Opleiding Informatica

AI and Avatars: The Effects of
Gender Appearance, Collaboration and Disclosure on Motivation

Boele Visser

Supervisors:
A.H. Zohrehvand & P.W.H. van der Putten

BACHELOR THESIS

Leiden Institute of Advanced Computer Science (LIACS)
www.liacs.leidenuniv.nl

8/8/2023

ABSTRACT

Does perceived AI gender impact users' motivation? Increasingly, interactions with an AI agent are taking place in the virtual world where AI embodies an Avatar, which could be represented with a specific gender. Past research has documented that the perceived gender of an AI influences people's interaction with the AI. In this research, we draw from studies on AI and avatars to theorise the effect of AI gender on users' motivation. To answer the question, we conducted an experiment inside a game with over 72,500 participants. Participants solved consecutive search problems as a game, either alone or together with an AI. Both the AI appearance and AI disclosure differed between groups of participants. We measured how likely they are to quit the game midway through the experiment and how intensely they played. Results showed that presence of another avatar decreases the chance of quitting the experiment by 3.3 percentage point (pp) and makes users play less intensely by 13.0 seconds per level compared to playing alone, while disclosure of the other avatar as AI increases the chance of quitting by 3.7pp and makes users play more intensely by 1.5 seconds per level compared to playing together without disclosure. We did not find significant effects that playing together with a disclosed AI has an effect on the chance of quitting or intensity of play. We find that the AI appearing masculine makes participants play 2.2 seconds per level less intense. This effect is moderated by the gender difference between the user's avatar and the AI avatar. Further research could replicate these results in non-gaming virtual worlds and investigate the causes of these effects.

TABLE OF CONTENT

INTRODUCTION

The use of Artificial Intelligence (AI) is increasing in society, with many companies integrating AI tools, such as ChatGPT, into their products. This raises the need for research on how the results of AI tools are communicated to users, as the way in which the results are communicated can affect how they are used (Adam et al., 2021). In addition, more and more communications with others happen in virtual worlds, where users are often represented by avatars. Since AI does not have an inherent gender, the gender of the avatar representing it could have an impact on human-AI collaboration motivation & performance. In this thesis, we will look into how the collaboration with AI and gender appearances of (AI) avatars affect the motivation of users inside the virtual worlds.

Research has shown that gender affects motivation, with men and women having different motivational drivers, both in the real and in the virtual world (Meece et al., 2006; Stefko et al., 2017; Veltri et al., 2014). Besides their own gender, the gender of the other in communication has an influence on motivation (Conti et al., 2001). These gender differences in motivation arise from the different social and cultural expectations, experiences and roles that are associated with being male or female (Meece et al., 2006). Gender roles and stereotypes can shape an individual's beliefs, values and behaviours, including their motivation (Meece et al., 2006). These differences between the composition of groups are not only present in the real world, but also in virtual worlds (Lee et al., 2014). The gender which an AI takes also has an influence on people's motivation in textual interactions (McDonell & Baxter, 2019). This suggests that the beliefs about gender are not only formed about other humans but also about AI.

Building on this research, the influence of AI gender appearances in virtual worlds on user motivation presents an interesting opportunity for research. The findings of this study could be applied to design more effective AI systems, such as educational AI avatars in virtual worlds. The findings could also be used to shed insights on human behaviour when the goal is to change the behaviour, for example when making gaming more inclusive. This study aims to contribute to this field by combining AI and virtual worlds to describe how the gender of an AI could be used to motivate users of virtual worlds.

To research this, we performed an experiment in the game Find The Button. Find The Button is a game on Roblox (a virtual world), in which players have to perform a simple task, which can be done together with others. Inside this game, over 72,500 participants played 15 levels of a simple search task. The participants either played alone or together with an AI avatar, from which the gender appearance was randomly chosen. We measured how intense participants played the experiment by dividing the time before leaving the experiment divided by the amount of levels completed. We used this measurement to compare the difference in motivation between different treatment groups, to see what effect the AI gender appearance has on motivation in this particular context.

THEORETICAL BACKGROUND

Motivation is defined as "the degree and type of effort that an individual exhibits in a behavioural situation" (Perry & Porter, 1982). There are two types of effort: intrinsic motivation and extrinsic motivation (Ryan & Deci, 2000). Intrinsic motivation occurs when an individual finds something interesting or enjoyable, while extrinsic motivation occurs when an individual is motivated by the outcome of a task. While intrinsic motivation is more about the process, extrinsic motivation is more about the outcome.

There are various ways to alter motivation, as suggested by Ryan & Deci (2000). To change intrinsic motivation, the process can be changed, such as by making the task more challenging. To change extrinsic motivation, the outcome can be changed, such as by offering a higher reward for completing the task.

Effective ways to alter motivation vary among different age groups. According to a literature review by Sekhar et al. (2013), motivation drivers vary among different age groups. For instance, training, job security, and recognition are key motivators for individuals early in their careers. However, as individuals progress in their careers, training becomes less effective in motivating them. Stead (2009) found that motivation is achieved less through progression and personal growth opportunities and more through autonomy when individuals become older.

Additionally, some research claims that gender also plays a role in motivation. For example, Stead (2009) found that workspace flexibility is a stronger motivator for older females than job security or recognition. Stefko et al. (2017) discovered that women are significantly more demotivated when there is no job security and when they perceive their manager as behaving unfairly.

Not only could an individual's gender influence their motivation, but the gender of others they interact with may also have an impact. In a study among children, Conti et al.

(2001) found that boys had a higher level of intrinsic motivation when the task of creating art was competitive, while girls had a lower level of intrinsic motivation in the competitive task. Additionally, Conti et al. found that both intrinsic and extrinsic motivation were higher when boys were in a group with other boys, as opposed to groups consisting of both boys and girls. These findings suggest that the gender composition of groups can influence motivation.

Social interactions with AI

Social interactions are not only possible with another human, but also with artificial intelligence (AI). An AI is perceived to be a social actor by humans, and the characteristics of the AI have an influence on how people interact (Kim et al., 2019). People trust AI based on their ability to think, plan and act. (Puranam & Vanneste, 2023).

When introducing an AI in an interaction, the framing used to introduce the AI influences how the aid of an AI is accepted. Mols (2020) asked participants to accept or reject advice given by an AI. The results showed that knowing how the algorithm was used previously, self-confidence about the aid and information about the creators of the AI does have a direct effect on the acceptance, while information about social usage of the AI has a negative effect on acceptance. Participants accepted the aid more when the perceived quality of the AI was higher, while the acceptance was lower when the task was perceived more complex.

Within the interaction itself, the form factor of the AI changes how participants conformed with the AI. Schreuter et al. (2021) conducted a study to investigate the impact of the form factor of an AI on the conformity of participants to a virtual assistant. Participants were asked to answer 20 general knowledge questions. After answering a question, participants received assistance from the AI in the form of either a voice clip or a written hint. Participants were then asked for their final answer on the question. The results showed that participants were more likely to conform to the AI when it gave voice assistance as opposed to textual assistance. The conformance was even greater when the AI had a human voice instead of a robotic voice.

In these interactions, the perceived gender of the AI influences the compliance with the AI. McDonnell & Baxter (2019) conducted a study in which they found that the gender cues of a chatbot, an AI that provides textual responses to users, influenced user satisfaction and compliance. This effect was even greater when the interaction involved gender-stereotypical subjects, such as mechanics.

Therefore, research has shown that there are various factors that influence how people interact with an AI. In the next section, we will look into another way of social interactions: inside virtual worlds.

Interactions in a virtual world

Virtual worlds are expanding beyond gaming communities to offer courses, stores, and financial services, allowing more aspects of everyday life to take place online (Verhagen et al., 2011). Verhagen et al. (2011) found that both intrinsic and extrinsic motivation, measured by factors such as escapism, visual attractiveness, perceived usefulness, and entertainment value, impact users' attitude towards using virtual worlds.

In the virtual world, users are typically represented by avatars, which are graphical representations of themselves created using computer technology (Holzwarth et al., 2006). Users can design their avatars to represent their actual appearance or their desired identity in the virtual world (Bélisle & Bodur, 2010).

The appearance of an avatar can influence the behaviour and motivation of the user, especially when the user can identify with their avatar. For example, Li & Lwin (2016) found that participants who felt their avatar represented their own body had higher enjoyment while playing exergames, which increased their motivation to exercise in the future. Similarly, Chen et al. (2022) found that players of the game Glory of the King had higher motivation in a competition when their avatar identity was similar to their desired identity. Baylor (2011) also claims that choosing an avatar that a user can identify with in terms of gender, ethnicity, age, and perceived competency can help achieve better learning results when using an avatar for educational purposes.

The appearance of an avatar can also influence how a user is perceived by others in virtual worlds. Bélisle & Bodur (2010) showed that the appearance of a user's avatar in terms of body and clothing influenced how the user was perceived by other avatars. Additionally, Van Der Land et al. (2015) found that users felt more connected to others in a group when they could identify with their own avatar and when the other avatars in the group looked similar to their own avatar.

In addition to the appearance of an avatar, gender can also influence how a user is perceived and behaves in virtual worlds. Veltri et al. (2014) found that men tend to use avatars of the opposite sex more often than women, and achievement and control are more important for men than it is for women. Women on the other hand are more likely to prioritise the visual appearance of their avatars and to seek friendships in online games.

Relationship between users' own avatar gender and motivation

Individuals tend to behave and think in ways that align with the stereotypes of their virtual avatar (Yee & Bailenson, 2007). A previous study by Lee et al. (2014) has shown that this effect also applies to gender stereotypes. Participants were asked to perform 20 arithmetic tasks in a virtual environment, where they were represented by an avatar. This avatar did not necessarily have the same gender as themselves. The strongest performance and motivation was seen when a participant with a male avatar competed against two female avatars, regardless of the participants gender.

We have seen that gender plays a role in how people are motivated. This gender difference is appearing both in education (Meece et al., 2006) and in the workplace (Stefko et al., 2017). Gender differences in how people are motivated carry on in virtual worlds, where men and women are motivated by different techniques (Veltri et al., 2014). Furthermore, research has shown that there are also gender differences in motivation to engage with virtual worlds themselves. For example, Leonhardt and Overå (2011) found that male students use video games up to 5 times more than female students. Considering the gender differences in video game usage, we believe that men are more motivated to play video games.

Since research suggests that users tend to behave like the stereotypes of their virtual avatar and men are more motivated to play video games, we suggest the following hypothesis:

H1: Participants with more masculine looking avatars exhibit higher intensity of play.

Relationship between playing with an AI avatar and motivation

Users do not have to experience a virtual world alone - they can also experience a virtual world together with someone else. The option to cooperate with other players motivates players to use a virtual world: Featherstone and Habgood (2019) found a 217% increase in app usage for an educational video game when the app added a feature where you could asynchronously cooperate together. Synchronous cooperation also motivates players: Peng & Crouse (2013) found that players of video games are motivated to play an exergame again more when they were playing together (in cooperation) with another player in the same physical space than when they were playing alone. When the AI behaves completely like a player, there should be no difference between the interaction with a player and with an AI, thus the effects of cooperating with another human would carry over to cooperating with an AI. Therefore, we suggest the following:

H2a: Participants who play with an AI avatar (treatment group, regardless of disclosure) have a lower chance of quitting the game before completing the experiment than those who play alone (control group).

H2b: Participants who play with an AI avatar (treatment group, regardless of disclosure) have a higher intensity of playing than those who play alone (control group).

Relationship between AI avatar gender and motivation

Conti et al. (2001) found that the gender of others who someone interacts with influences your motivation. Some research claims that men are motivated by competition and they believe they outperform others (Conti et al., 2001; Thaler, 2021), so men will want to keep ahead of others and are motivated to do so quickly. Other research claims that women behave more prosocially (Soutschek et al., 2017), so we believe women want to support others and continue to keep interacting longer, which will make them play longer.

Since an AI is perceived to be a social actor by humans (Kim et al., 2019), the characteristics of the AI determine how humans are interacting with the AI (Kim et al., 2019) and the perceived gender of an AI having an influence on people's motivation during textual interactions (McDonnell & Baxter (2019)), the gender effects will carry on when interacting with an ai. We therefore suggest the following:

H3: Participants playing with an AI avatar that appears more masculine have a higher intensity of playing.

Relationship between AI disclosure and motivation

When interacting with an AI, the user could know that the one they are interacting with is an AI or not (which implies it is a human). This AI disclosure has an effect on how users perceive an AI system (Tong et al., 2021; Keppeler, 2023). Tong et al. (2021) found that employees who get performance feedback created by an AI get a negative perception of this feedback when they know the feedback is created by an AI, which makes them less motivated by the feedback. Keppeler (2023) found that job candidates are less interested in an offer when their application is reviewed by an AI system and therefore are less motivated. This let us believe that AI disclosure causes a lower motivation of users, and therefore we suggest the following:

H4a: Chances of quitting the game before completing the experiment are positively associated with participants knowing that they are playing with an AI avatar, in comparison to playing without the AI being disclosed.

H4b: Intensity of play is negatively associated with participants knowing that they are playing with an AI avatar, in comparison to playing without the AI being disclosed.

The moderating role of avatar similarity

The motivation of a player is higher when players can identify with their own avatar (Lenggenhager et al., 2007; Chen et al., 2022). Players also prefer to create avatars which relate to their personal identity (Mazlan & Bakar, 2013). Players in virtual worlds are also more connected with others in a group when other avatars in the group look similar to their own avatar (Van der Land et al., 2015). We therefore suggest the following:

H5: The similarity between the AI avatar's gender appearance and the participant's avatar gender appearance moderates the effects of AI avatar gender appearance on motivation.

METHODS

In this section, we will discuss the methodology of the experiment we performed. We preregistered this experiment at the Open Science Foundation on 19-5-2023.[1]

Organisational setting

Roblox[2] is an online gaming platform that allows users to create, share and play their own virtual experiences. The experiences on the platform differ in genre, from roleplaying to racing and sports to puzzles; the only limit on the games is the imagination. Players can purchase virtual items using Robux, the virtual currency of Roblox. They can buy clothing for their avatar, access to private servers or power-ups or new features in games. With over 150 million monthly active users (MAU) and over 40 million different experiences, Roblox is one of the most popular online gaming platforms in the world.

Find The Button is a game on the platform Roblox[3]. Find The Button was created about 3 years ago and has grown ever since. The idea behind Find The Button is simple: players get placed in a level (a room), and their task is to find the button inside that room. Once they find the button, they click on it, and they progress to the next level. Currently, the game has over 900 different levels with themes ranging from mediaeval to wild-west and sci-fi to season-bound themes such as the holidays, with new levels and themes being added monthly. Players can experience the levels themselves, or together with friends.

The player base of the game is very diverse. Players come from around the world, with different age groups, language and devices. This makes it a good fit to research the impact of an avatars' gender on motivation on a very diverse player base.

---

[1] The preregistration is available at this link: https://osf.io/vf35a/?view_only=c827c9f421c34426bc52e557d6f7006a
[2] Roblox can be accessed via https://roblox.com
[3] Find The Button can be accessed via https://www.roblox.com/games/2665326799/Find-The-Button . Before you can access the game, you need to create a free account on Roblox.

Participants

Participants are players of the game Find The Button. Players could join the experiment by selecting a special mode called 'Science Mode'. Only the data of plays during the first two weeks are used, although it is possible to join the 'Science Mode' after the experiment' duration to be able to review the procedure. After participating in the experiment, players receive a cosmetic item and some in-game virtual currency.

Participants are able to participate in the experiment multiple times, but only their first participation will be taken into account for the analysis.

Due to privacy regulations of Roblox, it is not possible to review the age and gender of the participant itself. However, data about age is available on a generic level of the game, where we could see the distribution of players in age groups. It is most likely that the participants follow the same demographics as the game itself. Players from the game come from around the world, which makes the participants very diverse and eliminates the WEIRD society bias from the data (Henrich et al., 2010).

Procedure

A detailed view of the information shown and asked to participants are displayed in Appendix A.

When participants join the experiment, they get an information screen that they are going to participate in an experiment. One third of the participants are in the control group and will participate alone. The other two-third are in the subject group and will participate in an AI avatar. Within the subject group, one half was told they are playing with an AI (so one-third of all participants), while the other half (also one-third) was not told they are playing with an AI.

After the explanation, participants select their avatar. Participants were completely free to select the avatar they want; the question asked was ‘Select your avatar’. The avatars differ in gender appearances. The avatars were always displayed in the same order, from left to right. Participants play with their selected avatar during the experiment, and participants are asked to rate this avatar on a 5-point scale ranging from ‘very feminine’ to ‘very masculine’.

After they select the avatar, they play the ‘Science Mode’, which consists of 15 stages. In this mode, players cannot see other participants, except the AI avatar controlled by the computer when participants are in the subject group. This AI avatar is also called an NPC, a non-player avatar.

The appearance of the NPC is one of the 5 different avatars shown during the self-selection process. The appearance of the NPC is selected randomly out of the five preset avatars, regardless of the avatar chosen by the player. The NPC wanders around the room to preset positions and pretends to be searching for the button. When the participant finds the button, the NPC teleports to the next level within a few seconds. Therefore, the NPC is always playing the same level as the participant, which gives the participant the feeling that the NPC is playing with them.

The difficulty for all participants is the same; the level design is the same for all participants, as is the behaviour of the NPC. The only variables that are changed are playing alone vs. playing with an NPC, the disclosure of the NPC being an AI and the (gender) appearance of the NPC.

Measures

We proxy motivation based on how intensely players engage with the experiment and how likely it is that they complete all the stages. We measure *Intensity of Play* by calculating the time before leaving the experiment divided by the number of levels completed. A lower value of *Intensity of Play* therefore implies that the participant spends less time on a level on average, and therefore engages more intensely with the experiment. The motivation of the participants is measured by using a performance-based behavioural measure. Touré-Tillery & Fishbach (2014) explained that performance-based behavioural measures indicate both a higher outcome-focused motivation (extrinsic motivation) and process-focused motivation (intrinsic motivation).

In addition, we use a dummy variable to indicate if a participant has completed all 15 stages. The time before ending the participation is also measured - participation can end by either completing the 15 stages, or leaving earlier. The speed in which players complete the experiment therefore indicates if the avatar appearance influences either the intrinsic process-focussed motivation or extrinsic outcome-focussed motivation.

To get a more reliable measure for the gender of the avatars, we asked each participant to rate their avatar on a 5-points scale, with 1 being 'very feminine' and 5 being 'very masculine'. This gives an average score for each avatar between 1 and 5. To simplify the analysis, we calculated the average rating score of the avatar. Given that the average of our 1-5 scale is 3, we mapped the avatar to either feminine when the average score is below 3 or to masculine when the average score is above 3. Due to the expected size of our experiment, it is highly unlikely the average would be exactly 3. If this is the case however, the avatar would be treated as if it would have no gender and be excluded from the analysis about gender.

Analytical strategy

When participants stopped participating in the experiment (either by leaving earlier on or completing the experiment), the data was saved in a datastore in Roblox. To export the data to outside Roblox, we used the rblx-open-cloud module created by TreeBen77 (2023). This is a wrapper that handles all api requests to Roblox. On top of this module, we made our own code that lists over every key in the datastore and stores the data of the key in a csv file. To anonymise the dataset, the keys are stored in a separate csv file.

To test the hypotheses, we used an Ordinary Least Squares (OLS) and Probit model from the statsmodels package (Seabold & Perktold, 2010) in Python. The independent variable and dependent variables for each hypothesis are listed in Table 1, including the model used for each dependent variable. For H5 (the moderating effects of the gender differences), we created an multiple OLS model with independent variables AI avatar masculinity, the absolute difference between the AI avatar and participant avatar gender score (which results in either a 0 or 1 - similar or different - due to our simplification) and an interaction term which consists out of the AI avatar masculinity multiplied with the absolute difference.

IV → ↓ DV	Participant avatar masculinity	AI avatar presence	AI avatar masculinity	AI disclosure	Avatar gender differences
Chance of Completion (Probit)		H2a		H4a	
Intensity of Play (OLS)	H1	H2b	H3	H4b	H5

Table 1: Independent and dependent variables per hypothesis.

Ethics

Before joining the experiment, participants are informed about its details. The experiment is non-intrusive and participation is voluntary, with the option to leave at any time without giving a reason.

During the experiment, participants in a specific treatment group are not told whether they are interacting with another human or an AI, as this knowledge could alter their behaviour. However, they are informed of this fact directly after participation. The AI is programmed to never harm participants, but rather to encourage and help them to complete the task while remaining friendly at all times.

No personal data is gathered about the user. While the age of the individual participants is not known, all users of roblox (or their guardians if they are minors) have accepted the terms of use of Roblox Corporation (2023). This terms of use includes responsibility of the guardians for the activities of minor users on Roblox and within its experiences, including our experiment. Academic research is not forbidden according to the Roblox terms of use.

Participants' data is stored based on their player id, which cannot be traced to a real-life person. During analysis, player ids and data are separated, ensuring that it is never known which data belongs to which participant.

After participation, participants were informed about the nature of the experiment in a short debrief. In this debrief, participants were told in which group they participated (control group, AI without disclosure on beforehand or AI with disclosure on beforehand). Besides that, participants were notified that they received a reward and were prompted to continue playing different levels in the game.

The reward participants received consisted of 100 coins and a free pet, which was exclusively available for the experiment. Although the pet was exclusive to the event, similar pets were available for 300 coins inside the game. This makes the equivalent reward in coins approximately 400 coins. At the time of the experiment, it was possible to purchase 175 coins for 25 Robux (approximately €0.38), which makes the monetary value of the reward coins approximately €0.86 when you would purchase the coins. However, coins and other pets are also available when playing the game without making purchases, and only 0.5% of the players make purchases inside the game. This makes it very difficult to tell how the participant would value the rewards exactly, but it is likely that the perspective of a reward encouraged players to participate in our experiment.

RESULTS

During the 2-week time window, 126,013 participants started the experiment. From this group, 8488 participants (6.74%) completed all levels of the experiment.

In Figure 1 the progress of participants over the experiment is displayed. In Figure 2, the time spent in the experiment before either leaving or completing the experiment is displayed. The mean time spent in the experiment is 249.4 ($\sigma = 287.6$), while the median time spent is 142 seconds. The mean time including the standard deviation is plotted in the error bar displayed in Figure 3.

53,099 (42.14%) participants left the experiment during the first level. This implies that participants did not finish a level. Since they did not finish a level, participants never completed a task. This data is left out in our analysis, because there is no indication if participants did not understand the task well enough or that their motivation was affected by the treatment they were subjected to.

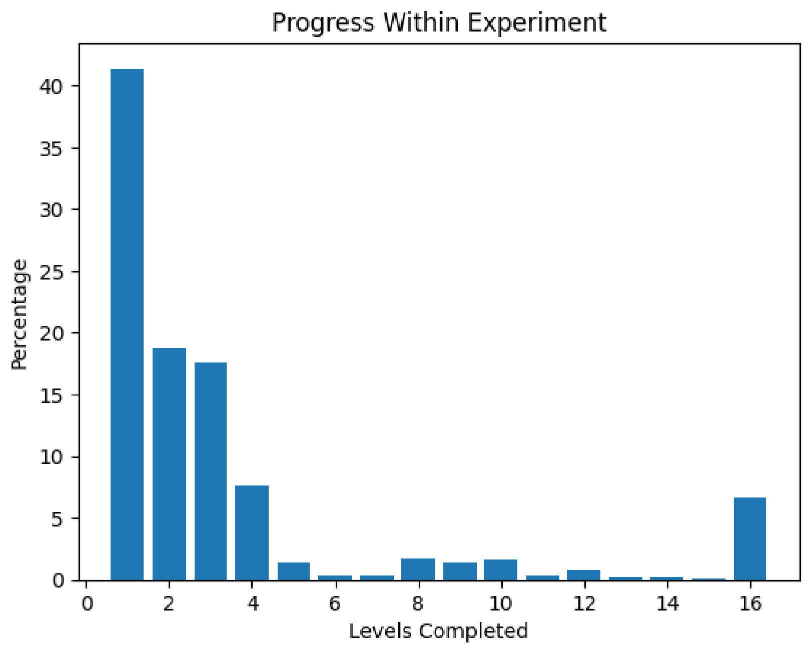

Figure 1: Progress of participants in the experiment.

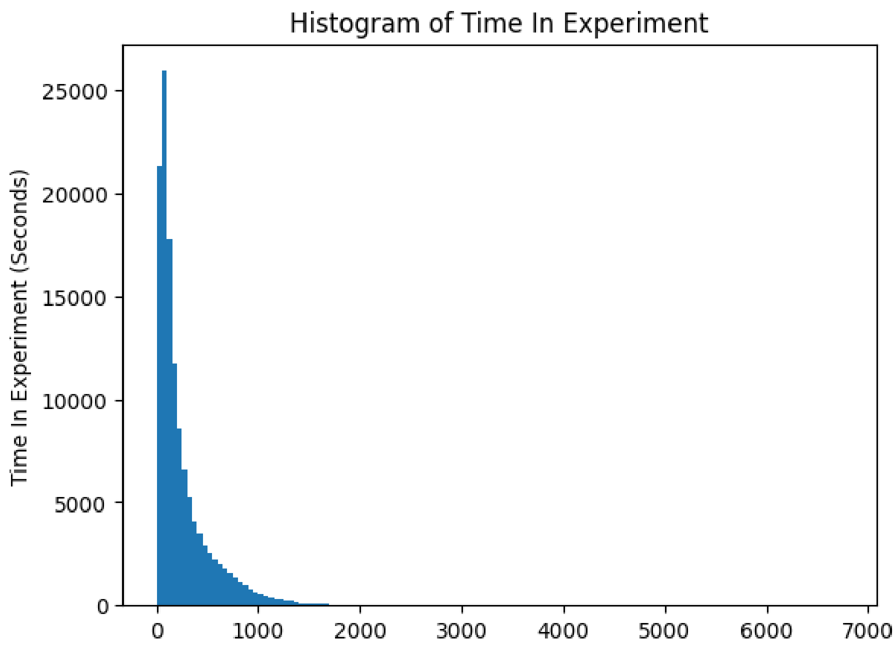

Figure 2: Distribution of the time in the experiment.

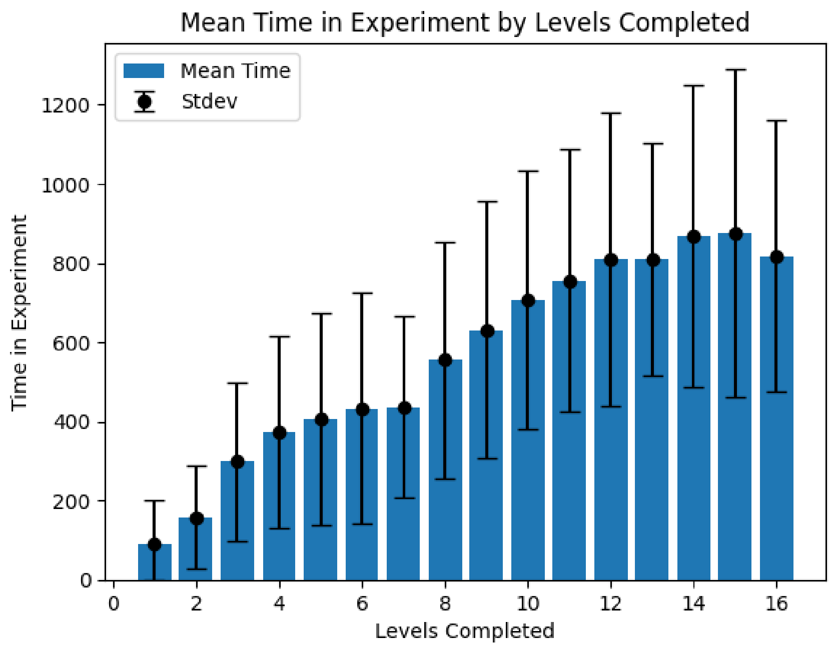

Figure 3: Mean time in experiment grouped by levels completed.

Based on previous observations within Find The Button, it is not possible to complete a stage within 5 seconds when you did not play the level before (based on loading times, time to search the button and the time to click on it). If a participant completed a stage within this

5-seconds timeframe, it is likely that the participant used some kind of cheating or did know the location of the button beforehand. Therefore, we left out the data of players from which the Intensity of Play (the average time spent per level) is below 5.0 seconds. This was the case for 11 (0.01%) participants. This makes the total sample for analysis consist of 72,903 participants, of which 8488 (11.64%) completed the experiment.

In Table 2, an analysis of the difference between participants who did and did not complete the experiment could be found. 11.6% of the participants completed all levels of the experiment, with the average participant completing 3.67 levels (when excluding participants who did not complete a single level).

	Completed	Not-completed
Number of entries	8480	64,423
Levels completed	$\mu = 16.0, \sigma = 0.0$	$\mu = 3.67, \sigma = 2.47$
Intensity of Play	$\mu = 51.18, \sigma = 21.35$	$\mu = 86.53, \sigma = 62.65$

Table 2: Analysis of variables between completed/not-completed experiments.

As mentioned in the methodology, there is no information available about the age of individual participants. However, if we assume the participants in the experiment follow the same demographics as the game, nearly 50 percent of the participants are 13 or younger, with an additional 17 percent being aged between 13 and 17. There are no other game-level demographics available.

Out of the participants, 23,079 (31.66%) participants were in the control group, 24,894 (34.15%) participants played together with an AI and did not know they were playing with an AI and 24.930 (34.20%) participants played together with an AI and knew they playing with an AI. The distribution of both the self-chosen avatar and the assigned AI avatar of the groups who played with an AI are listed in Table 3. Over 50% of the participants chose the first avatar as their own avatar.

AI avatar → ↓ Participant avatar	Avatar 1	Avatar 2	Avatar 3	Avatar 4	Avatar 5	
Avatar 1	7502	7293	7552	7322	7388	37,057 (50.83%)
Avatar 2	1078	1036	1114	1033	1025	5286 (7.25%)
Avatar 3	1762	1669	1634	1821	1726	8612 (11.81%)
Avatar 4	2596	2677	2699	2578	2622	13,172 (18.07%)
Avatar 5	1770	1753	1711	1776	1766	8776 (12.04%)

Table 3: Distribution over avatar types

To assess the gender of the avatars, we took the mean of all rankings given for that particular avatar. If the average avatar score was under 3, it was categorised feminine, while it was categorised masculine when the average avatar score was over 3. The categorization of all avatar types are listed in Table 4.

Avatar	Average score	Categorization
1	1.903 ($\sigma = 1.26$)	Feminine
2	2.026 ($\sigma = 1.17$)	Feminine
3	3.509 ($\sigma = 1.27$)	Masculine
4	3.721 ($\sigma = 1.42$)	Masculine
5	4.051 ($\sigma = 1.41$)	Masculine

Table 4: Categorization of avatars

Hypothesis Tests

We first tested H1 (*Participants with a more masculine looking avatar exhibit a higher intensity of play*) using linear regression. We used an ordinary least squares regression to estimate how the categorisation of the masculinity of participants' avatars affected their intensity of play. In Figure 4, the mean intensity of play grouped by the participant's chosen avatar is displayed. In Figure 5, the mean intensity of play grouped by the masculinity categorisation (masculine/feminine participant avatar) is displayed.

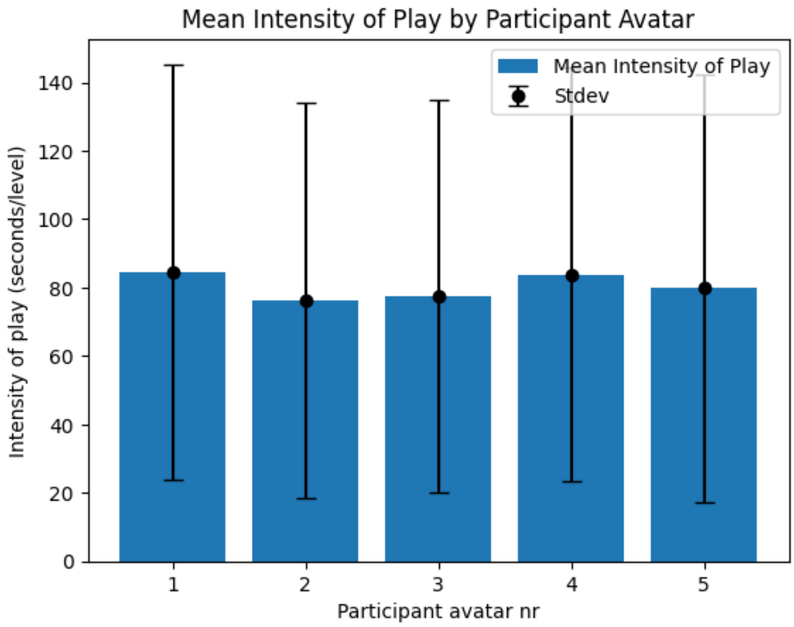

Figure 4: Mean intensity of play by participant avatar number.

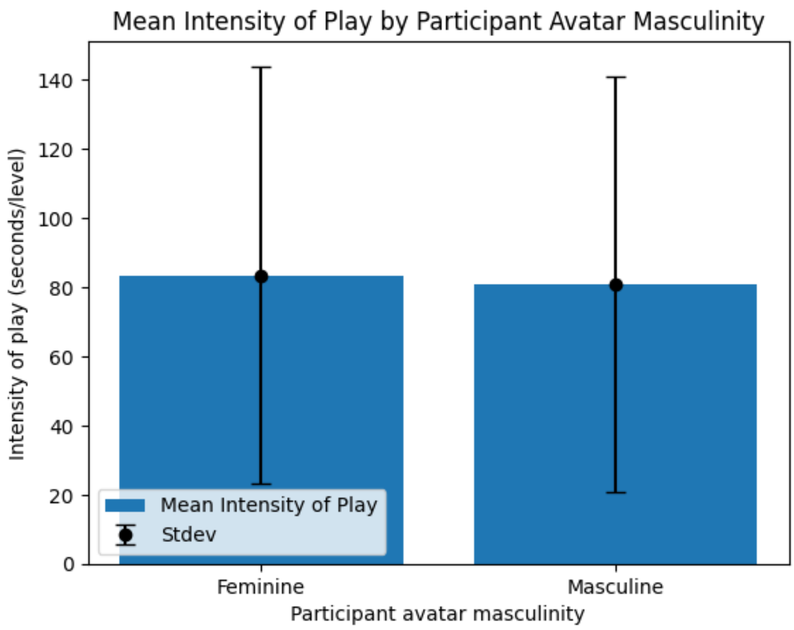

Figure 5: Mean intensity of play by participant avatar masculinity.

We found a significant effect (R^2 = 0.000, F(1, 72901) = 32.57, p < 0.001, 95% CI [-3.476, -1.699]) for intensity of play. Participants' *Intensity of Play* decreased 2.6 seconds per level when their avatar was masculine. This is equivalent to 0.04 standard deviations. This means that participants whose avatar appearance was more masculine play more intensely, and therefore we accept H1.

We then tested the hypotheses about playing with an AI (regardless of the disclosure) versus playing alone. To test H2a (*Participants who play with an AI avatar (treatment group) have a lower chance of quitting the game before completing the experiment than those who play alone (control group)*), a probit model was calculated to estimate the probability of completing the experiment (the amount of participants that completed all levels divided by the total amount of participants, given the treatment type) based on the presence of another avatar. In Figure 6, the rate of completion grouped by the treatment type is displayed.

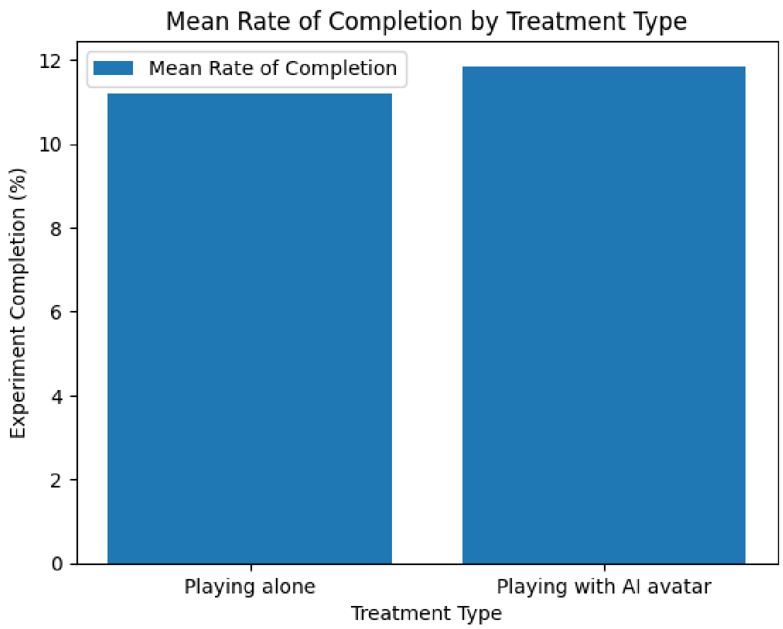

Figure 6: Rate of completion by treatment type.

We found a significant effect (Pseudo R^2 = 0.000, p = 0.013, 95% CI [0.007, 0.058]) that the presence of AI enlarges the chance of the participant completing the experiment. Participants' chance of completion increased 3.3 percentage points when participants played together with an AI avatar. Since a higher chance of completion implies a lower chance of quitting the experiment, we accept H2a.

We conducted an ordinary least squares regression analysis to estimate the relationship between the presence of an AI avatar (treatment group) and the intensity of play (*H2b: Participants who play with an AI avatar (treatment group) have a higher intensity of playing than those who play alone (control group)*). In Figure 7, the intensity of play grouped by the treatment type is displayed.

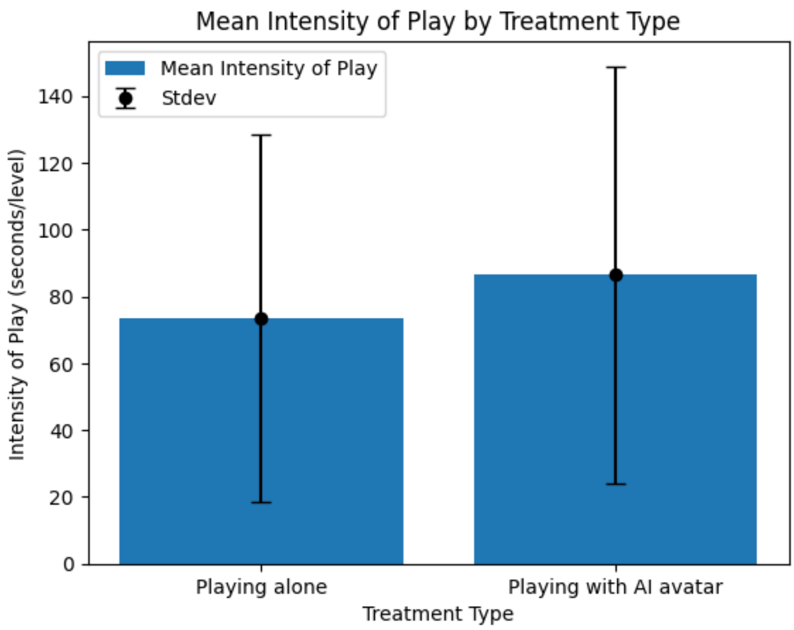

Figure 7: Intensity of play grouped by treatment type.

We found a significant effect (R^2 = 0.01, $F(1, 72901)$ = 734.4, $p < 0.001$, 95% CI [12.033, 13.909]) for intensity of play. Participants' *Intensity of Play* increased by 13.0 seconds per level when playing together. That is equivalent to 0.22 standard deviations. This means that participants who played together with another (AI) avatar played less intensely. This is the inverse of what we expected, and therefore we reject H2b.

To test the influence of the effect of the gender appearance of the AI avatar, we conducted an ordinary least squares regression analysis (H3: *Participants playing with an AI avatar that appears more masculine have a higher intensity of playing*). In Figure 8, the intensity of play grouped by the AI avatar masculinity is shown.

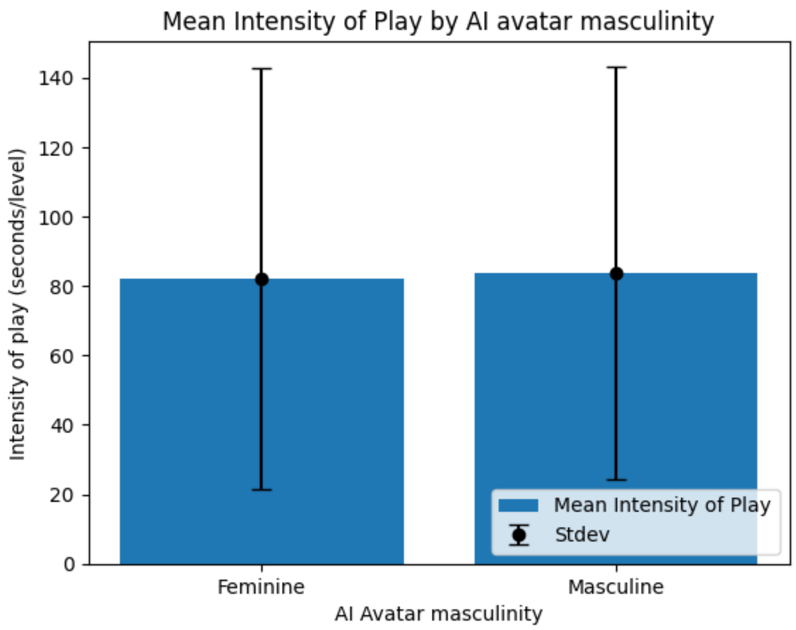

Figure 8: Intensity of Play by AI avatar masculinity.

We found a significant effect for intensity of play ($R^2 = 0.000$, $F(1, 72901) = 10.03$, $p = 0.002$, 95% CI [0.841, 3.570]). Participants' *Intensity of Play* increased by 2.2 seconds per level when playing together with a more masculine AI avatar. That is equivalent to 0.04 standard deviations. This means that participants who played together with an AI avatar that appeared more masculine played less intensely. This is the inverse of what we expected, and therefore we reject H3.

Fourth, we tested if participants knowing they are playing with an AI has an influence on motivation, in comparison to playing together with an AI while not knowing the avatar is an AI. To test H4a (*Chances of quitting the game before completing the experiment are positively associated with participants knowing that they are playing with an AI avatar*), a probit model was calculated to estimate the probability of quitting based on the disclosure of playing with an AI. In Figure 9, the mean rate of completion grouped by the treatment type is displayed.

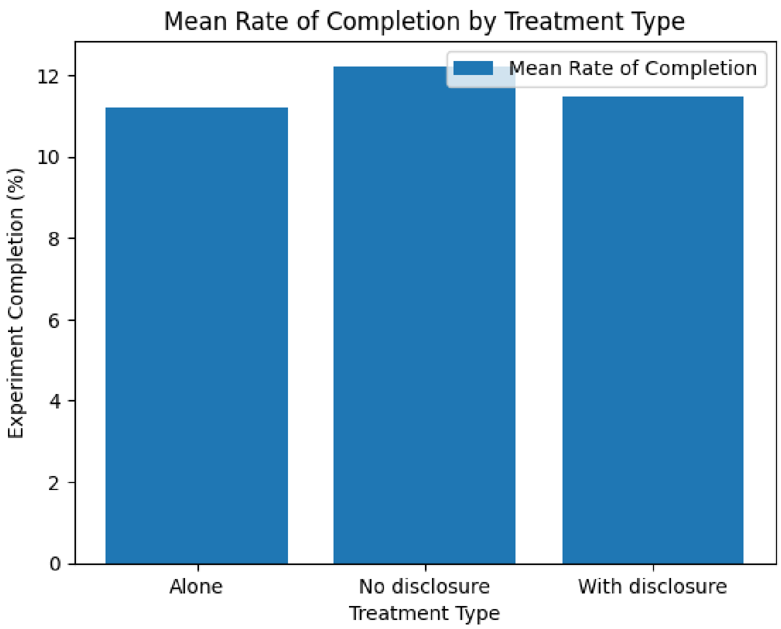

Figure 9: Rate of completion by treatment type.

We found a significant effect (Pseudo $R^2 = 0.000$, $p = 0.010$, 95% CI [-0.066, 0.009]) that the disclosure of AI lowers the chance of the participant completing the experiment in comparison to playing without disclosure. Participants' chance of completion decreased by 3.7 percentage points when participants played together with an AI avatar. Since a lower chance of completion implies a higher chance of quitting the experiment, we accept H4a.

In addition, we found a significant effect (Pseudo $R^2 = 0.000$, $p = 0.001$, 95% CI [0.022, 0.080]) that playing together with a non-disclosed AI increases the chance of participants completing the experiment in comparison to playing alone. This means that playing together with a non-disclosed AI decreases the chance of quitting the experiment in comparison to playing alone.

However, we did not find a significant effect (Pseudo $R^2 = 0.000$, $p = 0.370$, 95% CI [-0.016, 0.043]) that playing together with a disclosed AI increases the chance of completing the experiment. This means that playing together with a disclosed AI does not significantly decrease the chance of quitting the experiment in comparison to playing alone.

To test the hypothesis about the effect of participants knowing that they are playing with an AI avatar on their intensity and duration of play, we conducted an ordinary least squares regression analysis (*H4b: Intensity of play is negatively associated with participants knowing that they are playing with an AI avatar*). In Figure 10, the intensity of play grouped by the treatment type is displayed.

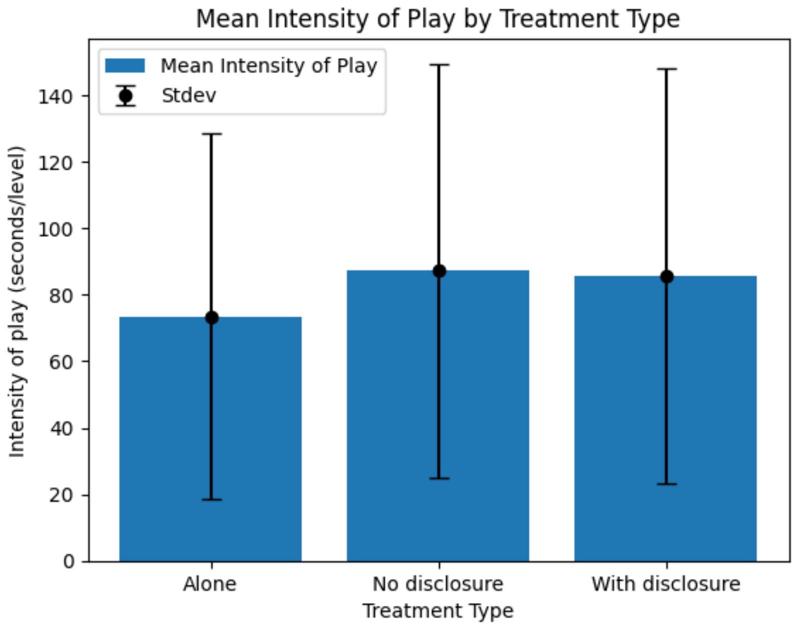

Figure 10: Intensity of play by treatment type.

We found a significant effect ($R^2 = 0.000$, $F(1, 49822) = 6.957$, $p = 0.008$, 95% CI [-2.569, -0.379]) for intensity of play. Participants' *Intensity of Play* decreased by 1.5 seconds per level when the other avatar was disclosed as AI. That is equivalent to 0.02 standard deviations. This means that participants who played together with an AI avatar that was disclosed as such, play more intensely. This is the inverse of what we expected, and therefore we reject H4b.

In addition, we found a significant effect ($R^2 = 0.013$, $F(1, 47971) = 648.8$, $p < 0.001$, 95% CI [12.654, 14.762]) that playing together with a non-disclosed AI increases the value of *Intensity of Play* by 13.7 seconds per level, in comparison to playing alone. This is equivalent to 0.23 standard deviations. This means that playing together with a non-disclosed AI makes participants play less intensely in comparison to playing alone.

We did not find a significant effect ($R^2 = 0.000$, $F(1, 48007) = 0.803$, $p = 0.370$, 95% CI [-0.003, 0.008]) that playing together with an AI avatar disclosed as such increases the value of *Intensity of Play*. Therefore, playing together with a disclosed AI avatar does not make participants play more or less intensely, in comparison to playing alone.

Finally, we tested H5 (H5: *The similarity between the AI avatar's gender appearance and the participant's avatar gender appearance moderates the effects of AI avatar gender appearance on motivation*). To do this, we conducted a multiple regression analysis using ordinary least squares to estimate the relationship between the similarity between the AI avatar's gender appearance and the participant's avatar gender appearance and their intensity and duration of play.

Before we conducted the multiple regression analysis, we calculated the gender difference for each participant by taking the absolute difference between the AI avatar and participant avatar gender score. This results in a 0 (similar) or 1 (different) as gender difference. Besides that, we calculated the gender interaction by multiplying the AI avatar masculinity with the absolute difference.

For the intensity of play, we found a significant model ($R^2 = 0.004$, $F(1, 5891) = 6.907$, $p < 0.001$). Participants' predicted intensity of play is equal to 52.407 - 0.0991 (AI Avatar Masculinity, p = 0.921) - 2.302 (Gender difference, p < 0.001) + 3.778 (Gender interaction, p = 0.007). Therefore, the participant's avatar gender appearance being different from the AI avatar gender appearance decreases the value of *Intensity of Play.*

This makes participants play more intensely when the gender appearance of the participants' avatar and the AI avatar is different. Since the difference is the inverse of the similarity (due to our simplification of masculinity to 0 or 1), participants played less intensely when both avatars appeared as the same gender. Therefore, we accept H5.

In Table 5, a quantitative overview of the effects is displayed. All significant models are displayed in bold.

IV → ↓ DV	Participant avatar masculinity	AI avatar presence	AI avatar masculinity	AI disclo-sure	Avatar gender differences
Chance of Completion (Probit)		**H2a: 3.3pp**		**H4a: -3.7pp**	
Intensity of Play (OLS) (seconds / level)	**H1: -2.6 s/level**	**H2b: 13.0 s/level**	**H3: 2.2 s/level**	**H4b: -1.5 s/level**	**H5: -2.3**

Table 5: Quantitative overview of effects.

DISCUSSION

This study researched various independent variables that could be used by an avatar controlled by Artificial Intelligence (AI) to improve users' motivation. Our findings show that in our experiment and game, the masculinity of the users' avatar makes participants play more intensely. Furthermore, playing together with another avatar controlled by AI significantly affects users' motivation. Besides that, the masculinity of the AI's avatar significantly increases the intensity of play. When users learn that the other avatar is controlled by AI, users had a higher probability of not completing the experiment and played more intensely. Finally, the difference between the user' gender appearance and AI's avatar gender appearance moderates the effect of the AI's avatar gender appearance and makes users play more intensely.

Contributions to existing literature

Previous studies have shown that cooperating with others can increase motivation (Featherstone & Habgood, 2019; Peng & Crouse, 2013). Our research shows that playing together with an avatar controlled by an AI lowers the chance of players quitting the experiment. Our findings suggest that the existing theory, which states that playing with someone else increases motivation, can be extended to include playing with an avatar controlled by AI. Future research in this field could focus on replicating our findings in other games and bringing these effects to non-gaming virtual worlds, to see if our findings are also applicable to other settings, such as virtual workspaces.

Besides that, the literature indicates that disclosing the use of AI in physical settings can lower user motivation (Keppeler, 2023; Tong et al., 2021). Our research supports this finding in virtual worlds, where we found that disclosing that an avatar is controlled by AI increases the likelihood of users quitting the experiment early in comparison with playing together with an AI without disclosure. This suggests that the disclosure itself, rather than the use of AI, may reduce motivation. Our findings also suggest that the 'AI-bias', where users are less motivated when interacting with an AI-controlled avatar compared to a human-controlled avatar, extends to virtual worlds. Further research could explore the causes and implications of this bias in virtual worlds in order to reduce the effects of this bias.

Our literature review found that players feel more connected to others when other avatars look similar to their own (Van der Land et al., 2015). Our research extends these findings to AI avatars in gaming virtual words, where we found that the gender difference between the participant's avatar and the AI avatar moderated the effect of the AI avatar's gender appearance on motivation. These findings suggest that in our experiment and game, gender effects that are claimed to have been observed in the real world may also manifest in virtual worlds. As virtual worlds become increasingly integrated into everyday life, it is worth researching if other common gender research findings are also applicable in these settings. Further research could investigate why the perceived gender of the AI avatar influences motivation, as this thesis focussed on demonstrating the existence of this effect rather than explaining the causes.

Practical implications

Companies that create virtual worlds, such as Meta, could use this research when designing avatars in the virtual world. If the sole goal is to maximise player intensity, it would be best to use a masculine avatar for the AI and to not tell users that they are interacting with an AI, in order to motivate users in their virtual world.

The results of this study also shows the need for diversity in virtual worlds to the creators of virtual worlds. We have seen that the similarity between the participants' avatar gender appearance and the AI avatar gender appearance moderates the effects of the AI avatar gender appearance on motivation. Therefore, preset genders of both the own avatar gender and the AI avatar gender can limit the effect of the AI avatar on motivation. Diversity in both avatars, such as being able to choose the avatars from a diverse list or being able to freely create the avatars, could help the user in motivation within the virtual world. Beyond gender, ensure the avatars communicate attributes players are looking for. The avatars in our experiment were designed to be picked based on gender, but there might be other attributes that participants selected those avatars on. More diverse avatars on those other attributes, regardless of gender, can also influence the motivation. It requires more research to investigate what other attributes might affect the motivation of players.

Besides the effects for the creators of virtual worlds, this research might also be of interest for the users themselves. When motivation for a task is in the users' own interest, such as when learning a new language, it would be better to do that task with an AI avatar that has a different gender appearance than the users' avatar. This way, the user will be motivated more to do and continue that task and perform this task more intensely. This will help them in achieving the goal of the task quicker.

Limitations

One limitation of this research is the potential influence of the demographics of the experiment on the reliability and generalisability of the results. Specifically, the age distribution of the participants, as inferred from the totality of the game Find The Button, could affect their behaviour during the experiment. Nearly 50 percent of the players of the game are 13 or younger, with an additional 17 percent being aged between 13 and 17. The younger age of participants might have caused participants to behave more playful than adults would do. Besides their own behaviour, they might have been influenced by the treatments in our experiments in a different way than adults would have done. To address this limitation, future research should recruit participants off-site to be able to note their age individually and investigate if the effects on motivation by our task and treatments differ for adults.

Another limitation is the quality of the AI used in the experiment. The AI avatar moved to preset points and did not interact with participants in another way, such as talking or following them. This could have let participants recognize that the avatar was controlled by an AI and behave differently as a result. To address this limitation, future research could use a more naturally-behaving AI.

42% of the participants failed to complete the first level of the task, suggesting that the task was not clear to the users. At the time of the experiment, there was no tutorial available in the game. Despite a short explanation, participants might have struggled to understand how to complete the task in the experiment. To address this limitation, future experiments could include a short tutorial to ensure that participants fully understand the task before participating in the experiment.

Besides the understanding of the task, the task and its setting might have had an influence on the outcomes of the study. The task was a simple search task, which could be seen as a more mathematical task, and the levels were themed around science.This particular task and setting could have led to a higher base motivation of some groups of participants, which could have worked through to the measurements. Future research could be changing the task and setting, while keeping the rest of the methodology the same, to investigate if the task and setting has an influence on the effects we found.

Finally, the selection of avatars was limited to only five avatars, which restricted the gender expressions in the avatars and diversity in skin colour. A participant raised this concern, which did not feel represented by the available avatars due to skin colours. We decided to not include diverse avatars, because participants should choose based on gender, not on other avatar features. However, this non-diverse set of avatars could have led to dissatisfaction about the experiment and a decreased motivation regardless of the treatment type for some participants. To address this limitation, future research could include a wider range of avatars to choose from or allow players to use their own avatar created on Roblox.

CONCLUSION

This thesis aimed to describe how AI avatars inside virtual worlds could be used to increase the motivation of a user. We conducted a field study inside the game Find The Button where we were able to control variables in the virtual world. We found that users playing with a more masculine looking avatar play more intensely and that playing together with an AI avatar lowers the chance of quitting the experiment before completing and make users play less intense. When this AI avatar appears more masculine, users play less intensely. The AI avatar being disclosed as an AI makes users have a higher chance of quitting the experiment and playing more intensely. Finally, the difference between the participant' gender appearance and AI gender appearance moderates the effect of the AI gender appearance. Further research could include replicating these results to other games & non-gaming virtual worlds and researching what the causes are for the effects we found.

REFERENCES

Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, *31*(2), 427-445.

Baylor, A. L. (2011). The design of motivational agents and avatars. Educational Technology Research and Development, 59, 291-300.

Bélisle, J. F., & Bodur, H. O. (2010). Avatars as information: Perception of consumers based on their avatars in virtual worlds. Psychology & Marketing, 27(8), 741-765.

Chen, S., Yi, Z., Wang, X., Luo, Y., & Liu, Y. (2022). Competitive game motivation and trait aggression among Chinese adolescent players of Glory of the King: The mediating role of avatar identification and game aggression. Aggressive behavior, 48(6), 563-572.

Conti, R., Collins, M. A., & Picariello, M. L. (2001). The impact of competition on intrinsic motivation and creativity: considering gender, gender segregation and gender role orientation. Personality and Individual Differences, 31(8), 1273–1289. https://doi.org/10.1016/S0191-8869(00)00217-8

Featherstone, M., & Habgood, J. (2019). UniCraft: Exploring the impact of asynchronous multiplayer game elements in gamification. *International Journal of Human-Computer Studies*, *127*, 150-168.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. Behavioral and brain sciences, 33(2-3), 61-83.

Holzwarth, M., Janiszewski, C., & Neumann, M. M. (2006). The influence of avatars on online consumer shopping behavior. Journal of marketing, 70(4), 19-36.

Keppeler, F. (2023). No thanks, dear AI! Understanding the effects of disclosure and deployment of artificial intelligence in public sector recruitment. *SSRN Electronic Journal*.

Kim, A., Cho, M., Ahn, J., & Sung, Y. (2019). Effects of gender and relationship type on the response to artificial intelligence. Cyberpsychology, Behavior, and Social Networking, 22(4), 249-253.

Lee, J. E. R., Nass, C. I., & Bailenson, J. N. (2014). Does the mask govern the mind?: Effects of arbitrary gender representation on quantitative task performance in avatar-represented virtual groups. Cyberpsychology, Behavior, and Social Networking, 17(4), 248-254.

Lenggenhager, B., Tadi, T., Metzinger, T., & Blanke, O. (2007). Video ergo sum: manipulating bodily self-consciousness. Science (New York, N.Y.), 317(5841), 1096–1099. https://doi.org/10.1126/science.1143439

Leonhardt, M., & Overå, S. (2021). Are there differences in video gaming and use of social media among boys and girls?—A mixed methods approach. International journal of environmental research and public health, 18(11), 6085.

Li, B. J., & Lwin, M. O. (2016). Player see, player do: Testing an exergame motivation model based on the influence of the self avatar. Computers in Human Behavior, 59, 350-357.

Mazlan, M. N. A., & Bakar, M. A. A. (2013). Students' perception of self-presentation towards avatar. Procedia-Social and Behavioral Sciences, 97, 361-367.

McDonnell, M., & Baxter, D. (2019). Chatbots and gender stereotyping. Interacting with Computers, 31(2), 116-121.

Meece, J. L., Glienke, B. B., & Burg, S. (2006). Gender and motivation. Journal of school psychology, 44(5), 351-373.

Mols, E. (2020). Algorithmic authority: motives behind relinquishing decision-making to artificially intelligent systems. Thesis Master Media Technology, LIACS, Leiden University

Peng, W., & Crouse, J. (2013). Playing in parallel: the effects of multiplayer modes in active video game on motivation and physical exertion. *Cyberpsychology, behavior and social networking*, *16*(6), 423–427. https://doi.org/10.1089/cyber.2012.0384

Perry, J. L., & Porter, L. W. (1982). Factors affecting the context for motivation in public organizations. The Academy of Management Review, 7(1), 89–98. https://doi.org/10.2307/257252 (AMR)

Roblox Corporation. (2023). Roblox Terms of Use. Retrieved July 4, 2023, from https://en.help.roblox.com/hc/en-us/articles/115004647846-Roblox-Terms-of-Use

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. Contemporary educational psychology, 25(1), 54-67.

Schreuter, D., van der Putten, P., & Lamers, M. H. (2021). Trust me on this one: conforming to conversational assistants. *Minds and Machines*, *31*, 535-562.

Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In 9th Python in Science Conference

Sekhar, C., Patwardhan, M., & Singh, R. K. (2013). A literature review on motivation. Global business perspectives, 1, 471-487.

Soutschek, A., Burke, C. J., Raja Beharelle, A., Schreiber, R., Weber, S. C., Karipidis, I. I., ... & Tobler, P. N. (2017). The dopaminergic reward system underpins gender differences in social preferences. *Nature Human Behaviour*, *1*(11), 819-827.

Stead, N. J. (2009). The effect of age and gender on motivation to work. In 8th Industrial and Organisational Psychology Conference (IOP) (p. 129-133)

Stefko, R., Bacik, R., Fedorko, R., Gavurova, B., Horvath, J., & Propper, M. (2017). Gender differences in the case of work satisfaction and motivation. Polish Journal of Management Studies, 16(1), 215-225.

Thaler, M. (2021). Gender differences in motivated reasoning. *Journal of Economic Behavior & Organization*, *191*, 501-518.

Tong, S., Jia, N., Luo, X., & Fang, Z. (2021). The Janus face of artificial intelligence feedback: Deployment versus disclosure effects on employee performance. *Strategic Management Journal*.

Touré‐Tillery, M., & Fishbach, A. (2014). How to measure motivation: A guide for the experimental social psychologist. Social and Personality Psychology Compass, 8(7), 328-341.

TreeBen77. (2023). rblx-open-cloud. GitHub. https://github.com/TreeBen77/rblx-open-cloud

Van Der Land, S. F., Schouten, A. P., Feldberg, F., Huysman, M., & van den Hooff, B. (2015). Does avatar appearance matter? How team visual similarity and member–avatar similarity influence virtual team performance. Human Communication Research, 41(1), 128-153.

Vanneste, B., & Puranam, P. (2023). Artificial Intelligence, Trust, and Perceptions of Agency.

Veltri, N., Krasnova, H., Baumann, A., & Kalayamthanam, N. (2014). Gender differences in online gaming: A literature review.

Verhagen, T., Feldberg, F., van den Hooff, B., Meents, S., & Merikivi, J. (2012). Understanding users' motivations to engage in virtual worlds: A multipurpose model and empirical testing. Computers in Human Behavior, 28(2), 484-495.

West, S., & Uhlenberg, D. (1970). Measuring motivation. Theory Into Practice, 9(1), 47–55. http://www.jstor.org/stable/1476060

Yee, N., & Bailenson, J. (2007). The Proteus effect: The effect of transformed self-representation on behavior. *Human communication research*, *33*(3), 271-290.

APPENDIX

A. Methodology in-game

In this appendix, we show what the experiment looked like for participants. When a participant selected they wanted to play the experimental levels (called Science Mode), they first saw the starting screen (Figure A1). On this screen, they could read information about the experiment, what data is being used and what reward they will get when they participate. They could only click the ‘Continue’ button when 10 seconds have passed - this was done to ensure participants read the information. When the participant was in the control group or the AI was not disclosed, the text ‘However, you will play this mode together with a NPC player controlled by the computer!’ was left out.

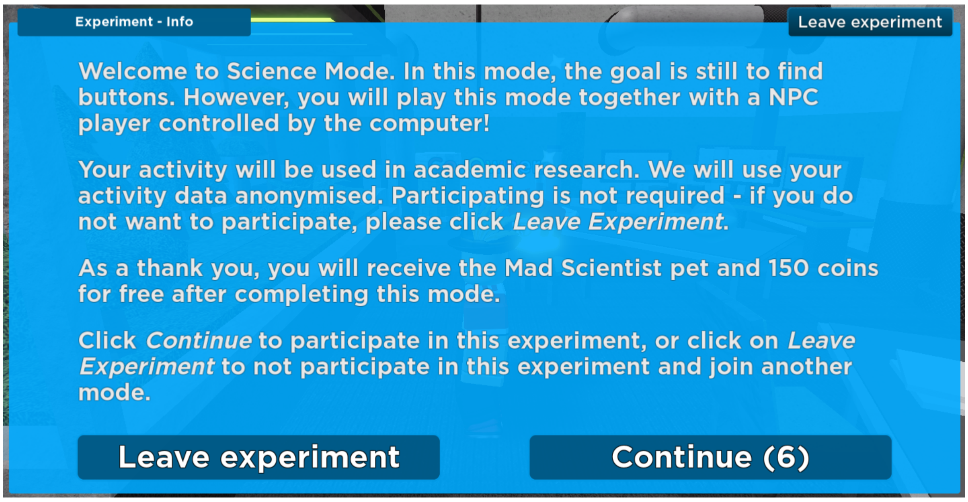

Figure A1: Starting screen

After they pressed ‘Continue’, the participant came on the avatar selection screen (Figure A2). On this screen, they could select the avatar they want to play with. They could only continue to the next screen when they selected an avatar.

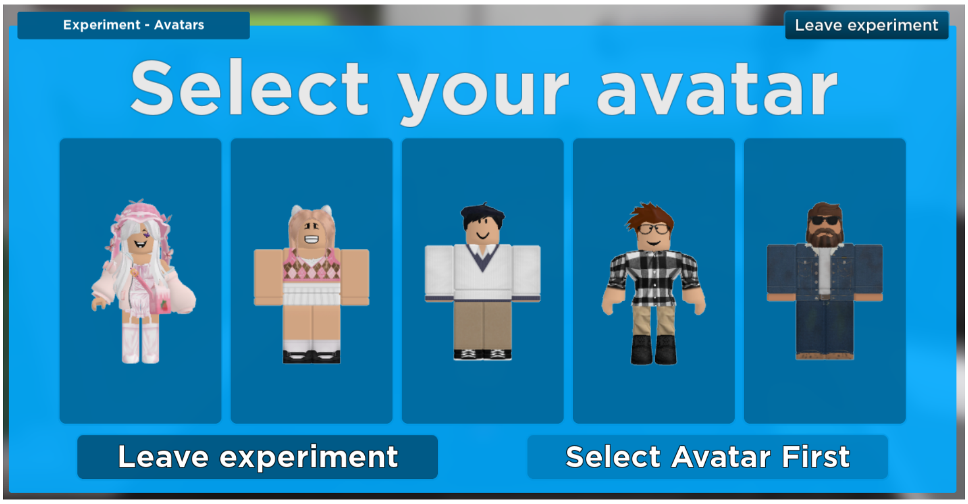

Figure A2: Avatar selection screen

When the participant selected an avatar, they were asked to rate the avatar on the gender on a separate screen (Figure A3). On this screen, both the avatar they selected and the question is visible. The question was 'The avatar is…', where participants could answer on a 5-point scale with 1 being 'Very feminine' and 5 being 'Very masculine'. The participant was only able to continue the experiment when they answered the question.

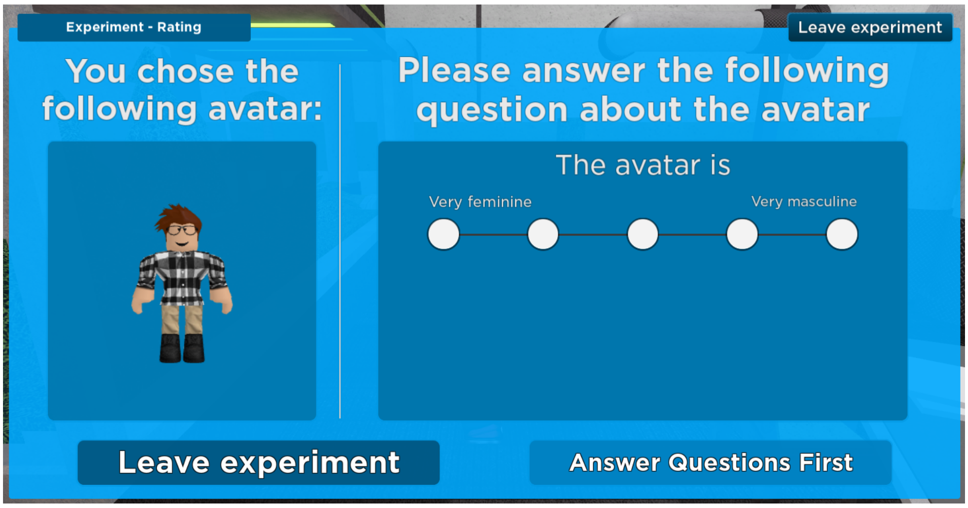

Figure A3: Avatar rating screen

After the participant answered the question, the experiment started. The avatar of the participant was morphed to the avatar they selected, and if the participant was in the treatment group, the AI avatar was created and was randomly morphed to one of the five avatars. They then played the levels in the experiment. An example of a level could be seen in Figure A4. Inside this figure, the avatar in the middle is the avatar chosen by the participant, while the avatar in the top left is the avatar of the AI. The task is to find the button, which is displayed in the image on the bottom (a round green circle). The square-like pet following the player character is the reward participants got for participating in the experiment.

Figure A4: Participant playing a level.

When the participant left the experiment before completing, they did not see any final screen - it was not possible to show them the information due to technical reasons (Roblox does not allow you to show anything when users closed the game). However, when the participant did complete the experiment, the final screen was shown (Figure A5). When the participant was in the control group, the text ‘This experiment was about AI avatars - you played together with an AI!’ was replaced with the text ‘You were in the control group’.

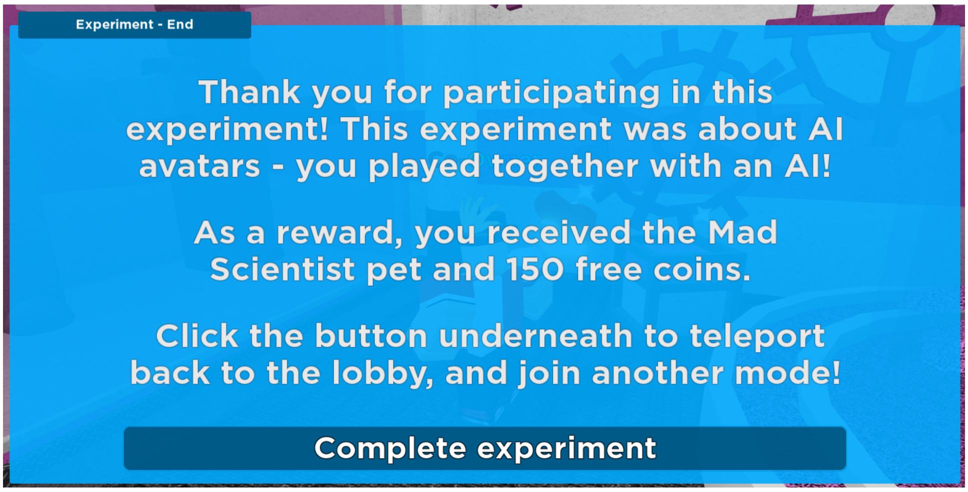

Figure A5: Final screen when you completed the experiment.

Due to the way the data was saved, there was no way to reward participants when they did not complete the experiment in the first attempt, but did in another attempt. To make sure the exclusive reward was available to every player, we decided to give rewards when players ended their participation - not just when they finished the experiment.