Influence of Synthetic Voices on User Risk-Taking Tendencies

Jibran Khalil
University of Austin at Texas
Austin, Texas 78712
Email: jibrankhalil@utexas.edu
Phone: (469) 403-9218

Kevin Li
University of Austin at Texas
Austin, Texas 78712
Email: kxl4126@utexas.edu
Phone: (469) 268-8934

Jack Si
University of Austin at Texas
Austin, Texas 78712
Email: jacksi@utexas.edu
Phone: (210) 762-8141

***Abstract*—Using a web-based text-based story in tandem with varying synthetic narrators, we evaluated the influence of synthetic voices on the numbers of risk user's would take. The experiment was constructed to store the respondent's gender, randomly assign them a male or female narrator, and present a constant number of risk-taking opportunities. We acquired 115 gathered responses, accrued from internet boards, friends, and peers, and saw a respondent distribution of 52 males and 63 females. 61 people played the game under the male narrator and 54 under the female narrator. Bolstered by a low p-value of 0.005 indicating statistical significance, our experiment provides strong evidence that the female narrator causes people to take more risks, regardless of their gender.**

I. Introduction

With rapid steps in AI and robotics, synthetic voices have begun emerging in the unlikeliest of fields. From video games to voice assistants, studies have predicted the speech and voice recognition market to blossom to over $31 billion dollars by 2025 [1]. Major tech corporations, such as Amazon and Apple, are vying for control of the voice market, and as a consequence the number of digital voice assistants are predicted to triple to 8 billion by 2023 [1]. With the market ready to erupt, we have tackled a valuable question that might prove invaluable to competing companies. Our experiment is focused on identifying whether a correlation exists between varying synthetic narrators and the risk-taking decisions made by the listener.

This research will allow the public and such companies to better consider the practical implications of different voices on user behavior. It follows that any organization concerned with how their technology may influence people can use this information to better achieve their unique needs. This research also acts as potential proof of the impact, or lack thereof, of a robot's voice on influencing real decisions that humans take in their daily lives. Such implications are extremely important to consider when dealing with any aspect of human robot interaction, whether that be in research or in society.

II. Background

Recent research and studies over synthetic voice have focused on analyzing marketability, consumer trends, and other consumer topics. However, there do exist few studies that share similarities with ours in their attempt to analyze specific emotional responses to different voice types (pertaining to both natural and synthetic). Charles R. Crowell's "Gendered Voice and Robot Entities: Perceptions and Reactions of Male and Female Subjects" used a blend of distractor math tasks and surveys to analyze response times, perceived friendliness, and perceived reliability of different voices [2]. Similarly, the study "The Influence of Voice Pitch on the Evaluation of a Social Robot Receptionist" analyzed the effects of varying voice types on attractiveness [3]. These respective studies, alongside multiple empirical reports, have all chosen to analyze their data through gender [2][3][4][5]. With clear distinctions between male and female reactions and with research to serve as foundational knowledge, we decided to also analyze our data by gender. However, from here our study deviates entirely. Instead of concentrating on emotional responses from listeners in active conversation with synthetic voices, we chose to analyze the influence that a synthetic narrator has towards a risky situation. No study has both tested influenced risk-taking by synthetic voices and placing the voice as a narrator rather than a direct conversational partner. With our experiment, we will try to detect whether synthetic voices can actually catalyze users to a certain action, rather than just prompt some sort of emotion.

The study "'I Don't Believe You': Investigating the Effects of Robot Trust Violation and Repair", despite not focusing on synthetic voice, contained structural components that were used in the construction of our study [6]. Within that respective study, there was the use of a tablet game that the robot and user would play. Responses made by the user based on certain robot actions were tabulated and analyzed. After the game was over, the player would answer a questionnaire, but valuable data was also embedded within the activity. Our study attempts to do something similar, where rather than implementing a survey, we would gauge influenced decisions by the activity itself. With the need to reduce all forms of bias, such as sampling error and response bias, we opted to allow the respondent's actions to identify possible influence from the narrator rather than a survey [7].

III. Method

A. Website

To distribute the story to participants, we developed a website with a story containing risks embedded in it. When opening the website, the user would be randomly assigned either a male or female narrator. Next, the user would input their demographic data. Finally, the game would begin; at each stage of the story, the narrator would read out the text for that particular moment and the player would be given options to select, some of which contained risks. Depending on the selected options, the story could diverge into different paths. After selecting an option, the narrator would read out the text for the next stage of the story and so forth until the user completed the game. To ensure the narrator's audio remained constant across different browsers, devices, and operating systems, we prerecorded each narration snippet and manually adjusted the volume of the audio so that the website would not rely on the browser's text to speech implementation to read out text. Additionally, the website utilized cookies to ensure each player could only play the game once.

In addition, music was added to the website. The background music was held constant regardless of narrator, age, or gender and was chosen to be subtle and relating back to the current point in the plot. Music has been proven to increase immersion, and has been seen implemented within audio books and video games alike [8].

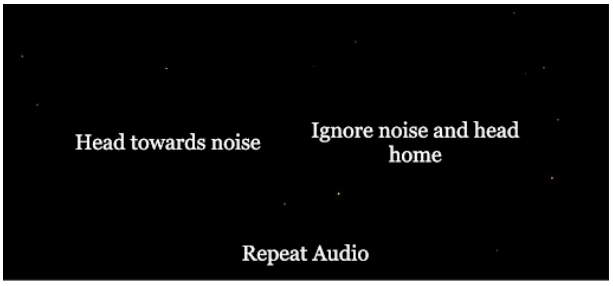

Fig. 1. One of the many decisions presented within the story

B. Database

Each option the user selected had an option ID. After finishing the game, the user's selected option ID's would be uploaded to a database along with their demographics and gender of the narrator. Figure 2 displays example nodes of data collected.

IV. Experimental Setup

A. Text-based story

The story was carefully constructed to provide the same number of opportunities to take risks and the same number of general decisions regardless of what path the user took. Throughout the writing process, we defined a risk as when a player opts to take an uncertain situation that could have unknown consequences instead of the potentially safer route. We modeled all risks off of this definition and remained consistent on the magnitude of each risky decision. As shown in Figure 1, a risk within the story would be heading towards an unknown commotion, while the safe alternative would be ignoring the noise and heading home. To offset the continuous risk-inducing decisions, We also included filler decisions throughout the story. This served to not only aid in obscuring the study's intentions, but to also help further immerse the listener.

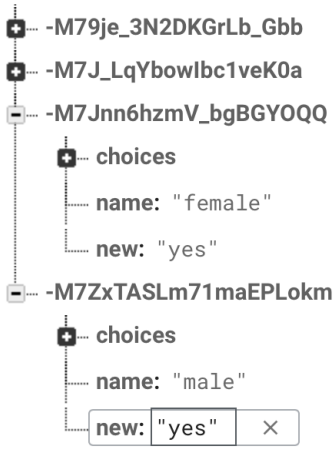

Fig. 2. Data for each respondent including choices made and gender (demographic data is included in choices).

B. Procedure

While dispersing the story link through Reddit, GroupMe, and Piazza, we made sure to provide clarity and a simple tutorial on the interface. Afterwards, the participants would then enter into the website. From here the narrator gathered the individual's age and gender, before embarking on the story. After the participant went through the entire story their decisions would be logged onto the website's database.

We intentionally crafted a simplistic model to obtain data due to the distance between us and the respondents. To minimize potential data inaccuracies and miscommunication between parties, the experiment simply consisted of the demographic collection and the text-based story.

V. Results

As a person steps through the game, each decision they make are stored in an array. When they finish the game and submit their results, the array of choices they make are submitted to the database, After responses are submitted, we can conduct analysis on the data by downloading the data as a JSON file and using Python to parse through the data. The demographic data was extracted from the first couple of elements in the array. As only some of the decisions were actually classified as potential risks, we would then iterate through each individual's decision and count how many risks they took in their run through the game. After storing the number of risks each individual took, we would average each demographic group's data. For example, as shown in Fig.3,

	Male Participants	Female Participants	Average Number of Risks
Male Narrator	2.147059	2.074074	**2.114754**
Female Narrator	3.0	2.6	**2.814815**
Average Over Gender	2.539683	2.326923	2.443478

Fig. 3. Table showing the average number of risks taken by male and female participants in relation to male and female narrators

the average number of risks taken by male participants under a male narrator is 2.147059 and the average number of risks taken by everyone under the male narrator is 2.114754. These averages are what were used in the significance tests, namely the male narrator average and female narrator average.

A. Participants

A total of 115 random individuals played through the story. Respondents consisted of friends, relatives, individuals on online boards and peers from the University of Texas at Austin. Participants were randomly assigned to the female or male narrator resulting in 54 people playing under the female narrator and 61 playing under the male narrator. The gender distribution saw a slight shift towards female participants with 52 of the participants being male and 63 being female.

B. Respondent Decisions

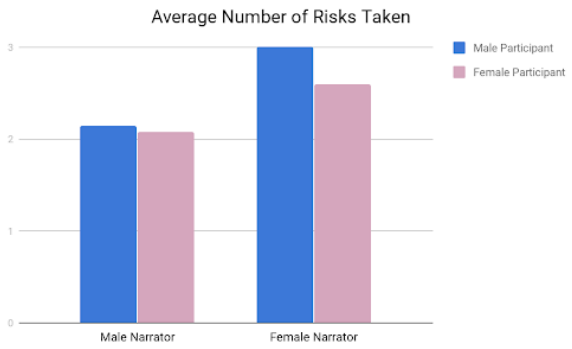

Fig. 4. Graph showing the average number of risks taken by male and female participants in relation to male and female narrators

As referenced in both Figure 3 and Figure 4, a noticeable uptick in average risks taken can be seen with the female narrator regardless of gender. Male participants see a 39.7 percent increase in risks taken, while female participants see a 25.4 percent increase in risks taken. Across both male and female participants, there is a 33.1 percent increase in risks taken. A statistically significant result would provide strong evidence that the gender of the narrator impacts a player's risk taking tendency. We used a two-sample t-test to test for statistical significance. The two sample means were the male narrator average and female narrator average, 2.814815 and 2.114754, respectively. The two sample sizes were the number of male participants and number of female participants, 52 and 63, respectively. We tested with 113 degrees of freedom and found that with a t-statistic of t = 2.714 and a p-value of p = 0.005. This shows that our result is statistically significant and provides strong evidence that the gender of the narrator does indeed impact risk taking tendencies. This result is also corroborated by the fact that both male and female participants had a higher tendency to take risks under the female narrator.

VI. Discussion

Our study showed a correlation between female narrators and increased risk taking. This paired alongside statistical tests that indicates statistical significance, shows not only an exciting trend, but the potential of an even deeper study. Most previous studies focus on how people feel about certain synthetic voices, while our result demonstrates the potential of robot voices to catalyze actions within individuals based on the actual configuration of the voice. This side of human robot interaction seems to be sparsely explored and we hope that such a result can inspire a new examination of the possibilities of synthetic voices. Our statistically significant result indicates that this phenomenon can be applied to other areas of human robot interaction that already occurs in day to day life. The different influences of different synthetic voices can allow providers of such services to tailor their products in a way that best suits their needs or the customer's desires. With the popularization of products including technologies like voice assistants, the ability to influence user's actions with the configuration of synthetic voices can have great impacts on business profitability.

In recent years, studies have emerged analyzing consumer behavior with regards to risk-taking theory. Risk-taking theory suggests that most consumers decide to buy a product under some degree of uncertainty about a given brand [9][10]. They must take on a perceived risk whenever shopping, which ties in beautifully to corporate profitability and the value in our study's data. Decision-making occurs throughout our daily life, but more importantly weighing the risk-return of our actions. Risk management and weighing risk with return not only occurs with our finance, but also in social interactions and conversation [11]. As an HRI study, the possibility that our data can indicate the possibility of potentially shaping conversation and dialogue solely on the voice of the robot is an exciting prospect.

VII. Conclusion

In this study, we embedded a text-based story within a website to identify whether a possible correlation existed between varying synthetic voices and the risk-taking decisions made by the listener. By modeling our structure on similar studies [2][3][6], we fixed valuable data sources within the story and chose to analyze acquired data through the split lens of gender. The analysis based on male and female participants led to the discovery of exciting trends, while also corroborating previous findings from foundational studies. Across both genders, the female narrator induced a higher average number of risks taken (Figure 3 and Figure 4). As a baseline correlation has been established, further studies are now possible with expansions in certain areas previously impossible. As the current pandemic situation hopefully is slowing down, future studies can take place in person and simulate more life-like consequences. Text-based stories, though with the right tools and plot line can bring high levels of immersion, typically cannot compete with perceived tangible situations with real-life consequences. As a result, despite strong evidence indicating the influence of synthetic voices on risk-taking further refinements and adjustments must be made.

Our study's results have demonstrated advantages with certain synthetic voices in regards to situations that either demand or decry risky behavior. With many important actions pivoting on a risk-return basis, the influence on decision-making has thousands of applications paired with the growing voice market. Corporations that are invested in voice assistants or other synthetic voice partners would find massive potential in regards to influencing consumers through their choices in voice type. However, keeping this all in mind, caution must be used when researching and designing voices with the intent of unknowing influence.

References

[1] G. V. Research, "Voice and speech recognition market size, share trends analysis report, by function, by technology (ai, non-ai), by vertical (healthcare, bfsi, automotive), and segment forecasts, 2018 - 2025," 2018.

[2] C. Crowell, M. Scheutz, P. Schermerhorn, and M. Villano, "Gendered voice and robot entities: perceptions and reactions of male and female subjects," 10 2009, pp. 3735–3741.

[3] A. Niculescu, B. Dijk, A. Nijholt, and S. See, "The influence of voice pitch on the evaluation of a social robot receptionist," 11 2011.

[4] E. Moriuchi, "Okay, google!: An empirical study on voice assistants on consumer engagement and loyalty," *Psychology Marketing*, 01 2019.

[5] T. Ammari, J. Kaye, J. Tsai, and F. Bentley, "Music, search, and iot: How people (really) use voice assistants," *ACM Transactions on Computer-Human Interaction*, vol. 26, pp. 1–28, 04 2019.

[6] S. S. Sebo, P. Krishnamurthi, and B. Scassellati, ""i don't believe you": Investigating the effects of robot trust violation and repair," in *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2019, pp. 57–65.

[7] J. Krosnick, *Assessing the Accuracy of Survey Research*, 10 2017, pp. 3–5.

[8] J. Fu, "The influence of background music of video games on immersion," *Journal of Psychology Psychotherapy*, vol. 05, 01 2015.

[9] E. Maziriri and T. Chuchu, "The conception of consumer perceived risk towards online purchases of apparel and an idiosyncratic scrutiny of perceived social risk: A review of literature," *International Review of Management and Marketing*, vol. 7, pp. 257–365, 07 2017.

[10] N. Lim, "Consumers' perceived risk: Sources versus consequences," *Electronic Commerce Research and Applications*, vol. 2, pp. 216–228, 09 2003.

[11] P. Drew, "Conversation analysis and social action," *Journal of Foreign Languages*, vol. 37, pp. 2–20, 05 2013.