THE PENNSYLVANIA STATE UNIVERSITY
MILLENNIUM SCHOLARS PROGRAM

DEPARTMENT OF AEROSPACE ENGINEERING

THE EFFECTS OF GENDER AND PLACEBIC MESSAGING ON ROBOT-GUIDED EMERGENCY EVACUATION

CIERA MCFARLAND
SPRING 2021

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Aerospace Engineering

Reviewed and approved* by the following:

Alan Wagner
Assistant Professor of Aerospace Engineering
Thesis Supervisor

Katie Fitzsimons
Assistant Professor of Mechanical Engineering
Reviewer

* Signatures are on file in the Millennium Scholars Program office.

ABSTRACT

Evacuation robots have the potential to prevent dangerous bottlenecking at exits during emergencies, potentially saving lives. However, these robots must compel a person to follow their instructions. The goal of this study was to consider how the gender of a robot's verbalizations and the nature of its message impacts whether or not people follow it rather than a crowd of people. This experiment utilized sounds in both the robot's messaging and the environment to create a more realistic scenario. Participants followed a robot through an office space within a simulation. During the emergency, the robot played a male or female audio message asking if the person wanted to follow it with no explanation, a placebic explanation, or a helpful explanation. Simulated people ran one way while the robot moved in the opposite direction. A siren sounded, and a timer counted down to increase the sense of urgency. While it was believed that any explanation would cause more people to follow the robot, no message seemed to have a greater influence than any other. Contrary to other studies, male participants did not favor either robot, while female participants slightly favored the male-voiced robot. The presence of audio in multiple places had a significant influence over behavior, causing a greater number of people to ignore the robot's guidance than was expected.

TABLE OF CONTENTS

LIST OF FIGURES iii

LIST OF TABLES iv

ACKNOWLEDGEMENTS v

Chapter 1 Introduction 6

Chapter 2 Related Work 8

Chapter 3 Simulation Setup 10

- Introduction Phase 13
- Navigation Phase 13
- Task Phase 15
- Emergency Phase 16
- Post-Simulation Survey Phase 18
- Audio Check Phase 18

Chapter 4 Results 20

- Statistical Analysis 21
- Placebic Messaging 21
- Gender 22
- Evacuation Time 25
- False Claims 26

Chapter 5 Discussion 27

- The Effect of Audio 27
- Gender Analysis 27
- Evacuation Indecision 28

Chapter 6 Conclusion 29

Appendix A Post-Simulation Survey Questions 30

Appendix B 2x2 Fisher's Exact Test Results 32

BIBLIOGRAPHY 33

LIST OF FIGURES

Figure 1. Simulation Overhead Map. The meeting room that the participant is led to is at the center. The two primary exits are seen along the left side, equidistant from the meeting room. 11

Figure 2. Evacuation Robot in Simulation Environment. This is the opening image of the main simulation environment and how the robot appears to a user. Its arms wave in the direction it is moving in. 12

Figure 3. Practice Environment. This is the opening scene of the practice environment. The text is the only direction given, and this phase is untimed. 13

Figure 4. Navigation Phase Instructions. These instructions provide experiment expectations for the participant. 14

Figure 5. Robot Guidance Throughout Navigation Phase. The robot is guiding the participant in a circuitous path to the meeting room by going through another room first. Its arms appear slightly bent as it points in the direction it is moving. 14

Figure 6. Meeting Room Environment. Arrival at the meeting room is indicated through visual cues, such as the large table with chairs, as well as the robot verbally thanking the participant for following them to the meeting room and directing them to read the documents on the table. 15

Figure 7. Emergency Notification. Prior to hearing any audio cues, this message notifies the participant of the emergency and emphasizes the goal of leaving the building by any means. 16

Figure 8. Emergency Evacuation. This is what the participant sees if they start to follow the robot from the meeting room. While the timer counts down, the robot moves toward an unseen right exit and the people move toward the unseen left exit. Unknown to the participant, these exits are equidistant from the meeting room. 17

Figure 9. Survey Questions Page 1. This screen appears after the simulation to gather information from the participant about their motivations and demographic data. 18

Figure 10. Audio Check Phase. This screen appears after both the Introduction and Post-Simulation Survey phases. The participant has three chances to guess the correct color to show that they likely heard the audio during the experiment 19

Figure 11. Exit Routes by Condition. When considering robot voice gender and message, the results are fairly similar across all trials. About a third of participants follow the robot. 22

Figure 12. Exit Routes by Robot Voice Gender. When considering only robot voice gender, there is little evidence to suggest a robot voice gender preference. 23

Figure 13. Exit Routes by Human Gender. When considering just human gender, a similar percent of men and women chose to follow the robot. 24

LIST OF TABLES

Table 1. Percentage of People that Followed the Robot Voice Gender Analysis. This data suggests that women may prefer a male-voiced robot, while men show no preference. .24

Table 2. Evacuation Times by Condition and Exit Route. These times demonstrate that those who follow the robot take longer to exit than those who follow the crowd............................25

Table 3. 2x2 Fisher Exact Test Results. These tests confirm the results from the Chi-squared tests; there are no significant differences in exit route based on message and voice gender. ...32

ACKNOWLEDGEMENTS

I would like to thank Dr. Alan Wagner for his supervision and support of this project, as well as Mollik Nayyar, who designed the simulation and was incredibly helpful throughout this process. I am also grateful for Dr. Katie Fitzsimons' thorough and helpful editing. Additionally, I would like to thank my brother, Torin McFarland, who provided his voice for the male-voiced robot experiments.

This material is based upon work supported by the National Science Foundation under Grant No. CNS-1830390. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Chapter 1

Introduction

Proper emergency evacuation procedures are designed to maximize the number of people who are able to safely escape a building. However, most building evacuation plans do not adequately account for how a human will respond during an actual emergency [1]. This is especially true when considering that a plan must consider group behavior. This experiment used simulation to study the question of whether or not emergency evacuation robots could lead people to safety while avoiding exit bottlenecking. If so, this would decrease evacuation times, which would likely save lives in time-sensitive emergencies such as fires. Similar experiments have shown that people tend to follow a crowd over a robot [2], but that written explanations of a certain length and nature can increase the number of people who follow the robot [3].

This preliminary study took that idea a step further by considering how a gendered audio message affected the evacuation behavior of male and female participants. A voice can lend a sense of anthropomorphism to a robot and help it be more persuasive and trustworthy in the eyes of a participant [4]. As shown by Eyssel et al., people generally feel a sense of closeness to a robot of the same gender. However, when a robot compels a human to respond a certain way, men are more likely to trust and listen to a female robot [5]. We therefore expected that more men would follow the female-voiced robot. Women do not usually show this gender bias, so their responses were predicted to be less biased toward either of the robot voices.

Three messages were also tested in these trials. One had no explanation for why the person should follow the robot. This message served as a control. The other two had placebic and helpful explanations respectively for why the person should follow the robot. These explanations were also meant to instill trust after the robot makes a guidance mistake prior to the emergency. Previous studies have shown that providing any reason for why someone should perform an action, regardless of its value, is more compelling than no reason [6]. We also witnessed this effect during our own robot simulation experiments [3]. We predict a similar effect in this study. This simulation included a siren noise to complete the audio-intensive environment. This was intended to make the situation feel more real and produce more instinctual responses that could resemble what would occur in a real-life emergency.

In the following five chapters, this document will provide a review of Related Works, outline the Simulation Setup, present the Results, provide a Discussion of those results, and give final thoughts in the Conclusion. Appendix A includes further information from the participant survey, and Appendix B expands the statistical analysis.

Chapter 2

Related Work

This experiment is based on creating trust between a human and a robot after that robot has made a mistake. One study [7] has shown that human-machine trust can be influenced by the user's expectations before the experiment has even started. People with high expectations for a machine experience greater changes in trust in response to differing machine performance. This suggests that preconceptions about robots can affect how people respond to the robot after a mistake.

For explanations, it has been seen that only certain types of mistakes can be forgiven after hearing an explanation. In [8], a promise is able to help compensate for untrustworthy behavior, but not untrustworthy behavior that includes deception. Promises could increase the rate of trust recovery in the short-term, but did not have more long-term influence on trust recovery than demonstrating trustworthy behavior. For a robot, this would mean that it can make mistakes, but should not deliberately mislead a user. Apologies will not entirely compensate for a lie, and, in an emergency, there is little time to display enough trustworthy behavior to regain trust.

In robot-based experiments, mitigation strategies can help reduce the impact of robot mistakes on a person's trust in the system [9]. If people know the robot may struggle with a task, they are more likely to forgive a mistake. Lee et al. also showed that an apology was the most effective way to convince a user of the robot's capabilities. In our experiment, the robot's helpful message attempted to display trustworthy behavior, but did not directly acknowledge its mistake.

In addition to the findings in [4], it is also thought that people believe humanoid robots are more capable [12]. That same study demonstrated a user preference for taller robots, as well as ones they perceived as having extroverted personalities. Of the four possible robot combinations of tall, short, humanoid, and non-humanoid featured in [12], short and non-humanoid was the least preferred. While the study presented in this dissertation features a small, non-humanoid robot, it may still be perceived as extroverted because of its human speech.

A series of experiments that utilized emergency evacuation robots in a real-life simulated emergency demonstrated that, when given the option to follow a robot or find one's own way, a single participant will almost always follow the robot [10]. This occurs even when the robot has behaved in an untrustworthy manner or demonstrates odd behavior given the emergency situation. This suggests a sense of overtrust in robots when there are no other humans around to either help make decisions or lead the participant to another exit. However, real situations involve other people and unexpected factors. People may still expect the robot to be a good guide for them, and can struggle to understand when it has failed at its task [11].

Previous virtual studies have built a strong foundation for this research and the methods used [2, 3]. It has been shown that having the robot directly guide the participant is more effective than simply pointing them in the appropriate direction. Additionally, when utilizing written messages, the robot was more likely to be followed if it provided an explanation. Both placebic and helpful explanations compelled more people to follow the robot than a request with no explanation. Using a long message with extensive information was not more persuasive than a short message, supporting the use of short explanation in emergencies due to the timing concerns of such situations.

Chapter 3

Simulation Setup

The experimental design consisted of 2x3 factorial experiment with gendered audio voices as one factor and explanatory messages with different types of information as the second factor. The two independent variables were robot voice gender and explanatory message type. There were two robot voice genders, male and female. The three types of explanatory messages were No Explanation, Placebic Explanation, and Helpful Explanation, also referred to as Messages 1, 2, and 3 respectively. The exact messages are provided below in Emergency Phase. The dependent variables were exit route and evacuation time. This experiment allowed for three types of exit routes, Follow Robot, Follow Crowd, and Other. Follow Robot and Follow Crowd refer to specific exits in the experiment, while Other accounted for any other exit used or a failure to finish. Evacuation time tracked the time to exit for those who successfully did so.

This experiment was run in a Unity simulation, the overhead map of which is seen below (Figure 1).

Figure 1. Simulation Overhead Map. The meeting room that the participant is led to is at the center. The two primary exits are seen along the left side, equidistant from the meeting room.

This simulation represented an office space. The bottom center of the image shows the door where the participant entered. The room at the very center was the meeting room, which was the starting point for the emergency. The door highlighted in green was the Follow Crowd exit, and the door at the end of the top left hallway was the Follow Robot exit.

The robot deployed in the simulation was a Turtlebot with two PhantomX Pincher arms on top (Figure 2).

Figure 2. Evacuation Robot in Simulation Environment. This is the opening image of the main simulation environment and how the robot appears to a user. Its arms wave in the direction it is moving in.

120 subjects participated in the experiment. Study participants were recruited using Amazon Mechanical Turk. This online service pays users for successfully providing data to online experiments and tracks their approval rating from the experiments they complete. Filters were used on Amazon Mechanical Turk to only allow users with high experiment approval rates and no previous interactions with this lab's simulations. The data from ten subjects was removed because of submission errors. Hence, data from 110 participants was obtained. The group as a whole was 58.2% male and 40.9% female. The average age was 38.6 years. The majority of participants had either Some College or a 4-Year College Degree (Bachelor). There were six unique phases in the experiment: Introduction, Navigation, Task, Emergency, Post-Simulation Survey, and Audio Check. Each phase is further described below.

Introduction Phase

The experiment began with a screen that thanked the subject for participating. They were told the research was about "testing methods to help people find their way in office buildings". They were then placed in a simple environment to practice using the keyboard commands that allowed them to move through the simulation (Figure 3).

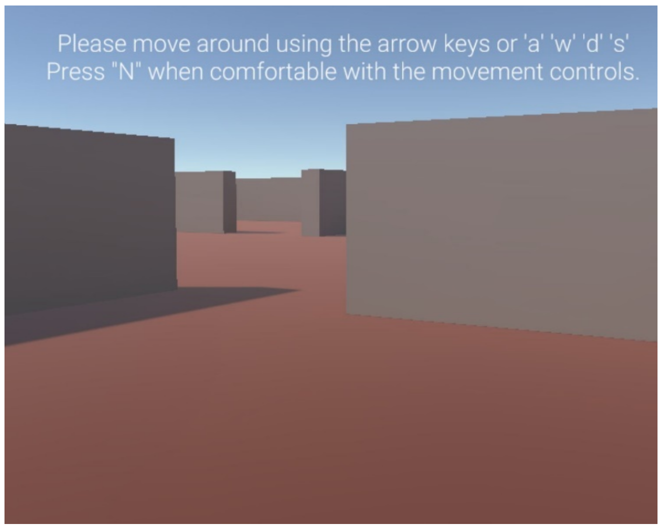

Figure 3. Practice Environment. This is the opening scene of the practice environment. The text is the only direction given, and this phase is untimed.

This phase of the experiment was untimed, and it only ended when the participant decided to move to the next phase of the experiment. There was no robot during this initial phase, so the focus was solely on gaining an ability to move through the environment.

Navigation Phase

The Navigation Phase was preceded by a page of instructions (Figure 4).

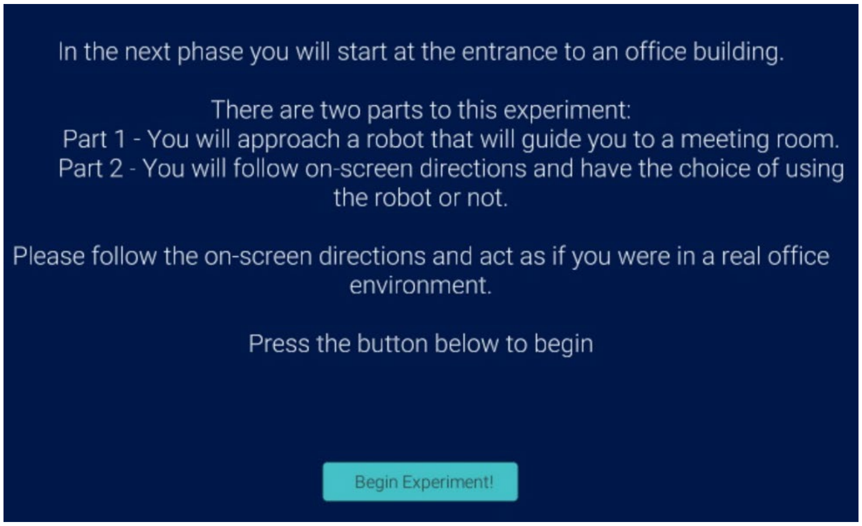

Figure 4. Navigation Phase Instructions. These instructions provide experiment expectations for the participant.

During the Navigation Phase, the on-screen directions instructed the participant to follow the robot to a meeting room. The robot did not provide any verbal instructions, and it led the participant through the office building, stopping if it ever got too far ahead of the person. It took an intentionally circuitous route, at times doubling back or moving in circles to demonstrate a clear mistake while guiding the participant (Figure 5).

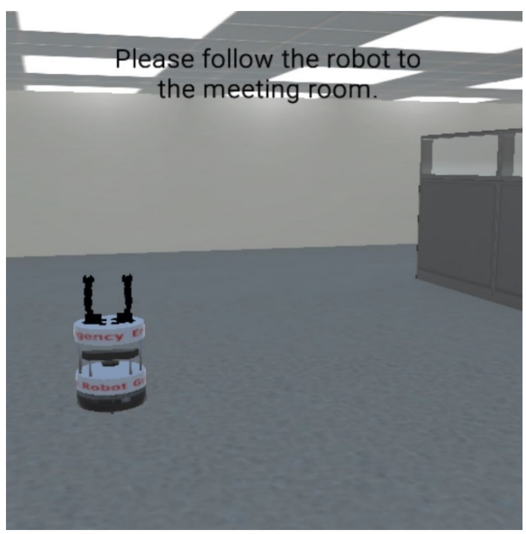

Figure 5. Robot Guidance Throughout Navigation Phase. The robot is guiding the participant in a circuitous path to the meeting room by going through another room first. Its arms appear slightly bent as it points in the direction it is moving.

The participant was ultimately led to the meeting room. When the participant entered the meeting room, their position was temporarily locked and the next phase started.

Task Phase

To begin the Task Phase, a male or female voice, depending on the condition, stated "Thank you for following me to the meeting room. Please read the documents on the table to proceed". This was the same message used in [3]. The participant was then free to approach the table (Figure 6).

Figure 6. Meeting Room Environment. Arrival at the meeting room is indicated through visual cues, such as the large table with chairs, as well as the robot verbally thanking the participant for following them to the meeting room and directing them to read the documents on the table.

When the participant reached the table, a survey appeared asking them to enter their favorite color and their opinion on the robot's performance. This was an attention check that also gathered initial impressions of the robot's skills. The phase ended when the participant submitted their response.

Emergency Phase

The Emergency Phase began with a written message informing the participant that there was a fire in the building and they must escape quickly in order to live (Figure 7).

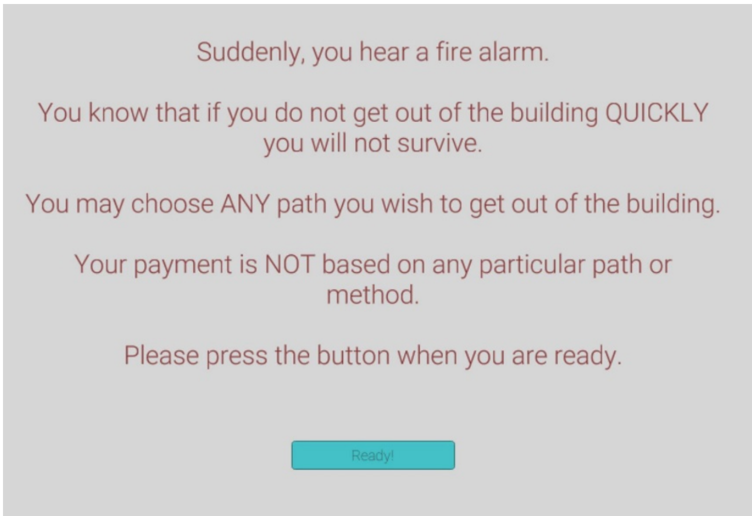

Figure 7. Emergency Notification. Prior to hearing any audio cues, this message notifies the participant of the emergency and emphasizes the goal of leaving the building by any means.

When the participant confirmed they had read the message, the robot then stated one of the following messages in either a male or female voice.

Message 1 (No Explanation): Excuse me, would you like to follow me?

Message 2 (Placebic Explanation): Excuse me, would you like to follow me because I am a robot?

Message 3 (Helpful Explanation): Excuse me, would you like to follow me because I know the closest exit?

This resulted in six total conditions, referred to as F1, F2, F3, M1, M2, and M3 accordingly. Message 1 acted as a control condition since it provided no additional information to the participant for making their decision. Message 2 provided irrelevant information to examine the influence that providing any explanation had on the participant's decision to follow the robot. Finally, Message 3 provided additional valid and useful information which should influence the participant's decision to follow the robot.

The participant was unable to move while the message was being delivered, but regained their control at the conclusion of the message. An audio siren noise was heard in the background. A crowd of non-player characters ran to an unseen exit on the left while the robot moved toward an unseen exit on the right (Figure 8). Unknown to the participant, both exits were equidistant from their starting point. Other exits were available throughout the building, including the door they started at.

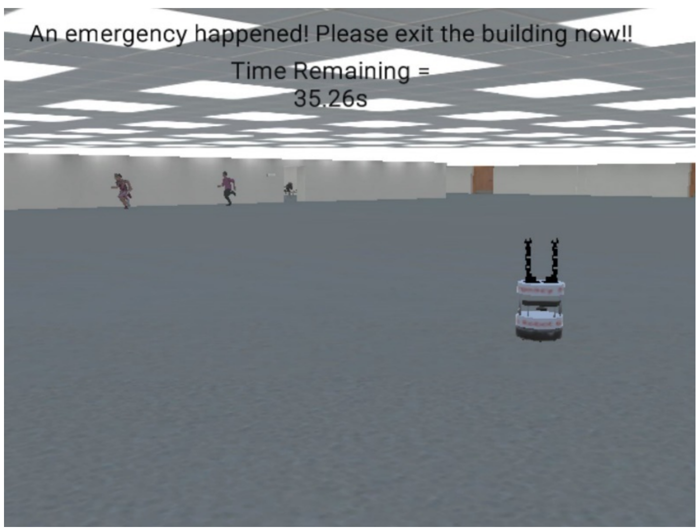

Figure 8. Emergency Evacuation. This is what the participant sees if they start to follow the robot from the meeting room. While the timer counts down, the robot moves toward an unseen right exit and the people move toward the unseen left exit. Unknown to the participant, these exits are equidistant from the meeting room.

A 40-second timer counted down to the conclusion of the Emergency Phase. Those who did not escape through any exit by the end of the 40 seconds received a message stating that they did not finish. They were then passed on to the Post-Simulation Survey Phase. This option to complete the survey after not finishing was added after F2 due to issues with participants not finishing. Those who did find exits were passed on to the next phase as soon as they exited.

Post-Simulation Survey Phase

In this phase, the participant answered several questions about the robot and their choice of exit, as well as justifications for their decision (Figure 9).

Figure 9. Survey Questions Page 1. This screen appears after the simulation to gather information from the participant about their motivations and demographic data.

They also provided demographic data, such as gender, which was crucial for the analysis of this experiment. The survey collected data about their feelings toward technology as well as firefighters to better understand what affects trust in these situations. The full transcript of these questions can be found in Appendix A.

Audio Check Phase

Functional audio was essential for a subject to participate in this experiment. Therefore, it was necessary to check that the participant could actually hear the audio. This phase occurred twice, after both the Introduction Phase and the Post-Simulation Survey Phase. Seven colors were listed. The participant was prompted to listen and select the randomly chosen color that was stated aloud. The participant was also able to replay the audio of the chosen color as needed (Figure 10).

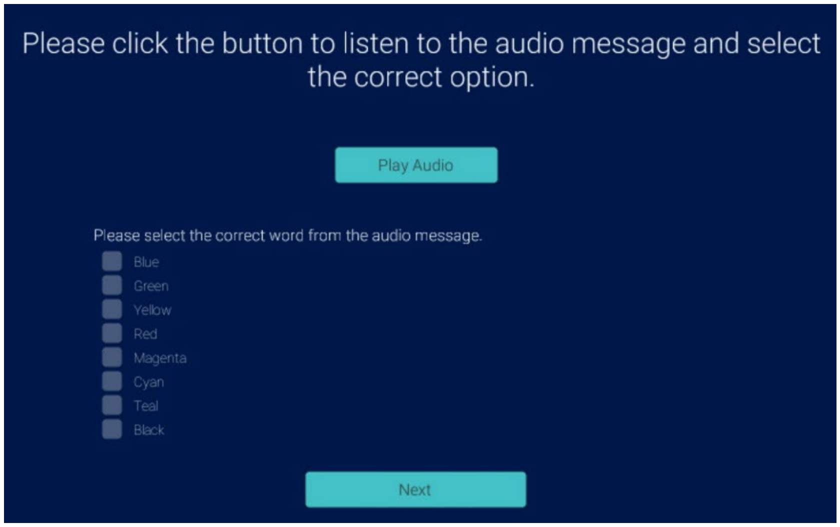

Figure 10. Audio Check Phase. This screen appears after both the Introduction and Post-Simulation Survey phases. The participant has three chances to guess the correct color to show that they likely heard the audio during the experiment.

They had three opportunities to select the color that was chosen. Each failed response warned the participant about how many remaining opportunities they had to identify the correct color. If all three attempts were incorrect, the participant was locked out of continuing the experiment and instructed to close their browser. Performing this task twice attempted to ensure that the participant had functional audio for the duration of the experiment.

Chapter 4

Results

A participant's evacuation decision was classified as Follow Robot or Follow Crowd once a participant successfully came within a certain distance of the robot's or the crowd's exit point. Subjects that were unable or unwilling to sufficiently indicate their decision were classified as Other. These cases are not included in the total when considering the significance of a certain message, robot voice gender, human gender, or average evacuation times as they were all a result of participants not finishing. They are displayed on the graphs and included in the total number of participants simply to describe all 110 runs. The ability to gather data from participants who did not finish was not added until after condition F2 due to issues with participants being unable to find an exit in a timely fashion. There are not enough cases of this nature to determine if robot voice gender, human gender, or message had an effect on whether or not people finished. Participants who exited through a different exit would also be counted as Other and included in the analysis, but no one successfully did so. There is also one case using Unlisted for gender because the participant did not provide identification as male or female. This is not counted in any gender-based analysis, but is included in the general study of whether participants followed the robot or the crowd because they successfully completed the experiment. Future experiments may include a Non-Binary option to be more gender-inclusive and potentially avoid such cases.

Statistical Analysis

The effects of robot voice gender, human gender, and message were further analyzed using a Chi-squared test of association and $\alpha = 0.05$. Each test checked the significance of the previously mentioned factors by comparing the number of participants in each group who followed a robot instead of the crowd. This test was chosen for its ease of evaluating associations in a 2x3 experiment of two categorical variables. Given the nature of a preliminary study, most tests included some expected counts less than 5. While no test proved significance, the two tests with low expected counts were cross-checked against a series of 2x2 Fisher's Exact tests to confirm the lack of significance found. The results of these tests can be found in Appendix B.

The evacuation time data was analyzed using a hypothesis test for difference in means. Given the lack of significance found for the message and robot voice gender conditions, the evacuation times for Robot Follow and Crowd Follow were compared across all six trials. While $n = 6$ is small for hypothesis testing, the effect is apparent enough that this was deemed acceptable for purposes of a preliminary study.

Placebic Messaging

When considering each group of 15-20 participants as a whole, roughly a third of the group chose to follow the robot as illustrated in Figure 11 (M = 30.26%, SD = 8.97%).

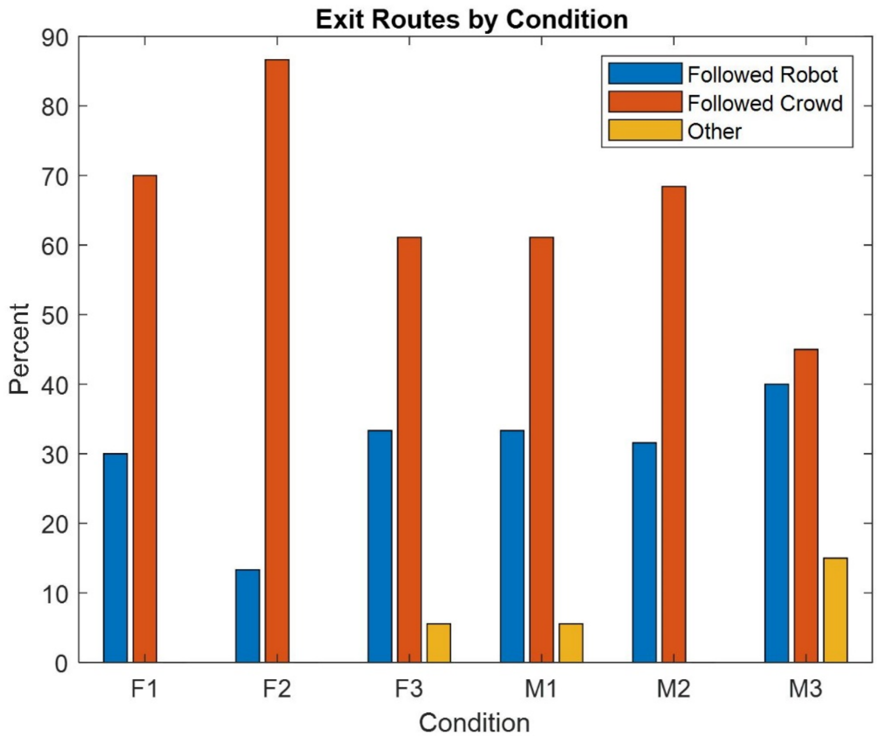

Figure 11. Exit Routes by Condition. When considering robot voice gender and message, the results are fairly similar across all trials. About a third of participants follow the robot.

The F2 group saw a smaller percentage of people follow the robot. Additionally, the M3 group saw a slightly higher rate of people following the robot, but also had more Other exits (3 of 20). Ignoring voice gender, the message contents does not appear to have a significant effect over the number of participants who chose to follow a robot (X^2 (2, N=105) = 2.418, p = 0.299).

Gender

Gender was considered from two different angles, the gender of the robot's voice and the gender of the participants. The effect of robot voice gender is shown in Figure 12, which

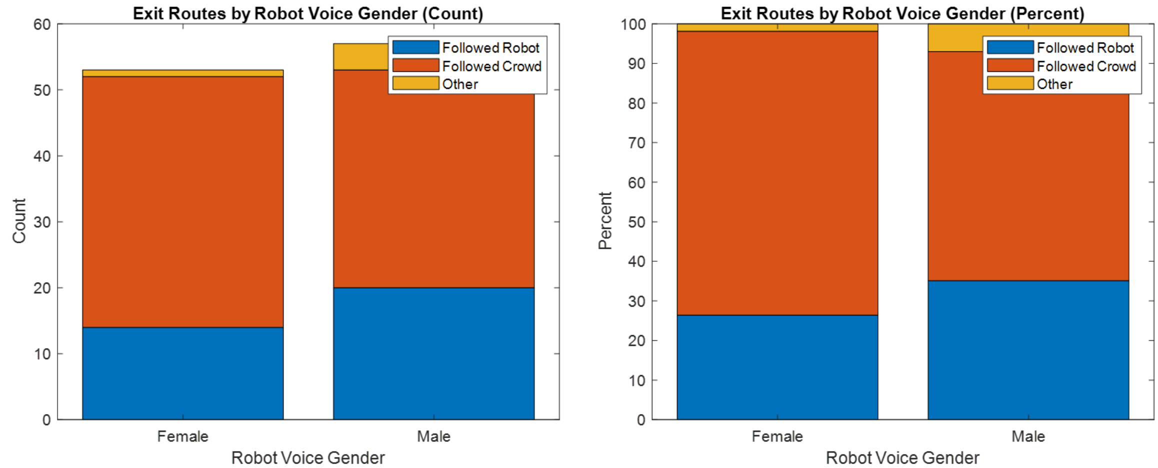

Figure 12. Exit Routes by Robot Voice Gender. When considering only robot voice gender, there is little evidence to suggest a robot voice gender preference.

differentiates all cases strictly by robot voice gender rather than message condition. When looking at the gender of just the robot voice, a similar percent of participants chose to follow the robot with the female voice (*M* = 25.56%, *SD* = 10.72%) and the robot with the male voice (*M* = 34.97%, *SD* = 2.57%). This data is depicted in both the count and the percent to show the effects of the small sample size, especially for the robot with the female voice.

When considering the robot's voice gender and message, the message did not have a significant effect for the robot with the female voice (X^2 (2, N=52) = 2.110, p = 0.348) or the robot with the male voice (X^2 (2, N=53) = 0.979, p = 0.613). Figure 13 considers the effect of human gender on whether or not a participant would choose to follow the robot, regardless of the gender of the robot's voice or the message.

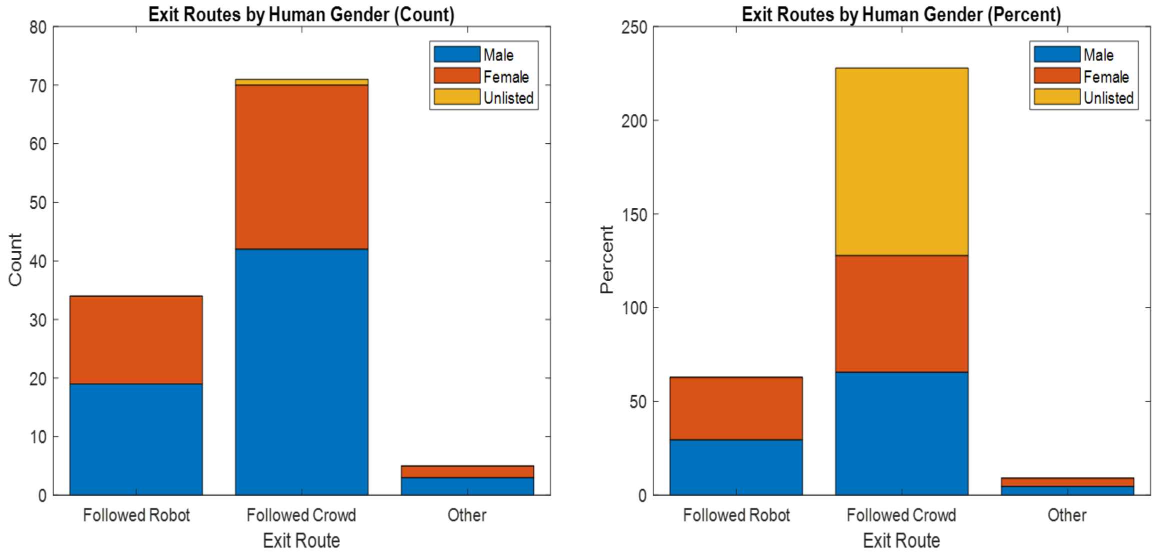

Figure 13. Exit Routes by Human Gender. When considering just human gender, a similar percent of men and women chose to follow the robot.

When ignoring both the gender of the robot's voice and the robot's message, roughly a third of male participants (*M* = 31.00%, *SD* = 6.01%) and female participants (*M* = 35.24%, *SD* = 23.38%) opted to follow the robot. These trials were majority male participants (in total, 64 men, 45 women, 1 unlisted), but the effect was seen for both men and women in spite of the disparity in number of male versus female participants.

Finally, we consider the breakdown of women and men following a robot with a male or female voice regardless of message (Table 1). The fraction of each gender that followed the robot in all three of its runs is provided for both robot voices. This table only considers the Follow Robot and Follow Crowd conditions where a gender was listed by the participant.

Table 1. Percentage of People that Followed the Robot Voice Gender Analysis. This data suggests that women may prefer a male-voiced robot, while men show no preference.

Followed Female-Voiced Robot		Followed Male-Voiced Robot	
Female: 4/21 (19.05%)	Male: 10/31 (32.26%)	Female: 11/22 (50.00%)	Male: 9/30 (30.00%)

The table suggests that women are more likely to follow a robot with a male voice, while men are not particularly biased toward either robot.

Evacuation Time

The mean and standard deviation of evacuation time for each condition and the two primary exit routes are presented in Table 2.

Table 2. Evacuation Times by Condition and Exit Route. These times demonstrate that those who follow the robot take longer to exit than those who follow the crowd.

Simulation	Robot Follow Evacuation Time (s)	Crowd Follow Evacuation Time (s)
F1	*M* = 30.19, *SD* = 5.78	*M* = 25.51, *SD* = 6.99
F2	*M* = 26.24, *SD* = 6.51	*M* = 23.63, *SD* = 5.31
F3	*M* = 35.07, *SD* = 7.43	*M* = 21.48, *SD* = 7.93
M1	*M* = 34.23, *SD* = 6.39	*M* = 26.15, *SD* = 6.38
M2	*M* = 30.49, *SD* = 7.95	*M* = 21.55, *SD* = 5.19
M3	*M* = 30.62, *SD* = 7.93	*M* = 25.75, *SD* = 6.42

Table 2 shows that the Robot Follow group always has a higher average evacuation time than the Crowd Follow group for the same message and robot voice gender. When considering only the evacuation route, the average evacuation times for Robot Follow (M = 31.14, SD = 3.18) and Crowd Follow (M = 24.01, SD = 2.12) differ by 7.13 seconds. A hypothesis test for this difference (H_0: $\mu_R - \mu_C = 0$, H_a: $\mu_R - \mu_C \neq 0$, n = 6) yields a p-value of less than 0.001. The people that do choose to follow the robot exit more slowly than those that do not follow the robot.

False Claims

Each group also had a few instances of false claims. A claim was considered false when a participant claimed that they chose the robot's guidance even though motion tracking data shows that they switched routes halfway through the emergency, when they claimed that they chose the robot's guidance when they only ever followed the crowd, or when they claimed that they did not choose the robot's guidance when they followed the robot. As the number of false claims was low, and many were due to people switching routes, these claims seem to stem from the subject incorrectly characterizing their decision rather than from social desirability bias. For the case of switching routes, the question is likely perceived as unclear because the person did follow the robot and then changed directions, which was not an option provided in the survey.

Chapter 5

Discussion

The Effect of Audio

These trials included two audio elements that were not featured in previous experiments with this simulation: the robot's audio messages and the siren noises heard during the Emergency Phase. This was done to present the information in a more familiar format and heighten the subject's feelings of urgency. All groups reported nearly 100% motivation to exit the building during the Emergency Phase. However, while previous results showed the robot's message could influence the participant, this effect was entirely lost in the audio experiments. Between the running crowd and the loud sirens, the participants seemed essentially overwhelmed by the situation. They likely resorted to a more instinctual response of herd mentality, resulting in a similar proportion of people following the crowd each time.

Gender Analysis

While these trials suggest some link between human decision-making and robot voice gender, it is uncertain if this trend would have significant effects in a real emergency. While individual women seemed more likely to follow the male-voiced robot, the behavior of the F2 group may have also increased the disparity. That was the only group where no women followed the robot at all, even though six participated.

Evacuation Indecision

Despite the high motivation, two types of evacuation indecision were seen in the data. The first were incidents where participants essentially froze in place for brief periods of time. This was seen infrequently, but may have increased some of the evacuation times. The more common form of evacuation indecision was participants who started out following the robot and ended up turning back to follow the crowd instead. Several instances of this occurred in each condition with some participants nearly making it to the Follow Robot hallway before turning back.

In the Evacuation Time data, those who started following the robot and ultimately exited with the crowd were counted as Follow Crowd. Participants who followed the robot almost always took a direct path to the exit while participants following the crowd took both direct routes and long routes when they turned away from following the robot. However, even though Follow Crowd participants travelled a greater distance on average to reach their exit due to doubling back, they still have faster average evacuation times. This suggests participants who follow the robot are either hesitating before following the robot or perhaps the robot is simply too slow in the simulation. Given the rare occurrence of participants freezing at the decision point, it is likely the robot slowed down the participants.

The robot is able to move as fast as the human in the simulation. However, it cannot go further ahead of the person than a certain fixed distance. When the robot stops due to it being too far ahead of the participant, it cannot always start moving again quickly enough to reach full speed as the person catches up. This could result in people moving slower until they see the exit for themselves. However, the high standard deviation for evacuation time data could also suggest varying levels of navigational skill among participants.

Chapter 6

Conclusion

This study began with two primary hypotheses, one concerning gender, the other placebic messaging. Regarding gender, it was assumed that women would not show a bias toward either robot voice while men would be biased toward a female-voiced robot. The reverse effect occurred with men showing little bias and women being slightly biased in favor of a male-voiced robot. As this is only a preliminary study, larger and more extensive trials would need to be conducted in the future to confirm this effect and fully explore the cause behind it.

For placebic messaging, it was believed that Message 1 would compel the smallest proportion of people to follow the robot, while Messages 2 and 3 would perform better. All three messages had a similar success rate regarding how many people followed the robot. While this is a contradiction of previous studies with this simulation, it may be an indication that further study is needed in audio-intensive environments to see if commonly held beliefs, such as the value of a placebic explanation, actually hold true in an emergency situation. This study may also indicate simple explanations that do not acknowledge mistakes may be insufficient to regain human-robot trust in a more realistic environment. The influence of audio throughout the simulation is clear, so robots must learn how to effectively utilize it in order to make a difference.

Appendix A

Post-Simulation Survey Questions

The following is a list of all questions presented in the Post-Simulation Survey Phase. For multiple-choice questions, the answers are listed. Questions without answers listed should be assumed to be open-ended.

1. How did you find your way or attempt to find your way out of the building?
2. After the emergency began, did you intend to use guidance from the robot?
    a. Yes
    b. No
    c. Why or why not?
3. After the emergency began, were you motivated to find the exit as quickly as possible?
    a. Yes
    b. No
    c. Why or why not?
4. In the emergency phase, did you believe the robot would find the exit quickly?
    a. Yes
    b. No
    c. Why or why not?
5. I trusted the robot when I made my choice to follow or not follow the robot in the emergency.
    a. Yes
    b. No
    c. Trust was not involved in my decision

   d. Why or why not?

6. What is your gender?
   a. Male
   b. Female
7. What year were you born in?
8. What country do you live in?
9. What is your occupation?
10. Which of these best describes the highest level of education you have completed?
    a. Less Than High School
    b. High School
    c. Some College
    d. 2-Year College Degree (Associate)
    e. 4-Year College Degree (Bachelor)
    f. Masters Degree
    g. Doctoral Degree
    h. Professional Degree (MD, JD)

The following questions provided a scale from 1-7 and directed the participant to "Please rate your agreement with the follow statements".

1. I am comfortable with using new technology.
2. I believe firefighters are trustworthy guides in a fire emergency.

Appendix B

2x2 Fisher's Exact Test Results

To ensure the lack of significance found by the Chi-squared tests of association were accurate, each condition was tested using three 2x2 Fisher's Exact tests. Fisher's Exact Test is a more precise test of independence that is accurate for any sample size. The p-values of these tests are presented below. As the Fisher's Exact test is done for 2x2 situations, three tests were necessary per case to cover all message combination conditions. In each case, we are considering the number of people who follow a certain exit route, whether that is crowd, female-voiced robot (F Robot), or male-voiced robot (M Robot).

Table 3. 2x2 Fisher Exact Test Results. These tests confirm the results from the Chi-squared tests; there are no significant differences in exit route based on message and voice gender.

	(F) Robot/ Crowd	(M) Robot/ Crowd
Msg 1/Msg 2	0.419	1
Msg 2/Msg 3	0.229	0.495
Msg 1/Msg3	1	0.728

From this table, it is clear that the conclusions found by the Chi-squared tests of association were accurate; the messages and robot voice gender did not have significant effects on a participant's choice of exit route.

BIBLIOGRAPHY

[1] E. D. Kuligowski, "Modeling Human Behavior during Building Fires," National Institutes of Standards and Technology, Technical Note 1619, 2008.

[2] M. Nayyar and A. R. Wagner, "Effective Robot Evacuation Strategies in Emergencies," *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2019, pp. 1-6, doi: 10.1109/RO-MAN46459.2019.8956307.

[3] M. Nayyar, Z Zoloty, C. McFarland and A. R. Wagner, "Exploring the Effect of Explanations During Robot-Guided Emergency Evacuation," in *Lecture Notes in Computer Science vol 12483, International Conference on Social Robotics, November, 2020.* Wagner A.R. et al., Eds. Springer, Cham., 2020, pp. 13-22, doi: 10.1007/978-3-030-62056-1_2.

[4] F. Eyssel, L. de Ruiter, D. Kuchenbrandt, S. Bobinger and F. Hegel, "'If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism," *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2012, pp. 125-126, doi: 10.1145/2157689.2157717.

[5] M. Siegel, C. Breazeal and M. I. Norton, "Persuasive Robotics: The influence of robot gender on human behavior," *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2009, pp. 2563-2568, doi: 10.1109/IROS.2009.5354116.

[6] M. Eiband, D. Buschek, A. Kremer and H. Hussman, "The Impact of Placebic Explanations on Trust in Intelligent Systems", in *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow, Scotland, UK, May 4-9, 2019*, Association for

Computing Machinery, New York, NY, USA, Paper LBW0243, pp 1-6, doi: 10.1145/3290607.3312787.

[7] V. L. Pop, A. Shrewsbury and F. T. Durso, “Individual Differences in the Calibration of Trust in Automation”, *Human Factors*, vol. 57, no. 4, pp. 545–556, 2015, doi: 10.1177/0018720814564422

[8] M. E. Schweitzer, J. C. Hershey and E. T. Bradlow, “Promises and lies: Restoring violated trust”, *Organizational Behavior and Human Decision Processes*, vol. 101, no. 1, pp. 1-19, 2006, doi: 10.1016/j.obhdp.2006.05.005.

[9] M. K. Lee, S. Kiesler, J. Forlizzi, S. Srinivasa and P. Rybski, "Gracefully mitigating breakdowns in robotic services," *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2010, pp. 203-210, doi: 10.1109/HRI.2010.5453195.

[10] P. Robinette, W. Li, R. Allen, A. M. Howard and A. R. Wagner, "Overtrust of robots in emergency evacuation scenarios," *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2016, pp. 101-108, doi: 10.1109/HRI.2016.7451740.

[11] P. Robinette, A. R. Wagner and A. M. Howard. “Investigating Human-Robot Trust in Emergency Scenarios: Methodological Lessons Learned”, in: *Robust Intelligence and Trust in Autonomous Systems, April, 2016.* Mittu R., Sofge D., Wagner A., Lawless W, Eds. Boston: Springer, doi: 10.1007/978-1-4899-7668-0_8.

[12] M.L. Walters, K.L. Koay, D.S. Syrdal, K. Dautenhahn and R. Te Boekhorst, “Preferences and perceptions of robot appearance and embodiment in human-robot interaction trials”, in: *Proceedings New Frontiers in Human-Robot Interaction, symposium at the AISB09 convention, Edinburgh, Scotland, UK April 8-9, 2009*, K. Dautenhahn, Eds.

Ciera McFarland
cqm1300@gmail.com

EDUCATION

Bachelor of Science in Aerospace Engineering, Pennsylvania State University, 2021
Thesis Title: *The Effects of Gender and Placebic Messaging on Robot-Guided Emergency Evacuation*
Thesis Supervisor: Alan Wagner, Ph.D.

RESEARCH EXPERIENCE

Research Assistant, Pennsylvania State University (REAL Lab), January 2020-Present
Determined how a robot's placebic messages and perceived gender affect human evacuation in a Unity-simulated emergency for undergraduate thesis.
Observed human response to robot's placebic messages of varying length during a Unity-simulated emergency.

Research Assistant, Georgia Institute of Technology, Summer 2019
Designed, fabricated, and programmed a hexapod robot capable of demonstrating different gaits.
Compared gait data to simulated models.

Research Assistant, Pennsylvania State University (BioRob-InFlu Lab), January 2018-October 2019
Collected flight speed data from flies exposed to visual stimuli in a controlled setup.
Analyzed how fly flight speed was affected by wing damage.

Research Assistant, North Carolina State University, Summer 2018
Fabricated composite plates, created sensor arrangements, and measured ultrasound signals traveling between piezoelectric sensors and fiber Bragg grating sensors.

PRESENTATIONS

C. McFarland, et al, "Damage Detection on Composite Plates Using Fiber Bragg Grating and Piezoelectric Ultrasonic Sensors," presented poster at 33rd American Society for Composites Technical Conference, Seattle, WA, 2018.

M. Nayyar, et al, "Exploring the Effect of Explanations during Robot-Guided Emergency Evacuation," presented at 12th Int. Conf. on Social Robotics, 2020.

ADDITIONAL EXPERIENCE

Senior Contributing Writer for SnoQap, February 2019-February 2020
Wrote seventeen articles on relevant technical and social issues in STEM.