THESIS

USING GENDER SWAP IN VIRTUAL REALITY FOR INCREASING EMPATHY AGAINST STEREOTYPE THREATS

Submitted by

Zahra Borhani

Department of Computer Science

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Summer 2020

Master's Committee:

Advisor: Francisco R. Ortega

J. Ross Beveridge
Benjamin A. Clegg

ABSTRACT

USING GENDER SWAP IN VIRTUAL REALITY FOR INCREASING EMPATHY AGAINST STEREOTYPE THREATS

The stereotypes associated with women in computer science are potential barriers that prevent female students from developing an interest in this field. This problem persists when attempting to establish a career after graduating. This project shows a tool that potentially increases empathy using avatar gender-swap in a virtual reality setting that simulates a job interview experience. The virtual environment includes two avatars, one for the interviewee and one for the interviewer. The objective is to understand the effects of virtual embodiment and the potential to increase empathy towards the opposite sex by participating in a job interview task simulated in virtual reality when the avatar gender is swapped. The participants should perform a job interview task under three different conditions, microaggression stereotype threat, direct stereotype threat, and no threat. This thesis will showcase all the necessary tools required to accomplish this goal and provide a path forward for a user experiment.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Francisco R. Ortega, for his support and ensuring that we followed through with this project to completion. I also would like to thank my M.S. thesis committee member, Dr. J. Ross Beveridge and Dr. Benjamin A. Clegg for offering their time and providing helpful advises. Finally, I would like to acknowledge the National Science Foundation (NSF) and the following NSF awards: IIS-1948254 and BCS-1928502.

DEDICATION

I would like to dedicate this thesis to my amazing parents, who have been constant cheerleaders through every academic and personal endeavor in my life. Thanks mom and dad for always believing me and for encouraging me to strive for my dreams.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

The existing gender gap in Science, Technology, Engineering, and Math (STEM) majors, especially in Computer Science, is one of the current challenges that threaten a high percentage of the world's population formed by women. The stereotypes associated with this area are a significant barrier that prevents female students from developing an interest in these fields [9]. Statistics in 2015 showed women made up to 57 percent of the workforce in the world, while only 25 percent of them are working in computing occupations [10].

The under-representation of women is getting even worse in computer science majors. As you can see in Figure 1.2, the female share of computer science degrees is declining from 28% to 18%. However, the male share of computer science degrees is increasing over time and reached more than doubled (about 7% in 2017). This deduction is not the case for all majors, as it can be seen in Figure 1.2. This figure shows how much worse the computer science major is at bringing in women compared to the other majors, and how steep the decline has been in recent decades. Therefore, it is crucial to act fast on this issue.

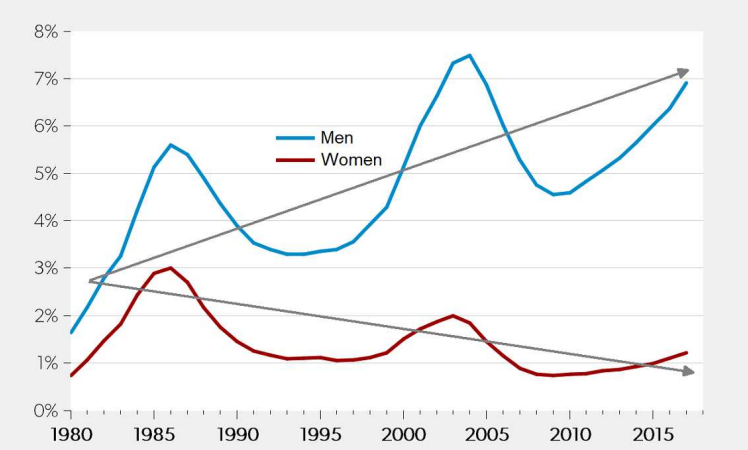


Figure 1.1: Gender share comparison of bachelor's degrees in computer science, 1980-2017 by American Enterprise Institute (AEI) [1].

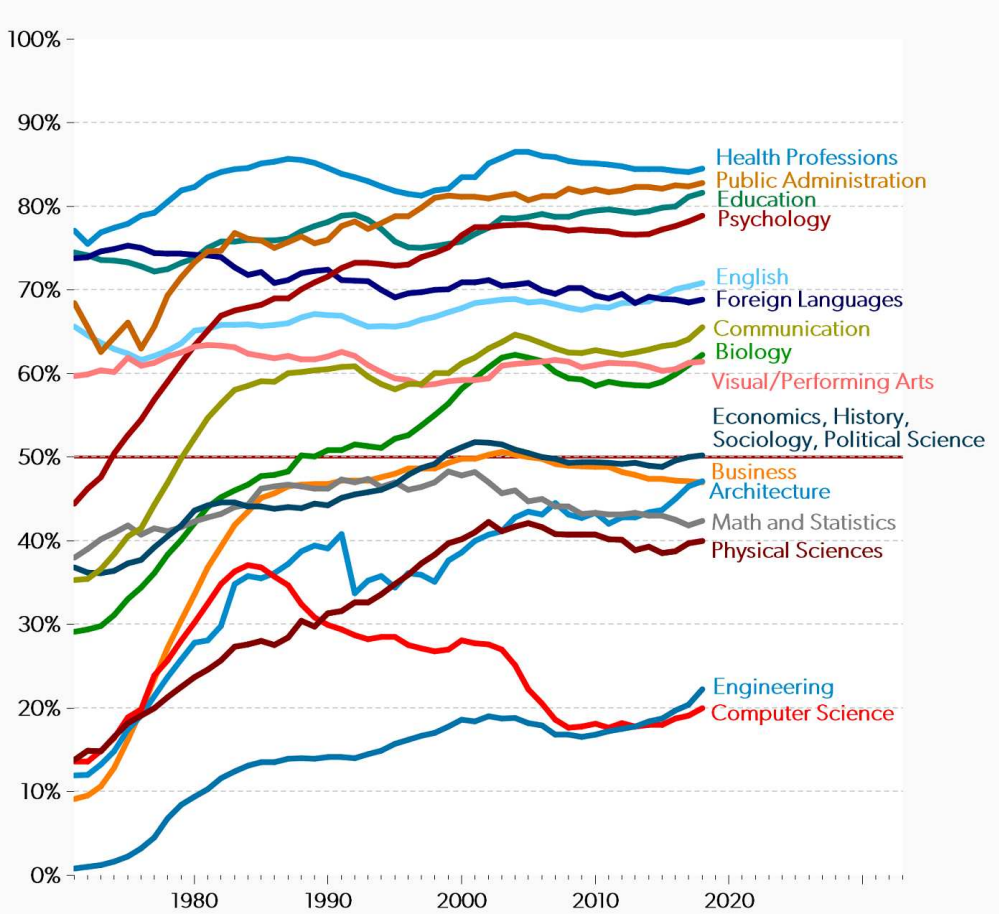


Figure 1.2: Female share of bachelor's degrees by major, 1971 to 2018 by AEI [2].

Research has shown that females and males' reaction to negative feedback in specific environments is different, and women are more responsive to negative feedback than men [11]. Stereotypes associating this area with men is a significant barrier that prevents female students from developing interests in this field. Meanwhile, research has shown that diversity has a high impact on innovation and creativity in society [10]. The positive effect of females in different areas provides another reason to consider this issue and act on it. Research has shown that experiencing or observing a stereotype threat can increase empathy among individuals [12]. Increasing empathy might be useful to decrease the severity of the stereotype threat if we can simulate this experience for a large group of people.

Previous research showed that virtual reality (VR) is useful for experiencing situations that we cannot simulate in the real world (e.g., race, age, or gender) [9, 13, 14]. In this thesis, a system is introduced that helps to investigate if experiencing stereotypical behaviors will change the participants' behavior or perspective toward the opposite gender. In other words, swapping the participant's gender in a VR environment while experiencing a stereotype threat condition. We want to study the effect of this gender-swapping on users' behavior when they experience sexist behavior. Gender swapping refers to using an avatar with an apparent gender opposite to the gender the user specifies as their own [13]. A previous study has shown that using a customized male (vs. female) avatar has a positive effect on the scores of participants on a math test. The participants gave the test in the presence of the stereotype that women are not as good as men at math [14]. In another study, Peck et al. manipulated an avatar's gender to investigate the effect of gender swap on the performance of users in a stereotype threat environment [9]. Their results supported that stereotype threat can be induced in an immersive virtual environment using a gender swap experiment. In a 3×6 factorial experiment, the experiment conditions were manipulated, and the order of tasks to observe the influence of gender-swap on empathy when participants experience a sexist behavior. This project can provide the basis for similar studies, and we can apply it to other stereotype threats such as race, age, or the ones that involve LGBTQ groups.

1.1 Major Contribution

This research's contribution is a system that can potentially increase the empathy of the opposite gender in computer science, toward female students, by using gender-swapped avatars in VR. In particular, this system was design using an iterative process by using a questionnaires, a pilot study, and design concepts from human-computer interaction. The contributions include understanding the tool and design guidelines to develop these types of systems, as well as a pilot study that provided additional design criteria for the design of this tool. Finally, we conducted two questionnaires to confirm the final design of the system. The system is completed for an experimenter and is ready to run in human-subject studies.

Figure 1.3: Reality-Virtuality continuum created by Paul Milgram. Figure adapted from [3].

1.2 Concepts

In this section, we define some of the terms used in the rest of this thesis. Understating these concepts will help the reader to understand the succeeding chapters better.

1.2.1 Reality-Virtuality continuum

In 1995, Paul Milgram and Fumio Kishino designed the reality-virtuality continuum [3]. Figure 1.3 shows this continuum. The reality–virtuality continuum encompasses all possible compositions of real and virtual objects:

- **Reality:** The real (physical) world that we are living in.
- **Augmented Reality:** The real world where virtual objects are added to it. These virtual objects might be in the form of digital 2D or 3D objects, images, animation, data, etc.
- **Augmented Virtuality:** Real objects, images, or data are added to the virtual world [15]. An example of augmented virtuality is presented in the movie "Valerian and the City of Thousand Planets", directed by Luc Besson.
- **Virtual Reality:** All the objects in this environment are syntactic and entirely digitally created [3].

1.2.2 Uncanny Valley Theory

Uncanny Valley theory suggests that humanoid objects (such as avatars) whose appearance resembles actual human beings provoke uncanny or strangely familiar feelings of eeriness in ob-

servers. The main reason is due to the subtle differences between what is real and what is not. [16]. For example, consider an avatar that is very similar to a human; however, the avatar's movements differ from an actual person. This theory describes as the humanoid object becomes more similar to a human, people get more comfortable with that, but this is reversed after a certain threshold, and continues to decrease dramatically until the humanoid object becomes an (almost) true replica of a human. This section is called the uncanny valley.

1.2.3 Stereotype threat

The term stereotype threat has been introduced in [17] as the phenomenon in which members of negatively stereotyped groups under-perform in stereotype salient situations. Examples of this threat are abundant in the real world, such as underperformance of females in STEM majors. Other examples include the ones that involve race, age, religion, appearance, or LGBTQ groups [18]. Here the stereotype threats are divided to two groups:

- **Microaggressions/implicit stereotype threat:** This refers to daily behaviors or sayings that might insult or create a hostile environment for a minority or endangered group of the society. This type of threat might be very subtle and implicit, and they are often unintentional, in contrast to the direct stereotyping, which is explicit and intentional [18, 19].

- **Direct/explicit stereotype threat:** These stereotyping behaviors are often explicit, intentional, conscious, intentional, and more offensive. Explicit stereotype threat targets an entire societal group or an individual belonging to that group and intended to hurt the victim. This type of stereotypes in [18] are called macroassaults.

1.2.4 A Wizard-of-Oz Experiment

The phrase Wizard-of-Oz experiment refers to an experiment wherein an experimenter (the wizard) simulates a theoretically intelligent computer application. In a Wizard-of-Oz experiment, the subject interacts through an interface with a human/wizard simulating the behavior of a sys-

tem [20]. One of the typical applications of this methodology is to investigate human-computer interaction in systems under development [21].

1.3 Related Work

As previously mentioned, members of negative stereotype threat groups underperform in stereotype salient situations. Various research studies showed that there are different groups of people in danger of being stereotyped such as Women in STEM [22–24], followers of certain religions [25], and people from ethnic and racial minority groups in academia [26]. An example of this is microaggression stereotypes towards Muslim Americans [27]. There are a variety of research studies on stereotype threats, their causes, and their impacts.

Brown studied the effect of stereotype threat on the performance of Black undergraduate women in math [28] and his results confirmed the negative effect of stereotype threat on the participants' performance. Denise Sekaquaptewa et al. in [29] used an interview as an approach to investigate how implicit stereotyping may cause discrimination against black people. Interview tasks have been a repeating theme for studying stereotyping behaviors and are shown to be useful to capture implicit or unconscious stereotyping behaviors.

Gender stereotyping is one of the main concerns in stereotyping behaviors. Spencer et al. showed that when women performed math because they were afraid of being judged by the negative stereotype that women have weaker math ability [30], they tended to perform worse. This specific type of stereotype is called a predicament stereotype threat, which is accompanied by a stressful feeling of becoming embarrassed. It has been suggested that the anxiety may cause disruption in women's math performance.

A similar study showed how fear of underperformance in math, science, and engineering (MSE) among women causes them not to appear as qualified as they are in reality [31]. Specifically, this research investigated how a male instructor's nonverbal sexist behavior (for example, sitting with an open body posture relatively close to the participant's avatar) may affect a virtual

classroom. The results collected from 76 female college students showed students underperformance in their classes compared to a nonsexist instructor in a virtual classroom [31].

In a related study [32], participants were asked to read the transcripts of three different job interviews of a female applicant for different stereotypical conditions (i.e., salient, hostile, and benevolent). Participants were then asked to evaluate and rate the interviewee and interviewer on different perspectives: applicant competence, applicant likeability, applicant hireability, and interviewer favorability. The researchers were interested in how female applicants were affected and whether they were able to appear as competent applicants during the interview. Their results confirmed benevolent sexism harms women in the workplace.

Surprisingly, many of these stereotypes have become the norm and common belief among the women themselves. For example, a study on female gamers [33] revealed that a majority of female players have shown a preference toward gender-swapping when choosing their avatar. One possible explanation is that the females are treated differently; therefore, preferring male avatars.

Virtual reality is getting very popular in stereotype threat studies and experiments because it provides an opportunity to experience these situations as they might occur in the real world [9, 34, 35]. Besides, actual tests might be expensive, unsafe, or impossible to simulate in real life. One of the essential factors in simulating a real-world experiment is the people shown in this world in the form of avatars.

There are various research studies on the impact of the appearance of virtual avatars on how humans perceive them [34, 36]. One example of changing a physical characteristic is the work by Jorge and Khan [34]. They showed that avatar appearances and agent body size effected on how participants act during physical activity in the virtual environment. Their results showed that when the participant uses an obese avatar, they act at a slower pace and act more like an obese person than when they use a regular-sized avatar [34]. Another use-case of the ability to make changes to the size of the avatar is for patients with anorexia nervosa (AN) who have a persistent distorted experience of the size of their body [35, 36]. Virtual Reality experiences can help them to have a better perception of their body size.

Slater et al. showed that first-person perspective of a life-sized virtual human female body that is assigned to a male participant is sufficient to generate the illusion of a body transfer [35]. They concluded that the male participant showed signs of body-ownership over the assigned avatar. This gender swapping is an interesting experiment that can be conducted in a virtual environment that might be impossible in the real world. Ben Kuchera found that one of the male participants who saw himself as a woman in virtual reality was amazed by the fact that he could see the world from a woman's point of view [37]. This immersive experience and feeling of embodiment show another potential use-case of this gender-swapping for the people who are willing to change their gender so that they can get acquainted with and prepared for how they will feel in the future.

The gender swapping can easily be achieved in a virtual world by assigning an avatar with the opposite gender to the participant's own. Similar kinds of experiments can be extended to race, age, and other minority groups that can be distinguished by appearance features. A previous research has shown that white person that experienced embodiment in a black avatar showed a reduction in their implicit racial bias [38].

Peck et al. investigated the effect of gender body-swap illusion on working memory and stereotype threat [9]. They showed participants (all female) under-performed in a determined task if they perceived the stereotype threat condition by being told that women do not perform as well as men in this specific task before they start the experiment. Underperforming got worse when the job became more complicated. This decrease in performance might be because they believed their relatively worse performance was inherent to their gender. Peck at al. also showed that when the participants performed the same tasks with a male avatar, while they believed that men do better in the tasks, they performed better than when they embody a female avatar.

In many of the studies mentioned above, the authors investigated the effects of gender swapping on the performance of the participants. However, none of them study the use-case of gender-swapping on increasing empathy against stereotype threats. That is one of the main reasons behind this thesis.

Chapter 2

Tools and Technologies

This chapter explains the tools and technologies that were used to design our software and experiments. The tools used for each task are explained along with other similar tools that can be used for that task.

2.1 Designing an avatar

One of the most essential human activities is social interaction. In-person meetings is one of the traditional ways of social interaction that allow people to communicate more clear. People can use speech, gaze, facial expressions and body language to communicate. The capability of reading facial expressions and body language often help to understand messages faster and easier. It might be possible to replace speech by texting but it is hard to replace facial expressions and body language by text. Also, research has shown face-to-face interactions increase the percentage of engagement compared to other types of communications such as voice message or text [39].

One of the most amazing features of virtual reality is simulating face-to-face interactions using avatars and agents. It might not be completely the same as a real life face-to-face interaction but it is the most similar one compared to any other media, such as text messages. Users can share the same virtual space and see objects in their real size and shape. Also, the body language, gazing and facial expressions can be expressed through the avatars as long as speech and text.

The design of an avatar has an important role on the degree of immersion for the user. There are multiple tools that can be used for creating a virtual avatar. Since a job interview environment was designed for the experiment, this study is interested in humanoid avatars which have full-body shape and facial expressions that are designed to resemble the ones in a real person. 3D humanoid avatars can be divided into cartoon-like and photo-realistic avatars [40]. Figure.2.1 shows a humanoid avatar that is created using MakeHuman tool [4] verses a cartoon-like avatar from google-poly website [5]. There are a variety of software for creating 3D characters that can be

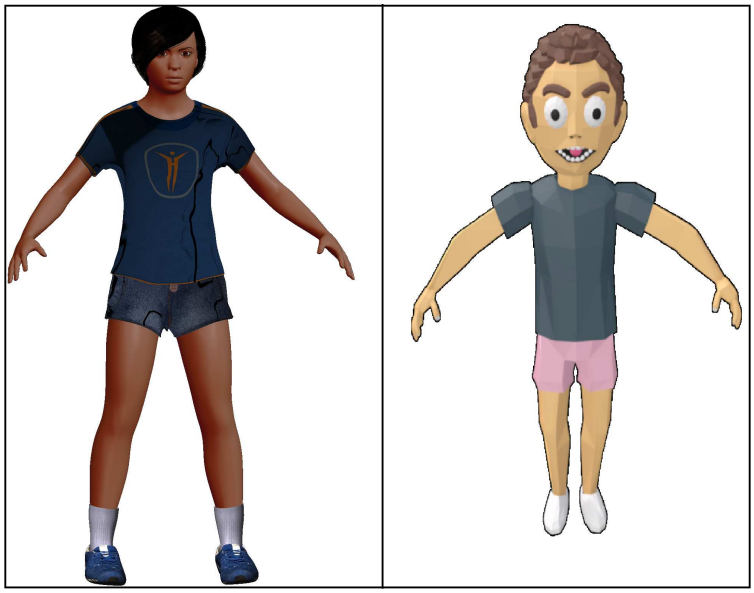


Figure 2.1: Photo-realistic avatar [4] vs Cartoon-like avatar [5].

used in different projects such as graphic projects, as well as games and 3D user interface projects. Four different tools for creating avatars Were investigated four tools, in order to find the proper one for the experiment.

Make Human

MakeHuman is a free and open-source tool that uses 3D morphing technology for creating 3D humanoid avatars. All characters are photo-realistic and resemble the body-shape of a human. It includes multiple standard avatars than can be modified using feature sliders. Sliders for modification of avatars can be divided to two parts:

- **General modifications:** These modifications affects on all avatar organs such as age, weight, gender, muscle, height, or race (African, Asian and Caucasian)
- **Focused modification:** The sliders are designed to modify a specific part of the body such as the length of arms, the distance of fingers, the scale of torso and the shape of head.

Adobe Fuse CC

Adobe Fuse CC is another software that allows all designers to create highly customized high-poly humanoid characters. You can modify different parts of avatar's body similar to MakeHuman. In addition, you can modify different parts of an avatar's face such as cheeks, teeth, ears, eyes and lips. One of the most useful features of Adobe Fuse is blendshapes that allows the designer to create facial expressions. Also, facial hair can be added to the avatars, which is a feature that is missing in makehuman.

Character Creator 3

Character Creator 3 [7] is another software for creating 3D realistic looking humanoid characters. Unfortunately it is not freely available. In addition to the features that previous tools provide, it allows the user to control some scene objects, such as lighting, and camera view. All avatars are fully rigged and have morphs for facial expressions. One of the biggest limitation of this software is the limited clothing options that it provides for the avatars.

Microsoft Rocketbox Avatar

The Microsoft Rocketbox Avatar [8] is a library which can be accessed through GitHub. Avatars are ready-to-use and the user cannot modify any feature of the avatar's body or its clothing without using a third-party software such as Blender. Microsoft Rocketbox Avatar includes 115 different avatars that can be divided into three groups: adults, children, and professions. All avatars are fully rigged but it does not support blendshapes which are used for controlling the facial expressions. One of the features of this library is the diversity of characters for various professions, religions, races, and ages.

Comparing tools

This section compares some features of the aforementioned tools for creating 3D humanoid avatars. All avatars are photo-realistic but the quality of them might vary. Figure 2.2 shows exam-

Figure 2.2: Avatars created with MakeHuman, Adobe fuse, Character creator and Microsoft rocketbox master (in order from Left to right) [4, 6–8].

ples of avatars from these tools. These avatars are created by Makehuman, Adobe Fuse, Character Creator 3, and Microsoft Rocketbox Avatar, from left to right accordingly.

2.1.1 Daz3D

Daz3D [41] is a freely available software for modeling and rendering realistic characters, clothes, and scenes. This tool is not used in this study. However, it can be used to alleviate some of the shortcomings of the earlier-introduced tools. For example, Rocketbox lacks blendshapes, however, Das3D gives the the designer the ability to modify facial expressions (among other animations) of the avatar that is imported via Rocketbox. Daz3D can import the avatars from Rocketbox (since they are fully rigged) and create blendshapes. Then, Daz3D allows the file to be exported in the desired format. Therefore proper avatars can be created by a combination of multiple tools to accomplish the desired task, for example, a designer may create an avatar with Makehuman in combination with and Daz3D.

Table 2.1: Comparing different tools for crating 3D avatars

Features \ tools	MakeHuman	Adobe Fuse	Character Creator	Rocketbox
Free	Yes	Yes	No	Yes
Ease of use	Easy	Easy	Medium	Easy
Avatar quality	Low	medium	High	High
Facial bones	Yes	No	No	No
Blend shapes	No	Yes	Yes	No
Modification	Body/face	Body/face	Body/face	-
Standard avatars	20	16	14	114
Clothing	limited	limited	limited	1
Profession clothing	No	No	No	Yes
Pose	18	2(T/A pose)	9+32 hand gestures	1
Facial hair	No	Yes	Yes	Yes
Importing to unity	Easy	Hard	Easy	Easy
Animation	None	2000(Mixamo)	None	None
Exporting formats	dae,fbx,obj, stl	fbx	obj, fbx	fbx
Avatar modification	Body/organ	Body/organ/face	Body/organ/face	None

2.1.2 Maxiamo

Maxiamo [42] is a software that works with Adobe Fuse. It has the ability to Auto-rig and animate the avatars that are made with Adobe Fuse. For this purpose, a '.obj' format of avatar needs to be uploaded to Mixamo. Mixamo includes more than 2000 short animations, such as sitting, standing, walking, dancing, and running, among others. In this study, Maxiamo is used for creating some of the avatars' animations.

2.2 Major Technologies

2.2.1 Unity

Unity is a popular game engine for developing high-quality 2D or 3D applications and games [43]. Unity supports HTC Vive and its trackers. A graphical scene of Unity is composed of multiple game objects. Each object can have multiple features such as material, physics, animators. Unity operates with object-oriented scripting languages such as C#. Scripts give the ability to interact with the objects and creating movements. This study used Unity for simulating the virtual reality environment.

2.2.2 HTC Vive Head Mounted Display

HTC Vive [44] is a virtual reality headset developed by HTC and Valve Corporation, which provides the ability to experience the virtual worlds. Here are some of the reasons for choosing HTC Vive over other head-mounted displays available on the market:

- HTC Vive offers one of the most immersive virtual reality experience in today's market.
- HTC Vive offers a reasonable setup for translating human movements to the virtual world using two included sensor bases that should fix at different points.
- HTC Vive trackers can be added to the main device to facilitate full-body tracking.
- HTC Vive controllers use a USB cable for charging, which is better than using batteries. It gets more important when it is used for academic experiments that might last for several hours.
- HTC Vive works well with Unity, the software used to develop the virtual environment for this experiment.

Of course, there are some disadvantages when compared to Oculus headsets. It is easier to learn and work with Oculus head-mounted displays, and they are less expensive.

Chapter 3

Method

The design of the environment is fundamental when creating immersive experiences. In our experiment, which is a simulated job interview task, the design of avatars has an essential role because users are almost exclusively interacting with the avatars. If the avatar appropriately designed, interacting with an avatar in virtual reality can be amazing. In our project, users cannot have a proper feeling of being immersed in a virtual experiment if their avatar or interviewer's avatar does not have humanoid behavior. For this purpose, different features, depending on the roles of the avatars, are implemented. In general, these features include animations, full track body movements, eye gazing, lip-syncing, voice audio, and facial expressions.

The design of the avatars is divided into two parts in this experiment: the interviewer's avatars and interviewee's avatars. Some features are the same in the design of both such as eye movement and blend shapes, but because of the different roles they have, some of the features are specific to one of them.

3.1 Interviewer's Avatar

Adobe Fuse was used for creating the interviewer's avatar that has five main features: animations, eye gazing, lip-syncing, facial expressions, and pre-recorded audio. The following sections explain these features in detail:

3.1.1 Eye Gaze Communication

The eyes are one of the most critical parts of the body for social interactions. Vertegal has shown that when somebody is either listening or speaking, they kept eye contact with the audience most of the time [45]. However, this is not the case when they are thinking. These results confirm that eye gaze is an excellent predictor of conversational attention. When the speaker talks or

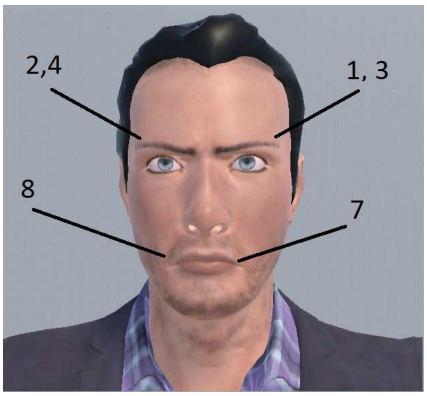
looking at an object, the audience keeps their eyes on that object as well [40]. It is hard to create the feeling of being alive for avatars when their eyes do not have any movement.

The eyes of the interviewer configured so that they have approximately natural movements, and their gaze is on the interviewee most of the time while the participants are speaking. It makes the interviewer's avatar more realistic and helps the user feel like they are participating in a real interview procedure. For this purpose, an available asset from Unity Asset Store called Realistic Eye Movement was used [46]. This asset makes characters more lifelike by animating the avatar's eyes, head, and eyelids. This asset provides the designer with multiple options to control the avatar's eye movements, looking around randomly, at the player, or at points of interest that are pre-defined.

In our experiment, the target of the avatar's head used as an object of interest, therefore the interviewer, most of the time, focuses on the interviewee's avatar. It lets the designer control the size of the points of interest, the distance of the player notice, the maximum and minimum time of look, head weight, eyelids, and microsaccades. Figure 3.3 shows the interviewer's eye movements.

3.1.2 Facial Expressions

The face of the avatar is one of the parts of its body that has an essential role in expressing feelings. The participant continuously observes the face during the interview. Research has shown that facial expressions help create a more immersive environment. The human face has 43 different muscles, making a human face capable of producing more than $10,000$ expressions [47]. Koda, in [48], used 12 expressions for the human face: happy, sad, approving, disapproving, proud, ashamed, grateful, angry, impressed, confused, remorseful, and surprised. However, this study focuses on six basic emotions that are more frequent and universally recognized: joy, surprise, fear, anger, disgust, and sadness [49, 50]. Adding these facial expressions to the interviewer's avatar, help us to make it more realistic. It also helps to express different feelings, which is very useful for simulating stereotyping behaviors.



Blendshapes
1: BrowsDown Left
2: BrowsDown Right
3: BrowsIn Left
4: BrowsIn Right
5: EyesWide Left
6: EyesWide Right
7: Frawn Left
8:Frawn Right

Figure 3.2: The angry face and its corresponding blendshapes

Unity uses the Skinned Mesh Renderer component to implement animations for the facial bones. Skinned Mesh Renderer component has a sub-component called blendshapes that can modify facial parts of avatar, such as blinking, raising an eyebrow, opening the mouth, frowning, and other facial movements. Various facial expressions made by combining multiple blend-shapes of this kind. As an example, suppose Unity has the blend-shapes shown in Figure 3.1, and the facial expression of anger aims to be created. Figure 3.1 shows an angry face and its corresponding blend-shapes that should be modified to get the desired facial expression. The blendshaps that should be modified for this purpose are BrowsIn Left, BrowsIn Right, BrowsDown Left, BrowsDown Right, Frawn Left, and Frawn Right blend-shapes. The table shows the corresponding numbers presented in the figure to these blend-shapes.

Once the blend-shapes that have to be modified to express a facial expression are recognized, the severity of the feeling controlled by changing the amounts of those blend-shapes.

Similarly, we have designed the the other five basic facial expressions. Figure 3.4 shows these facial expressions that were built for the avatar and can all be initiated using the keyboard which is controlled by the experimenter. Chapter 3 talks about this part in more details.

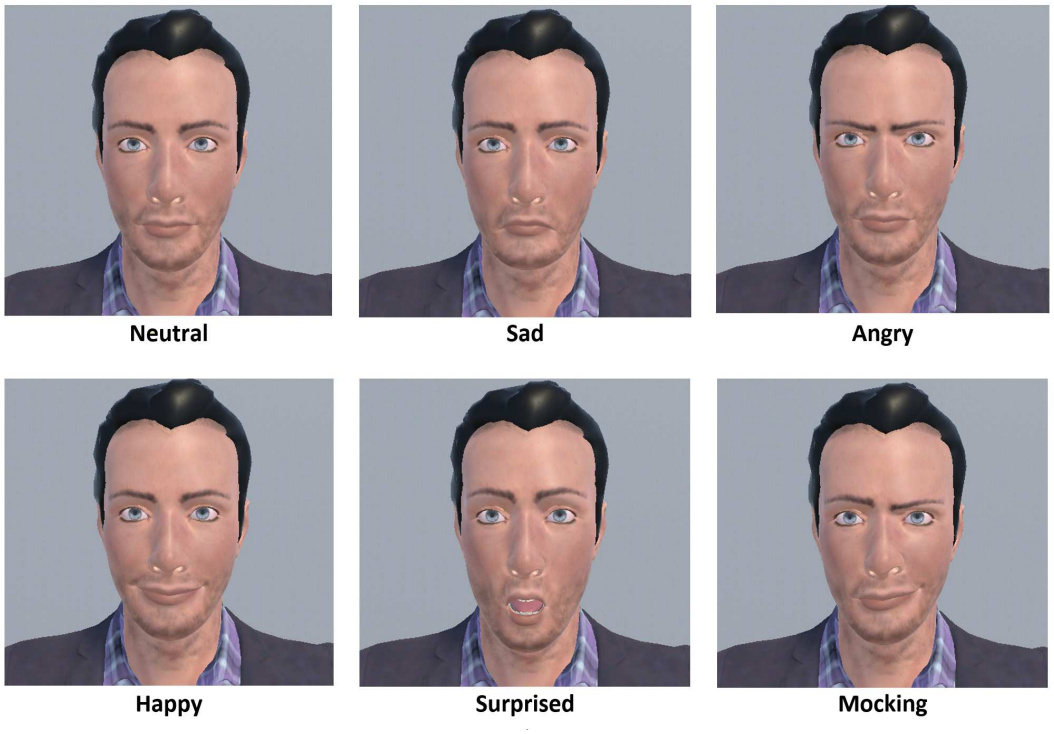


Figure 3.4: Different facial expressions that are designed for the interviewer's avatar.

Figure 3.3: Realistic eye movements

3.1.3 Lip Syncing

To make the avatar more realistic, one of the features that is added is lip-syncing when the avatar is talking. To implement this feature, a modified pre-built asset [51] was used. This asset adds movements to lips such that they synchronized with the interviewer's voice. This module

uses frequencies of the playing sound to control the movements of the lips. Some thresholds allow us to tune the number of movements for each sound. It uses the blend shapes related to the mouth and face of the avatar to make the desired movements. These values need to be changed when working with different voices because of a change in the frequencies; however, since the speech of the interviewer pre-recorded, this does not become a problem in our work. The script of this asset was modified to be applicable to this work.

3.1.4 Animations

Several different animations (series of purposed body movements) are used for the interviewer's avatar. These animations are imported from Maxiamo and can control using a controller component in Unity. These animations include walking, sitting, turning the direction while walking, changing the position from a standing pose to a sitting pose, and other animations. Some of the pre-built animations were modified because the original ones did not precisely match the expectations. In general, the main animations are used at the beginning of the experiment, the interviewer's avatar enters the room, walks forward toward his chair, sits on the chair, and starts the interview.

3.1.5 Avatar's Speech

A script is attached to an empty object in a scene that contains the pre-recorded audio. In general,42 pre-recorded voices were used, including the interviewer's questions, the sound of his footsteps during the walking phase at the beginning, different sounds for the laugh and humming. Some of these pre-recorded audios are inputs of the lip-syncing module that controls the movements of the lips.

During the interview, the interviewer asks a question and then waits for a response from the participant. Since we can not estimate the amount of time that the participant requires to answer the questions, a Wizard-of-Oz controller controls the interviewer's avatar. The Wizard-of-Oz system controlled by the experimenter controls facial blend-shapes of the interviewer's avatar and speech. To be consistent among the different runs of the experiment for different participants, the blend-

shapes are used for each question (which includes the facial expressions) grouped. They can trigger by pressing a specific key on the keyboard.

3.2 Interviewee's avatars

Interviewee's avatars, similar to interviewer's avatars, are created using Adobe Fuse. The following subsections explains the implementation of specific features to the interviewee's avatars.

3.2.1 Full body tracking

To create a more expressive avatar, a more immersive environment, and increase the feeling of body-ownership by the participant, the interviewee's avatar tracks the participant's movements. For this purpose, we used an HTC Vive Head Mouthed display, two controllers, and three trackers. Each of these six devices were used to track a specific part of the human body.

To begin, the SteamVR plugin package [52] was imported into Unity. This asset is in charge of loading 3d models for VR controllers, handling their inputs, and estimating the actual movements by using the movements in the controllers [53].
After adding the "CameraRig" object and removing the main camera from the scenes, the user can see the virtual environment from a first-person perspective. Also, three controller objects used as targets for the left foot, right foot, and hip. The targets are attached to the corresponding body-part of the avatar using a script attached to these targets. The script has an input called devices that should be valued with the correct device (HMD, controllers, and trackers) to target the corresponding part of the participant's body (devices fastened to the body-parts before the start of the experiment). The typical ordering of these devices in the main controller are as follows:

In the future, the hand controllers will be replaced by two other trackers because it is more realistic for the participant not to hold two controllers during the interview.

Another asset used in this project for tracking is the Final IK asset [54] which is available on the Unity asset store (this is not a free asset). Final IK coordinated arms and the legs' movements,

Table 3.1: HTC Vive devices and their corresponding numbering

Typical Vive Devices Numbering	
Device number	Actual device
Device 1	Base station a
Device 2	Base station b
Device 3	Left Hand controller
Device 4	Right Hand controller
Device 5	Body Tracker 1
Device 6	Body Tracker 2
Device 7	Body Tracker 3

although the trackers connected to only wrists and feet. It also has Locomotion, which is an amazing feature that helps an avatar to move around more realistically by mimicking humans' real movements. Besides, it gives the ability to control the weight, feet distance, step threshold, angle threshold, velocity, the maximum amount of stretch in the legs, and other features. Specifically, this system is interested in the VRIK module of Final IK, which needs to be valued by the targets that have been created to control the avatar. Figure 3.5 shows how the avatar tracks the users movements.

3.2.2 Height and arm length calibrations

The height and arm lengths of participants varies from one person to the other. So that the height and the length of arms were calibrated using a script that is attached to the avatar. To manipulate the length of the left and right legs, the "Leg Length MIP" component was changed. Similarly to change the lengths of left and right arms, the "arms-length MIP" component was modified. Since making the appropriate change to each of these modules for each participant is time-consuming, a script was provided that gets the height of the participant as input and sets the rest of variables to match the values of all these components with the body of the participant. Figure 3.6 shows how the height of avatar changes based on the user's height.

Figure 3.5: Tracking the user's movements by the avatar. The above row shows the avatar as the user see it in the mirror.

3.2.3 Facial expressions

This system did not consider facial expressions and eye movements of the interviewee's avatar due to technical limitations in capturing the participant's facial expressions and eye movements. Though, a Wizard-of-Oz system can still control facial expression. The facial expressions were used here are same as the ones that were introduced in the previous sections for the interviewer's avatar. A similar script to the ones used before is used here. However, the difference is that the facial expressions are controlled using a Wizard of oz system by the experimenter during the experiment (they grouped as was the case for the interviewer's avatar). For example, when the experimenter observes that the participant is smiling, presses a specific key in the keyboard, reflects the corresponding facial expression in the avatar.

Figure 3.6: Calibrating the height of avatar based on the height of user.

3.3 Designing the Scene

3.3.1 Room

A simple room with a wooden floor, including two bookshelves, a table, two chairs, a painting, a plant, two mirrors, two lamps, and a laptop as an interview room was designed. All assets are free assets that are imported from the Unity asset store. Figure 3.7 shows the interview room.

3.3.2 Mirror

This study used two mirrors from Vive Stereo Rendering Toolkit [55] and modified some features to apply to our experiment. A mirror was added to the front of the interviewee's position (instead of the bookshelf), and another one was added to the right-hand position of the interviewee's chair.

Figure 3.7: VR interview room

3.4 Wizard-of-Oz

In this study, an experimenter acts as a wizard by controlling the avatar by pressing a specific key on the keyboard. Figure 3.8 shows the Wizard-of-Oz control device. The keyboard keys are split under three groups by distinct colors: interview questions, facial expressions, and pre-recorded sounds.

3.5 Questionnaires

Two questionnaires were designed for this evaluating the correctness of the experiment. The first one verified that the facial expressions of the interviewer's avatar were correct. The second one verified that the interview questions designed for each of the stereotyping conditions correctly simulated the three conditions used (i.e. direct, microaggression, and control). Additional details about these two questionnaires, as well as the analysis of the responses provided by the participants, are provided in Chapter 5. The questionnaires themselves are available in the appendices.

3.5.1 Gender Stereotype Questionnaire

Since a standard questionnaire was not found for applying sexist behaviors, different questions were collected from different stereotypical papers [56–61]. We modified these questions to apply to this experiment. This questionnaire has 35 interview phrases and sentences, including two

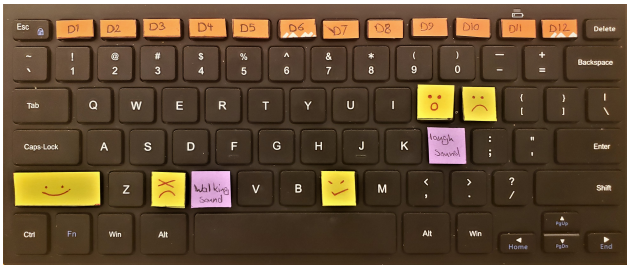


Figure 3.8: The Wizard-of-Oz system

questions itself; the first question asks about the degree of the stereotyping of the question, and the second one asks about the frequency of encountering that question in the real world. Questions are ordered randomly in the questionnaire. The questions are mentioned in the 6. A text-box is provided for the users to report if they have experienced or observed a stereotypical situation before. This questionnaire presented in Appendix B.

3.5.2 Facial Expressions Questionnaire

A standard questionnaire from a previous study [62] was used for evaluating the correctness of facial expressions.This questionnaire has two parts; one that shows the pictures of the interviewer avatar with different facial expressions, and asks participants to select the proper label for the expression from a limited number of options. The other part shows the same pictures again but asks the participant to write the first expression that they think of by seeing the picture. This questionnaire is shown in Appendix A.

Chapter 4

Initial Experiment

The initial design of the experiment was slightly different. The initial results showed no agreement on the feeling of avatar embodiment by the users, but results confirmed the emergence of empathy in the users. The following sections explain the differences and the reasons behind changing the setup in the main experiment:

4.1 Interviewee's avatar

The interviewer's avatar was created using Adobe Fuse, the same as the one was used in the main experiment. However, the interviewee's avatar was created using Oculus's basic tools for creating custom avatars.

Problem: The avatars were built with this tool did not have a complete body (they did not have legs and arms), but their head and hands were synchronized with the participant's movements. Figure 4.1 shows the avatar in the virtual environment. However, the issue with this initial implementation was that the initial results showed there was no agreement on the feeling of avatar embodiment by the users, and this might affect the feeling of empathy in our experiment.

Solution: The main experiment uses full-body avatars instead of half-body avatars. Adobe Fuse was used to design these new avatars.

4.2 Head-mounted Displays

In the initial experiment, an Oculus Rift S head-mounted display was used for experiencing the virtual reality environment. After deciding to import the full-body avatars in the new experiment, the necessity of tracking the feet had emerged .

Problem: Oculus Rift S was not able to track the whole body movements without a third-party device.

Solution: Different options could help to solve this problem:

- **Kinect:** Kinect works with Oculus and can track body movements. However, the problem is that Kinect has a limited line of sight and is not appropriate for our experiment that requires the participants to sit or walk with their avatars.

- **HTC vive trackers:** Vive trackers can track full-body movements. Since Vive trackers do not work with Oculus Rift S, HTC Vive HMD was decided to be used instead of Oculus Rift S.

4.3 Mirror

In the initial experiment, a free asset called mirror [63] was downloaded from the Unity asset store. Also, the mirror was placed on the right side of interviewee's chair.

Problem: The mentioned mirror did not work with HTC Vive properly. It makes a shadow or a second copy of everything. Also, the mirror's quality was not good, and there were some blinking lines in the mirror. Also, some participants expressed that they were not able to see their avatar's body clearly during the interview and that was a potential contributor to the fact that they did not have the feeling of body ownership.

Solution: A provided mirror from Vive Stereo Rendering Toolkit [55] was imported and some features of that was modified to apply to the new experiment. The current mirror has a higher quality. Also, another mirror was added to the front of the interviewee's position (instead of the bookshelf). With these modifications, the participants can see themselves during the interview without turning their heads.

Figure 4.1: The virtual environment contains the half body avatar

4.4 Controlling the facial expressions

For controlling the blend-shapes of interviewer's avatar, in the initial experiment, a Wizard-of-Oz system was used. This system is controlled by the experimenter during the experiment and the facial expression was chosen at the run-time.

Problem: The system was controlled by a human, so it was not behaving the same for all the participants, which could introduce bias in the final results.

Solution: The facial expressions for each question were automated in the final design. The experimenter only controls the time that each question should be asked. The group of facial expressions that accompanies that question automatically is triggered that makes the behavior of the interviewer's avatar consistent in all the experiments.

4.5 Interview questions

The questions used for the initial experiment were mostly directly showing stereotyping. This stereotyping behavior was very explicit.

Problem: The feedback from the participants showed that asking all these direct questions in a row made the interviewer's behavior very aggressive and less frequent.

Solution: In the final design, the direct questions were modified to be more realistic. Also, a new condition called microaggression was added to the final design. In microaggression conditions, the questions have stereotyping, but they are very subtle.

Chapter 5

Survey Results

This chapter breaks the survey results in three sections: gender stereotyping, facial expressions, and subjective response.

5.1 Questionnaire on Gender Stereotypes

This section analyzes the responses collected from the "Questionnaire on Gender Stereotyping". The questions are a 5-point Likert scale. In total, 56 participants answered this questionnaire (15 female and 41 male participants).

Figure 5.1 shows the age range of the participant and the ethnicity of the participants for the surveys accordingly. As the figures show, most of the participants (more than 70%) have an age range of 20 to 25 years. The average age of all the participants is 25.7 years. Also, it shows that most of the participants (almost 60%) are White. Asian group contains the second biggest group of the participants. Also, it shows that participants from the African-American and also Hispanic ethnicity are present among the participants.

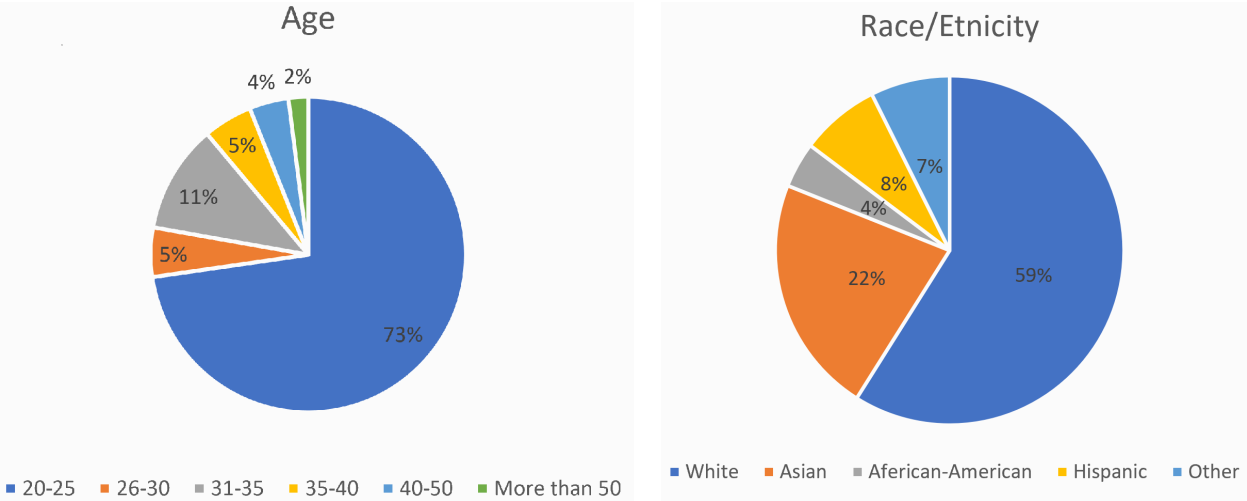


Figure 5.1: Age and race of participants

The severity agreement plot shows the amount of agreement for the questions of the interview in the three conditions. This plot aims to check the appropriateness of the questions of three interview conditions. The questions of each interview can simulate the stereotyping condition for the participants. In this plot, the vertical axis shows the numbers assigned to the levels of severity, ranging from strongly disagree (1) to strongly agree (5). To aggregate the answers, the median of their answers for each of the questions was computed. Then, for each of the interviews, the average of the median scores for its constituent question was computed. The procedure has been repeated for male and female participants separately. Figure 5.2 shows the results for both.

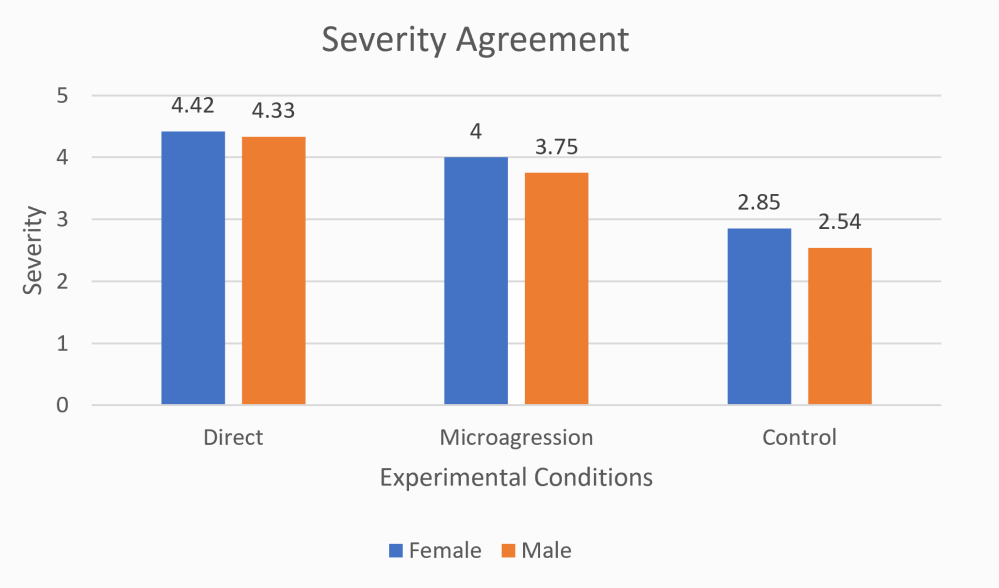


Figure 5.2: Severity of agreement

As the figure shows, there is an overall agreement for the severity of the stereotyping for the questions designed for the three conditions; participants have felt a higher degree of stereotyping for the questions of the direct condition, and less for the microaggression questions. Almost no stereotyping felt for the control condition. The analysis showed that for each of the conditions, female participants had felt a higher degree of stereotyping, which is probably due to a higher

sensitivity toward stereotyping, which seems reasonable because of how they background they have had.

Figure 5.3 aggregates the answers of the participants to the frequency questions. The purpose of these questions is to see if the questions of the interviews are similar to what they have observed in the real-world. As the results show, the control questions have been observed more frequently as opposed to the questions of the other two conditions. Also, the frequency of the questions of the direct condition and microaggression condition is noticeable and above "Sometimes." This figure shows that the questions are similar to those that participants have observed in the real world, and are appropriate to simulate a real interview.

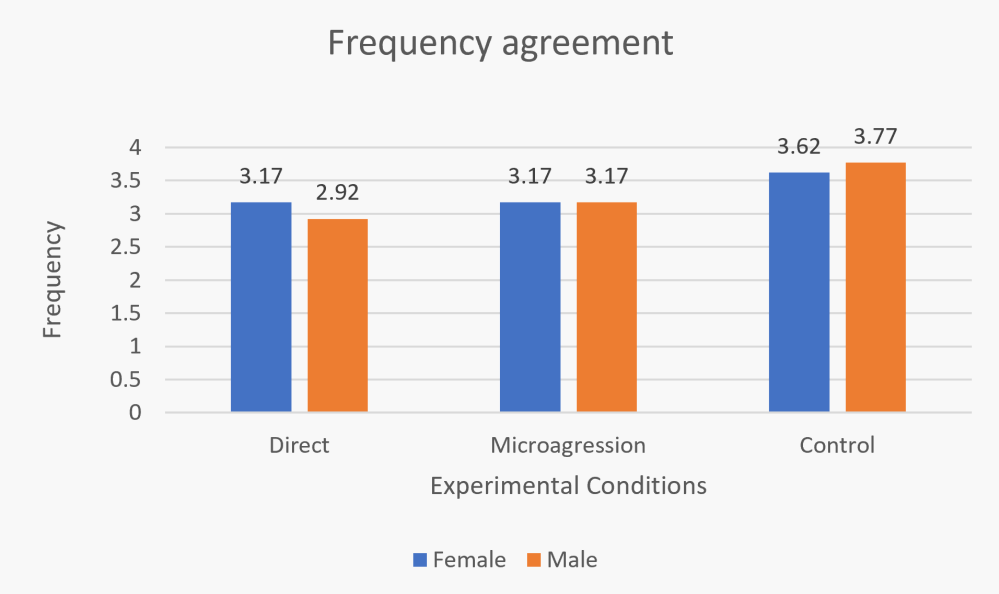


Figure 5.3: Frequency of agreement

Another noticeable observation is that some questions were reported to be rarely seen by male participants. In contrast, the female participants have chosen a much higher frequency of observing those questions in the real world that confirms that although these problems exist, men might have encountered them much less than women. The following examples show some of these questions:

- Q6) To be honest, We were not used to employing female programmers because I do not have a good experience with our female employees.
- Q13) Unfortunately, I doubt a female programmer can be as productive as a man. Can you convince me that I am wrong?
- Q35) I know women might have a problem with staying longer at night. Is it the same for you?

Also the participants were asked to explain one of the experiences they had with stereotyping behaviors. Among 15 female participants, one responded with no experience with this kind of behavior, eight left it blank, and six responded to this question. Also, some of the male participants reported that they had observed these kinds of behaviors. The following section mentions some of the interesting responses to this question:

- "Yes, in a job interview for a software engineer position, they brought up the maternity leave policy. I did not ask about it, but they assumed that since I am a woman, I will be pregnant at some point."
- "Yes, people do not believe I could pick computer science and think I cannot make it."
- "Yes- I am a female in computer science. I have had classmates expect less of me in my work, such as coding assignments. I had also had a few classmates hit on me when the situation was inappropriate."
- "While explaining output of a program for a specific input that was asked to be run despite having run it on my own previously, I was interrupted by a male and told to "just do what I tell you."
- "I have seen it a lot in CS classes, lots of guys think girls cannot code as well."

5.2 Facial Expression Questionnaire

This section reports the responses gathered from the "Questionnaire on Facial Expressions". Similar to the previous questionnaires, questions are 5-point Likert scale. In total, 64 participants answered this questionnaire. The range of the age of participants is 20 to 71. The average age was 25.5. The race of the participants includes White, Asian, African-American, and Hispanic.

Figure 5.4 shows the summary of how much participants agree to the facial expression that is designed in the interviews by the interviewer's avatar. As the figure shows, except for the "Fearful", the agreement is always above 80% and, most of the time, above 90%. The level of agreement for the facial expressions shows satisfactory enough to consider them for our experiments. The "Fearful" facial expression, which shows a low level of agreement, was eliminated from the expressions used in our experiment.

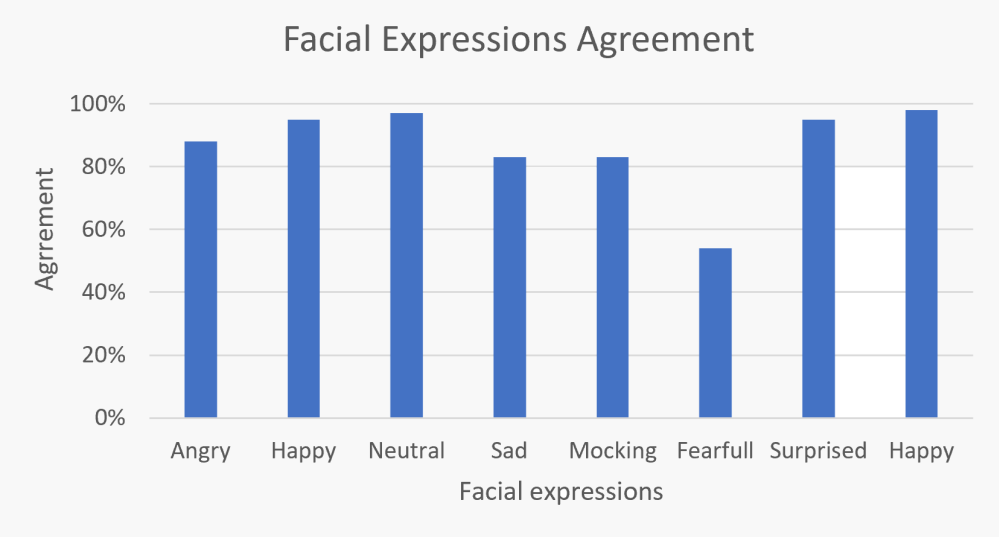


Figure 5.4: The agreement percentage of facial expressions by participants

Chapter 6

Experiments

This chapter explains how the experiment is designed. The main goal is to evaluate the effects of a virtual avatar gender-swap on increasing empathy by experiencing stereotype threat for a job interview task.

Figure 6.1: The Experiment Scene.

6.1 Final Experimental Design

Three different scenarios were implemented for three different experimental conditions. The gender of the interviewee avatar and the type of threat varies among these scenarios. The following shows the three experimental conditions that exist in our experiment:

- **FNT:** Female avatar while there is no threat
- **FMST:** Female avatar under a microaggression stereotype threat
- **FDST:** Female avatar under a direct stereotype threat

The experiment is a within-subjects experiment, and each participant will perform all of these three conditions (FNT, FMST, and FDST). The interviewer's avatar changes for each scenario to

make the difference in behaviors more believable. Questions that the interviewer asks and facial expressions vary in these scenarios to simulate the real world's actual threat. The design of the environment and the taken time is the same.

In stereotype threat conditions, the interviewer performs a sexist behavior in two different ways: microaggression stereotype threats and direct stereotype threats. The sexist behavior are shown in two different ways: the interview's questions and the facial expressions that are shown as a reaction to the interviewee's response (or while the interviewer is asking the questions).

6.2 Avatars

In general, five avatars were designed for this experiment—three avatars for the interviewer's role and two avatars for the interviewee.

Since each participant will experiment with all the three conditions, the same avatar can not be used for all of the conditions. Therefore, three different avatars for three different experimental conditions were designed. All the avatars are used for the interviewer's role wear suits and none of them has glasses on so that the participant can see the eye movements and facial expressions. Figure 6.2 shows these interviewer's avatars.

The interviewee's avatar is always female. For this experiment, both men and women will be recruited. If the participant is male, using the female avatar causes gender-swapping. Therefore, the stereotype threat situation happens in both gender-swapping and salient situations for male participants. On the other hand, if the participant's gender is female, the avatar's gender is the same as the participant's gender. Also, the interviewee's avatars were designed in two races, White and African- American. So that, the avatar with the same participant's race will be used. If neither of the races matches the participant's race, they will be asked to choose the one that they are more comfortable with. This feature helps to increase the feeling of embodiment in the participants. Figure 6.3 shows the avatars that are designed for the interviewee's role.

Figure 6.2: Interviewer's avatars.

Figure 6.3: Interviewee's avatars.

The experiment follows a 3×6 factorial design. The independent variables are experimental conditions (FSMT, FDST, and FNT), and the order is randomly chosen from one of the 6 permutations that can be made with these three experimental conditions. All variables are within-subjects.

6.3 Equipment

This section explains the equipment that are used for this experiment. Figure 6.4 shows this equipment.

- **HTC Vive Head-Mounted Display (HMD):** This headset has a diagonal FOV of 110 degrees and 1080* 1200 per eye resolution in each eye.

- **Two HTC VR Base Stations:** Base stations are used for tracking Vive trackers, Vive headset, and controllers to provide an immersive room-scale virtual reality experience. These base stations put in the height of 2 meters and a distance of 4 meters from each other (by considering the straight line that connects them and constitutes the diagonal of the virtual environment). Each base station has to be connected to different networks; as an example, network b is used by base station one, and base station two connects to network c).

- **Two Stands:** They are used to steadily hold the base stations at the desired height because the base stations do not work if used in a lower height or precarious position.

- **Two HTC Vive Controllers:** These controllers are used to track the avatars' hands.

- **Three Vive Trackers:** This experiment used three trackers for tracking the left and right feet and the hip.

- **Three Dongles:** Three dongles were used to connect the three trackers. This experiment requires multiple dongles because when the controllers are connected, Vive trackers need the dongles to work at the same time as the controllers. The dongles are kept at least 45 cm (18 in) away from the computer and in a fixed position.

- **External USB Hub:** Since the used computer has only three USB ports available, while seven USB ports are needed, this system requires an external USB hub.

- **Two TrackStraps and One TrackBelt:** They are used to fasten the trackers to the body parts.

- **Chair:** An actual chair is used to simulate the chair that the participant is sitting on during the interview.

- **Laptop:** The environment was implemented in Unity and rendered on an Alienware laptop with core-i7 3.8GHz processor CPU, 16GB of RAM, and NVIDIA GeForce GTX 2070 graphics card.
- **Camera:** The whole procedure will be recorded using a camera. For this purpose, an action camera, brave 4, is used.

Figure 6.4: Equipment

For making the experiment process more user friendly, some stickers on trackers were used, so that it becomes easy to distinguish left and right hands (feet) from each other. Figure 6.5 shows it.

Figure 6.5: Making the experiment user friendly

Figure 6.6 shows a user wearing headset and trackers. As you can see in the picture, three trackers are used for the left foot,the right foot and the hip.

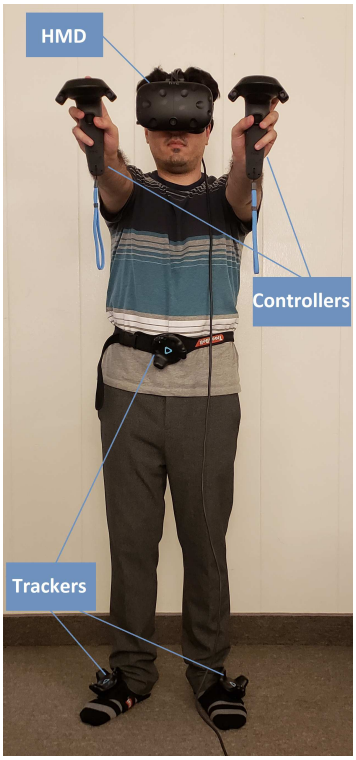


Figure 6.6: A user wearing headset and trackers

6.4 Interview questions

The following section shows the interview questions for three experimental conditions. Since a single standard design was not found for our interview questions, which simulates a stereotype threat condition, the questions from different references were collected and modified to be applicable to our purpose. Also, to make sure that the questions are appropriate and reflect what we require, we prepared a survey. This survey shows the degree of stereotyping of each of these questions and the frequency of which they are used in the real world. The results of this survey will be discussed in the results section.

Table 6.1: Interview's questions for FMST condition

I) Microaggressions/ Implicit stereotyping
Verbal
1) Hi! My name is David! Thank you for applying for this position. There are no female programmers in our company. Could you please introduce yourself? [56]
2) Our company is a small computer science company. The purpose of this company is to provide smart home related applications.
3) We were not used to employing female programmers but fortunately we have started to employ female programmers this year [57].
4) Do you think women are interested in this kind of positions?
5) My wife had difficulty in coding, too, but she worked hard to get ahead in her work. [58]
6) I know women can work as good as men.
7) Are you interested in coding? [59]
8) You are too pretty! Have you ever thought to become a model? [60]
9) Are you married? - [Yes] => You are amazing, I can't believe you can do both. - [No] => I understand, it is really hard to do both
10) No worries, I don't think of you as a regular female employee. You look very strong to me [58].
11) We have some benefits for women. Did you know that you, as a woman, can get extra off days? You-as a potentially pregnant person- can have more days off which is, of course, a good thing for you [61].
12) Later, If you have a problem in your project, there are a lot of men that can help you with your problems [57].
13) Sometimes we need our employee to stay late, do you think you can do that?
14) Thank you for applying for this position. We send you an email as soon as we decide. Bye.
Behavioral
1) Interrupting the participant before she finishes her response. (ex: Q6)
2) Using Smiley and Mocking face more

Table 6.2: Interview's questions for FMDT condition

II) Direct/ Explicit stereotyping
Verbal
1) Hi! Finally, a female applicant. I have to tell you there are no, really, there are no female programmers in our company. My name is David. Could you please introduce yourself?
2) Our company is a small computer science company. The purpose of this company is to provide smart home-related applications.
3) To be honest, We were not used to employing female programmers because I don't have a good experience with our female employees.
4) I think some people have difficulty with this kind of job position. Is it not the same for you?
5) Often women have difficulty in coding, but I think you can work hard to get ahead in your work.
6) Unfortunately, I doubt a female programmer can be as productive as a man. Can you convince me that I am wrong?
7) Do you have any experience with coding?
8) You are too pretty to be a good programmer. [Mocking face]
9) Are you married? - [Yes] => I can't believe you can do both. - [No] => It is easy to guess, it is really hard to do both.
10) No worries, you might be stronger than other females.
11) We have some benefits for women. Did you know that you, as a woman, can get extra off days because of pregnancy person- can have more days off.
12) I understand females might have problems in coding. If you have a problem with your project later, you can ask your colleagues to help you.
13) Sometimes we need our employee to stay late, do you think you can do that?
OK, Thank you. We will send you an email about the results. Bye.
Behavioral
1) Interrupting the participant while expressing sorry we are out of time (ex Q8)
2) Using mocking and angry face more

Table 6.3: Interview's questions for FNT condition

III) Normal/ Control
Verbal
1) Hi! My name is David. Could you please Introduce yourself?
2) Our company is a small computer science company. The purpose of this company is to provide smart home related applications.
3) Every year we employ female/male programmers to help us improve the company
4) I think some people have difficulty with this kind of job position. Is it not the same for you?
5) I had difficulty in coding, too, but I worked hard to get ahead in my work.
6) I know you can work well based on your resume.
7) Are you good at coding?
8) You look smart.
9) N/A
10)No worries! I think you have a good background.
11) We have some benefits for our employees. Did you know that our employees can get extra off days based on special situations?
12) Later, If you have a problem in your project, there are a lot of men that can help you with your problems.
13) Just as a warning, sometimes we need our employees to work longer in a day.
14) Thank you for applying for this position. We send you an email as soon as we decide. Bye.
Behavioral
1) More smiley face
2) Not interrupting the interviewee
3) Having eye contact with the interview when he/she talks (most of the time)

6.5 Procedure

The following sections show the procedure of the experiment and how it will be conduct it in the future.

Each experiment is divided into four tasks (the order of the last three ones are chosen randomly):

- Getting familiar with interviewee's avatar
- Giving the interview using a female avatar under microaggression stereotype threat condition (FMST)
- Giving the interview using a female avatar while there is no threat(FNT)
- Giving the interview using a female avatar under direct stereotype threat condition (FDST)

All the participants will be tested individually in the NUILab of the Computer Science department at Colorado State University. Upon arrival at the location of the experiment, each of the participants will be greeted by the experimenter and will fill and sign a written consent form. Then, participants will fill in a demographic questionnaire as well as an empathy pre-questionnaire. The experimenter explains the tasks, and the participants will become aware of the procedure or the experiment and how the virtual environment simulates an in-person interview for a job. Then, the experimenter explains how to wear trackers and helps the participant wear the trackers correctly.

The experimenter should record the whole experiment for each of the participants using a pre-setup camera. The amount of delay in the participants' response and the time it takes for them to answer a question will be kept track of using a smartphone.

In the beginning, after running the environment, the participants will be asked to take a look at the virtual environment. There is a mirror on the wall at the right-hand side of the avatar. They will watch themselves in the mirror while performing simple tasks (moving the head to the right and left, moving the hands). Participants will see the avatar moves as they move the body parts that help them get acquainted with their avatar.

Afterward, the experimenter will ask the participants if they are ready for the interview to start. Once participants consent, the experimenter will ask them to sit on the chair and will press the specific key on the keyboard to initiate the entrance of the male interviewer from a door at the side of the room, which is behind the chair used for the interviewee's avatar. As the interviewer enters the room, the participant hears the sound of walking plays through the speakers of HMD to make the interviewee aware of his entrance. Once the interviewer reaches the chair in front of the interviewee, the interviewer starts his pre-recorded greeting. Using the input through the keyboard from the experimenter,

The interviewer will ask the pre-recorded questions using the keyboard controlled by the experimenter once the timing is correct. As explained before, the avatar performs the facial expressions automatically while asking the questions.

For each experimental condition, the interviewer's avatar will ask the prepared questions in the specified order. In general, each experimental condition includes 13 statements and questions that are about 4 minutes in total. After finishing the first trial, participants will answer an avatar embodiment questionnaire. Then, there will be a break between the first and second experimental conditions.

For the second experimental condition, participants experience a similar environment, but the interviewer's avatar, along with his questions and facial expressions, change based on the stereotypical conditions. They will take a 5-minutes rest and perform the third condition experiment, for which the interviewer's avatar and questions will change accordingly.

In the end, after finishing the third experiment, the participant will fill out an empathy post-questionnaire.

The total experiment time depending on the length of the participants' answers, can vary. In general, we estimate that each experimental condition task will last for ten minutes for each participant. In total, performing the whole experiment (including the time that it takes for them to get the instructions and filling out the questionnaires) will last for 50 minutes. Some participants will be compensated with extra credits.

6.6 Measurements

It is hard to measure stereotyping and empathy because they are qualitative, and also people are unwilling to admit their negative attitudes [29] what they think about specific social groups. For this purpose, two measurements were considered to evaluate the stereotyping degree of the questions, the feeling of body-ownership, and their effect on increasing empathy.

6.6.1 Response Time

In each trial, the interviewer asked some questions, and the interviewee should answer the question after that. The experimenter will record the response time for answering each question as a measure of stereotyping behavior. The hypothesize is that participants will answer the questions that represent stereotyping behavior with more delay.

6.6.2 Questionnaires

The subjective evaluation of the experiment will be assessed using the questionnaires. For this purpose, four questionnaires that each one has a specific purpose were designed:

Embodiment Questionnaires

These questionnaires aimed to evaluate the feeling of body-ownership that users perceive during the experiment by both female and male avatars. This study uses the standard questionnaire in [9, 64]. For this purpose, we modified the provided questionnaire to be able to apply to our experiment. We eliminated the questions that did not apply to our study context. Also, we added questions 9 and 11 that seem to be proper and relevant for our purpose. Authors in [9,64] designed the questionnaire for a hand ownership study. Therefore, the questionnaire was modified to apply to our experiment for a full-body avatar.

Participants will fill in the questionnaire after two experimental conditions (FMST/FDST and FNT). The questions are shown in Table 6.4. The scores of each response are given on a five-point Likert scale (-2 = strongly disagree; 2 = strongly agree). For the reversed questions, $-i$ was used, where i is the answer of the participants to the question.

Empathy Questionnaires

A standard questionnaire was used and modified to apply to our purpose [65]. The empathy pre-questionnaire will be given to participants before they start performing all tasks, and participants will answer a post-questionnaire after performing the tasks. Since the questionnaires are similar, only the table for the post-questionnaire is shown here. Similar to the embodiment questionnaires, a five-point Likert-type scale is used for these questions as well. A previous study by Toronto University divided the questions into two groups:

- **Understanding Others:** The questions of this section aimed to measure the degree of empathy felt by the users.
- **Communicate Understanding:** The questions aimed to measure the effect of the experiment on the interviewee and how the interviewee elicits a response to sexist behavior.

Table 6.4: Embodiment Questionnaire

Embodiment Questionnaire
1. I felt as if the virtual body was my own body.
2. It seems as if I had more than one body.
3. I felt as if my body was turning virtual.
4. I felt as if I could control the movements of the virtual body.
5. The virtual body was obeying my will, and I can make it move just like I want it.
6. I felt as if the virtual body was controlling my will.
7. It seemed as if the virtual body had a will of its own.
8. I felt as if the virtual body was controlling me.
9. I felt that I was wearing different clothing than my real clothing.
10. I felt that the body in the virtual world was not me.
11. I felt comfortable with the gender of avatar

Table 6.5: Empathy Post-Questionnaire

Empathy Formative post-Questionnaire
Understand Others
1. I was able to see things from females' point of view
2) When I didn't agree with Interviewer's point of view about women, I expressed my disagreement
3) When I disagree with Interviewer's perspective about women , it was hard for me to understand his perspective.
4) I considered females' circumstances when I was talking to the interviewer
5) I tried to imagine how I would feel if I were a woman.
6) When a female has programming problems, I won't make it generalize to all women.
7) I think I can understand women's circumstances better now.
8) I will consider females' circumstances when I am talking to them.
Communicate Understanding
1) When I see a woman who feels bad about her programming skills, I will try to encourage her.
2) I will tell my friends things like "we should have more women in computer science" or "companies should start hiring more female programmers"
3) When I know one of my female friends is not very confident about programming skill, I will try to talk to them about it.
4) When I see someone making a stereotype situation, I will express my disagreement.
5) When I know one of my female friends is not very confident about programming skill, I will try to talk to them about it
6) When a woman feels bad about her programming skills, I will try to encourage her
7) I will try to convince my friends that women must have the opportunity to work in software companies.

Chapter 7

Limitations

There are some limitations in this work that will be addressed in the future. This section mentions some of those limitations:

- Fewer female participants answered the Gender Stereotype Threat questionnaire compared to the number of male participants.
- The avatars are only designed for two races, White and African- American, that might affect the feeling of an embodiment for the participants of other races.
- Our tool does not recognize the facial expressions and eye movements of participants. That might also affect the feeling of embodiment. With the current technology that were used for our experiment, this was not possible, but some headsets have eye-movement detection. However, detecting the facial expression still might not be practical because the headset covers the face.
- The interviewee's avatars do not have lip-syncing, and it might affect the feeling of embodiment if the users notice there is no lip movement while they are talking and looking at the mirror.

Chapter 8

Conclusion and Future Work

In this project, a tool that can potentially increase the empathy of the opposite gender in computer science was designed. To evaluate the effect of gender-swap illusion in virtual reality for increasing empathy, we designed an environment where users can experience an interview under stereotype threat conditions. We designed appropriate questionnaires to evaluate the impact of the experiment on the users. Our set-up still suffers from some shortcomings in the design that causes a decrease in the feeling of embodiment, and because that can be effective on the results for the empathy, we will address them in the future.

Also, two surveys were designed to verify the appropriateness of the interview's question for each of the conditions and correctness of facial expressions for the interviewer's avatars.

In the future, We will look for more female participants and analyze the gander stereotype questionnaire again, although the results are already confirming what we expected. Also, we will design more avatars for various ethnicity for the interviewee's avatar.

Bibliography

[1] American enterprise institution-comparing gender share in science. https://www.aei.org/. Accessed July 10, 2020.

[2] American enterprise institution- female share by major. https://www.aei.org/. Accessed July 10, 2020.

[3] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, volume 2351, pages 282–292. International Society for Optics and Photonics, 1995.

[4] Makehuman. http://www.makehumancommunity.org/. Accessed July 10, 2020.

[5] Google poly. https://poly.google.com/view/aNby0pcw595/. Accessed July 10, 2020.

[6] Adobe fuse cc. ttps://www.adobe.com/products/fuse.html. Accessed July 10, 2020.

[7] Character creator 3. https://www.reallusion.com/character-creator/. Accessed July 10, 2020.

[8] Microsoft rocketbox master library. https://github.com/microsoft/Microsoft-Rocketbox/. Accessed July 10, 2020.

[9] Tabitha C Peck, My Doan, Kimberly A Bourne, and Jessica J Good. The effect of gender body-swap illusions on working memory and stereotype threat. *IEEE transactions on visualization and computer graphics*, 24(4):1604–1612, 2018.

[10] C Ashcraft, B McLain, and E Eger. Women in tech: The facts 2016 update. see what's changed and what hasn't. *NCWIT. Available from: https://www. ncwit. org/sites/default/files/resources/ncwit_women-in-it_2016-full-report_final-web06012016. pdf*, 2016.

[11] Adriana D Kugler, Catherine H Tinsley, and Olga Ukhaneva. Choice of majors: Are women really different from men? Technical report, National Bureau of Economic Research, 2017.

[12] Jane M Scholl and Steven R Sabat. Stereotypes, stereotype threat and ageing: implications for the understanding and treatment of people with alzheimer's disease. *Ageing & Society*, 28(1):103–130, 2008.

[13] Jing-Kai Lou, Kunwoo Park, Meeyoung Cha, Juyong Park, Chin-Laung Lei, and Kuan-Ta Chen. Gender swapping and user behaviors in online social games. In *Proceedings of the 22nd international conference on World Wide Web*, pages 827–836. ACM, 2013.

[14] Rabindra Ratan and Young June Sah. Leveling up on stereotype threat: The role of avatar customization and avatar embodiment. *Computers in Human Behavior*, 50:367–374, 2015.

[15] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.

[16] Chin-Chang Ho and Karl F MacDorman. Revisiting the uncanny valley theory: Developing and validating an alternative to the godspeed indices. *Computers in Human Behavior*, 26(6):1508–1518, 2010.

[17] Claude M Steele and Joshua Aronson. Stereotype threat and the intellectual test performance of african americans. *Journal of personality and social psychology*, 69(5):797, 1995.

[18] Ronald A Berk. Microaggressions trilogy: Part 3. microaggressions in the classroom. *The Journal of Faculty Development*, 31(3):95–110, 2017.

[19] Derald Wing Sue, Christina M Capodilupo, Gina C Torino, Jennifer M Bucceri, Aisha Holder, Kevin L Nadal, and Marta Esquilin. Racial microaggressions in everyday life: implications for clinical practice. *American psychologist*, 62(4):271, 2007.

[20] Niels Ole Bernsen, Laila Dybkjær, and Hans Dybkjær. Developing interactive speech systems. In *Designing Interactive Speech Systems*, pages 61–85. Springer, 1998.

[21] Christoph BENZM ULLER, Armin Fiedler, Malte Gabsdil, Helmut Horacek, Manfred PINKAL, Jrg SIEKMANN, Dimitra TSOVALTZI, Bao Quoc VO, and Magdalena WOLSKA. A wizard-of-oz experiment for tutorial dialogues in mathematics. 2003.

[22] Ruth Van Veelen, Belle Derks, and Maaike Dorine Endedijk. Double trouble: How being outnumbered and negatively stereotyped threatens career outcomes of women in stem. *Frontiers in psychology*, 10:150, 2019.

[23] Christine R Starr, Barrett R Anderson, and Katherine A Green. "i'ma computer scientist!": Virtual reality experience influences stereotype threat and stem motivation among undergraduate women. *Journal of Science Education and Technology*, 28(5):493–507, 2019.

[24] Adrian Opre and Dana Opre. The gender sterotype threat and the academic performance of women's university teaching staff. *Journal for the Study of Religions and Ideologies*, 5(14):41–50, 2010.

[25] Iris Wigger. Anti-muslim racism and the racialisation of sexual violence:'intersectional stereotyping'in mass media representations of male muslim migrants in germany. *Culture and Religion*, 20(3):248–271, 2019.

[26] Maykel Verkuyten, Jochem Thijs, and Nadya Gharaei. Discrimination and academic (dis) engagement of ethnic-racial minority students: a social identity threat perspective. *Social Psychology of Education*, 22(2):267–290, 2019.

[27] Kevin L Nadal, Katie E Griffin, Sahran Hamit, Jayleen Leon, Michael Tobio, and David P Rivera. Subtle and overt forms of islamophobia: Microaggressions toward muslim americans. 2012.

[28] Arielle A Brown. *Stereotype threat and Black women's math performance: Do racial identity and experimenter race matter?* PhD thesis, University of Illinois at Urbana-Champaign, 2019.

[29] Denise Sekaquaptewa, Penelope Espinoza, Mischa Thompson, Patrick Vargas, and William von Hippel. Stereotypic explanatory bias: Implicit stereotyping as a predictor of discrimination. *Journal of Experimental Social Psychology*, 39(1):75–82, 2003.

[30] Steven J Spencer, Claude M Steele, and Diane M Quinn. Stereotype threat and women's math performance. *Journal of experimental social psychology*, 35(1):4–28, 1999.

[31] Felix Chang, Mufan Luo, Gregory Walton, Lauren Aguilar, and Jeremy Bailenson. Stereotype threat in virtual learning environments: Effects of avatar gender and sexist behavior on women's math learning outcomes. *Cyberpsychology, Behavior, and Social Networking*, 22(10):634–640, 2019.

[32] Jessica J Good and Laurie A Rudman. When female applicants meet sexist interviewers: The costs of being a target of benevolent sexism. *Sex Roles*, 62(7-8):481–493, 2010.

[33] Zaheer Hussain and Mark D Griffiths. Gender swapping and socializing in cyberspace: An exploratory study. *CyberPsychology & Behavior*, 11(1):47–53, 2008.

[34] Jorge Peña, Subuhi Khan, and Cassandra Alexopoulos. I am what i see: How avatar and opponent agent body size affects physical activity among men playing exergames. *Journal of Computer-Mediated Communication*, 21(3):195–209, 2016.

[35] Mel Slater, Bernhard Spanlang, Maria V Sanchez-Vives, and Olaf Blanke. First person experience of body transfer in virtual reality. *PloS one*, 5(5), 2010.

[36] Anouk Keizer, Annemarie van Elburg, Rossa Helms, and H Chris Dijkerman. A virtual reality full body illusion improves body image disturbance in anorexia nervosa. *PloS one*, 11(10), 2016.

[37] Ben Kuchera. Being someone else: How virtual reality is allowing men and women to swap bodies, 2014.

[38] Sarah Lopez, Yi Yang, Kevin Beltran, Soo Jung Kim, Jennifer Cruz Hernandez, Chelsy Simran, Bingkun Yang, and Beste F Yuksel. Investigating implicit gender bias and embodiment of white males in virtual reality with full body visuomotor synchrony. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–12, 2019.

[39] Nancy K Baym, Yan Bing Zhang, and Mei-Chen Lin. Social interactions across media: Interpersonal communication on the internet, telephone and face-to-face. *New Media & Society*, 6(3):299–318, 2004.

[40] Interactive 3d characters social virtual reality course. https://www.coursera.org/. Accessed July 10, 2020.

[41] Daz3d. https://www.daz3d.com/. Accessed July 10, 2020.

[42] Maxiamo. https://www.mixamo.com/. Accessed July 10, 2020.

[43] Unity. https://unity.com/. Accessed July 10, 2020.

[44] Vive. https://www.vive.com/. Accessed July 10, 2020.

[45] Roel Vertegaal, Robert Slagter, Gerrit Van der Veer, and Anton Nijholt. Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 301–308, 2001.

[46] Realistic eye movements. https://assetstore.unity.com/. Accessed July 10, 2020.

[47] The nature of things. https://www.cbc.ca/natureofthings/. Accessed July 10, 2020.

[48] Tomoko Koda and Toru Ishida. Cross-cultural study of avatar expression interpretations. In *International Symposium on Applications and the Internet (SAINT'06)*, pages 7–pp. IEEE, 2006.

[49] Paul Ekman. Basic emotions. *Handbook of cognition and emotion*, 98(45-60):16, 1999.

[50] Magali Batty and Margot J Taylor. Early processing of the six basic facial emotional expressions. *Cognitive Brain Research*, 17(3):613–620, 2003.

[51] Gerard Llorach, Alun Evans, Josep Blat, Giso Grimm, and Volker Hohmann. Web-based live speech-driven lip-sync. In *2016 8th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, pages 1–4. IEEE, 2016.

[52] Steamvr plugin. https://assetstore.unity.com/packages/tools/integration/steamvr-plugin-32647. Accessed July 10, 2020.

[53] Valve software. https://valvesoftware.github.io/. Accessed July 10, 2020.

[54] Final ik. https://assetstore.unity.com/. Accessed July 10, 2020.

[55] Mirror steamvr. https://assetstore.unity.com/packages/templates/systems/steam-vr-playmaker-toolkit-63425. Accessed July 10, 2020.

[56] Amy J Anderson, Bernadette Sánchez, Christine Reyna, and Hector Rasgado-Flores. " it just weighs in the back of your mind": Microaggressions in science. *Journal of Women and Minorities in Science and Engineering*, 26(1), 2020.

[57] Derald Wing Sue. Microaggressions, marginality, and oppression. *Microaggressions and marginality*, pages 3–22, 2010.

[58] Madonna G Constantine, Laura Smith, Rebecca M Redington, and Delila Owens. Racial microaggressions against black counseling and counseling psychology faculty: A central challenge in the multicultural counseling movement. *Journal of Counseling & Development*, 86(3):348–355, 2008.

[59] Laurie A Rudman and Eugene Borgida. The afterglow of construct accessibility: The behavioral consequences of priming men to view women as sexual objects. *Journal of Experimental Social Psychology*, 31(6):493–517, 1995.

[60] Microaggressions-be careful what you say. https://www.npr.org/. Accessed July 10, 2020.

[61] Sabine C Koch, Stefan Konigorski, and Monika Sieverding. Sexist behavior undermines women's performance in a job application situation. *Sex roles*, 70(3-4):79–87, 2014.

[62] Christian Becker-Asano and Hiroshi Ishiguro. Evaluating facial displays of emotion for the android robot geminoid f. In *2011 IEEE workshop on affective computational intelligence (WACI)*, pages 1–8. IEEE, 2011.

[63] Mirror. https://assetstore.unity.com/packages/tools/network/mirror-129321. Accessed July 10, 2020.

[64] Mar Gonzalez-Franco and Tabitha C Peck. Avatar embodiment. towards a standardized questionnaire. *Frontiers in Robotics and AI*, 5:74, 2018.

[65] Research collaboration. http://www.researchcollaboration.org/. Accessed July 10, 2020.

Appendix A

Facial Expressions Questionnaire

Facial Expressions

Thank you for taking the time to answer these questions. The purpose of this survey is to collect information about the correctness of the facial expression of the avatars that are used in my experiment. Your information are confidential. The required fields are shown with an asterisk.

Name *

First Name Last Name

Email *

example@example.com

Age

Gender *

- Female
- Male
- None-binary
- Other

Race/Ethnicity *

- White
- Native american
- Asian
- African american
- Hispanic
- Non-Resident Alien (of any race or ethnicity)

Part One

Please assign labels to the pictures, if you find a suitable label. Then, select how strong you think the label matches the picture.

Label *

Matching Percentage *

Label *

Matching Percentage *

Label *

Matching Percentage *

Label *

Matching Percentage *

Label *

Matching Percentage *

Label *

Matching Percentage *

Back Next

Label *

Neutral
Sad
Angry
Happy
Fearful
Surprised
Mocking
None of these labels

Label *

Matching Percentage *

Label *

Matching Percentage *

Back Next

Part Two

You are requested to type one word into the text field below the picture if you think there is no proper label for that picture in the previous page. If for some picture you have no idea, you might leave the corresponding text field blank. Anything that comes to your mind when looking at the pictures is fine. There are no 'correct' choices.

F1

F2

F3

F4

F5

F6

F7

F8

Back Next

F9

Back Next

Thank you for your time and participation

Do you want to be considered for the next surveys/experiments?

- Yes
- No
- Maybe

Comments

Submit Test!

Appendix B

Gender Stereotype Questionnaire

Demographic

Questionnaire on Gender Stereotyping

This is an anonymous survey and nobody will know your identity. Please answer the questions honestly. Thank you in advance!

Gender

- ○ Male
- ○ Female
- ○ None-binary
- ○ Others

Race/ ethnicity

- ○ African american
- ○ White
- ○ Asian
- ○ Native american
- ○ Hispanic
- ○ Prefer not to say
- ○ Others

Age

Major/Occupation

Description and goal

The purpose of this survey is to validate the degree of stereotyping of questions of a job interview.

The following questions are going to be used in a simulated computer science job interview experiment in virtual reality.

Each question has two parts:
Part A: You are asked to determine the degree of stereotyping.
Part B, you are asked to determine the frequency of that the question in real world .
(Some questions are designed for a women, if you are a man, put yourself in a woman's shoe)

Part 1

Q1) Hi! My name is David! Thank you for applying for this position. There are no female programmers in our company. Could you please introduce yourself?

Q1-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q1-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q2) Hi! Finally a female applicant. I have to tell you there are no, really, there are no female programmers in our company. My name is David. Could you please introduce yourself?

Q2-A) This is a stereotypical question.

- ○ Strongly agree

- ◯ Somewhat agree
- ◯ Neither agree nor disagree
- ◯ Somewhat disagree
- ◯ Strongly disagree

Q2-B) How frequent do you think this happens in real life?

- ◯ Very frequently
- ◯ Frequently
- ◯ Sometimes
- ◯ Rarely
- ◯ Never

Q3) Hi! My name is David. Could you please Introduce yourself?

Q3-A) This is a stereotypical question.

- ◯ Strongly agree
- ◯ Somewhat agree
- ◯ Neither agree nor disagree
- ◯ Somewhat disagree
- ◯ Strongly disagree

Q3-B) How frequent do you think this happens in real life?

- ◯ Very frequently
- ◯ Frequently
- ◯ Sometimes
- ◯ Rarely
- ◯ Never

Part 2

Q4) Every year we employ female/male programmers to help us improve the company.

Q4-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q4-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q5) We were not used to employing female programmers but fortunately we have started to employ female programmers this year.

Q5-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q5-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q6) To be honest, We were not used to employing female programmers because I don't have a good experience with our female employees.

Q6-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q6-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Part 3

Q1) I think some people have difficulty with this kind of job position. Is it not the same for you?

Q7-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q7-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently

- ○ Sometimes
- ○ Rarely
- ○ Never

Q8) I think women have difficulty with these kind of positions. Is it not the same for you?

Q8-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q8-B) How frequent do you think this happens in real life8

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q9) Do you think women are interested in these kind of positions?

Q9-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q9-B) How frequent do you think this happens in real life9

- ○ Very frequently

- ◯ Frequently
- ◯ Sometimes
- ◯ Rarely
- ◯ Never

Part 4

Q10) My wife had difficulty in coding, too, but she worked hard to get ahead in her work.

Q10-A) This is a stereotypical question.

- ◯ Strongly agree
- ◯ Somewhat agree
- ◯ Neither agree nor disagree
- ◯ Somewhat disagree
- ◯ Strongly disagree

Q10-B) How frequent do you think this happens in real life?

- ◯ Very frequently
- ◯ Frequently
- ◯ Sometimes
- ◯ Rarely
- ◯ Never

Q11) I had difficulty in coding, too, but I worked hard to get ahead in my work.

Q11-A) This is a stereotypical question.

- ◯ Strongly agree
- ◯ Somewhat agree
- ◯ Neither agree nor disagree
- ◯ Somewhat disagree
- ◯ Strongly disagree

- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q13-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q14) I know women can work as good as men.

Q14-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q14-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q15) I know you can work well based on your resume.

Q15-A) This is a stereotypical question.

- ○ Strongly agree

- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q15-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Part 6

Q16) Are you good at coding?

Q16-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q16-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q17) Are you interested in coding?

Q17-A) This is a stereotypical question.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

Q17-B) How frequent do you think this happens in real life?

- Very frequently
- Frequently
- Sometimes
- Rarely
- Never

Q18) Do you have any experience with coding?

Q18-A) This is a stereotypical question.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

Q18-B) How frequent do you think this happens in real life?

- Very frequently
- Frequently
- Sometimes
- Rarely
- Never

Part 7

Q19) You look smart.

Q19-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q19-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q20) You are too pretty to be a good programmer.

Q20-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q20-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q21) You are too pretty! Have you ever thought to become a model?

Q21-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q21-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Part 8

Q22) Are you married?
- [Yes] => I can't believe you can do both.
- [No] => It is easy to guess, it is really hard to do both.

Q22-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q22-B) How frequent do you think this happens in real life?

- ○ Very frequently

- ◯ Frequently
- ◯ Sometimes
- ◯ Rarely
- ◯ Never

Q23) Are you married?
- [Yes] => You are amazing, I can't believe you can do both.
- [No] => I understand, it is really hard to do both

Q23-A) This is a stereotypical question.

- ◯ Strongly agree
- ◯ Somewhat agree
- ◯ Neither agree nor disagree
- ◯ Somewhat disagree
- ◯ Strongly disagree

Q23-B) How frequent do you think this happens in real life?

- ◯ Very frequently
- ◯ Frequently
- ◯ Sometimes
- ◯ Rarely
- ◯ Never

Part 9

Q24) No worries, you might be stronger than other females.

Q24-A) This is a stereotypical question.

- ◯ Strongly agree
- ◯ Somewhat agree
- ◯ Neither agree nor disagree
- ◯ Somewhat disagree

- ○ Strongly disagree

Q24-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q25) No worries, I don't think of you as a regular female employee. You look very strong to me.

Q25-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q25-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q26) No worries! I think you have a good background.

Q26-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree

- ◯ Neither agree nor disagree
- ◯ Somewhat disagree
- ◯ Strongly disagree

Q26-B) How frequent do you think this happens in real life?

- ◯ Very frequently
- ◯ Frequently
- ◯ Sometimes
- ◯ Rarely
- ◯ Never

Part 10

Q27) We have some benefits for our employees. Did you know that our employees can get extra off days based on special situations?

Q27-A) This is a stereotypical question.

- ◯ Strongly agree
- ◯ Somewhat agree
- ◯ Neither agree nor disagree
- ◯ Somewhat disagree
- ◯ Strongly disagree

Q27-B) How frequent do you think this happens in real life?

- ◯ Very frequently
- ◯ Frequently
- ◯ Sometimes
- ◯ Rarely
- ◯ Never

Q28) We have some benefits for women. Did you know that you as a woman can get extra off days? You -as a potentially pregnant person- can have more days off. Do you plan to give birth

to a child in the next 2 years?

Q28-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q28-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q29) We have some benefits for women. Did you know that you as a woman can get extra off days? You-as a potentially pregnant person- can have more days off which is of course a good thing for you.

Q29-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q29-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely

Part 11

Q30) Later, If you have a problem in your project, there are a lot of men that can help you with your problems.

Q30-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q30-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q31) If you later have a problem in your project, don't worry, you can count on your colleagues.

Q31-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q31-B) How frequent do you think this happens in real life?

- ○ Very frequently

- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q32) I understand females might have problems in coding. If you have a problem in your project later, you can ask your colleagues to help you.

Q32-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q32-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Part 12

Q33) Sometimes we need our employee to stay late, do you think you can do that?

Q33-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q33-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q34) Just as a warning, sometimes we need our employees to work longer in a day.

Q34-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q34-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Q35) I know women might have problem with staying longer at night. Is it the same for you?

Q35-A) This is a stereotypical question.

- ○ Strongly agree
- ○ Somewhat agree
- ○ Neither agree nor disagree
- ○ Somewhat disagree
- ○ Strongly disagree

Q35-B) How frequent do you think this happens in real life?

- ○ Very frequently
- ○ Frequently
- ○ Sometimes
- ○ Rarely
- ○ Never

Part 13

Have you experienced or seen an stereotype threat situation in your life? If yes could you please explain it.

Thank you for taking this survey.

If you are enrolled in CS464 course, please send an email to faezeborhani@gmail.com with the code that you will get after submitting this survey.