Analyzing the Effect of Avatar Self-Similarity on Men and Women in a Search and Rescue Game

Helen Wauck\textsuperscript{1}, Gale Lucas\textsuperscript{2}, Ari Shapiro\textsuperscript{2}, Andrew Feng\textsuperscript{2}, Jill Boberg\textsuperscript{2}, Jonathan Gratch\textsuperscript{2}

\textsuperscript{1}University of Illinois at Urbana-Champaign, \textsuperscript{2}USC Institute for Creative Technologies

\textsuperscript{1}Urbana, IL, \textsuperscript{2}Playa Vista, CA

wauck2@illinois.edu, \{lucas,gratch,feng,boberg,shapiro\}@ict.usc.edu

Figure 1. Screenshots of participants playing the search and rescue game with photorealistic avatars. Left: Participant stops and brings up the map (upper right corner) to figure out which hallway to take. Right: Participant runs toward a lever to open a blast door and rescue the gems inside.

ABSTRACT

A crucial aspect of virtual gaming experiences is the avatar: the player’s virtual self-representation. While research has demonstrated benefits to using self-similar avatars in some virtual experiences, such avatars sometimes produce a more negative experience for women. To help researchers and game designers assess the cost-benefit tradeoffs of self-similar avatars, we compared players’ performance and subjective experience in a search and rescue computer game when using two different photorealistic avatars: their own self or a friend, and when playing either a social (rescuing people) or a nonsocial (rescuing gems) version of the game. There was no effect of avatar appearance on players’ performance or subjective experience in either game version, but we also found that women’s experience with self-similar avatars was no more negative than men’s. Our results suggest that avatar appearance may not make a difference to players in certain game contexts.

ACM Classification Keywords

H.5.0. Information Interfaces and Presentation: General

Author Keywords

avatars; photorealism; identification; games; gender; player experience; performance

INTRODUCTION

As technological advances over the last decade have produced increasingly realistic graphics for video games and other immersive virtual experiences, one’s avatar - the player’s virtual self-representation [3] - has been subject to increased attention. The research literature is replete with examples of how avatars -ification, the player’s ability to identify with their avatar’s visual and behavioral characteristics [34], can benefit gaming experiences. Identification with player-customized avatars ranging from cartoonish Nintendo Miis [22] to more realistic human characters [9] has been shown to improve subjective player experience in laboratory settings, and players’ self-reported experiences with a wide variety of commercial games indicate the benefits of avatar identification “in the wild” [38], even if the avatar does not look much like the player [20, 23].

Now that we have the technological capability to produce more realistic avatars in games and training simulations, perhaps players could benefit even more from an avatar that actually looks just like them: a virtual doppelganger [4]. However,
implementing high quality self-similar avatars comes with a significant time and monetary investment, even in this age of increasing graphical capabilities. Are there really measurable benefits for the player that would make such an investment worth the time and money?

Due to the newness of academic research on avatar appearance, the research literature is rather sparse and does not provide a clear answer to this question. Some studies have found positive effects of using a self-similar avatar on performance or subjective experience [40, 16, 21, 9, 28, 30], while many others report mixed effects due to gender differences [15, 2, 26]. In addition, only one of these studies used extremely high fidelity, photorealistic avatars. The results of this study were mixed; self-similar avatars had no effect on performance in a maze running game, but had benefits for subjective experience, particularly for male players [26].

The focus of the present study is to contribute to a more cohesive body of work on the effects of high fidelity self-similar avatars in gaming to help game designers evaluate the costs and benefits of implementing them. Building off Lucas et al.’s work [26], we use a third person search and rescue computer game to assess the effect of using photorealistic self-similar avatars on performance and subjective experience. We extend Lucas et al.’s work and build upon it in three main ways.

First, rather than comparing a self-similar avatar to a stranger’s avatar as Lucas et al. did, we compare a self-similar avatar to an avatar of the participant’s close friend to avoid confounding the effects of familiarity and self-similarity. Second, we use a game with a more realistic game environment (graphics, music, sound effects, and narrative) than the maze task in Lucas et al.’s work to see if avatar appearance makes more of a difference to players when realistic avatars are matched with a realistic environment to produce a more cohesive, immersive gaming experience. Third, we examine the moderating effect of in-game social elements by comparing the effects of avatar appearance between two different versions of our search and rescue game: one where the player rescues people (social) and one where the player rescues gems (nonsocial) (see Figure 1).

We conducted a study with 33 friend pairs (66 participants) comparing performance (number of people or gems rescued) and subjective experience (engagement, enjoyment, and frustration) between participants playing as themselves and those playing as their friend. We analyzed the social nature of the game (rescuing people versus gems) as a moderator of avatar self-similarity effects. Given the prevalence of gender differences in self-similar avatars’ efficacy in previous work [2, 26], we also analyzed gender as a potential moderating variable.

Surprisingly, we found no effect of avatar appearance on performance or subjective experience in both the people and gems versions of the game. We did find an interaction between gender and avatar appearance for performance, but the interaction vanished when controlling for players’ prior gaming experience. We also found that gender did not moderate the effect of avatar appearance on subjective experience either, a positive result given that previous research has found that women sometimes respond more negatively to self-similar avatars than men [2, 26, 17]. Overall, our results suggest that avatar appearance may not have much of an effect on players’ performance and subjective experience, at least in the kind of game we studied. We discuss recommendations for game designers based on our findings and offer suggestions for future research that might be more likely to detect an effect of avatar appearance if one exists. Our results contribute to the growing body of work on self-similar avatars and provide insights for game designers about the cost-benefit tradeoff of implementing them.

BACKGROUND AND RELATED WORK

We describe the current state of research on self-similar avatars and how our work builds upon it. The earliest studies of players’ experience with their in-game avatars examined player identification with the avatar rather than the player’s actual physical similarity. In identifying with their avatar, the player temporarily alters their self-concept to include some of their avatar’s characteristics and in some sense “becomes” that character [20, 23]. Players that are able to identify with their avatar generally report a more positive gaming experience, both in games where the avatar can be customized by the player and in games where this is not possible [20, 38, 9].

The Benefits of Self-Similar Avatars

Bolstered by the above findings, more recent research has aimed at determining if actual physical similarity by itself can promote player identification resulting in better performance or subjective experience with a specific task. Some results in this direction of research have been positive; for instance, Birk et al. found that identifying with one’s customized avatar increased autonomy, immersion, invested effort, enjoyment, and positive affect in an endless runner task [9]. In the domain of health research, a series of 3 studies found that participants who saw their own avatar lose or gain weight based on physical activity levels engaged in more voluntary exercise during the study than those who did not see their avatar gain or lose weight or who witnessed weight change on a different avatar’s body. In addition, participants reported exercising more in the 24 hour period following the study when they witnessed their own avatar running than when they saw someone else’s running or their own avatar standing still [16]. The benefits of self-similar avatars also extend to education; in a study by Parmar, students learned programming concepts better in a virtual reality learning environment when they used self-similar, embodied avatars [30].

Even just seeing a static image of their avatar can affect people’s behavior and subjective experience in a task. Hersfield et al. found that participants who were shown an aged version of their self-similar avatar engaged in more prudent long-term financial decision making [21]. A study by Vasalou et al. found that participants who saw a customized avatar representation of themselves while playing a social dilemma game with another person online reported higher levels of private self-awareness, which can encourage self-disclosure [27], than those who did not see an avatar [40]. In the context of education, a study by Mazlan found that students given customizable human avatars were more motivated to learn on an online course platform [28]. Taken together, these results suggest that using avatars that are similar to the player may.
benefit players’ performance and subjective experience on various tasks in laboratory or educational settings, and these benefits may even carry over into players’ daily lives.

**What About Photorealistic Self-Similar Avatars?**

If less photorealistic self-similar avatars benefit performance and subjective experience, how much more could users benefit from an even more realistic avatar? The only research to date that has examined the effects of photorealistic self-similar avatars is a study by Lucas et al. where participants completed a maze navigation game using either a self-similar avatar or a stranger’s avatar, both of which were generated using body scanning technology. The results of this study were mixed; while there were no effects of avatar appearance (self versus stranger) on participants’ performance in the game, participants felt more connected and engaged and liked their avatar more when it was their own self than when it was a stranger. In addition, men enjoyed the game more when playing as their own avatar, whereas women enjoyed it more with a stranger’s avatar [26]. While photorealistic avatars provide exciting possibilities for enhancing players’ performance and subjective experience, one study cannot give us a clear sense of the costs and benefits associated with implementing them. Our work seeks to provide more guidance in this area by building upon Lucas et al.’s work. Like Lucas et al., we examine the effect of photorealistic self-similar avatars on player performance and experience in a game. However, we extend and improve upon this work in several important ways, all of which will be discussed in the following sections.

**Gender Can Moderate the Effects of Self-Similar Avatars**

In certain contexts, gender can have a surprising amount of influence on the performance and subjective experience effects of self-similar avatars. For instance, Fox et al. found that after observing their avatar eat candy and carrots in a virtual reality simulation, men who reported high levels of presence in the virtual experience were more likely to eat candy than men reporting low levels of presence, whereas women with high presence levels were less likely to eat candy than women with low presence levels [15]. In another study comparing interventions for alleviating public speaking anxiety, men had less state anxiety and reported more perceived communication competence after watching a self-similar avatar give a speech than when they were asked to simply imagine giving the speech, but for women, seeing their own avatar give the speech produced more anxiety and less perceived competence than the imagination exercise [2]. And Lucas et al. found that men enjoyed a maze navigation game more when playing as their own self than when playing with a stranger’s avatar, but women enjoyed the game less playing as their own self than with the stranger’s avatar [26].

These negative effects for women observed across multiple studies suggest that more work is needed to understand why and under what conditions these gender differences appear. One possible explanation may be that women were more sensitive to the flaws in the computer-generated representations of themselves than men and did not like seeing themselves portrayed in a less flattering light. The avatars used in Fox et al.’s and Aymerich-Franch and Bailenson’s work were self-similar, but not photorealistic. In addition, the body scanning methods used to generate avatars in Lucas et al.’s study rendered hands and hair poorly and used male animations for movement, regardless of the participant’s gender. Our work improves avatar fidelity compared to these previous studies in two main ways: we use photorealistic avatars coupled with better hand and hair rendering than Lucas et al.’s, and we give female avatars female animations. These improvements allow us to determine whether the avatar’s degree of fidelity is indeed a cause of gender differences in the perception of self-similar avatars.

**Social Tasks and Self-Similar Avatars**

As far as we are aware, no previous work has studied the social nature of the task as a moderator of the effects of avatar appearance on performance or subjective experience. However, Lucas et al. provides a compelling case for studying the interaction between avatar appearance and the social nature of the task the avatar is used for [26]. Since seeing a self-similar avatar has been shown to heighten users’ self-awareness in online social interactions [27], giving players their own self as an avatar may prime them to care more about a task that is more social in nature. We contribute the first study analyzing how the addition of social elements to a task affects performance and subjective experience with a self-similar avatar. Our study uses two versions of a search and rescue game: one where the task is to rescue people (a social task), and another where the task is to rescue gems (a nonsocial task). In addition, we design a new game for our study that is more immersive and realistic than the maze game in Lucas et al.’s work, since they reported that their more crudely designed game environment may have masked avatar appearance effects due to low player engagement and immersion.

The goal of the present study is to build off Lucas et al.’s study and other previous work, investigating the effect of a photorealistic self-similar avatar on performance and subjective experience and the role of gender and in-game social elements as potential moderators of these effects.

**HYPOTHESES**

Given the benefits of self-similar avatars found in other research studies [40, 16, 21, 9], we hypothesize that our photorealistic avatars and our immersive, realistic game setting will likewise reveal benefits to performance and subjective experience when playing as a self-similar avatar:

**H1:** Participants will rescue more people or gems when playing as their own avatar than when playing as their friend’s.

**H2:** Participants will have a better subjective experience with their own avatar than with their friend’s avatar such that they report more engagement, more enjoyment, and less frustration while playing the navigation game.

Based on the research literature suggesting that reducing anonymity increases social self-awareness [27], we believe that the addition of a social element will benefit players with a self-similar avatar:

**H3:** Adding a social element to the game (rescuing people instead of rescuing gems) will moderate the effect of avatar appearance on performance and subjective experience such that
participants playing as a self-similar avatar will rescue more people than gems and report more engagement, enjoyment, and less frustration when rescuing people.

Lastly, given the gender differences in response to self-similar avatars observed in previous work [15, 2], particularly the findings of Lucas et al. [26], we predict that female friend pairs will be more self-conscious about their physical appearance than male friend pairs and will therefore have a more negative experience than male pairs with the game:

**H4:** Female pairs will have a worse subjective experience (less engagement, less enjoyment, and more frustration) and find their avatar less attractive than male pairs.

**METHODS**

To answer our research questions, we conducted an experiment where we invited pairs of friends to come to our lab, scanned their bodies to generate photorealistic avatars, and then had them play a search and rescue game with one of the two scanned avatars. Both members of each pair played as the same avatar (one played as their own self, while the other played as their friend). We used a friend’s avatar instead of a generic avatar or stranger’s avatar to avoid confounding the effects of avatar familiarity and similarity. Half of the friend pairs in our study were assigned to play a social version of our game where they rescued people, and the other half played a nonsocial version of the game where they rescued gems.

**Generating Photorealistic 3D Avatars of Participants**

We created photorealistic avatars for participants using body scanning technology and post-processing software on the fly. While a 3D avatar can be constructed using traditional 3D techniques in combination with photographic body scans, our method utilizes only automated and near-automated processes. Such automated methods require no expertise and allow us to create participants’ avatars under tight time constraints during the study. For a more detailed discussion of avatar generation methods, see Feng et al. 2015 and 2017 [13, 14].

First, we capture and reconstruct the 3D body model from the participant. The 3D body model is captured using a photogrammetric cage [37]. The cage consists of 100 raspberry Pi computer and camera modules (Figure 2A). The capture process requires the participant to stand still inside the cage in an "A-pose" (see Figure 2B for an example) for several seconds. During capture, all 100 cameras take photos of the participant simultaneously from different directions. These photos are then used to reconstruct the 3D body model using photogrammetry software (Agisoft Photoscan). The capture and reconstruction take about 10 minutes for each participant.

The scanned character model also requires proper rigging structure in order to be animated in our game. We utilized the auto-rigging method proposed in [13] to produce a controllable, 3D avatar automatically from the scanned model. The method utilizes a 3D human model database to generate a morphable model to automatically fit a 3D human scan. Once the morphable model is constructed to fit the 3D human scan, we can transfer the location of skeletal bones, as well as the skinning deformation information, onto the scan as summarized in Figure 2B. The quality of the skinning and bone location is similar to that of the original rigging, which can be performed once by a professional 3D rigger. After the autorigging stage, the 3D skinned avatar can be incorporated into the game.

**The Game**

We designed a timed search and rescue game for our study. At the beginning of the game, players receive a text transmission from a radio operator telling them that they have 10 minutes to save either 24 people (in the social condition) or 24 gems (in the nonsocial condition) from a collapsing mine. The operator shows the player a map of the mine, indicating that 6 rooms marked in red contain the people or gems to be saved (4 per room) and that the player should navigate to these rooms. Once the player reaches one of the 6 rooms, they must pull a nearby lever to open the room’s door, rescuing all 4 people or gems inside. Examples of rescued people and gems are shown in Figure 2C. In the social condition, the people the player must rescue are photorealistic nonplayer characters built using the same technology as described in the previous section. Once the lever is pulled and the door to their room opens, they raise their hands briefly in celebration and then disappear. Likewise, in the nonsocial condition, the gems are displayed briefly after the door to their room opens and then disappear.

The game ends when the player pulls all 6 levers or when the 10 minute time limit is up - whichever comes first. At the end of the game, the player is shown a screen indicating how well
they did in terms of the number of people or gems rescued. If
they manage to pull all 6 levers (rescuing all people or gems),
the word "VICTORY" is displayed at the top of the screen. If
the player fails to pull all 6 levers, they see the words "GAME
OVER" instead of "VICTORY."

A screenshot of the game interface is shown in Figure 1. The
player holds down the W key to move forward, and can hold
down the A and D keys to turn left or right, respectively, while
they are standing still or moving forward. Pressing the M key
displays the map. When players are close to a lever, they can
pull it by pressing the spacebar to open a nearby blast door
and rescue the people or gems trapped in the room behind it.
The player sees a front view of their avatar in the lower left
corner of the screen throughout the game.

During early pilot tests, we showed the player the map at all
times during the game; they did not need to press a key to
display it. However, this made the task too easy since most
players were able to pull all the levers well within the 10
minute time limit. Thus, we decided to change the way we
implemented the map in three main ways to make the game
harder and therefore avoid ceiling effects when analyzing
player performance. First, the map is only displayed when
the player holds down the M key. Second, the player cannot
move and display the map at the same time. Third, the map is
a static image and thus gives no indication of where the player
is on the map at any time. Another round of pilot testing
demonstrated that with these changes, the game’s difficulty
level was hard enough to avoid ceiling effects, but not so hard
that players could not find and pull any levers.

In order to make sure participants had a chance to familiarize
themselves with the controls, we had them play a short tutorial
before the main game. The tutorial explained the controls
and then gave the player free rein to explore a large room
where they could try out each of the controls: moving, turning,
pulling a lever, and displaying the map.

**Study Procedure**

We recruited same gender friend pairs for our study via a
Craigslist ad. Participants were required to be between the
ages of 18-75 to avoid vision issues, motion sickness, and low
technology readiness, which are more prevalent over age 75.

Upon arrival at the testing site, each participant read and signed
an informed consent form, which allowed them to choose how
their data would be shared, if at all, after the study. Participants then filled out
a pre-game survey, which included questions about gender,
previous video game experience, and friendship closeness.

We asked participants about previous video game experience
because more familiarity with similar types of video games
might improve their performance in our game. To gain a com-
prehensive understanding of each participant’s prior gaming
experience, we asked them whether they had ever played video
games, and if so, whether they currently played, how often
they played, for how long they had been playing, how good
they thought they were at video games, and what specific
genres were their favorite.

We asked each member of a pair to rate the closeness of their
friendship with each other because different levels of friend-
ship closeness could affect how participants responded to play-
ing with their friend’s avatar and might therefore moderate the
effect of avatar appearance on performance or subjective expe-
rience. To measure friendship closeness, we used a modified
18 question version of Polimeni et al.’s Friendship Closeness
Inventory [32]. Since several questions in this inventory asked
about specific hobbies or activities people participated in with
their friends, we omitted these questions in order to avoid
inflating the scores of people who simply happen to have a lot
of hobbies or enjoy a lot of different types of activities.

After completing the pre-game survey, the experimenter
scanned participants one at a time using the method described
in the previous section to construct a 3D avatar of one person
in the pair to be used in the game. While both participants
in each pair were scanned, each ended up playing the game
using the same avatar. One person played as themselves, while
the other person played as their friend. The experimenter ran-
domly assigned each person to play as either themselves or
their friend by directing them to sit down at whichever of two
identical computers they liked. The computers differed only in
that one was labeled "A" and the other "B." Whoever sat at the
computer marked "A" played as themselves, and whoever sat
at "B" played as their friend. Participants did not know which
avatar their friend was using. We randomly assigned each pair
to play either the social version of the game (rescuing people)
or the nonsocial version (rescuing gems).

Once scanning was complete, the experimenter instructed the
pair to have a seat in a corner of the room and wait while
their 3D models were processed and put into the game. The
experimenter provided magazines for participants to read and
allowed them to use their phones while they waited, but asked
them not to talk to each other from this point on until the end
of the study to avoid the possibility of participants discussing
how they rated each other on the friendship closeness survey
and to avoid introducing unnecessary noise into the data via
varying conversation topics. Participants did not know that
only one of their 3D models would be used for the game.
Processing the 3D model, rigging it for animation, and placing
it into the game took about 30 minutes.

As soon as the processing and game setup phase was com-
plete, the experimenter set up screen capture software and
directed participants to sit back down at the computers, put
on headphones, and click "Play" to start the game. In order
to prevent participants from being distracted while playing
the game, they sat with their backs to each other and were
separated by a folding wooden screen. As participants played
through the game’s tutorial to familiarize themselves with the
controls, the experimenter observed them to make sure they
understood all the controls and helped guide them through the
tutorial if necessary. Participants were given as much time as
they needed to try out all the controls in the tutorial.

Once they felt comfortable with the controls, participants
clicked a button to proceed to the main game’s start screen.
The experimenter then explained to participants that they
would be alone in the room while playing the game; the ex-

After they finished the game, participants completed a post-game survey asking about their subjective experiences. A complete list of all dependent measures used in this study are shown in Table 1. During gameplay, we recorded three in-game measures: number of levers pulled (our main measure of performance), time spent moving, and time spent reading the map. We measured time spent moving and looking at the map to help understand player strategies while playing the game. Differences in strategy could help explain any differences in performance we might observe.

A complete list of all dependent measures used in this study is shown in Table 1. During gameplay, we recorded three in-game measures: number of levers pulled (our main measure of performance), time spent moving, and time spent reading the map. We measured time spent moving and looking at the map to help understand player strategies while playing the game. Differences in strategy could help explain any differences in performance we might observe.

After they finished the game, participants completed a post-game survey while the experimenter waited outside the room. The post-game survey asked participants about their subjective experience with the game and their avatar - specifically, their levels of engagement, enjoyment, frustration, emotional response, avatar attractiveness, and avatar identification. Emotional response questions asked participants about the extent to which they cared what happened to the people or gems they were rescuing and were adapted from a measure used in previous work [26]. Avatar identification was used as a manipulation check for avatar appearance. Once both participants were finished with the survey, they rang the bell. The experimenter returned to the room, paid each participant, and explained that both of them had been playing as the person sitting at computer “A.” The study took each pair about an hour to complete. Each pair received $60 for completing the study ($30 for each participant).

<table>
<thead>
<tr>
<th>Measure</th>
<th># of Subscales</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levers pulled</td>
<td>1</td>
<td>1-6</td>
</tr>
<tr>
<td>Time spent moving</td>
<td>1</td>
<td>Percentage</td>
</tr>
<tr>
<td>Time spent reading map</td>
<td>1</td>
<td>Percentage</td>
</tr>
<tr>
<td>Enjoyment [36]*</td>
<td>7</td>
<td>7-pt Likert</td>
</tr>
<tr>
<td>Emotional Response [26]</td>
<td>3</td>
<td>5-pt Likert</td>
</tr>
<tr>
<td>Frustration</td>
<td>1</td>
<td>7-pt Likert</td>
</tr>
<tr>
<td>Avatar Attractiveness [19]</td>
<td>2</td>
<td>7-pt Likert</td>
</tr>
<tr>
<td>Avatar Identification [39]</td>
<td>8</td>
<td>7-pt Likert</td>
</tr>
</tbody>
</table>

Table 1. List of all dependent measures analyzed. All subjective experience measures except frustration included multiple subscales. Scores on each subscale were averaged to calculate the overall score for each measure. *We used the interest/enjoyment subset of scales from the Intrinsic Motivation Inventory to measure enjoyment.

RESULTS
In total, we recruited 38 pairs of same gender friends for the study (76 people). We excluded a total of 5 pairs from our analysis: 2 pairs due to technical difficulties, 1 pair due to failed attention checks in the pre-game survey, 1 pair including someone who had participated in our game’s pilot testing, and 1 pair including a participant who skipped the tutorial and therefore did not understand how to play the game. Excluding these pairs from our analysis left us with 33 pairs (66 participants): 17 pairs in the social condition and 16 pairs in the nonsocial condition. There were a total of 17 female pairs and 16 male pairs, with more male pairs in the social condition (10 versus 7) and more female pairs in the nonsocial condition (10 versus 6). All 66 participants reported having played games before participating in the study, and 46 participants (70%) reported that they played games currently.

Since we used friend pairs for this study, we analyzed our results at the pair level rather than by individual participant, treating each pair as one subject and making avatar appearance (self versus friend) a within-subjects variable. This approach is commonly used in social psychology studies involving friend pairs, since two friends can be considered more similar to each other than two strangers [41]. With our sample size of 33 pairs, the probabilities of detecting small, medium, and large effects were 0.19, 0.75, and 0.99, respectively.

Our primary analysis consisted of two sets of mixed effects ANOVAs. To answer H1, H2, and H3, we performed four 2x2 ANOVAs with avatar appearance (self versus friend) and social condition (rescuing people versus rescuing gems) as fixed effects and friend pair as the random effect for each of the following dependent measures: levers pulled, engagement, enjoyment, frustration, and avatar identification. To answer H4, we performed four 2x2x2 ANOVAs with avatar appearance, social condition, and gender as fixed effects and friend pair as the random effect for each of the following dependent measures: engagement, enjoyment, frustration, and avatar attractiveness. We also performed ANOVAs with avatar identification, friendship closeness, and emotional response as the dependent variables as manipulation checks for avatar appearance.

Manipulation Checks
Participants reported identifying more with their own avatar ($\mu = 4.81$) than their friend’s ($\mu = 2.65, d = 1.37, F(1,31) = 30.86, p < 0.001$), indicating that the avatar appearance manipulation was successful in promoting participants to identify more with their own avatar than their friend’s. To confirm that participants had indeed brought a friend to the study and not just a random acquaintance, we looked at each pair’s friendship closeness ratings. Since our measure of friendship closeness had 18 subscales, we first averaged each participant’s ratings across all subscales. An inspection of the averaged scores revealed that each friend pair tended to rate each other’s friendship similarly (median score difference = 0.88, max difference = 2.11), we averaged each pair’s ratings together to get one score. Friendship scores were skewed high (median = 5.75), confirming that most pairs were close friends.

No Avatar Effects - But No Gender Effects, Either
Only 3 participants were able to pull all 6 levers within the time limit. Scores were approximately normally distributed, with a mean and median of 3 levers pulled, indicating that there were no floor or ceiling effects (the task was neither too hard nor too easy). Contrary to H1, there was no main effect of avatar appearance on levers pulled ($F(1,29) = 2.20, p = 0.15$). Neither was there an effect of avatar appearance.
on reported levels of engagement (all \( p > 0.50 \)), enjoyment (all \( p > 0.18 \)), or frustration (all \( p > 0.28 \)), contrary to our expectations in \( \text{H2}. \)

Contrary to \( \text{H3} \), adding a social element to the game did not moderate the effect of avatar appearance on performance or on any measure of subjective experience. There was no interaction between avatar appearance and levers pulled (all \( p > 0.12 \)), avatar appearance and enjoyment (all \( p > 0.81 \)), avatar appearance and engagement (all \( p > 0.38 \)), or avatar appearance and frustration (all \( p > 0.12 \)). Our results did not support \( \text{H4} \) either. There was no effect of gender on engagement \( (F(1, 29) = 2.38, p = 0.13) \), enjoyment \( (F(1, 29) = 0.57, p = 0.46) \), frustration \( (F(1, 29) = 0.20, p = 0.66) \), or avatar attractiveness \( (F(1, 29) = 1.77, p = 0.19) \).

Although subjective experience (engagement, enjoyment, emotional response, and frustration) did not differ between experimental conditions, subjective experience overall was positive. Emotional response and enjoyment were moderately high (\( \mu = 3.73, \text{median} = 4 \) and \( \mu = 5.56, \text{median} = 6 \), respectively), engagement was slightly below average (\( \mu = 2.79, \text{median} = 2.90 \)), and frustration was moderately high (\( \mu = 3.36, \text{median} = 3.5 \)). While the game provoked an emotional response in general and most pairs seem to have had an overall positive experience playing, the degree of emotional response and enjoyment they reported was not tied to their avatar’s appearance or whether they were rescuing people or gems.

**Mediation Analyses**

Given that the data did not support any of our hypotheses, we conducted a series of follow-up analyses to determine the extent to which our results could be explained by mediating variables such as gender, prior gaming experience, and emotional response to the game.

**Gender Differences in Performance**

While we found no gender differences in subjective experience, we thought that gender differences might have masked the effects of avatar appearance on performance or other in-game behaviors. We therefore performed three 2x2x2 mixed effects ANOVAs with avatar appearance, social condition, and gender as the fixed effects for the following dependent measures: levers pulled, time spent moving, and time spent using the map. There was a main effect of gender on levers pulled \( (F(1, 29) = 4.83, p = 0.036) \) such that male pairs pulled more levers (\( \mu = 3.34 \)) than female pairs (\( \mu = 2.62, d = 0.53 \)), qualified by an interaction between gender and avatar appearance \( (F(1, 29) = 5.32, p = 0.028) \). Males playing as their friend pulled more levers (\( \mu = 3.81 \)) than females playing as their friend (\( \mu = 2.53, d = 0.92, t(30.70) = -2.63, p = 0.013) \), but when playing as themselves, levers pulled did not differ between males (\( \mu = 2.88 \)) and females (\( \mu = 2.71, t(28.07) = -0.37, p = 0.71 \)). No other effects were significant (all \( p > 0.06 \)).

There was also a main effect of gender on time spent moving \( (F(1, 29) = 6.03, p = 0.020) \) such that male pairs spent more time in motion (\( \mu = 68\% \)) than female pairs (\( \mu = 58\%, d = 0.74 \)), but time spent moving was not affected by avatar appearance \( (F(1, 29) = 0.36, p = 0.55) \), and there were no other main effects or interaction effects (all \( p > 0.10 \)). In addition, there were no significant main effects of avatar appearance \( (F(1, 29) = 2.72, p = 0.11) \) or gender \( (F(1, 29) = 1.37, p = 0.25) \) on time spent reading the map.

**Prior Game Experience Mediates Gender Differences**

To understand the reasons for men’s improved performance relative to women when using their friend’s avatar, we looked at prior gaming experience as a possible mediating variable since prior research has shown that men and women have different gaming habits [43]. We reran our performance and behavior analyses for gender, but this time controlling for each of our prior gaming experience measures using ANCOVAs. Since all 66 participants reported having played video games before, we omitted this variable from our analyses, leaving us with 5 prior gaming experience measures to analyze: whether participants currently played video games, how often they played, for how long they had been playing, how good they were at playing video games, and what genres they preferred.

Current gamers (\( n = 46 \)) pulled more levers (\( \mu = 3.28 \)) than people who did not currently play games (\( n = 20, \mu = 2.25, t = 0.89, F(1, 42.9) = 5.85, p = 0.020) \). In addition, the main effect of gender on levers pulled became non-significant when controlling for current gaming experience \( (F(1, 34.4) = 1.40, p = 0.25) \). However, controlling for current gameplay habits did not eliminate the interaction between gender and avatar appearance for levers pulled \( (F(1, 28.1) = 6.13, p = 0.020) \).

The remaining measures of prior game experience were answered only by those who reported playing games currently (\( n = 46 \)). There was no significant effect of how often participants reported playing games \( (F(1, 21.9) = 0.0019, p = 0.97) \), how long they had been playing games \( (F(1, 35.9) = 1.94, p = 0.17) \), or how good they thought they were at playing games \( (F(1, 23.3) = 0.0095, 0.92) \) on the number of levers pulled.

Lastly, we looked at participants’ game genre preferences as another potential mediating variable for men’s better performance when playing as their friend. On the post-game survey, the game genre preference question allowed participants to select as many items as they wanted from a list of 15 different game genres (see Figure 3). Since those who reported not currently playing games were not asked this question, we report results only for current gamers (\( n = 46 \)), 29 of whom were male and 17 of whom were female.

We performed latent class analysis (LCA) to see if game genre preferences tended to cluster based on gender. To test this gender-based hypothesis, we fixed the number of latent classes at 2. We allowed the LCA estimation algorithm to run until convergence, and estimated the model with different starting values 100 times to increase the likelihood of finding the global maximum. Our final LCA model converged with a maximum log-likelihood value of -291.15 (AIC = 644.29, BIC = 700.98, \( \chi^2 = 29329 \)). We assigned each participant to the class that the LCA model predicted for them in order to do further analysis.

As Figure 3 shows, members of Class 1 selected the Puzzle, Platformer, Music, and Family genres more often than members of Class 2, while members of Class 2 selected the Role-Playing, Adventure, Fighting, Action, and Sports genres more often than members of Class 1. As we expected,
We hypothesized that the presence of a social element - rescuing people - might make participants more emotionally invested in the social condition, but if participants were not more emotionally invested in the social condition, then this might explain why players’ experience when using a self-similar avatar was not affected by social condition. To investigate this, we performed a 2x2x2 mixed effects ANOVA with avatar appearance, social condition, and gender as the fixed effects and emotional response as the dependent variable. There was no effect of avatar appearance, social condition, or gender on emotional response (all $p > 0.13$), and no interaction effects (all $p > 0.26$), confirming that our social condition did not have the desired effect of making people care more about the search and rescue task.

Emotional Response to the Game

We hypothesized that the presence of a social element - rescuing people - might make participants more emotionally invested in our search and rescue game. But if participants were not more emotionally invested in the social condition, then this might explain why players’ experience when using a self-similar avatar was not affected by social condition. To investigate this, we performed a 2x2x2 mixed effects ANOVA with avatar appearance, social condition, and gender as the fixed effects and emotional response as the dependent variable. There was no effect of avatar appearance, social condition, or gender on emotional response (all $p > 0.13$), and no interaction effects (all $p > 0.26$), confirming that our social condition did not have the desired effect of making people care more about the search and rescue task.

DISCUSSION

Ultimately, none of our four hypotheses were supported in this work. Avatar appearance did not have any main effect on performance (contrary to H1) or subjective experience (contrary to H2). Furthermore, the social content of the game (rescuing people versus rescuing gems) did not moderate the effect of avatar appearance on performance or subjective experience (contrary to H3). Likewise, gender did not moderate the effect of avatar appearance on subjective experience (contrary to H4). However, gender and prior gaming experience affected performance independent of avatar appearance, indicating the importance of analyzing these variables as moderators in video game studies.

Replication of Previous Findings

Our work builds primarily on a study by Lucas et al. that analyzed players’ performance and subjective experience with a photorealistic self-similar avatar in a maze running game. Lucas et al. found that men enjoyed the game more when using their own avatar, but women enjoyed it less when using their own avatar, often commenting that they wished their avatar looked more attractive [26]. We did not observe any of these gender differences in our study. This is likely due to the higher fidelity of the avatars used in our experiment. Lucas et al. used a handheld body scanner for avatar scanning, which often produced poor results on participants’ hands and hair. Women may have been more affected by poor representations of their hands and hair than men. In addition, Lucas et al. used only male animation skeletons for rigging avatars, so the female avatars may have moved in a less natural way. In this light, the lack of gender differences we observed when participants used a self-similar avatar is a good thing. It may be that when avatars are sufficiently photorealistic and move in a natural
way, both women and men can enjoy playing as themselves without feeling too self-conscious about their bodies.

Overall, Lucas et al. found that self-similar avatars conferred subjective experience benefits (although no performance benefits) compared to photorealistic avatars of a stranger: participants reported enjoying the game more and feeling more connected and engaged with their own avatar. Like Lucas et al., we did not find an effect of avatar appearance on performance. Unlike Lucas et al., however, we did not find any differences in subjective experience between participants who used their own avatar and someone else’s. It is possible that the elements we added to our game to make it feel more game-like and immersive - more realistic environments, sound effects, music, and a story premise - may have made participants focus less on their avatar’s appearance than in Lucas et al.’s study and more on other aspects of the game, reducing the impact avatar appearance might have had on their subjective experience.

In the following sections, we discuss some additional reasons why we may not have seen an effect of avatar appearance on performance or subjective experience in this study.

**Game Was Fun, But Lacked Social Interaction**

Given that players reported high levels of enjoyment and emotional response overall, it is likely that the parts of the game that were fun and emotionally engaging for the participants were not the parts we manipulated in our experiment. Lots of elements contribute to an entertaining, positive gaming experience: story, music, visuals, sound effects, and more. And we made sure to incorporate these features into our game in order to make it realistic and immersive enough that our experimental manipulations might make a difference in players’ performance and subjective experience. While it seems we succeeded in making the game enjoyable, our manipulations of avatar appearance and social condition may not have been a good match for the search and rescue game we designed.

First of all, the "social" version of our game may have been only superficially social. Most of the games that allow players to use a self-similar avatar are story-driven role-playing games such as the popular *Elder Scrolls* [7], *Mass Effect* [8], and *Fallout* [6] series. In these kinds of games, the player is interacting with other characters or players frequently, and their social interactions can have direct consequences for gameplay. Nearly every character interaction in these games allows players to decide whether to civilly persuade and reason with another character or attack them, to abide by the laws of the land or break them, and through this chain of interaction decisions, the player develops a reputation with other characters in the game, which affects the extent to which they will be willing to either help or hurt the player.

Games that do not allow for this kind of social interaction, like our search and rescue game, may not benefit from customized avatars since the player’s social reputation is not at stake. Because we were attempting to keep all other aspects of the game consistent between the social and nonsocial conditions, the player’s interaction was limited to seeing the rescued people raise their arms in celebration for a couple of seconds and then disappear, just as each rescued gem was revealed for a few seconds before disappearing. In addition, several participants told us after the study was over that they paid attention to their avatar’s appearance for the first minute or so of the game, but then ignored it as they became focused on the search and rescue task. The emphasis our game placed on collecting items within a time limit rather than on social interaction likely prevented avatar appearance from having an effect on participants’ performance and subjective experience in both the social and nonsocial conditions.

A sensible direction for future research would be to test the effect of self-similar avatars in a game where players are required to interact with non-player characters or other human players directly. Self-similar avatars might have even more of an effect if players’ choices about how to interact with other characters have real in-game consequences, as in role-playing games like *Elder Scrolls* and *Mass Effect*, where players’ social reputation (and criminal record) is at stake.

**The Avatar as Actual Self Versus Ideal Self**

Another potential reason why we found no effect of avatar appearance may have been our use of photorealistic rather than player-customized avatars. Avatar customization is the main method by which players can create self-similar avatars in commercial games like *Elder Scrolls* and *Mass Effect*, but players often do not choose to create an avatar that looks just like them. Instead, they create an "ideal" version of themselves - how they would like to be rather than how they actually are [20, 23]. Prior research suggests that gaming experiences are most motivating and emotional when players’ in-game perception of themselves aligns with their ideal self rather than simply their actual self [20, 22, 38, 33, 9]. As Trepte and Reinecke put it, "The player’s identification with the avatar seems to be crucial for experiencing entertainment, but is not necessarily tied to similarity" [38].

We used self-similar photorealistic avatars in our study instead of allowing participants to make an "ideal" version of themselves because we felt that the "actual self" avatars studied in this previous work may have been less effective than "ideal self" avatars simply because they were low fidelity representations. However, participants with a high degree of discrepancy between their ideal and actual selves may still prefer to use an avatar representing their ideal self over a highly photorealistic, accurate representation. Adjustments to self-similar avatars’ appearance can also lead to significant behavior changes, a phenomenon termed "The Proteus Effect" [44]. For instance, participants in a virtual discussion group who were given black-cloaked avatars developed more aggressive attitudes than those given white-cloaked avatars [31]. In addition, increasing an avatar’s height can increase people’s confidence in a negotiation task, while increasing an avatar’s attractiveness promotes more interpersonal intimacy in a virtual conversation with another human [44]. Such effects can be moderated by self-relevance, the extent to which a user finds their avatar relevant to their own sense of self [35].

Future work could investigate the interplay between the ideal self, self-relevance, and the Proteus Effect by allowing participants to tweak their photorealistic self-similar avatars’ appearance. The same animation tools we used to create the
Avatars can also adjust their height and weight [13]. Future work could also develop new tools to enable a photorealistic avatar’s clothing and hairstyle to be modified and compare players’ perceptions of customized (ideal self) versus non-customized (actual self) photorealistic avatars.

Virtual Reality Versus Traditional Video Games
Many of the studies that have found an effect of self-similar avatars on performance or experience studied these avatars in the context of virtual reality simulations [16, 15, 2], but our study used a traditional video game without virtual reality. Immersive virtual reality may increase players’ sense of presence by more closely mimicking the real world and the ways people interact with it [10, 12]. More specifically, virtual reality places the player inside a virtual environment rather than having them view it on a 2D screen, and the player likewise controls their avatar directly via body movements rather than by keyboard or button presses. This more embodied experience has been shown to heighten the sense of realism [1], and is especially important in collaborative virtual environments, where both verbal and nonverbal communication with other team members’ avatars is key [18, 5]. Avatar appearance may therefore make more of a difference in such contexts. To determine if virtual reality truly benefits more from self-similar avatars than traditional gaming media, future work could investigate the use of photorealistic self-similar avatars in a virtual reality context.

Are Self-Similar Avatars Worth the Effort?
The results of our study suggest that it is probably not worth investing time and money into developing highly photorealistic self-similar avatars for virtual experiences similar to our search and rescue game. However, the fact that we did not replicate the gender differences in experience seen in previous work [15, 2, 26] suggests that investing in the best technology available may be worth it to avoid negatively affecting women’s player experience in situations where a self-similar avatar is necessary. But our study is only the second to investigate photorealistic avatars in video games; more research is needed to understand how these avatars affect performance and experience in different kinds of games and virtual experiences. And while the technology used to capture avatars in our study is relatively inexpensive, players must be individually scanned, which may not be feasible for commercial games until more portable 3D scanning technology such as the iPhone TrueDepth sensor [25] become more widely distributed.

Our overall recommendation to designers of video games and other virtual experiences is to exercise caution when deciding whether to implement high fidelity self-similar avatars, especially for games like ours that have no social interaction component. We strongly urge designers not to think of avatar customization as a necessary feature of a fun, engaging virtual experience, but rather to consider what kind of experience they want to design, and whether avatar customization ties in with the purpose and structure of that experience.

Limitations
The primary limitation of this work is that we only studied the effects of self-similar avatars in one ten minute game. Longers
REFERENCES
http://elderscrolls.bethesda.net/en/skyrim
http://www.masseffect.com/


