Student responses were analyzed across each of the three conditions (NVR vs. IVR vs. EVR) to examine the effect of viewing metaphor on student opinions.

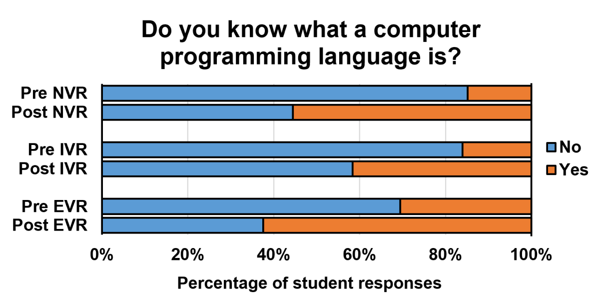


Fig. 6. Student responses to the question "Do you know what a computer programming language is?"

For the question *"Do you know what a computer programming language is?"* (Figure 6), the Wilcoxon signed-rank test indicated that overall student reports on knowledge of computer programming languages were significantly higher in the post-experiment responses as compared to pre-experiment, Z = -4.802, p <0.001. Across the conditions, student responses were significantly higher post-experiment as compared to pre-experiment for the IVR condition (Z = -2.646, p <0.01), EVR condition (Z = -2.84, p <0.01), as well as NVR condition (Z = -2.887, p <0.01).

For the question *"Do you see yourself as a computer programmer?"*, the analysis indicated that overall student reports on seeing themselves as computer programmers were significantly higher in the post-experiment responses as compared to pre-experiment, Z = -2.558, p = 0.01.

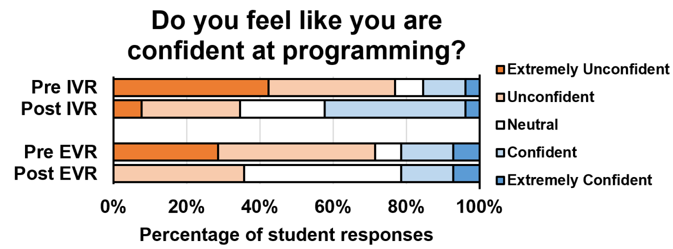


Fig. 7. Student responses to the question "Do you feel like you are confident at programming?"

For the question *"Do you feel like you are confident at programming?"* (Figure 7), the analysis indicated that overall student reports on programming confidence were significantly higher in the post-experiment responses as compared to pre-experiment, Z = -3.24, p <0.001. Across the conditions, student responses were significantly higher post-experiment as compared to pre-experiment for the IVR condition (Z = -2.747, p <0.01) and the EVR condition (Z = -1.966, p = 0.05). There were no significant differences found between pre and post in the NVR condition.

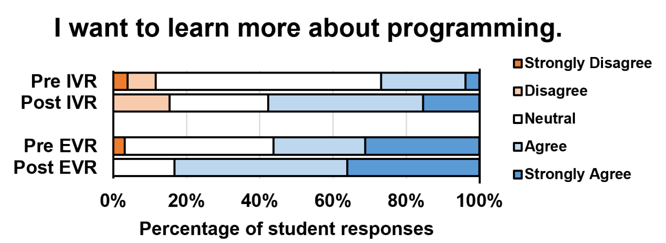


Fig. 8. Student responses to the question "I want to learn more about programming."

For the statement *"I want to learn more about programming"* (Figure 8), the analysis indicated that overall student reports on wanting to learn more about programming were significantly higher in the post-experiment responses as compared to pre-experiment, Z = -3.076, p <0.01. Across the conditions, student responses were significantly higher post-experiment as compared to pre-experiment for the IVR condition (Z = -2.295, p = 0.02) and the EVR condition (Z = -2.137, p = 0.04). There were no significant differences found between pre and post in the NVR condition.

5.4 Quantitative results: programming performance

The programming scores obtained from the final programmed performances of the students were analyzed against the factor of condition. Levene's test for homogeneity was significant, which meant that the data failed the assumption of homogeneity of variances. Therefore, a Welch one-way ANOVA was used for the analysis. Pairwise post-hoc tests for the three levels of condition were conducted using the Games-Howell method, which does not assume equality of variances. The descriptive statistics of these results are shown in Figure 9.

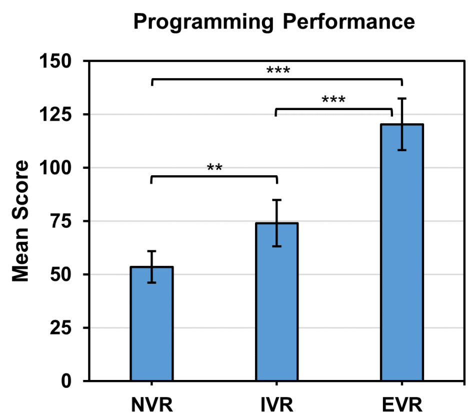


Fig. 9. Mean programming scores of the students for each of the NVR, IVR, and EVR conditions. *** indicates a significant statistical difference with $p < 0.001$; ** indicates a significant statistical difference with $p < 0.01$. Error bars represent 95% confidence interval.

The Welch's ANOVA analysis revealed a significant effect of condition, F(2, 55.43) = 45.4, $p < 0.001$, p. $\eta^2 = 0.51$. Overall, participants in the EVR condition (M=120.33, SD=35.75) scored significantly higher as compared to those in the IVR condition (M=74.02, SD=26.95), $p < 0.001$ and NVR condition (M=53.5, SD=19.21), $p < 0.001$. Further, participants in the IVR condition (M=74.02, SD=26.95) scored significantly higher than the participants in the NVR condition (M=53.5, SD=19.21), $p < 0.01$. Thus, students' programming performance significantly increased between each condition, from NVR (lowest) to IVR to EVR (highest).

5.5 Qualitative Results

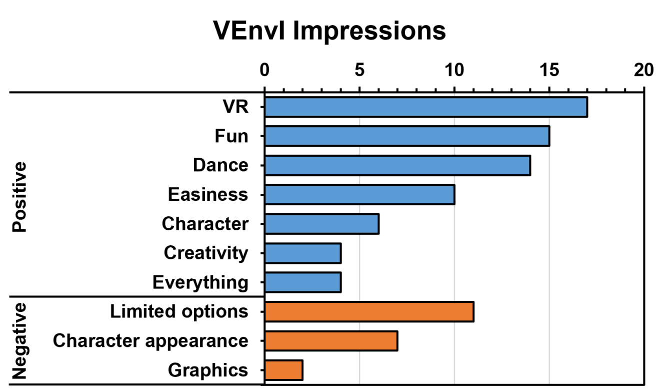


Fig. 10. Number of positive and negative impressions about VEnvI from coded qualitative responses.

The students were asked open-ended questions to gain further insight on their experiences within VEnvI. We conducted a thematic analysis on the qualitative debriefing data, guided by our research questions. In our analysis, we used elements of qualitative analysis such as in-vivo, process, and open coding [58]. Using NVivo 12 software, two researchers inductively coded the data separately, labeling emergent phenomena in the data to arrive at a codebook. We met regularly during the analysis process to discuss discrepancies in the applications of the codes, reexamine the codebook, and reflect on contradictory data. The analysis revealed themes in two overarching categories: positive and negative. The positive themes included "engaging VR experience", "learning programming with dance", and "ability to be creative". The negative themes included "lack of customization", "unpleasant visuals", and "technology limitations". The following sections dive deeper into these themes.

When talking about overall impressions of using VEnvI (see Figure 10), students shared more positive responses (77%) than negative (23%). Of the positive responses, 24% of the students attributed their positive experience to virtual reality (VR, headset, Oculus, virtual camera), 21% liked VEnvI because it was fun (it was fun/cool, enjoyed, liked it), 20% loved the dance aspect (dancing, moves, combine programming with dance), 14% liked the easiness (easy to use/learn, simple), 9% enjoyed it because of the character they programmed (character looked real/was cool, you get your own character), 6% liked the ability to create (creative, experiment, do anything), and 6% mentioned they loved everything about VEnvI. Of the negative responses, 55% wanted more choices (more body/hair/eye/clothing choices, more faces, more moves), 35% did not like the look of their character (scary, not real, not looking like how I want her to), and 10% wanted better graphics.

Overall, the students enjoyed the VEnvI experience and responded positively with quotes such as *"VEnvI was really fun and interesting"*, *"It was very cool and I learned a lot"*, and *"I probably wouldn't have even tried or signed up for this but it was worth it."*

Students rated the VR experience highly positive overall. *"VR stuff and the program made us feel like it was amazing"* said one student, *"looking into the virtual world was unbelievable"* said another. Students really liked the experience of using the Oculus Rift, stating *"I felt involved when we put on the virtual goggles"* and *"I wish I could have did the goggles more."* Students found the VR experience realistic, and mentioned *"VR is so similar to real life"* and *"I would use VENVI (to) escape from reality."*

Students reported high sense of presence with quotes such as *"I felt like I was actually there"*, and *"it felt like I was there with my character."* Students commented on the realism of the virtual environment with statements such as *"it was very vivid"* and *"it was like I teleported."*

Students also voiced concerns with the experience and suggested various improvements. One student stated, *"If you kept walking forward there is a limit of how far you can go, whereas in real life you can go as far as you want."* The cabling of the HMD will limit or obstruct the movement of the users unless a wireless solution is implemented. Many students wanted the ability to customize the environment, stating *"I wish you could change the scene"* and the need for *"more backgrounds like a city or a farm."* Students also desired interactivity with the virtual character, wondering *"if the character was actually able to have a conversation with."* Finally, some students wanted the character to look like themselves, with the desire for *"creating one to resemble me"* and to *"make a character and make them me."*

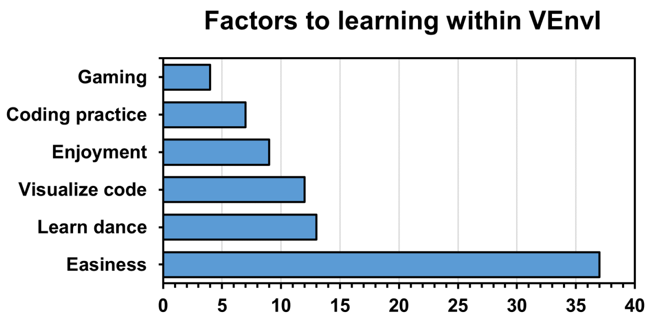


Fig. 11. Number of coded responses on factors influencing learning within VEnvI.

When talking about factors that affected their learning from using VEnvI (see Figure 11), 45% said VEnvI made it easy to learn and understand coding (easy to use, simple to program/code, simple interface, self-explanatory), 16% liked that they could learn to dance using VEnvI (learn to dance better, learn new moves, helps me dance), 15% mentioned they learned better because they could visualize the code (code in blocks, see my commands, visual learner, watching the character perform the code, can see what's going wrong), 11% learned better because they had fun (fun, cool, enjoyed), 8% liked that they could get coding practice (practice, can perfect coding, can make character go slow), and 5% loved to learn while playing a game (play/make games, good at video games, good game controls).

Students commented on the learning benefits of VEnvI by saying *"this was very educational"*, *"I like that it combines programming and dancing"*, *"it helped me learn computer programming"*, and *"I felt like I learned a little more about code."* Students found VEnvI user-friendly, stating *"I think it was pretty easy to use VEnvI"* and *"the words and things we learned were easy to understand."* Students believed VEnvI to be useful, one student stated that VEnvI *"will help with my future of being an engineer."*

6 Discussion

The primary research question of this study (RQ 1) was to determine the extent to which immersive viewing and virtual self-representation via a body-scaled co-located self-avatar within an educational virtual environment facilitated visual analysis and learning of computational thinking concepts and principles. Overall, our intervention did work in helping to learn abstract computational concepts across all conditions, where students scored significantly higher in the cognitive test post-intervention as compared to pre-intervention. H1 proposed that there would be an increasing effect on cognitive performance between the three conditions, from NVR to IVR to EVR. Our results revealed that immersive VR with self-avatars (EVR) within VEnvI achieved higher success in enhancing computational learning among the students as compared to the immersive condition without self-avatars (IVR), as well as the desktop-only condition (NVR) for certain levels of the cognitive taxonomies. This was denoted by the higher cognitive scores in the EVR condition under the remembering and understanding category of Bloom's taxonomy, and the multistructural domain of the SOLO taxonomy, as compared to both the IVR and the NVR conditions. Further, there were differences in the pre-test scores between students in the EVR and IVR conditions for the multistructural category of the SOLO taxonomy, thus partially reducing the impact of EVR over IVR in the multistructural domain. Therefore, H1 was partially supported by our data.

According to Bloom's taxonomy [47], the remembering and understanding category entails that the learner can retrieve and recall knowledge from long-term memory, and demonstrate comprehension through one or more forms of explanation. Further, according to the SOLO taxonomy [48], the multistuctural category entails that the learner can acquire and understand multiple cognitive concepts, though is not yet at the level of forming a high-level, coherent picture. Research has shown that embodiment within virtual reality improves the functioning of episodic and spatial memory and enhances recall [59], [60], [61]. Virtual self-avatars within VEnvI helped the students in understanding several independent concepts better, and enabled recalling and explaining them more successfully, as determined by the SOLO and Bloom's categories. This shows that for levels of cognition involving recall of several independent programming concepts and providing explanations, utilizing a co-located virtual self-avatar facilitates embodied cognition and leads to improved pedagogical benefits. Irrespective of the conditions, all students scored significantly higher overall post-experiment as compared to pre-experiment with respect to cognition, strongly supporting the VEnvI outreach program's efficacy overall in imparting computational thinking in middle school students.

This study also intended to objectively measure programming performance as part of the primary research question. The students' programmed performances were evaluated based on the blocks used by the student, the complexity of the program, and the program duration. A gradual improvement in programming performance was observed between the three conditions, where students in the IVR condition scored significantly higher than students in the NVR condition, and students in the EVR condition scored significantly higher than both IVR and NVR conditions. Therefore, H2 was fully supported. Overall, VR was able to enhance the programming performance among students. As revealed by recent research, virtual self-avatars in immersive experiences have the potential to significantly improve spatial perception and affordances in VR [18], [20], [24], [62]. Perhaps watching their characters perform the programmed choreographies through an egocentric first-person perspective, while being spatially grounded via their self-avatars, allowed the students to critically analyze their programs and think more creatively, resulting in increased variety, complexity, and duration of their programming performance in VEnvI.

A secondary research question of this study (RQ 2) involved studying the effects of immersion and virtual self-avatars on presence, and attitudes towards STEM and computer science. Both the EVR and the IVR metaphors within VEnvI were successful in immersing the students in the learning environment, supported by the high telepresence and social presence scores as compared to the NVR condition, partially supporting H3. Research shows that self-avatars enhance users' sense of presence by positively contributing to place and plausibility illusion, creating a sensation of being in a real place and experiencing the situation as actually occurring [21], [22], [23]. In this study, students experienced a sense of "being there" in the VEnvI environment and were able to socially connect with their virtual characters, for whom they authored programmed choreography in both the EVR and IVR metaphors. Students reported immersive VR to be more engaging, which, as results show, motivated them to actively learn the underlying programming concepts better.

Overall, students expressed higher confidence and desire to learn about computer programming after interacting with VEnvI. VR proved to be successful in grabbing the attention of middle school students. Both the EVR and the IVR conditions were significantly more successful in positively altering students' attitudes towards computer programming as compared to the NVR condition, as shown by the results. A desire to choose computing as a major in college was not shown to be significant different across any of the three conditions. This can be attributed to the short duration of the study, as each group of students interacted with VEnvI for only a six-week experiment period. A long-term intervention and follow-up interactions with the students are required in assessing the effects of VEnvI and the immersive embodied metaphor on altering students' desires to choose computer science as a career.

6.1 Limitations

Conducting research with voluntary participation of middle school children in a realistic classroom setting is extremely challenging. Complexities involve obtaining support from school administration and teachers for possible disruptions in planned curricular activities, and enrollment of assenting children with parents and guardians consenting to research participation. In this research, to attain a sample size of 90 participants, the study had to be conducted over multiple iterations. This possibly led to differences in the participant groups.

The quantitative analysis revealed one pre-test difference in cognitive test scores between the EVR and IVR condition for the multistructural level of the SOLO taxonomy. The participants in the EVR condition scored higher than participants in the NVR condition. In the post-test session, participants in EVR scored significantly higher than participants in both IVR and NVR. However, the pre-test difference between EVR and IVR may have reduced the validity of the finding that participants in EVR scored significantly higher than participants in IVR condition. However, the result that participants in EVR scored significantly higher than participants in NVR in SOLO's multistructural dimension is still valid.

Strong differences between conditions were only found in two of the six different categories of Bloom's and SOLO cognitive taxonomies. This is partly due to the short duration of the intervention, but also points to the design of the curriculum and the study. Future interventions will have to incorporate the learning of more complex and higher-order programming concepts, providing ample time for students to grasp and practice these concepts.

Within the study, the students programmed their choreographies, viewed them in VR, and then went back to make modifications to their program. One limitation of this study is that the between-subjects evaluation was only performed for the viewing phase. Self-avatars and immersion were not studied holistically, where all three phases of programming within VEnvI—program authoring, viewing, and re-authoring—would exist within the same perspective. A future study would explore the cyclical process of learning, authoring, viewing, and re-authoring, all within the same VR metaphor.

At the time we conducted our study with middle school students, we did not consider using an explicit embodiment questionnaire, which has recently been derived from the extensive research conducted on sense of embodiment. We wanted to ensure that we did not burden the middle school students with a multitude of questionnaires, as they could easily lose interest in the study. We therefore elected to use a minimum number of questionnaires. As with all our questionnaires, we always had to adapt the standardized questionnaires such as presence, etc., such that middle school students can comprehend the questions, and this was often challenging. We also did not measure body movement through tracking sensor readings, since the tracking method used in this study was approximate, with the primary purpose of engaging the students in the learning experience. Since running the study, new and affordable VR headsets, tracking systems, and cheap and easily wearable motion tracking technology have been developed that could be used to provide near-accurate tracking. Such novel tracking systems can measure the physical movement of each participant, which future iterations of this study can utilize.

Given the ambitious nature of this study, in this research thrust we did not focus on how self-avatar customization or personalization affected students' learning and impressions in VR learning activities. However, our current research establishes a baseline of how a body-scaled, co-located, gender-matched self-avatar grounds participants in VR educational activities and enables participants to enact actions in VR with content they programmatically create, thus facilitating learning of computational thinking concepts in this process. In future research, we hope to compare and contrast the participants' ability to customize or personalize their self-avatar as compared to a standardized self-avatar on learning, presence, and VR impressions.

7 Conclusions and Future Work

Minority populations, especially women in Science, Technology, Engineering, and Math (STEM) fields continue to remain underrepresented, and novel interventions such as VEnvI have immense potential to aid in alleviating this problem. VEnvI proved to be successful in developing computational thinking in middle school students and enhancing their interest and engagement in the field of computer science.

Utilizing co-located virtual self-avatars within VR enabled VEnvI to facilitate embodied cognition for learning abstract computer programming concepts, and improving programming performance among the students. VR has immense potential to captivate the minds of students, enhance their learning experience, and motivate a willingness to learn. VEnvI was able to achieve this by making the learning of computer programming concepts fun using immersive VR and merging it with dance. Students advocated for using the immersive self-avatar-based VR metaphor to visualize their programmed choreographies and stated, "I liked how it's easy to learn coding", and "definitely loved the virtual reality experience". Immersive VR within VEnvI was able to engage students and boost their interest in computer science education, as corroborated by the significantly higher cognitive scores as well as the anecdotal student responses.

Not much work has been done in studying middle school students' (especially 6th and 7th grade students) attitudes towards VR, namely through studying their sense of presence, the usability of the application, and enthusiasm towards VR in general, and especially in the field of education. One of the broader aims of this research is to advance the knowledge by examining how VR systems can be integrated into a technology-based STEM education curriculum to enhance the pedagogical outcomes, and test children's acceptance and use of such technology towards learning. This research provides evidence of a heightened sense of presence via immersive self-avatar-based VR among middle school students, and provides the VR community an insight into the minds of middle school students regarding their reactions to immersive VR for education.

This work also has a significant impact on the education research community. This research is one of the first that links active body movements, immersive viewing, and self-representation via self-avatars in VR to cognition. Theoretical research has argued for many decades that embodiment can affect cognition, but there have not been any empirical studies that provide evidence in the favor of such theories, especially in VR-based learning. Our work also suggests that educators and developers of educational VR simulations, who want to enhance knowledge and understanding or simultaneous acquisition of multiple abstract concepts, can do so by employing immersion and self-avatars in VR learning experiences. Thus, this work is pivotal in terms of providing evidence of improved learning through an immersive embodied system mapping the key affordances of dance to programming abstract concepts of computing (such as sequences, loops, parallel programming, etc.).

7.1 Future work

The limitations of this research discussed in subsection 6.1 present opportunities for future research. For instance, a study spanning the entire school year instead of the short duration of six weeks, with a more careful approach of enrolling a representative and homogenous participant pool, could lead to improved power of the results showing a stronger effect on student interest in the field of computing. Since the current study only performed a between-subjects evaluation of the viewing phase in each VR metaphor, a future study would involve the entire process of learning, program authoring, viewing, and re-authoring within the same perspective in each VR metaphor. Further, this research only examines a form of embodiment within immersive VR through co-located virtual self-avatars. Other ways of enabling active embodiment within VEnvI could involve authoring new movements for the virtual characters using body tracking, socially responsive virtual peers, virtual doppelgängers created using a 3D scan of the user as a virtual representation of themselves for creating programming movements, and a body-based programming metaphor by implementing an immersive, embodied, and interactive VR within VEnvI. These forms of embodiment also have considerable prospects for future research in facilitating embodied cognition. Finally, future work will also consider examining the body motions of users in EVR, IVR, and NVR conditions via unobtrusive motion tracking systems during their viewing and engagement experiences with their programmed choreographies in VR.

8 Acknowledgement

The authors wish to greatly thank the student participants and teachers who joined the empirical evaluations. This material is based upon work supported by the U.S. National Science Foundation (NSF INSPIRE Award) under Grant CNS 1344228.

References

[1] N. E. Seymour, A. G. Gallagher, S. A. Roman, M. K. O'Brien, V. K. Bansal, D. K. Andersen, and R. M. Satava, "Virtual reality training improves operating room performance: Results of a randomized, double-blinded study," *Annals of Surgery*, vol. 236, no. 4, pp. 458–464, 2002.

[2] H. G. Hoffman, "Virtual-reality therapy," *Scientific American*, vol. 291, no. 2, pp. 58–65, 2004. [Online]. Available: http://www.jstor.org/stable/26060647

[3] J. C. Chan, H. Leung, J. K. Tang, and T. Komura, "A virtual reality dance training system using motion capture technology," *IEEE Transactions on Learning Technologies*, vol. 4, no. 2, pp. 187–195, 2011.

[4] K. F. Hew and W. S. Cheung, "Use of three-dimensional (3-D) immersive virtual worlds in K-12 and higher education settings: A review of the research," *British Journal of Educational Technology*, vol. 41, no. 1, pp. 33–55, 2010.

[5] 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, pp. 29–40, 2014.

[6] T. J. Misa, *Gender codes: Why women are leaving computing*. John Wiley & Sons, 2011.

[7] National Science Board, "Science and engineering indicators," National Science Foundation, Arlington, Virginia, Tech. Rep. NSB-2018-1, 2018.

[8] K. Modi, J. Schoenberg, and K. Salmond, "Generation stem: What girls say about science, technology, engineering, and math," *A Report from the Girl Scout Research Institute. New York, NY: Girl Scouts of the USA*, 2012.

[9] 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 *2016 IEEE Virtual Reality (VR)*, 2016, pp. 131–140.

[10] S. Cooper, W. Dann, and R. Pausch, "Alice: a 3-D tool for introductory programming concepts," in *Journal of Computing Sciences in Colleges*, vol. 15, no. 5. Consortium for Computing Sciences in Colleges, 2000, pp. 107–116.

[11] J. Maloney, M. Resnick, N. Rusk, B. Silverman, and E. Eastmond, "The scratch programming language and environment," *ACM Transactions on Computing Education (TOCE)*, vol. 10, no. 4, p. 16, 2010.

[12] L. Shapiro, *Embodied cognition*. Routledge, 2019.

[13] K. Kilteni, R. Groten, and M. Slater, "The sense of embodiment in virtual reality," *Presence: Teleoperators and Virtual Environments*, vol. 21, no. 4, pp. 373–387, 2012.

[14] J. Llobera, M. V. Sanchez-Vives, and M. Slater, "The relationship between virtual body ownership and temperature sensitivity," *Journal of The Royal Society Interface*, vol. 10, no. 85, p. 20130300, 2013.

[15] D. Banakou, R. Groten, and M. Slater, "Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes," *Proceedings of the National Academy of Sciences*, vol. 110, no. 31, pp. 12 846–12 851, 2013.

[16] K. Kilteni, I. Bergstrom, and M. Slater, "Drumming in immersive virtual reality: the body shapes the way we play," *IEEE Transactions on Visualization & Computer Graphics*, no. 4, pp. 597–605, 2013.

[17] T. C. Peck, S. Seinfeld, S. M. Aglioti, and M. Slater, "Putting yourself in the skin of a black avatar reduces implicit racial bias," *Consciousness and cognition*, vol. 22, no. 3, pp. 779–787, 2013.

[18] B. J. Mohler, S. H. Creem-Regehr, W. B. Thompson, and H. H. Bülthoff, "The effect of viewing a self-avatar on distance judgments in an hmd-based virtual environment," *Presence: Teleoperators and Virtual Environments*, vol. 19, no. 3, pp. 230–242, 2010.

[19] E. A. McManus, B. Bodenheimer, S. Streuber, S. De La Rosa, H. H. Bülthoff, and B. J. Mohler, "The influence of avatar (self and character) animations on distance estimation, object interaction and locomotion in immersive virtual environments," in *Proceedings of the ACM SIGGRAPH Symposium on applied perception in graphics and visualization*. ACM, 2011, pp. 37–44.

[20] E. Ebrahimi, L. S. Hartman, A. Robb, C. C. Pagano, and S. V. Babu, "Investigating the effects of anthropomorphic fidelity of self-avatars on near field depth perception in immersive virtual environments," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2018, pp. 1–8.

[21] A. Steed, S. Frlston, M. M. Lopez, J. Drummond, Y. Pan, and D. Swapp, "An 'in the wild'experiment on presence and embodiment using consumer virtual reality equipment," *IEEE transactions on visualization and computer graphics*, vol. 22, no. 4, pp. 1406–1414, 2016.

[22] R. Schwartz and W. Steptoe, "The immersive vr self: performance, embodiment and presence in immersive virtual reality environments," in *A networked self and human augmentics, artificial intelligence, sentience*. Routledge, 2018, pp. 108–116.

[23] U. Schultze, "Embodiment and presence in virtual worlds: a review," *Journal of Information Technology*, vol. 25, no. 4, pp. 434–449, 2010.

[24] A. Bhargava, H. Solini, K. Lucaites, J. W. Bertrand, A. Robb, C. C. Pagano, and S. V. Babu, "Comparative evaluation of viewing and self-representation on passability affordances to a realistic sliding doorway in real and immersive virtual environments," in *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 2020, pp. 519–528.

[25] B. Day, E. Ebrahimi, L. S. Hartman, C. C. Pagano, A. C. Robb, and S. V. Babu, "Examining the effects of altered avatars on perception-action in virtual reality." *Journal of Experimental Psychology: Applied*, vol. 25, no. 1, p. 1, 2019.

[26] V. R. Lee, *Learning technologies and the body: Integration and implementation in formal and informal learning environments*. Routledge, 2014.

[27] L. W. Barsalou, "Grounded cognition," *Annual Review of Psychology*, vol. 59, pp. 617–645, 2008.

[28] S. Gallagher, *How the body shapes the mind*. Cambridge University Press, 2005.

[29] M. Wilson, "Six views of embodied cognition," *Psychonomic Bulletin & Review*, vol. 9, no. 4, pp. 625–636, 2002.

[30] S. Papert, *Mindstorms: Children, computers, and powerful ideas*. Basic Books, Inc., 1980.

[31] S. B. Daily, A. E. Leonard, S. Jörg, S. Babu, and K. Gundersen, "Dancing Alice: Exploring embodied pedagogical strategies for learning computational thinking," in *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*. ACM, 2014, pp. 91–96.

[32] A. E. Leonard and S. B. Daily, "The dancing alice project: Computational and embodied arts research in middle school education," in *Proceedings of the 45th ACM Technical Symposium on Computer Science Education*, 2014.

[33] A. E. Leonard, S. B. Daily, S. Jörg, and S. V. Babu, "Coding moves: Design and research of teaching computational thinking through dance choreography and virtual interactions," *Journal of Research on Technology in Education*, vol. 0, no. 0, pp. 1–19, 2020. [Online]. Available: https://doi.org/10.1080/15391523.2020.1760754

[34] L. Lin, D. Parmar, S. V. Babu, A. E. Leonard, S. B. Daily, and S. Jörg, "How character customization affects learning in computational thinking," in *Proceedings of the ACM Symposium on Applied Perception*, ser. SAP '17. New York, NY, USA: Association for Computing Machinery, 2017. [Online]. Available: https://doi.org/10.1145/3119881.3119884

[35] C. M. Armstrong, G. M. Reger, J. Edwards, A. A. Rizzo, C. G. Courtney, and T. D. Parsons, "Validity of the virtual reality stroop task (VRST) in active duty military," *Journal of Clinical and Experimental Neuropsychology*, vol. 35, no. 2, pp. 113–123, 2013.

[36] J. Bertrand, D. Brickler, S. Babu, K. Madathil, M. Zelaya, T. Wang, J. Wagner, A. Gramopadhye, and J. Luo, "The role of dimensional symmetry on bimanual psychomotor skills education in immersive virtual environments," in *2015 IEEE Virtual Reality (VR)*. IEEE, 2015, pp. 3–10.

[37] K. Johnsen, A. Raij, A. Stevens, D. S. Lind, and B. Lok, "The validity of a virtual human experience for interpersonal skills education," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2007, pp. 1049–1058.

[38] A. B. Raij, K. Johnsen, R. F. Dickerson, B. C. Lok, M. S. Cohen, M. Duerson, R. R. Pauly, A. O. Stevens, P. Wagner, and D. S. Lind, "Comparing interpersonal interactions with a virtual human to those with a real human," *IEEE transactions on visualization and computer graphics*, vol. 13, no. 3, pp. 443–457, 2007.

[39] M. C. Johnson-Glenberg, D. Birchfield, P. Savvides, and C. Megowan-Romanowicz, "Semi-virtual embodied learning-real world STEM assessment," in *Serious Educational Game Assessment*. Springer, 2011, pp. 241–257.

[40] M. C. Johnson-Glenberg, H. Bartolomea, and E. Kalina, "Platform is not destiny: Embodied learning effects comparing 2d desktop to 3d virtual reality stem experiences," *Journal of Computer Assisted Learning*, 2021.

[41] B. Dalgarno and M. J. Lee, "What are the learning affordances of 3-D virtual environments?" *British Journal of Educational Technology*, vol. 41, no. 1, pp. 10–32, 2010.

[42] R. Coulter, L. Saland, T. P. Caudell, T. E. Goldsmith, and D. C. Alverson, "The effect of degree of immersion upon learning performance in virtual reality simulations for medical education," *Medicine Meets Virtual Reality 15: in Vivo, in Vitro, in Silico: Designing the Next in Medicine*, vol. 125, p. 155, 2007.

[43] 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.

[44] D. Parmar, J. Bertrand, S. V. Babu, K. Madathil, M. Zelaya, T. Wang, J. Wagner, A. K. Gramopadhye, and K. Frady, "A comparative evaluation of viewing metaphors on psychophysical skills education in an interactive virtual environment," *Virtual Reality*, vol. 20, no. 3, pp. 141–157, 2016.

[45] C. Juan, F. Beatrice, and J. Cano, "An augmented reality system for learning the interior of the human body," in *IEEE International Conference on Advanced Learning Technologies*. IEEE, 2008, pp. 186–188.

[46] J. Madden, S. Pandita, J. Schuldt, B. Kim, A. S. Won, and N. Holmes, "Ready student one: Exploring the predictors of student learning in virtual reality," *PloS one*, vol. 15, no. 3, p. e0229788, 2020.

[47] L. W. Anderson, D. R. Krathwohl, and B. S. Bloom, *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Allyn & Bacon, 2001.

[48] J. B. Biggs and K. F. Collis, *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome)*. Academic Press, 2014.

[49] K. L. Nowak and F. Biocca, "The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments," *Presence*, vol. 12, no. 5, pp. 481–494, 2003.

[50] N. Dodgson, "Variation and extrema of human interpupillary distance," in *Stereoscopic Displays and Virtual Reality Syst XI*, vol. 5291. International Society for Optics and Photonics, 01 2004, pp. 36–46.

[51] E. A. McManus, B. Bodenheimer, S. Streuber, S. de la Rosa, H. H. Bülthoff, and B. J. Mohler, "The influence of avatar (self and character) animations on distance estimation, object

interaction and locomotion in immersive virtual environments," in *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*, ser. APGV '11. New York, NY, USA: Association for Computing Machinery, 2011, p. 37–44. [Online]. Available: https://doi.org/10.1145/2077451.2077458

[52] E. Ebrahimi, A. Robb, L. S. Hartman, C. C. Pagano, and S. V. Babu, "Effects of anthropomorphic fidelity of self-avatars on reach boundary estimation in immersive virtual environments," in *Proceedings of the 15th ACM Symposium on Applied Perception*, ser. SAP '18. New York, NY, USA: Association for Computing Machinery, 2018. [Online]. Available: https://doi.org/10.1145/3225153.3225170

[53] A. Bhargava, R. Venkatakrishnan, R. Venkatakrishnan, H. Solini, K. M. Lucaites, A. Robb, C. Pagano, and S. Babu, "Did i hit the door effects of self-avatars and calibration in a person-plus-virtual-object system on perceived frontal passability in vr," *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[54] C. Flavián, S. Ibáñez-Sánchez, and C. Orús, "Impacts of technological embodiment through virtual reality on potential guests' emotions and engagement," *Journal of Hospitality Marketing & Management*, vol. 30, no. 1, pp. 1–20, 2021. [Online]. Available: https://doi.org/10.1080/19368623.2020.1770146

[55] J. Isaac and S. V. Babu, "Supporting computational thinking through gamification," in *3D User Interfaces (3DUI), 2016 IEEE Symposium on*. IEEE, 2016, pp. 245–246.

[56] C. A. Zanbaka, B. C. Lok, S. V. Babu, A. C. Ulinski, and L. F. Hodges, "Comparison of path visualizations and cognitive measures relative to travel technique in a virtual environment," *IEEE Transactions on Visualization and Computer Graphics*, vol. 11, no. 6, pp. 694–705, 2005.

[57] E. Suma, S. Finkelstein, M. Reid, S. Babu, A. Ulinski, and L. F. Hodges, "Evaluation of the cognitive effects of travel technique in complex real and virtual environments," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, no. 4, pp. 690–702, 2009.

[58] A. Strauss and J. Corbin, *Basics of qualitative research*. Sage publications, 1990.

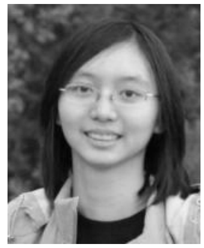
[59] C. Tuena, S. Serino, A. Gaston-Bellegarde, E. Orriols, D. Makowski, G. Riva, and P. Piolino, "How virtual embodiment affects episodic memory functioning: a proof-of-concept study," *Annu. Rev. Cyberther. Telemed*, vol. 15, pp. 98–103, 2017.

[60] C. Tuena, S. Serino, L. Dutriaux, G. Riva, and P. Piolino, "Virtual enactment effect on memory in young and aged populations: A systematic review," *Journal of clinical medicine*, vol. 8, no. 5, p. 620, 2019.

[61] C. Tuena, S. Serino, E. Pedroli, M. Stramba-Badiale, G. Riva, and C. Repetto, "Building embodied spaces for spatial memory neurorehabilitation with virtual reality in normal and pathological aging," *Brain Sciences*, vol. 11, no. 8, p. 1067, 2021.

[62] J. Zhao, T. Sensibaugh, B. Bodenheimer, T. P. McNamara, A. Nazareth, N. Newcombe, M. Minear, and A. Klippel, "Desktop versus immersive virtual environments: effects on spatial learning," *Spatial Cognition & Computation*, vol. 20, no. 4, pp. 328–363, 2020.

Dhaval Parmar is a Postdoctoral Research Associate in the Khoury College of Computer Sciences at Northeastern University. His research is focused on human centered computing, with interests in important topics such as STEM education and broadening participation among under-represented minorities. His current research involves applications in heath education and health behavior training, using relational virtual agents within virtual reality.



Lorraine Lin is a Ph.D. graduate from the School of Computing at Clemson University. Her research interests include studying avatar appearances and their effects on body-ownership and learning within virtual reality.

Nikeetha DSouza is a Postdoctoral Fellow in the Office of the Vice Provost for Diversity and Inclusion at Indiana University. Her research interests include the role of identity, culture, language, social relationships, justice, and equity in science, technology, engineering, and mathematics (STEM) education.

Sophie Jörg is an Associate Professor in the School of Computing at Clemson University. Her primary research interests are in computer graphics, especially in animation and perception of virtual characters. She is furthermore interested in virtual reality, machine learning, and human-computer interaction.

Alison E. Leonard is the Associate Professor of Arts & Creativity in the College of Education at Clemson University. Her research focuses on exploring how the arts play a role in education and teacher education, primarily dance's role and its possibilities in schools.

Shaundra B Daily is a professor of practice in Electrical and Computer Engg. & Computer Science at Duke University. Her research involves design, implementation, and evaluation of technologies, programs, and curricula to promote justice, equity, diversity, and inclusion in STEM. She is currently Co-PI of the Alliance for Identity-Inclusive Computing, Education and Workforce Director for the Athena AI Institute, Faculty Director of the Duke Technology Scholars Program, and co-leads the 3C Fellows program.

Sabarish V. Babu is an Associate Professor in the School of Computing at Clemson University. His research interests are in the design, development, and evaluation of virtual reality systems, specifically in the creation of interactive scenarios and user studies in immersive virtual environments for education and training. He is also a renowned expert in near-field perception and motor control in VR, virtual humans and 3D human-computer interaction.