How Character Customization Affects Learning in Computational Thinking

Lorraine Lin
Clemson University
lorrain@clemson.edu

Dhaval Parmar
Clemson University
dkparma@g.clemson.edu

Sabarish V. Babu
Clemson University
sbabu@clemson.edu

Alison E. Leonard
Clemson University
aleona2@clemson.edu

Shaundra B. Daily
University of Florida
shanib@ufl.edu

Sophie Jörg
Clemson University
sjoerg@clemson.edu

ABSTRACT

The ability to select or customize characters in educational applications and games has been shown to influence factors related to learning effects such as transfer, self-efficacy, and motivation. Most previous conclusions on the perception of virtual characters and the effect of character assignment in interactive applications have been reached through short, one-task experiments. To investigate more long-term effects of assigning versus customizing characters as well as explore perceptions of personal character appearance, we conduct a study in which sixth and seventh grade students are introduced to programming concepts with the software VEnvI (Virtual Environment Interactions) in seven one-hour sessions over two weeks. In VEnvI, students create performances for virtual characters by assembling blocks. With a between-subjects design, in which some of the students can alter their character and others are not given that possibility, we examine the influence of the presence or absence of character choice options on learning.

We hypothesize that students have higher learning outcomes when they can choose and customize how their character looks compared to when they are assigned a character. We confirm this hypothesis for a category of learning (Remember and Understand) and give insights on students' relationships with their character.

CCS CONCEPTS

• **Social and professional topics → Computer science education**; **K-12 education**; *Student assessment*; • **Human-centered computing** → *Empirical studies in collaborative and social computing*;

KEYWORDS

virtual character, character customization, pedagogical agent, character appearance

SAP '17, Cottbus, Germany

978-1-4503-5148-5/17/09…$15.00
DOI: 10.1145/3119881.3119884

ACM Reference format:
Lorraine Lin, Dhaval Parmar, Sabarish V. Babu, Alison E. Leonard, Shaundra B. Daily, and Sophie Jörg. 2017. How Character Customization Affects Learning in Computational Thinking. In *Proceedings of SAP '17, Cottbus, Germany, September 16–17, 2017,* 8 pages.
DOI: 10.1145/3119881.3119884

1 INTRODUCTION

From storybook fables to preschool shows, characters are some of the first teachers that people engage with. The importance of characters as more than an entertainment platform is gaining traction as fictional figures are becoming recognized for their appeal as instructors or assistants in educational, medical, and military applications [Johnsen et al. 2005; Kenny et al. 2007; Lane et al. 2013]. Previous research indicates that character appearance affects learning [Baylor 2005]. One advantage of virtual applications compared to real life is that the appearance of virtual agents or teachers can easily be altered to personal preferences. Studies on games have shown that allowing avatar options has positive impacts on players [Lim and Reeves 2009; Turkay and Kinzer 2014]. We therefore believe that providing character customization choices in an educational application might have a positive effect on student learning.

Most previous studies on the effect of character appearance or customization in interactive applications are relatively short interventions displaying agents with a rather static role [Baylor and Kim 2003; Baylor et al. 2003; Jin 2009]. To investigate the effects of character customization options on learning effects, we use the learning software VEnvI (Virtual Environment Interactions). VEnvI is a system and curriculum aimed at middle school students that combines computer science and dance concepts. It uses an embodied approach to teach students computational thinking in a motivating context [Daily et al. 2014, 2015; Parmar et al. 2016]. Teachers explain six basic programming concepts: sequences, loops, variables, conditionals, functions, and parallel programming; and perform dance activities with the students that illustrate the concepts.

In parallel, students are introduced to the VEnvI software, which has a drag-and-drop interface with code blocks similar to Alice [Kelleher et al. 2007], Scratch [Maloney et al. 2010], or Looking Glass [Gross et al. 2010] as illustrated in Figure 1. Students use the computer science concepts they learned to create their own dances for a virtual character. They are encouraged to leave their seats and dance with their characters at any time. VEnvI furthermore supports virtual reality. With a VR headset, users can observe their character and

Figure 1: Screen capture of the VEnvI interface. The virtual character can be seen on the top left. Move sequences and computational constructs, represented by blocks, can be chosen in the bottom left area and assembled to a dance in the drag and drop area on the right side.

dance beside it. Parmar et al. [2016] describe the software and curriculum in greater detail. The system can be obtained by contacting its researchers and developers [VEnvI 2017].

In our experiment, sixth and seventh grade students from a local middle school spent seven one-hour sessions over the course of two weeks using VEnvI to programmatically choreograph dance performances for virtual characters, which were either randomly assigned to them or which they customized at the beginning of the experiment. The main questions we aim to answer are:

- What are the effects of character customization vs. assignment on learning?
- How do students perceive their character? What relationship do they have to it?

2 RELATED WORK

2.1 Character Appearance Affects our Perception and Behavior

The appearance of a virtual character influences the way we perceive and judge it [McDonnell et al. 2012, 2009; Wellerdiek et al. 2015]. It also influences us when we interact with a virtual character or see ourselves as one [Kilteni et al. 2013; Lin and Jörg 2016; Plant et al. 2009; Ring et al. 2014; Rosenberg-Kima et al. 2008].

Interactive media allows a dimension of experience above viewing: by being able to control one character specifically or take on a social role in a game world, interactivity allows an override of the distance between player and character. Hefner et al. [2007] propose that people mainly experience interactive media through a monadic (identification-driven) connection with characters.

Identification with a character can further be split into two categories: players can create a character that is simply a replica of their actual self, or players can create a character that embodies an ideal image they would like to have of themselves. Jin [2009] found that players reported greater immersion during game-play if they were primed with hope and played the game using an avatar reflecting the actual self, or if they were primed with duty and then played the game using an avatar reflecting the ideal self.

2.2 Character Choice vs. No-Choice Conditions

Being able to interact or identify as a character provides positive effects in general to media experiences. Allowing users to select or customize characters has been found to result in further beneficial effects. In a World of Warcraft study by Lim and Reeves [2009], just having the option to select an avatar produces higher arousal, measured by skin conductance activity and heart rate, during gameplay. Turkay and Kinzer [2014] observed participants in a customization group (CG) versus a non-customization group (NCG) playing Lord of the Rings Online. They found that both avatar-based customization and time positively impacted players' identification and empathy with their avatars. Participants also rated hairstyle as the most important customization aspect, followed by hair color and eye color. Qualitative findings showed that many of the CG players customized characters to have some aspect of themselves, such as a skill (playing an instrument) or physical characteristic (hair, eyes, build.) The researchers suggest that, for educational games, avatar-based customization may increase students' motivation, and that if customization is available, players will identify with characters further and focus more on events related to their characters.

2.3 Effects on Learning

While the addition of a character to a learning system does not necessarily have a direct impact on learning [Baylor 2011], the ability to control character appearance via customization or selection, in contrast to simply having a character assigned, impacts factors associated with learning effects in users. Baylor conducted two studies to observe these factors. In the first study [Baylor et al. 2003], students were given the ability to select pedagogical agents that would give a presentation on coping with college life. African-American students rated their agent as "more facilitating of learning, more credible, more human like, and more engaging," than Caucasian students rated their agents. African-American students were also significantly more likely than Caucasian students to choose agents based on gender and ethnicity. Females were more likely to select cartoon-like agents than males, and females were significantly more likely than males to use personal experience for choosing their agents, while males were more likely to choose agents based on previous teachers. Research has shown that character appearance has effects even if characters are just assigned. In the second study [Baylor and Kim 2003], students were assigned agents to aid in learning basic instructional planning skills. Assigned male agents were overall more motivational than assigned female agents: working with male agents resulted in higher reported self-regulation and greater self-efficacy, and male agents were rated as more useful. Such a positive bias towards male agents, although the effects were small, was not found in the first study.

3 MATERIALS AND METHODS

The goal of our experiment is to investigate the effect of character customization on student learning and explore how students perceive the virtual character they work with. Our hypothesis is that students will have higher learning outcomes when they can choose and customize how their character looks compared to when they are assigned a character. To investigate this hypothesis, we used a between-subjects design: students were randomly assigned to a group where they had the option to customize their character (C) or to a group where they were assigned a character (A).

3.1 Customization and Assignment

The VEnvI software allows the user to customize their character using the Unity Multipurpose Avatar (UMA) system, an open-source avatar creation framework available in the Unity Asset Store [2017]. With the UMA system in VEnvI, participants can choose character gender, as well as alter the character's height, upper weight, lower weight, skin tone, clothing colors, and eye color. In a pilot test with the UMA system, participants commented on their characters mentioning that "it was creepy" and "it looked fake". These comments suggest that the characters provided were not appealing to the demography, which could have invalidated our results.

For our experiment, we therefore created a new character customization system and integrated it into VEnvI. An androgynous child model along with its outfits, accessories, and base textures were provided by Adobe Fuse CC. These 3D assets were rigged and skinned in Maya 2016. Adobe Photoshop CC 2017 was used to further edit textures for tinting when necessary. A customization menu and edits to the existing VEnvI code allowing swappable character accessories, textures, and tinting were implemented within the Unity3D game engine. Our customization system provides 12 skin tones, 5 eye colors, 12 hair styles, and 12 hair colors on a first screen; 12 top styles, 8 bottom styles, 8 shoe styles on a second screen; and 7 glasses styles on a third screen (see Figure 2 (a) and (b)). The clothing and accessories could be tinted in 11 colors. Options for all customizable items were limited to twelve or fewer so participants would not spend too much time in the customization menu, but chosen to allow as much room for individuality as possible. Skin, hair, and eye color options offered were chosen to be inclusive to a wide demography. Hair, top, and bottom styles aimed to be as balanced as possible, including a wide mix of female, male, and unisex styles. Characters for the assignment group (see Figure 2 (c)) were pre-designed to reflect a wide variety of demography and clothing styles using the same options. Figure 7 shows examples of characters created with our system.

3.2 Participants

We conducted our experiment at a middle school in four groups: sixth and seventh grade students from a graphics communication class, and students in the same two grades from a dance aerobics class. There was no overlap in students between the two classes. The study was offered as an opt-in activity for the graphics communication class and was mandatory for all students of the dance aerobics class. Informed consent was obtained from all students whose data was used in this study and from their parents following the guidelines approved by our Institutional Review Board. Two dance aerobics class students did not consent and their data was discarded. Further five students' data were excluded as they missed some of the sessions. This led to a total of thirty-six participants (4M, 32F) in our analysis between eleven and thirteen years of age. The distribution of students in each group can be seen in Table 1. Of those, 26 identified as Caucasian, 3 as African American, 1 as Arab, 1 as Asian, 1 as Indian, 1 as Hispanic, and 3 as multiracial. For each group, the experiment consisted of seven one-hour sessions within a time span of two weeks.

Table 1: Participant distribution in our four groups

	graphics communication	dance aerobics
sixth grade	9 (3M, 6F)	12 (all F)
seventh grade	5 (1M, 4F)	10 (all F)

3.3 Design and Measures

Our experiment used a between-subjects design, with eighteen students in the customization condition (C) and eighteen students in the assignment condition (A). The independent variable is the type of character assignment provided, customization vs. assignment. The main dependent variable is the learning effect, measured with a cognitive pre- and post-test. Factors possibly related to learning effect, such as enjoyment and motivation, and to the appearance of the characters and the relationships students had with them, were also included. In total, our measurements are as follows:

Cognitive Pre- and Post-Test. The cognitive test was administered twice, once at the beginning and once at the end of our experiment. It tests for knowledge on computer programming concepts such as sequences, loops, if-else statements, variables, and functions. The questions from the cognitive test are inspired by Bloom's taxonomy [Airasian et al. 2000; Bloom et al. 1984] and the SOLO [Biggs and Collis 2014] taxonomy. Both taxonomies are learning models that categorize educational objectives into different levels. Our cognitive test observes three of the six Bloom cognitive domain levels that build on each other: Remember and Understand, which is retrieving information from memory ("What is the purpose of a loop?"); Apply, which is taking known information and implementing it in a new situation ("Based on the blocks below, will the character perform the "Cha Cha" or "Clap?"); and Analyze, which is breaking down a complex problem into parts to better understand it ("What order should the block below be put into if I want the character to clap first, move to the right side, and then reverse twice?"). Our cognitive test also observes three of the five SOLO levels that build on each other: Unistructural, which is when students are able to answer one portion or aspect of a concept ("Look at the blocks below. When 'PLAY' is clicked, what will the character perform?"); Multistructural, which is when students grasp several aspects of a concept but treat them as independent components ("Give an example of a conditional in your life."); and Relational, which is when students are aware of the logic and steps they need to do to solve the problem ("Look at the blocks in Picture A and Picture B below. Would a character performing Picture A and another performing Picture B do the same thing?").

Experience Questions. Experience questions asked students about factors such as their confidence, interest in, and opinions on the importance of learning about dance and programming that might have changed due to experiencing VEnvI ("How likely are you to choose computing as a major in college?"). These questions were

(a) Customization screen 1 (b) Customization screen 2 (c) Assignment screen

Figure 2: Customization and assignment screens.

Table 2: Overview of experiment schedule

Session 1	Demographics, cognitive pre-test, experience questions, introduction to VEnvI, dance activity
Session 2	Sequences, dance activity, intro to VEnvI software and VR, character customization and questions
Session 3	Loops and parallelization combined with dance activities, implementation in VEnvI
Session 4	Variables, conditionals, and functions with dance activities, implementation in VEnvI
Session 5	Implementation of computing concepts in VEnvI
Session 6 and 7	Dance show, cognitive post-test, experience and character questions

asked twice, once at the beginning and once at the end of our experiment. In addition, questions that were relevant once students had completed VEnvI, measuring enjoyment, ease of learning, and motivation, were asked at the end.

Character Questions. Students were asked questions regarding their character at the start of and after the study. Additional questions that were relevant once students had completed VEnvI were asked at the end ("How much did you feel like dancing with the character?"). To examine if students tend to create characters similar to them, we took pictures of students together with their characters.

3.4 Procedure

The experiment took place over seven one-hour sessions spread over the course of two weeks following the VEnvI curriculum. An overview of the activities in each session is provided in Table 2. In the first session, after filling out an initial demographic questionnaire, participants were asked to answer experience questions and complete the cognitive pre-test to the best of their abilities. Participants were invited to a discussion of what they think research is, and we summarized what to expect over the next two weeks. Finally, students learned some of the choreography available in VEnvI with an instructional dance activity.

In the second session, participants were introduced to the VEnvI software. Students were randomly assigned to one of our two conditions. In all remaining sessions, students were seated in two rows facing each other with a large space in-between where the dancing activities took place. Students in the customization condition were seated in one row and students in the assignment condition were seated in the other row. This procedure was chosen to ensure that the students from the different conditions could not look at each other's screens and to avoid discussions amongst students that had customization options and students who did not have these choices. Students interacted with VEnvI on Apple Macbook Pro laptops.

Upon opening VEnvI, participants in the C group were given options to customize their character. Once they were satisfied, they could name their character and proceed to the main VEnvI programming interface. This process took about five minutes. Participants in the A group had no control over their characters' appearance. Upon opening VEnvI, they viewed a pre-designed character randomly assigned to them, and the character creation process only consisted of naming their character before proceeding to the main VEnvI interface (see Figure 2 (c).) After allowing participants a few moments with their character and the VEnvI software, we asked the character questions. Participants were also introduced to the Virtual Reality setup: from this session on, they had access to view their dances in Virtual Reality through the use of an Oculus Rift head-mounted display and a Microsoft Kinect tracker (see Figure 3.)

Figure 3: The Virtual Reality setup.

In the next sessions, students were introduced to programming concepts intermingled with illustrative dance activities. For example, researchers explained the concept of loops, then invited participants to illustrate the concept by repeating a number of dance moves as if they were looped. In the last two sessions, students presented the dances they programmed and filled out questionnaires.

4 RESULTS

We analyzed all questions we asked – the cognitive pre- and post-test, the experience questions, and the character questions – as well as the pictures of the students with their characters.

4.1 Cognitive Pre- and Post-Test

Participants' number of correct responses in the cognitive pre- and post-test were converted to a percentage score. The scores were normally distributed based on descriptive statistics analysis. We conducted a 2x2 mixed model repeated measures ANOVA with the cognitive pre- and post-test scores as a within-subjects repeated measures factor and the type of character choices provided – Customization (C) vs. Assignment (A) – as a between-subjects factor. We ensured Box's M test of equality of covariance matrices was not significant. Mauchly's test of sphericity was conducted to ensure error variance in groups of samples was found to be equivalent.

We found a main effect of session (pre vs. post) with $F(1, 34) = 28.178, p < 0.001, \eta^2 = 0.453$. Overall, on the cognition questionnaire, participants in the post-test session ($M = 36.90, SD = 16.33$) scored significantly higher than participants in the pre-test session ($M = 25.79, SD = 13.90$), $p < 0.001$, which shows that our VEnvI program overall was successful in teaching students concepts.

The results were separated into further categories for analysis: Remember and Understand, Apply, and Analyze questions according to Bloom's taxonomy; and Unistructural, Multistructural, and Relational questions based on the SOLO taxonomy.

As expected, main effects of session, comparing the results between the pre- and the post-test, were found for most categories, namely for Remember and Understand, Apply, Analyze, and Multistructural (see Table 3). In all categories, students scored significantly higher in the post-test session at the end of the VEnvI curriculum than in the pre-test session, which confirms that they learned some of the concepts we taught. We furthermore found significant results regarding the type of character assignment provided for the categories Remember and Understand and Analysis.

4.1.1 Remember and Understand. Our analysis revealed an interaction effect of session (pre vs. post) x type of character assignment (C vs. A) with $F(1, 34) = 6.54, p < 0.05$, and $\eta^2 = 0.161$ (see Figure 4.) Post-hoc comparisons with Bonferroni corrections showed that while participants in both groups improved significantly between the pre- and the post-test, participants in the Customization group increased their scores significantly more, by 25 percentage points on average from 8.33 ($SD = 17.15$) to 33.33 ($SD = 19.17$), than students in the Assignment group who only improved by 11.11 percentage points on average from 8.33 ($SD = 12.13$) to 19.44 ($SD = 13.71$) ($p < 0.05$ for all comparisons except for the results for the two groups in the pre-tests).

4.1.2 Analyze. In the Analyze category we found a main effect for the type of character assignment (C vs. A) with $F(1, 34) = 5.13, p < 0.05, \eta^2 = 0.131$. Participants in the Customization group ($M = 37.5, SD = 21.44$) scored significantly lower than participants in the Assignment group ($M = 56.94, SD = 29.46$) with $p < 0.05$ (see Figure 5 (a).) There was no interaction effect. Taking into account the main effect of session, this means that both groups improved in the questions in the Analyze category in a similar way, but students in the Assignment group were ahead of students in the Customization group to start with. Our two groups, even if the students have been randomly assigned to the conditions, therefore differed in this category from the beginning on.

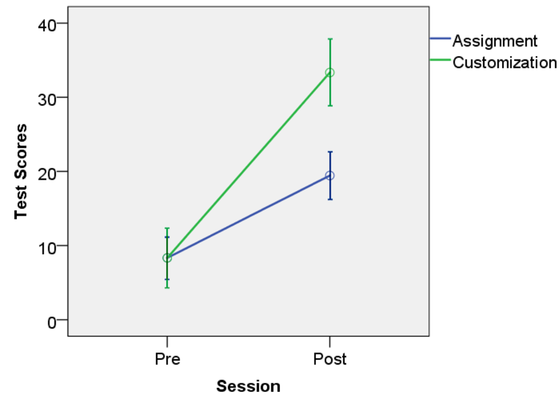


Figure 4: Interaction effect for Remember and Understand. Error bars represent the standard errors of the mean.

4.2 Experience Questions

To get insights on students' experience, we also conducted 2x2 mixed model repeated measures ANOVAs with the pre- and post-answers as a within-subjects repeated measures factor and the type of character assignment as a between-subjects factor. Again, we ensured that Box's M test of equality of covariance matrices was not significant, and Mauchly's test of sphericity was conducted.

We found main effects for four questions listed in Table 4. Students felt significantly more confident at dancing after the program and significantly more of them felt that they knew what a computer programming language is. However, on average they were less inclined to want to learn more about dance and programming after our program. While these results are not connected to our main questions, they give insight into the effectiveness of the VEnvI curriculum. Overall, 24 of the 36 participants gave positive comments to the VEnvI experience itself such as "It was fun."

4.3 Character Questions

The Character Questions included ratings but also open questions. We report our quantitative and qualitative findings below.

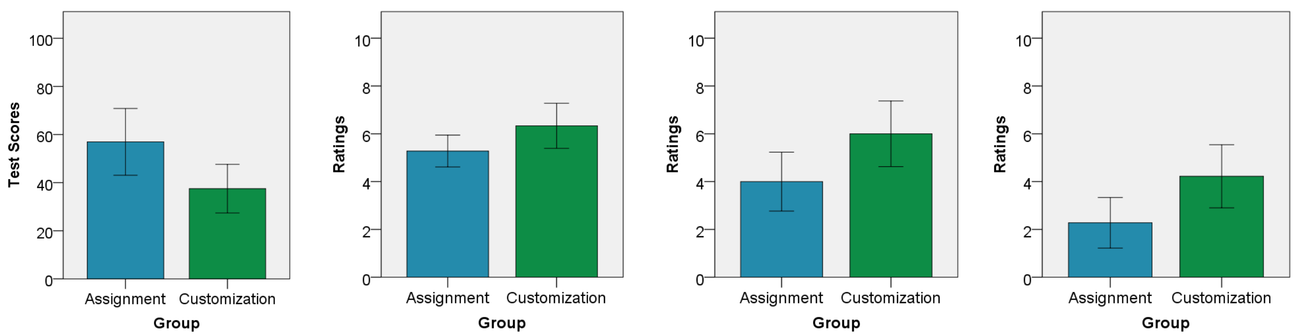
4.3.1 Quantitative Analysis. The question "How well do you like your character?", which was asked at the beginning and end of the program, was analyzed with a 2x2 ANOVA in the same way than the cognitive pre- and post-test and the experience questions. There was a main effect of session (pre vs. post) with $F(1, 34) = 5.97, p < 0.05$, and $\eta^2 = 0.149$. On average, participants liked their characters significantly less ($M = 5.24, SD = 2.41$) at the end of the VEnvI program than in the session during which the character was created ($M = 6.21, SD = 2.08$). The analysis also revealed a main effect of the type of character assignment for this statement

Table 3: Main effects of session in individual categories of the cognitive pre- and post-test. All effects showed increases in student scores.

Category	F-test, p-value, and eta-squared	Pre-test M and SD	Post-test M and SD
Remember and Understand	$F(1,34) = 44.20, p < 0.001, \eta^2 = 0.565.$	$M = 8.33, SD = 14.64$	$M = 26.39, SD = 17.87$
Apply	$F(1,34) = 5.79, p < 0.05, \eta^2 = 0.146.$	$M = 29.63, SD = 16.37$	$M = 36.73, SD = 21.87$
Analyze	$F(1,34) = 6.39, p < 0.05, \eta^2 = 0.158.$	$M = 40.28, SD = 33.42$	$M = 54.17, SD = 30.18$
Multistructural	$F(1,34) = 54.21, p < 0.001, \eta^2 = 0.615.$	$M = 8.79, SD = 12.90$	$M = 26.85, SD = 17.49$

Table 4: Significant results from the experience questions.

Statement	Scale	Pre Session M, SD	Post Session M, SD	p value	Direction
Do you feel like you are confident at dancing?	1-10	M = 5.07, SD = 3.83	M = 6.50, SD = 2.81	p <0.05	UP
I want to learn more about dance.	1-5	M = 4.21, SD = 0.98	M = 3.74, SD = 1.18	p <0.001	DOWN
Do you know what a computer programming language is?	0-1	M = 0.14, SD = 0.36	M = 0.51, SD = 0.51	p <0.001	UP
I want to learn more about programming.	1-5	M = 3.72, SD = 0.87	M = 3.18, SD = 0.76	p <0.001	DOWN



(a) (b) (c) (d)

Figure 5: Main effects of group (Customization vs. Assignment) were found (a) for the Analyze category of the cognitive pre- and post-test and for the questions (b) "How well do you like your character", (c) "Did the presence of others affect your decision to dance or not", and (d) "Did you feel like the character was yourself". Except for the first result, ratings were significantly higher in the Customization group compared to the Assignment group. Error bars represent the standard errors of the mean.

with $F(1,34) = 14.26, p < 0.001$, and $\eta^2 = 0.30$. Participants in the Customization group ($M = 6.33, SD = 2.00$) scored significantly higher than participants in the Assignment group ($M = 5.28, SD = 1.41$), $p < 0.05$ (see Figure 5 (b).)

One-way between-subjects ANOVAs were conducted on questions that were only asked at the end of the program. We found a main effect of the type of character assignment ($F(1,34) = 4.71, p < 0.05, \eta^2 = 0.122$) for the question "Did the presence of others affect your decision to dance or not?" Participants in the Customization group ($M = 6.0, SD = 2.91$) felt that the presence of others affected their decision to dance significantly more than participants in the Assignment group ($M = 4.0, SD = 2.61$) (see Figure 5 (c).)

For the question "Did you feel like the character was yourself?" the analysis also revealed a significant main effect of group ($F(1,34) = 5.30, p < 0.05, \eta^2 = 0.135$). Participants in the Customization group ($M = 4.22, SD = 2.80$) scored significantly higher than participants in the Assignment group ($M = 2.28, SD = 2.24$) (see Figure 5 (d).)

To investigate the type of character students created, we also asked "Did you create this character to be: yourself, a better version of yourself, a friend, a teacher, or other?" This question was inspired by Baylor et al.'s work [2004]. One participant's response was omitted because the question was left blank. Selection of the five choices was not evenly distributed in the Assignment vs. Customization groups ($p < 0.001$, Fisher's exact test). As can be seen in Figure 6, students did not identify as much with their characters if it was assigned to them: no one from that group selected "yourself" and only two students selected "a better version of yourself". None of the students chose to view their character as a "teacher." Students who chose 'other' saw their characters as dancers, "a famous dancer," random people, "a person on TV," or "Someone I am telling what to do. Like if I was the coach and they were the student."

4.3.2 Qualitative Findings. Seventeen of the students in the Customization group allowed us to take photos of them with their characters, see Figure 7 for examples.

Of these students, 15 created a character with the same gender as them, 13 chose a similar skin color, 6 chose a similar hair color, and 6 chose a similar hair style. All customizable characters were changed from the default template character, which had a medium skin tone, brown boy-styled hair, and a monochromatic outfit.

Students were furthermore asked what they liked and disliked about their character. Six positive comments from students in the

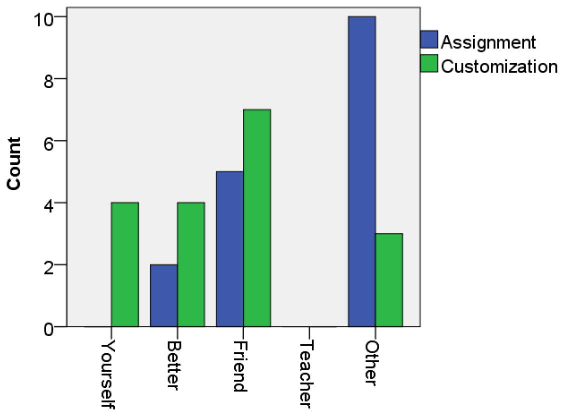


Figure 6: Scores for "Did you create this character to be."

Assignment group mostly complimented the character's existing outfit parts like shirt or shoes. All participants in the Assignment group commented that they disliked something about the character. Of these, 8 expressed a want to have a customization system, 2 would have liked a different gender, 1 would have liked a different skin tone, 2 wanted to identify more with the character ("I was kinda hoping it would look like me and it doesn't," "I can't relate to him"), and 7 gave negative comments about the avatar's appearance ("...he's not very fashionable," "...it's pretty ugly.")

Participants in the Customization group had 9 positive comments about the characters: 8 were about appearance and 1 about being able to customize. There were 12 negative comments: Of those, 3 were not directly about the character's appearance but about the VEnvI system ("Sometimes she's hard to control, like she'll go offstage"), and 9 expressing that they wanted more customization options ("I kinda wanna change her outfit" (participants are not able to return to the customization screen and edit their characters once they are in the VEnvI dance interface), "I don't like the hair. Needs more variety of accessories, styles, and colors.") One of the students in the Customization group said, "I tried to make it me but there wasn't any way to make it look like me."

Eight of the participants gave their characters personalities, 6 of them were from the Customization group. Personalities given were positive like "a goof," and "I gave it [classmate]'s personality."

5 DISCUSSION

In our results, we found multiple main effects of session, showing students' cognitive scores improved during the VEnvI program in general. Our two groups did not perform fully equally in our pre-test, but such results can happen despite randomly assigning students to the conditions. It neither confirms nor rejects our hypothesis of higher learning outcomes in the Customization group.

We found evidence that character customization has a positive effect on learning when looking at specific categories of learning, which confirms our hypothesis. Participants who were given the possibility to customize characters learned better on Remember and Understand questions than students who were assigned characters.

Several reasons for this result are possible. Participants in the Customization group might be more attentive to the lesson. They might have felt more empathy with their character as they were given choices, as found by Turkay and Kinzer [2014], or been more psychophysiologically alert just from having the ability to customize, as found by Lim and Reeves [2009]. Participants in the Customization group liked their character more than participants in the Assignment group. They identified more with their character and also were more likely to see their character as themselves, better versions of themselves, or friends. Supporting identification with the character, over half of the Customization group students chose a character with the same gender and similar skin tone to them. Customization group students might have felt more support from their character, leading to a better learning performance.

A further finding was that students liked their character less at the end of the study than shortly after customizing it. The short duration of the customization process, animation artifacts, or unrealistic motions could explain this result. The ability to edit a character at a later point might change this finding and increase students' satisfaction with their characters over time.

Interestingly, students in the Customization group felt that the presence of others around them affected their decision to dance more than students in the Assignment group. This result points toward a lower confidence of students in the Customization group. The comments in both groups range from neutral ("I don't care what people think") to negative ("...I didn't want to look stupid"). Of the 24 participants who commented, none of the participants said the presence of others on their decision to dance was positive.

5.1 Conclusion and Future Work

In summary, we conducted a study to see how the presence or absence of character customization influenced a learning system, as well as observed participant choice of character appearance. Our main finding was that participants with customizable characters learned better in the Remember and Understand level of Bloom's taxonomy. Participants with customizable characters also liked and identified with their characters more and were more influenced by the presence of their peers in deciding to dance.

Several further questions arise from this study: Our main finding only applies to a specific category of learning, Remember and Understand, and more studies will be necessary to understand why and if customization only improves learning effects in this category.

Previous studies have shown that some characters are more effective as pedagogical agents than others. However, given the ability to customize or select characters, students may create a character appearance that is not conducive to teaching. Future studies could investigate if people naturally build characters that would facilitate learning the most, or if it is more beneficial for learning in some cases to assign pedagogical agents.

In addition, we decided to limit our characters to virtual humans. Future work could assess learning effects on a wider range of human or non-human characters such as aliens, zombies, or animals.

Finally, our study is on specific software combining dance and computer programming concepts. More studies on different learning systems and topics would be needed to confirm if the learning effects we discovered can be generalized to other fields and systems.

Figure 7: Participants with their customized characters.

6 ACKNOWLEDGMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1344228.

REFERENCES

Peter Airasian, Kathleen Cruikshank, Richard Mayer, Paul Pintrich, James Raths, and Merlin Wittrock. 2000. *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives.* Pearson.

Amy L. Baylor. 2005. The Impact of Pedagogical Agent Image on Affective Outcomes. In *Proc. of Workshop on Affective Interactions: Computers in the Affective Loop, International Conf. on Intelligent User Interfaces.* ACM, San Diego, CA.

Amy L. Baylor. 2011. The design of motivational agents and avatars. *Educational Technology Research and Development* 59, 2 (2011), 291–300.

Amy L. Baylor and Yanghee Kim. 2003. The Role of Gender and Ethnicity in Pedagogical Agent Perception. In *Proc. of E-Learn: World Conf. on E-Learning in Corporate, Government, Healthcare, and Higher Education.* AACE, Phoenix, AZ, 1503–1506.

Amy L. Baylor and Yanghee Kim. 2004. *Pedagogical Agent Design: The Impact of Agent Realism, Gender, Ethnicity, and Instructional Role.* Springer, Berlin, Heidelberg, 592–603.

Amy L. Baylor, E Shen, and Xiaoxia Huang. 2003. Which Pedagogical Agent do Learners Choose? The Effects of Gender and Ethnicity. In *Proc. of E-Learn: World Conf. on E-Learning in Corporate, Government, Healthcare, and Higher Education.* AACE, Phoenix, AZ, 1507–1510.

John B. Biggs and Kevin F. Collis. 2014. *Evaluating the quality of learning: The SOLO taxonomy (Structure of the Observed Learning Outcome).* Academic Press.

Benjamin S. Bloom, David R. Krathwohl, and Bertram B. Masia. 1984. *Bloom taxonomy of educational objectives.* Pearson, Boston, MA.

Shaundra B. Daily, Alison E. Leonard, Sophie Jörg, Sabarish V. Babu, and Kara Gundersen. 2014. Dancing Alice: Exploring Embodied Pedagogical Strategies for Learning Computational Thinking. In *Proceedings of the 45th ACM Technical Symposium on Computer Science Education (SIGCSE '14).* ACM, 91–96.

Shaundra B. Daily, Alison E. Leonard, Sophie Jörg, Sabarish V. Babu, Kara Gundersen, and Dhaval Parmar. 2015. Embodying Computational Thinking: Initial Design of an Emerging Technological Learning Tool. *Technology, Knowledge and Learning* 20, 1 (2015), 79–84.

Paul A. Gross, Micah S. Herstand, Jordana W. Hodges, and Caitlin L. Kelleher. 2010. A Code Reuse Interface for Non-programmer Middle School Students. In *Proc. of the 15th International Conf. on Intelligent User Interfaces.* ACM, 219–228.

Dorothée Hefner, Christoph Klimmt, and Peter Vorderer. 2007. Identification with the Player Character as Determinant of Video Game Enjoyment.. In *Entertainment Computing-ICEC 2007. Lecture Notes in Computer Science*, Vol. 4740. Springer-Verlag, Berlin, Heidelberg, 39–48.

Seung-A Annie Jin. 2009. Avatars Mirroring the Actual Self Versus Projecting the Ideal Self: The Effects of Self-Priming on Interactivity and Immersion in an Exergame, *Wii Fit. CyberPsychology & Behavior* 12 (Dec. 2009), 761–765. Issue 6.

Kyle Johnsen, Robert Dickerson, Andrew Raij, Benjamin Lok, Jonathan Jackson, Min Shin, Jonathan Hernandez, Amy Stevens, and D. Scott Lind. 2005. Experiences in using immersive virtual characters to educate medical communication skills. In *IEEE Virtual Reality, 2005.* Bonn, Germany, 179–186.

Caitlin Kelleher, Randy Pausch, and Sara Kiesler. 2007. Storytelling Alice Motivates Middle School Girls to Learn Computer Programming. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems.* ACM, 1455–1464.

Patrick G. Kenny, Arno Hartholt, Jonathan Gratch, William Swartout, David Traum, Stacy C. Marsella, and Diane Piepol. 2007. Building Interactive Virtual Humans for Training Environments. In *Interservice/Industry Training, Simulation and Education Conference (I/ITSEC).* Orlando, FL.

Konstantina Kilteni, Ilias Bergstom, and Mel Slater. 2013. Drumming in Immersive Virtual Reality: The Body Shapes the Way We Play. In *2013 IEEE Virtual Reality (VR).* IEEE, Lake Buena Vista, FL, 597–605.

H. Chad Lane, Clara Cahill, Susan Foutz, Daniel Auerbach, Dan Noren, Catherine Lussenhop, and William Swartout. 2013. The Effects of a Pedagogical Agent for Informal Science Education on Learner Behaviors and Self-efficacy. In *Artificial Intelligence in Education (Lecture Notes on Computer Science)*, Vol. 7926. Springer, Berlin, Heidelberg, 309–318.

Sohye Lim and Byron Reeves. 2009. Being in the Game: Effects of Avatar Choice and Point of View on Psychophysiological Responses During Play. *Media Psychology* 12, 4 (2009), 348–370.

Lorraine Lin and Sophie Jörg. 2016. Need a Hand?: How Appearance Affects the Virtual Hand Illusion. In *Proceedings of the ACM Symposium on Applied Perception (SAP '16).* ACM, 69–76.

John Maloney, Mitchel Resnick, Natalie Rusk, Brian Silverman, and Evelyn Eastmond. 2010. The Scratch Programming Language and Environment. *Transactions on Computing Education* 10, 4, Article 16 (Nov. 2010), 15 pages.

Rachel McDonnell, Martin Breidt, and Heinrich H. Bülthoff. 2012. Render Me Real?: Investigating the Effect of Render Style on the Perception of Animated Virtual Humans. *ACM Trans. Graph.* 31, 4, Article 91 (July 2012), 11 pages.

Rachel McDonnell, Sophie Jörg, Jessica K. Hodgins, Fiona Newell, and Carol O'Sullivan. 2009. Evaluating the Effect of Motion and Body Shape on the Perceived Sex of Virtual Characters. *ACM Transactions on Applied Perception* 5, 4, Article 20 (Feb. 2009), 14 pages.

Dhaval Parmar, Joseph Isaac, Sabarish V. Babu, Nikeetha D'Souza, Alison. E. Leonard, Sophie Jörg, Kara Gundersen, and Shaundra B. Daily. 2016. 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).* 131–140.

E. Ashby Plant, Amy L. Baylor, Celeste E. Doerr, and Rinat B. Rosenberg-Kima. 2009. Changing middle-school students' attitudes and performance regarding engineering with computer-based social models. *Computers & Education* 53, 2 (2009), 209 – 215.

Lazlo Ring, Dina Utami, and Timothy Bickmore. 2014. *The Right Agent for the Job?* Springer International Publishing, Cham, 374–384.

Rinat B. Rosenberg-Kima, Amy L. Baylor, E. Ashby Plant, and Celeste E. Doerr. 2008. Interface Agents As Social Models for Female Students: The Effects of Agent Visual Presence and Appearance on Female Students' Attitudes and Beliefs. *Comput. Hum. Behav.* 24, 6 (Sept. 2008), 2741–2756.

Unity Asset Store. 2017. UMA – Unity Multipurpose Avatar. (accessed May 26, 2017). https://www.assetstore.unity3d.com/en/#!/content/13930

Selen Turkay and Charles K. Kinzer. 2014. The Effects of Avatar-based Customization on Player Identification. *International Journal of Gaming and Computer-Mediated Simulations* 6, 1 (2014), 1–25.

VEnvI. 2017. VEnvi Project Website. (accessed May 26, 2017). http://venvi.org/venvihome.html

Anna C. Wellerdiek, Martin Breidt, Michael N. Geuss, Stephan Streuber, Uwe Kloos, Michael J. Black, and Betty J. Mohler. 2015. Perception of Strength and Power of Realistic Male Characters. In *Proc. of the ACM SIGGRAPH Symposium on Applied Perception (SAP '15).* ACM, 7–14.