Sexism is a constant problem experienced by women in society. Sexism is defined as an individual's attitudes, beliefs, and behaviors that either reflect negative views of individuals based on their gender or support an unequal status of men and women (Swim & Hyers, 2009). This not only affects men and women during their everyday life (Swim et al., 2001), but also extends to forces of nature where sexism can take a deadly toll. This can be seen in research showing that people even judge hurricane risk based on stereotypical gender expectations: Jung et al. (2014) found that feminine named hurricanes were more deadly than masculine named hurricanes due to the lower perceived risk associated with feminine names leading to a lack of preparedness and presumed severity. Alarming research such as this supports the argument that sexism is still a major threat in society today. We focus on sexisms' impact on women in the present study due to the amount of psychological discomfort women experience daily from chronic sexism (Swim et al., 2001). Additionally, with the world becoming more dependent on technology and new technological creations arising each day, it can be assumed that the sexism prevalent in society and day-to-day activities would also surely encompass the technological realm. Thus, we aimed to investigate this technological aspect of sexism as well.

According to the United Nations, the reason for the large gender gap in digital skills is due to the patriarchal and gender inequality in society (United Nations, 2019). The gender gap in digital skills continues to widen as more technological advances are made and is becoming a more pressing issue that needs to be addressed (United Nations, 2019). Artificial intelligence (AI), a recent addition to society, is the use of methods based on humans and animals to solve complex problems (Coppin, 2004). AI assistants have grown in use and popularity, with many U.S. residents having access to them and utilizing them on a regular basis (Dellaert, 2020). Additionally, while most AI assistants' voice characteristics are changeable, an overwhelming

majority of AI 's, specifically voice assistants, have a default set female voice (Siegel et al., 2009). This is where we believe possible effects of sexism and negative stereotyping towards women come into play.

Recent research suggests that the more people engage in stereotyping, the more these stereotypes are reinforced and ultimately ingrained into society (Martin & Slepian, 2018). As AI voice assistants are now largely involved in individual's lives, AI's have become influential in day-to-day activities and decision-making for many people (Dellaert, 2020). It follows then that AI voice assistants can be influential in many other aspects of people's lives, including stereotyping and sexism.

AI Voice Assistants

Voice Assistant Technology. Since Apple first introduced Siri in 2011, artificial intelligence powered voice assistants (VA's) have become well-established features of mobile devices (Guzman, 2019). Following Siri, additional prominently used voice assistants include Amazon's *Alexa*, Google's *Google Assistant*, and Microsoft's *Cortana*. Recently, there has been a growth in voice-based technology, and many people are now communicating with voice assistants daily in the same way they would with other humans (Sundar et al., 2017). The inclusion of conversational AI voice agents as part of mobile devices is a recent development, and before their introduction, the average person never had experience interacting with AI technology that could respond to speech and exhibit human-like social cues (Guzman, 2019). Now, natural language processing allows individuals to talk to and receive replies from VA's in a way similar to their interactions with other people (McLean & Osi-Frimpong, 2019) .Thus, VA technology and individuals' interactions with them is an opportune and important area of

research given the limited understanding and investigation on the effects of interacting with VA's and their influence on humans.

Female and Male Gendered AI Roles. It is evident that female and male VA's are created with the stereotypical societal roles of men and women in mind (Stuko, 2019). Female voiced AI assistants are marketed to fulfill secretarial obligations as well as helping and service roles; They act as one's personal assistant and are meant to meet the needs of their owner (Piper, 2016). Male gendered AI, conversely, are largely made for machine learning and knowledge databases, such as the AI programs *Watson* and *Wolfram Alpha*, created to answer complex questions and perform quick calculations (Stuko, 2019). Additionally, though consumer research has shown that people generally prefer female voices over male ones (Griggs, 2011), the context in which the users experience these voices matters. For instance, female voiced computers created to perform a dominant role, such as giving commands or rating performance, were evaluated more negatively by users than male-voiced computers performing the same role (Nass et al., 2006). Thus, there is evidence that the tendency to gender stereotypes can happen even without the presence of another human being, extending to artificial intelligence and computerized machines.

Sexism & Attitudes Towards Women

Sexual harassment, gender-related humor, the glass ceiling, cat calling, gender discrimination, the pink tax, the gender pay gap, and more are all examples of sexist attitudes prevalent in society. Author Laura Bates, creator of the Everyday Sexism Project, when interviewed by The Takeaway podcast (2016) stated that sexism is "something that's occurring everyday" when asked why she started this massive project. The project now displays hundreds of thousands of women's personal experiences of sexism (everydaysexism.com). Despite

progressive efforts towards gender equality made in history, the Everyday Sexism Project demonstrates how devastatingly prominent sexism still is today.

Sexism can take many forms and ranges from more overt, hostile behaviors and statements to more covert, "benevolent" beliefs about the different roles of women and men in society. We will detail the different types below.

Hostile Sexism. Hostile sexism is rooted in the belief that men have a higher competence than women and are more deserving of higher status and power over women. With this belief also comes the fear that women, with their sexuality and femininity, are trying to steal power from men (Becker & Wright, 2011). Hostile sexism is often accompanied by power and dominance (Feather, 2004*)*, social dominance orientation (Sibley, Wilson, & Duckitt, 2007), and encouragement of unpleasant female stereotypes (Glick et al., 2000). Hostile sexism is easy to notice and identify because its' directly insulting and biased insinuations are often rejected in today's society.

Benevolent Sexism. Benevolent sexism, as defined by Glick & Fiske (1996), is a "set of interrelated attitudes toward women that are sexist in terms of viewing women stereotypically and in restricted roles, but that are subjectively positive in feeling tone (for the perceiver) and also tend to elicit behaviors typically categorized as prosocial or intimacy-seeking". Benevolent sexism, while often seen as chivalrous, is instead reinforcing the idea that women are weak and dependent upon men to protect them, provide for them, and keep them pure. Benevolent sexism has even been shown to be more indirectly harmful for women than hostile sexism. Dardenne, Dumont, & Bollier (2007) found that benevolent sexism was more detrimental to a woman's occupational performance than hostile sexism. Benevolent sexism has also been found to

undermine social change because it is less likely to be perceived as an issue, where hostile

sexism promotes social change because of its harsh nature (Becker & Wright, 2011).

Current Study

More research needs to be done investigating how interactions with female AI voices are

affecting stereotypes for women, as there is a gap concerning this area of research. The United

Nations (2019) likewise expressed a great need for more research into the gender disparities

present in society. The wide gap in gender disparities in reference to technology will continue to

grow as technological advances are made if more action and research is not taken to address this

problem. Thus, we aimed to investigate two relevant hypotheses tested in the present study: First,

we hypothesized that listening to a female AI voice assistant would increase sexism ratings,

specifically benevolent sexism, compared to listening to a male AI voice assistant. And secondly,

we hypothesized that listening to a female AI voice assistant would increase traditional attitudes

towards women compared to listening to a male AI voice assistant.

Study 1

Method

Participants

The sample consisted of 62 participants (18 male, 40 female, and 4 non-binary). The ages

surveyed ranged from 18 to 61 (M = 26.1), most of which are students from the University of

Texas at Tyler using the SONA program offered. Other participants were recruited through social

media outlets, online survey platforms, and SurveyTandem and r/samplesize on Reddit.

Measures

Demographic Questions. Using a self-created demographic questionnaire, participants were asked to indicate their age, ethnicity, gender, and if they own AI voice assistant products or have experience using such products.

Attitudes towards Women Scale (AWS). We measured traditional and pro-femenist attitudes towards women based on the Spence-Helmreich (1972) Attitudes towards Women Scale (AWS). Specifically, we used an abridged version of the scale consisting of 15 questions which is highly correlated with the original 55 questions version of the AWS for both males and females (Spence & Helmreich, 1978). The AWS has long been the standard measure of sexism employed by researchers, and this shorter version is currently the most used in research due to its more concise manner of measuring attitudes towards women.

The AWS presents participants with statements describing different attitudes about the roles women have in society and asks them to indicate their level of agreement from four answer choices, with A=agree strongly, B= agree mildly, C= disagree mildly, and D=disagree strongly. In scoring items, A=0, B=1, C=2, D=3 except for a couple of items where the scale is reversed. A total high score indicates a profeminist, egalitarian attitude, while a low score indicates a traditional, conservative attitude.

Ambivalent Sexism Inventory (ASI). To measure characteristics of hostile and benevolent sexism, we used the Ambivalent Sexism Inventory (Glick & Fiske, 1996). This inventory is based on a multidimensional understanding of sexism that emcomposses both negative and positive attitudes. Hostile sexism reflects the typically held conceptualization of sexism, consisting of overtly negative perceptions and stereotypes about women (Alport, 1954). Benevolent sexism is defined by Glick and Fish (1996) as sexist attitudes that consist of stereotypical views towards women and the roles they can occupy. However, these attitudes are

subjectively "positive" in intention and tend to elicit prosocial helping and protective behaviors towards women.

In order to measure these attitudes as well as levels of overall sexism, the ASI presents people with statements concerning men, women, and their relationship in society. Participants are asked to indicate how much they agree or disagree with each statement on a scale of 1-5, with 1=Disagree strongly and 5=Agree strongly. The measure of overall sexism can then be found by averaging their combined score for all the items in the inventory, and the two subscales of hostile and benevolent sexism can be calculated separately.

Helpfulness and Competence Questions. After we asked the participants questions from the two scales, we additionally presented them with self-created questions intended to measure how helpful and competent they felt Siri's answers were. Examples of these questions are as follows: How helpful were Siri's responses?; How competent was Siri at providing correct answers?; How satisfied were you with Siri's overall assistance? All the responses to these items were measured on 10-point likert scales.

Materials

We used the online survey platform Qualtrics to create our survey due to its user-friendly design and capability to embed audio clips that participants can click on and listen to. We also used recorded audio clips of Apple's *Siri* from an IPhone, using both the female and male voice features in our survey.

We additionally created a short quiz consisting of 8 fictitious questions and answers about a fake country for the participants to complete with Siri's assistance. We decided to create questions about a fictitious country so that the participants would have to rely on Siri to tell them the answers. This way, they could not simply disregard Siri's responses and use their own

knowledge to correctly complete the quiz. Examples of quiz questions are as follows: What year did King Richard begin his reign over the country of Estrana?; What color is Estrana's national flag?; What was Estrana's population in 2019?

Procedure

In order to recruit participants, the researchers posted the link to the survey on social media (including Facebook, Instagram, Reddit) and posted the survey on Sona.

Once participants followed the link provided, they were directed to an anonymous, online survey. The survey began with a brief explanation of what the participants would be doing during the study and assured them that all responses were for research purposes only and would not be shared or distributed outside of the research lab. Participants were also presented with the benefits and risks of the survey, their rights as a participant, and the contact information of the researchers if they had any concerns. After this, participants were asked if they consented to participate in the study. If the participants agreed, they were brought to a demographics page which they filled out with their age, ethnicity, gender, and if they had AI technology products of their own or had access to these products. Next, they were given a pre-test for their current attitudes towards women via the AWS and their current ambivalent sexism scale via the ASI. Next, participants were brought to an instructional page on how to properly interact with the VA and answer the trivia that followed. Participants were then directed to a set of 8 trivia questions. The set of questions created for the trivia quiz were about a fictional country. We chose to make the quiz about a set of fake trivia so as to make it impossible for participants to "know" the answer without asking for the assistance of the VA. Each question had an embedded gendered VA that would tell them the answer to the trivia question and the participants had to rely on their VA to provide them with the "correct" answer. It is important to note that there were two

conditions in which participants were randomly assigned: 1) female VA and 2) male VA. These conditions were counterbalanced and both conditions had identical trivia questions with 3 questions that the VA would give an "incorrect" answer for. Once participants completed the trivia portion of the survey, they received their grade and noticed that their VA had given them the correct answer to each question. After they completed the trivia portion and received their grade, they were asked to rate the helpfulness, competence, and overall satisfaction with their VA. Lastly, participants were given a post-test of the AWS and ASI, debriefed, and thanked for their time and participation.

<div align="center">

Results

</div>

Effects of Voice Assistant's Gender on Hostile and Benevolent Sexism

First, we tested our hypothesis that listening to a female AI voice assistant would increase sexism ratings, specifically benevolent sexism, compared to listening to a male AI voice assistant. Results of an independent sample T-test found that the gender of the Siri each participant listened to did not have an effect on how sexist participants scored on the benevolent and hostile sexism post-test, $F(57)=.275$, p=NS; $F(57)= 2.639$, p=NS (see Table 1). Our results demonstrated that there was not a significant difference between the pre-tests and post-tests in either condition.

Effects of Voices Assistant's Gender on Attitudes Toward Women

Next, we tested our second hypothesis that listening to a female AI voice assistant would increase traditional attitudes towards women, compared to listening to a male AI voice assistant. Results from an independent sample T-test found that, similar to the results of our first hypothesis, the gender of the Siri each participant listened to did not have an effect on their

attitudes toward women post-test, $F(57)=6.338$, p=NS (see Table 1). Our results indicate that

there was not a significant difference between the pre-tests and post-tests in either condition.

Effects of Voice Assistant's Gender on Perceived Usefulness

Finally, we decided to examine if participants would rate the male and female Siri

differently for their helpfulness, competence, and overall satisfaction (for compared means

between these variables see Figure 1). Using an analysis of covariance controlling for

pre-existing sexism, results revealed that neither condition of the Siri voice had a significant

effect on how helpful, competent, and satisfactory participants rated Siri to be, $F(1,59)=.261$,

p=NS; $F(1,59)= .266$, p=NS; $F(1,59)=.031$, p=NS (see Tables 2,3, & 4). These results imply that

there was no significant difference between how participants in either condition rated their Siri

on helpfulness, competency, and overall satisfaction; Regardless of the gender of Siri's voice,

participants rated Siri performance similarly.

<div align="center">**Discussion**</div>

The primary goal of this research was to investigate whether people projected their sexist

views about women onto female voiced artificial intelligence. Contrary to our first hypotheses,

our findings show that, compared to interacting with a male voice assistant, interacting with a

female voice assistant did not result in a significant increase in participants' benevolent sexism.

Additionally, contrary to our second hypothesis, the gender of Siri's voice that participants

interacted with did not have an effect on their traditional attitudes toward women. Furthermore,

we found no effect of the gender of Siri's voice on participant's ratings of Siri's helpfulness,

competence, and their overall satisfaction.

Surprisingly, our results did not support the existing literature that people perceive, treat,

and rate male and female AI differently (Griggs, 2011; Nass, et al., 1997; Piper, 2016). One

possibility for our inconsistent findings with the literature could be the highly skewed number of young people who participated in our study (M = 26.1). Since we did not have a larger amount of older people from different generations, we collected data from a young generation that, while sexism is still very much prevalent in society, are typically taught more pro-feminist ideals (Gilmore, 2001). Consequently, this younger generation has largely abandoned the traditionally held attitudes towards women, thus resulting in lower levels of sexism. Since older generations tend to display more overt sexist attitudes towards women, largely in part to their outdated views of gender roles, a sample more representative of such older ages would have likely resulted in greater changes and differences in sexist attitudes as a result of Siri's gender.

Additionally, the discrepancy between our results and previous research findings might be explained by the pretest sexism scales we had the participants complete. The questions asked in the pre-test could have given participants an indication of what our research was looking for. Since the pre-test asked specific questions related to sexist views, participants might have intentionally chosen to rate the female voiced Siri favorably in an effort to appear less sexist.

Limitations

A limitation of this study was that there were only 62 participants. While this number of participants was sufficient enough for us to perform our analyses, it would have been beneficial to have a larger sample size to improve the reliability and validity of our results. Also, our survey design was considerably long—as it took participants about 15 to 25 minutes to complete—and multiple participants complained about the length of the survey, as well as the repetitive questions, since they were asked to complete the AWS and ASI twice due to the pretest-posttest design of our study. Consequently, we likey lost many participants' attention the further they progressed into the survey, resulting in potentially negligent and inaccurate responses to our

sexism measures. Additionally, as mentioned previously, we had a younger skewed participant population. Thus, it is likely that we were unable to capture the traditional attitudes toward women typically held by older generations, and instead examined a population that has been heavily exposed toward pro-feminist attitudes and who has largely been habituated to voice assistant technology. Overall, the findings from our first study show that regardless of the participant's gender or Siri's gender, participants rated Siri's performance similarly when Siri provided them with incorrect answers to the quiz. This suggests that participants actually held more progressive attitudes towards women, contrasting past research indicating that people tend to rate female AI voices more critically than male AI voices (Nass et al., 1997).

The main limitation of study 1, we believed, was the content of the quiz questions themselves. We believe that participants were not deceived by the trivia questions about the fictitious country, and thus it was hard to generalize about the effects of the Siri gender manipulation on such an unbelievable task. Thus, for study 2 we created a real but difficult trivia study for the participants to complete, and also gave them the choice to opt in or out of Siri help after hearing the voice introduce the AI assistant.

Study 2 Methods

Participants

This sample consisted of 272 female participants, 64 male participants, 5 non-binary participants, and 1 participant that identified as "other" for a total of 342 participants. The ages ranged from 18 to 69 (M=23.6), most of which were students from The University of Texas at Tyler. The rest of the participants were found through social media outlets. Out of the 342 participants, 203 identified as White, 73 identified as Hispanic, 38 identified as Black, 19 identified as Asian, and 9 either identified as Native Hawiian, American Indian, or "other".

Measures

Demographic Questions. Participants completed the same questions as in Study 1.

Ambivalent Sexism Inventory (ASI). Participants will answer the same questions from this scale in Study 1. The scale has subscales measuring for hostile and benevolent sexism.

Attitudes Toward Women (ATW). Participants were given the same scale as in Study 1. This scale measured traditional versus pro-feminist attitudes.

Helpfulness and Competence Questions. Participants answered the same three questions as in Study 1. However, for participants that chose to opt out of using the help from Siri, they were not shown the questions on how helpful, competent, and satisfied they were with Siri.

Materials

The main addition to the materials in study 2 was a re-thinking of the nature of the quiz that participants would take and potentially need Siri's help for.

We created a small trivia quiz in the survey. The quiz consisted of random questions at a medium difficulty level. Examples of these questions include what is the sixth digit in Pi?; Area 51 is located in which state?; What year was McDonald's founded?; What was the original name of "Google"?; Where is the pineapple plant originally from?

Procedure

For the recruitment of participants, a link to the survey was sent out through the SONA program at The University of Texas at Tyler and other social media outlets.

Once the link was followed by participants, they were given the consent form to read, which stated the risks and advantages of the research along with their rights and the main

researcher's contact information for any questions. If they said yes to giving their consent for the

research, they were taken to the demographic questions asking for their age, ethnicity, gender,

and if they had any experience with Siri. Next, participants were given a chance to hear and

interact with the Siri version they randomly received, male or female, and asked if they would

like Siri's help on the trivia questions that appeared on the next page. If they chose yes, they

were shown five trivia questions and a recording of the male or female Siri telling them the

answers. However, two of said answers were incorrect. They were then shown their score, which

was a three out of five, on the trivia and presented with questions asking how competent, helpful,

and satisfying Siri was to them. If they chose no, the trivia questions were presented without the

Siri recordings and after they received their score, they were not given the questions about Siri's

performance. After the participants finished with the trivia, they were shown the post-test with

the Ambivalent Sexism Inventory and the Attitudes Toward Women Scale questions. Once

participants answered the post-test they were thanked for their contribution and exited the survey.

<div align="center">

Results
</div>

Siri's Gender and Asking for Help

The first we conducted was a chi-square test of independence. Results of the chi-square

test of independence found that people did not ask for help at different rates from male versus

female Siris ($\chi^2 = .33$, $p = .57$) (see Table 5 & Figure 2). We then conducted a chi-square test of

independence to see if the participants' gender had an effect on whether they asked for help from

Siri. Results of the chi-square test of independence found that people did not ask for help at

different rates depending on their gender ($\chi^2 = .13$, $p = .719$) (see Table 6 & Figure 3).

Effects of Gender on Perceptions of Siri

We next examined the effects of Siri gender and participant gender on perceptions of Siri's helpfulness using a 2 (Siri: male, female) X 2 (Participant: male, female) ANOVA. Results showed that there was not a significant effect of the interaction between Siri gender and participant gender on the helpfulness variable, $F (1, 192) = .68$, $p = .41$. The main effects were also non-significant (see Table 7).

Siri's Gender and Trivia Score

Finally, we examined whether Siri's gender was related to participants' scores on the knowledge quiz. We filtered out the participants who chose not to use Siri's help on the trivia quiz before preforming this test. An independent samples t-test showed that there was a significant effect of the Siri gender manipulation on quiz scores, $t (198) = -2.27$, $p = .024$. This demonstrated that participants scored higher with the female Siri then the male Siri.

Discussion

In an attempt to address the limitations of the first study, the second study followed a similar study design addressing the same hypotheses. In addition to the hypotheses from Study 1, we also hypothesized that participants would criticize the female Siri more since they were given the option to accept or deny Siri's help. However, participants did not significantly rate the male and female Siri's differently even when they chose to accept Siri's help. Participants also did not ask for help from either Siri at a significant difference. There was no difference in who asked for help depending on the participants' gender or based on Siri's gender either. However, there was a significant difference in the trivia score for participants with a female Siri then a male Siri. As previously stated, both versions of Siri gave two incorrect answer choices during the trivia quiz meaning that the participants were more likely to disregard the female Siri's wrong answer than

the male Siri's. This indicates that participants chose not to believe or listen to the female Siri as much as the male Siri, since the participants with the female Siri scored higher on the quiz.

Once again, most of our findings have not supported the existing literature that people treat gendered artificial intelligence with societal sexism imposed on humans. While we did a better job at recruiting participants for this second study, there was still a large number of younger participants and women. The mean for this new study (M=23.6) was lower than the first study's mean (M=26.1) indicating that this second study had a similar issue to the first study. This could be one of the reasons that our hypotheses were not supported considering younger generations are seen as more accepting and less sexist compared to other generations as previously stated (Gilmore, 2001). The same could be said about women since a majority of our sample was women.

Limitations

Another possibility for why our findings did not support the existing literature is that, due to the nature of our virtual survey, participants were unable to interact with Siri in the same way that one would normally interact with their voice assistant. Normally, people need to verbally ask Siri a question and then wait for a response. However, for our online survey, we uploaded pre-recorded Siri responses for participants to listen to, so they were unable to verbally ask Siri a question. This resulted in a one-sided experience that was not representative of typical back-and-forth interactions with Siri, which might explain for the lack of effect on participants' ratings and sexism levels.

A limitation of this study would be the lack of age diversity. Since most of our participants were college students, we were unable to collect data from more participants from other generations. Since the younger generations have grown up in a world with technology and

artificial intelligence, it is more reasonable to assume from our data that they have an easier time separating humans from technology than the generations who were older when artificial intelligence was created. Another limitation is the lack of racial diversity. Over half (59%) of the participants identified as "white" which means we did not collect data from an equal amount of different races. This makes it difficult to rationalize what the data proves onto all races when it was a majority white population.

Strengths

In response to the need of a future study after Study 1, our second trial had more strengths than weaknesses this time around. Our study has added extra literature to the small amount already circling. Now having done two trials, our data has more significance for the research community regarding how humans interact with AI. In addition to the increase in literature, our research also adds to the idea that younger generations have more progressive attitudes toward women and societal gender norms in general. This study demonstrated that regardless of Siri's gender or the participants gender, there is no significant difference in the rating of the Siri's or whether they ask for help from Siri. This adds to the ongoing question on how people treat AI and how they affect human views on others.

Future Research

Previous research suggests that humans will treat female voiced AI differently than male voiced AI, while our data does not support the prior literature, new research should focus on older generations. Doing research on each of the other generations would provide a better picture of rationalizing how most humans will interact with gendered AI. In response to gendered AI studies, there should be more information on how non-binary AI would affect how people use and interact with them. Also, since we did find significance between Siri's gender and

participant's trivia scores, it is important to continue research looking into why there was a significant difference. Finally, the next best way to use this information would be to study the long-term effects of gender voiced AI through a longitudinal study of participants interacting with AI. Knowing the long-term effects would provide information on how gender voiced AI will affect upcoming generations and the best way to help provide good experiences and relationships with AI.

Conclusion

Our research findings did not support our hypotheses; however, it did offer valuable information for a better way to construct a similar experiment moving forward. Our findings suggest that this type of research would be better conducted in person where participants can have live interactions with Siri in order to best elicit their most accurate feelings and attitudes. This project has great potential to be conceptually replicated and expanded upon to allow for greater data collection about the effects of voice assistants on the human person, and we hope our study serves to encourage further investigation into this timely and important area of research.

References

Alport, G. W. (1954). *The nature of prejudice*. Reading, MA: AddisonWesley.

Becker, J., & Wright, S. (2011). Yet Another Dark Side of Chivalry: Benevolent Sexism

Undermines and Hostile Sexism Motivates Collective Action for Social Change. *Journal*

of Personality and Social Psychology, 101(1), 62-77.

Coppin, B. (2004). *Artificial intelligence illuminated*. Boston: Jones and Bartlett.

Dardenne, B., Dumont, M., & Bollier, T. (2007). Insidious Dangers of Benevolent Sexism:

Consequences for Women's Performance. *Journal of Personality and Social Psychology,*

93(5), 764-779.

Deaux, K. & Lewis, L.L. (1984). Structure of gender stereotypes: Interrelationships among

components and gender label. *Journal of Personality and Social Psychology, 46*(5).

991-1004 .https://psycnet.apa.org/doi/10.1037/0022-3514.46.5.991

Dellaert, B.G.C., Shu, S.B., Arentze, T.A. et al. (2020). Consumer decisions with artificially

intelligent voice assistants. *Marketing Letters, 31,* 335–347.

https://doi.org/10.1007/s11002-020-09537-5

Feather, N. T. (2004). Value correlates of ambivalent attitudes toward gender relations.

Personality and Social Psychology Bulletin, 30, 3-12. doi: 10.1177/0146167203258825

Fitzpatrick, M.J., McPherson, B.J. Coloring Within the Lines: Gender Stereotypes in

Contemporary Coloring Books. *Sex Roles* 62, 127–137 (2010).

https://doi.org/10.1007/s11199-009-9703-8

Gilmore, S. (2001). Looking back, thinking ahead: Third wave feminism in the United States.

Journal of Women's History, 12(4), 215-221.

Glick, P. & Fiske, S. T. (1996). The ambivalent sexism inventory: Differentiating hostile and

    benevolent sexism. *Journal of Personality and Social Psychology, 70*(3),

    491-512.https://doi/10.1037/0022-3514.70.3.491.

Glick, P., Fiske, S. T., Mladinic, A., Saiz, J. L., Abrams, D., Masser, B., Lopez, W. (2000).

    Beyond prejudice as simply antipathy: Hostile and benevolent sexism across cultures.

    Journal of Personality and Social Psychology, 79, 763-775.

Griggs, B. (2011, October 21). *Why Computer Voices Are Mostly Female*. Retrieved from

    https://www.cnn.com/2011/10/21/tech/innovation/female-computer-voices/index.html

Guzman, A. L. (2019). Voices in and of the machine: Source orientation toward mobile virtual

    assistants. *Computers in Human Behavior*, *90*, 343-350.

    https://doi.org/10.1016/j.chb.2018.08.009.

Hwang, G., Lee, J., Oh, C. Y., & Lee, J. (2019). It sounds like a woman. *Extended Abstracts of

    the 2019 CHI Conference on Human Factors in Computing Systems*.

    https://doi.org/10.1145/3290607.3312915.

Jenen, J., Winquist, J., Arkkelin, D., & Schuster, K. (2009). Implicit attitudes towards feminism.

    Sex Roles, 60, 14-20. http://dio.org/10.1007/s11199-008-9514-3.

Jung, K., Shavitt, S. K., Viswanathan, M., & Hilbe, J. M. (2014). Female hurricanes are deadlier.

    Proceedings of the National Academy of Sciences, *111*(24), 8782-8787. DOI:

    10.1073/pnas.1402786111

McLean, G. & Osei-Frimpong, K. (2019). Hey Alexa … examine the variables influencing the

    use of artificial intelligent in-home voice assistants. *Computers in Human Behavior*, *99*,

    28-37. https://doi.org/10.1016/j.chb.2019.05.009.

Nass, C., Moon, Y., & Green, N. (1997). Are machines gender neutral? Gender-stereotypic

responses to computers with voices. *Journal of Applied Social Psychology, 27*(10),

864-876. https://psycnet.apa.org/doi/10.1111/j.1559-1816.1997.tb00275.x

Nations, U. (2019). I'd Blush If I Could: Closing Gender Divides in Digital Skills Through

Education. *Equals Global Partnership.* 13-145

Parker, K., Funk, C. (2017, December 14). *Gender discrimination comes in many forms for*

today's working women. Pew Research Center.

https://www.pewresearch.org/fact-tank/2017/12/14/gender-discrimination-comes-in-many

-forms-for-todays-working-women/

Piper, Allison M., (2016). Stereotyping femininity in disembodied virtual assistants. *Graduate*

Theses and Dissertations. 15792. https://lib.dr.iastate.edu/etd/15792

Rometty, G. (2016). Digital today, cognitive tomorrow. *MIT Sloan Management Review, 58*, p

2-4

Sibley, C. G., Wilson, M. S., & Duckitt, J. (2007). Antecedents of men's hostile and benevolent

sexism: The dual roles of social dominance orientation and right-wing authoritarianism.

Personality and Social Psychology Bulletin, 33, 160-172. doi:

10.1177/0146167206294745

Spence, J.T. & Helmreich, R. (1972). The attitudes toward women scale: An objective instrument

to measure attitudes toward the rights and roles of women in contemporary society.

Catalog of Selected Documents in Psychology, 2(66).

Spence, J.T. & Helmreich, R. (1978). *Masculinity and femininity: Their psychological*

    dimensions, correlates, and antecedents. Austin, TX: University of Texas Press.

Stuko, D. M. (2019). Theorizing femininity in artificial intelligence: A framework for undoing

    technology's gender troubles. *Cultural Studies*, *34*(4), 567-592.

    https://doi.org/10.1080/09502386.2019.1671469.

Sundar, S. S., Jung, E. H., Waddell, F. T., & Kim, K. J. (2017). Cheery companions or serious

    assistants? Role and demeanour congruity as predictors of robot attraction and use

    intentions among senior citizens. *International Journal of Human-Computer Studies*, *97*,

    88-97. https://doi.org/10.1016/j.ijhcs.2016.08.006.

Swim, J., Hyers, L., Cohen, L., & Ferguson, M. (2001). Everyday sexism: Evidence for its

    incidence, nature, and psychological impact from three daily diaries. *Journal of Social*

    Issues, 57(1), 31-53.

Williams, R. & Wittig, M. (1997). "I'm not a feminist but…": Factors contributing to the

    discrepancy between pro-feminist orientation and feminist social identity. *Sex Roles,*
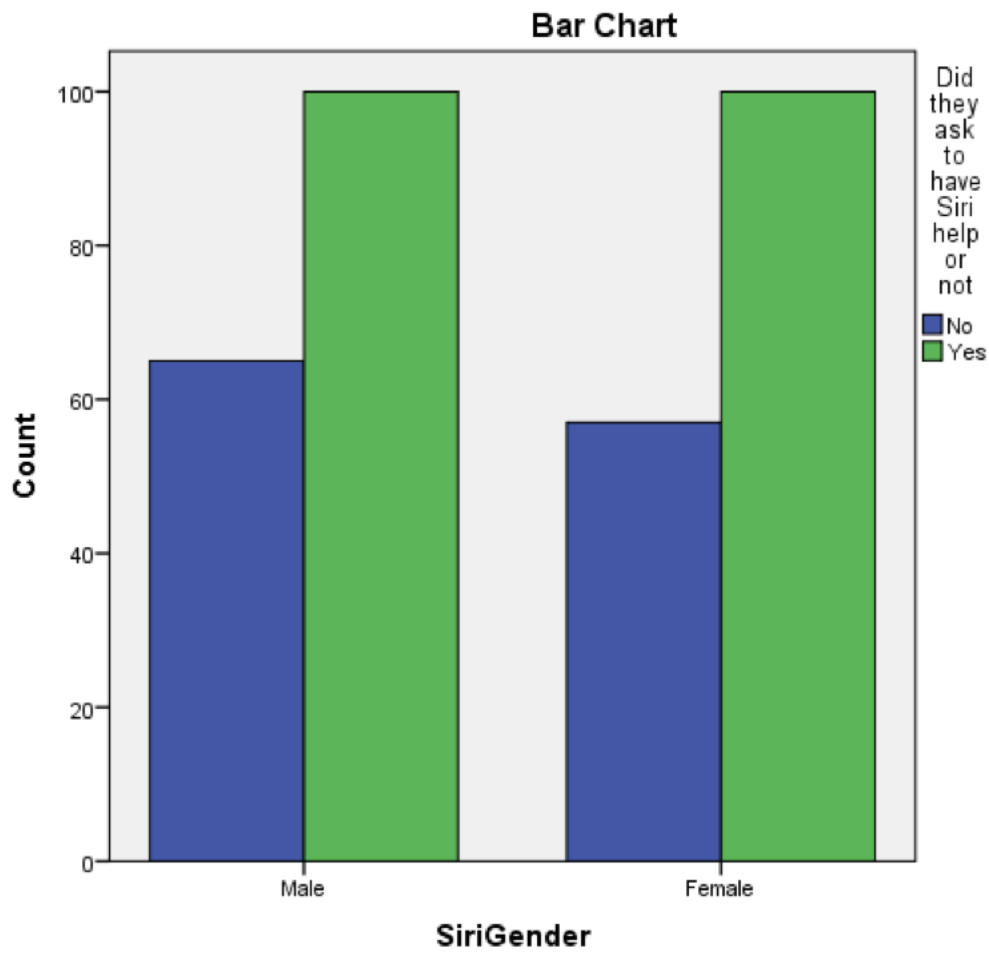
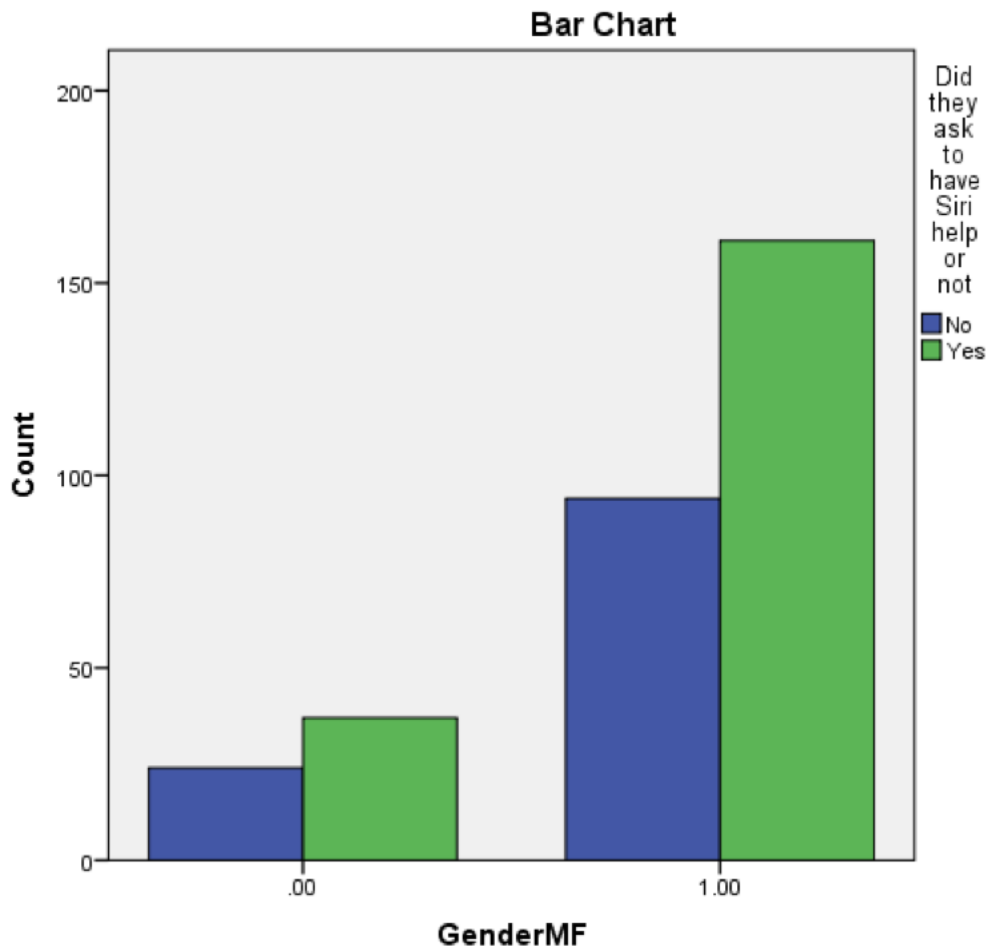   *37(11),* 885-903.

Figures

Figure 1
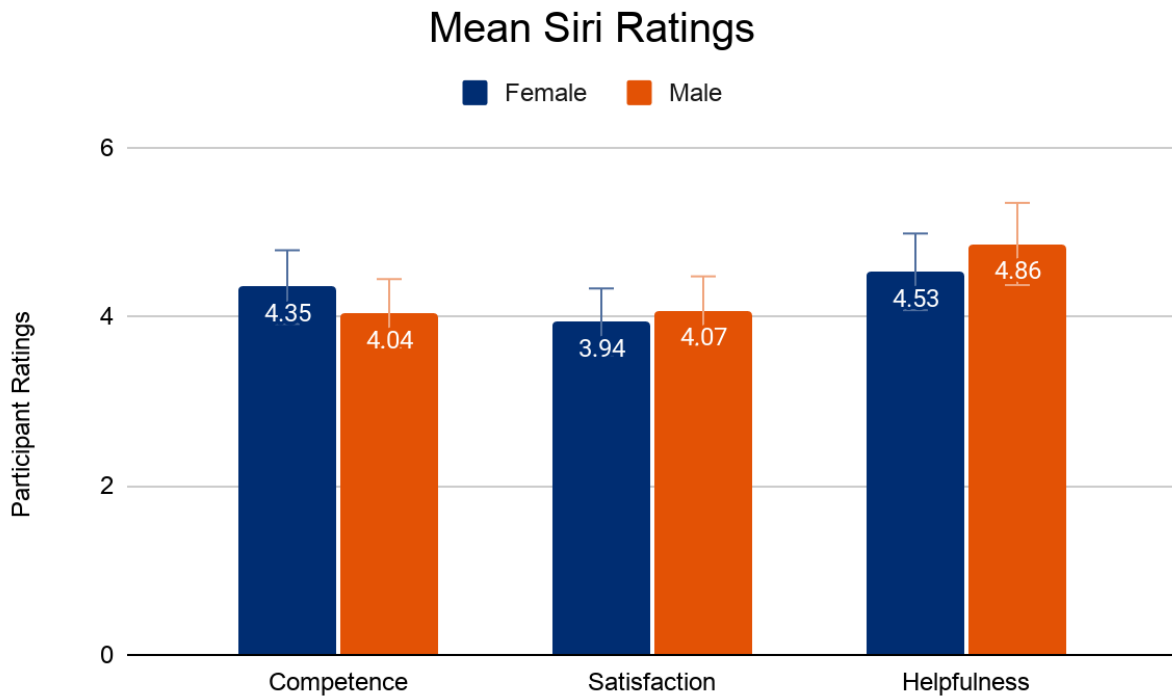
Note. Participant's rated Siri's usefulness on a scale of 1-10. The bars in blue represent the
average ratings of the female Siri while those in orange represent the average ratings of
the male Siri.

Figure 2

Note. This demonstrates the number of people that asked for Siri's help depending on which Siri

gender they received.

Figure 3

Note. This demonstrates the number of people who asked for Siri's help based on the

participant's gender (.00=Male, 1.00=Female).

Table 1

Differences in ASI and ATW Pre-test and Post-test

	F	Sig.
Hostile Sexism (ASI)	2.635	.110
Benevolent Sexism (ASI)	.274	.603
ATW Scale	6.338	.015

Note: Values are from an ANCOVA test for significance.

Table 2

T-test of Siri's Helpfulness Ratings

Source	df	Mean Square	F	Sig.
Pre-Test	1	3.255	.545	.436
Condition	1	1.558	.261	.612
Error	59	5.977		

Note: Participants rated how helpful they found their version of Siri; We controlled for their

pre-existing levels of sexism.

Table 3

T-test of Siri's Competence Ratings

Source	df	Mean Square	*F*	Sig.
Pre-Test	1	9.596	1.501	.225
Condition	1	1.701	.266	.608
Error	59	6.392		

Note: Participants rated how competent they found their version of Siri. We controlled for their pre-existing levels of sexism.

Table 4

T-test of Siri's Satisfaction Ratings

Source	df	Mean Square	*F*	Sig.
Pre-Test	1	8.783	1.325	.254
Condition	1	.204	.031	.861
Error	59	6.626		

Note: Participants rated how satisfied they were with their version of Siri. We controlled for their pre-existing levels of sexism.

Table 7

ANOVA Participant Gender x Siri Gender

Source	df	Mean Square	F	Sig.
GenderMF	1	.804	.158	.691
SiriGender	1	6.28	1.233	.268
GenderMF * SiriGender	1	3.445	.679	.411
Error	192			

Note. This demonstrates the effect of participant's gender on the perception of Siri.

Appendix A

Ambivalent Sexism Inventory Scale

1. No matter how accomplished he is, a man is not truly complete as a person unless he has the love of a woman.

2. Many women are actually seeking special favors, such as hiring policies that favor them over men, under the guise of asking for "equality."

3. In a disaster, women ought not necessarily to be rescued before men.

4. Most women interpret innocent remarks or acts as being sexist

5. Women are too easily offended.

6. People are often truly happy in life without being romantically involved with a member of the other sex

7. Feminists are not seeking for women to have more power than men.

8. Many women have a quality of purity that few men possess.

9. Women should be cherished and protected by men.

10. Most women fail to fully appreciate all that men do for them.

11. Women seek to gain power by getting control over men.

12. Every man ought to have a woman whom he adores.

13. Men are complete without women.

14. Women exaggerate problems they have at work.

15. Once a woman gets a man to commit to her, she usually tries to put him on a tight leash.

16. When women lose to men in a fair competition, they typically complain about being discriminated against.

17. A good woman should be set on a pedestal by her man.

18. There are actually very few women who get a kick out of teasing men by seeming

    sexually available and then refusing male advances.

19. Women, compared to men, tend to have a superior moral sensibility.

20. Men should be willing to sacrifice their own well being in order to provide financially for

    the women in their lives.

21. Feminists are making entirely reasonable demands of men.

22. Women, as compared to men, tend to have a more refined sense of culture and good taste.

Appendix B

Attitudes Toward Women Scale

1. Swearing and obscenity are more repulsive in the speech of a woman than a man.

2. Under modern economic conditions with women being active outside the home, men should share in household tasks such as washing dishes and doing laundry.

3. It is insulting to women to have the "obey" clause remain in the marriage service.

4. A woman should be free as a man to propose marriage.

5. Women should worry less about their rights and more about becoming good wives and mothers.

6. Women should assume their rightful place in business and all the professions along with men.

7. A woman should not expect to go to exactly the same places or to have quite the same freedom of action as a man.

8. It is ridiculous for a woman to run a locomotive and for a man to darn socks.

9. The intellectual leadership of a community should be largely in the hands of men.

10. Women should be given equal opportunity with men for apprenticeship in the various trades.

11. Women should be given equal opportunity with men for apprenticeship in the various trades.

12. Sons in a family should be given more encouragement to go to college than daughters

13. In general, the father should have greater authority than the mother in the bringing up of the children.

14. Economic and social freedom is worth far more to women than acceptance of the
    ideal of femininity which has been set up by men.

15. There are many jobs in which men should be given preference over women in
    being hired or promoted.

Appendix C

Study 1 Trivia Questions about a Fictitious Country

1. What year did King Richard begin his reign over the country of Estrana?

2. What is on Estrana's national flag?

3. What was Estrana's population in 2019?

4. What national animal represents Estrana?

5. What is Estrana's national flower?

6. What year was Estrana founded?

7. Who was Estrana's first monarch?

8. What is the capital of Estrana?

Appendix D

Helpfulness/Competence/Satisfaction Questions

1. On a scale of 1-10, how helpful would you rate Siri's overall performance?

2. On a scale of 1-10, how competent was Siri at providing you quiz answers?

3. On a scale of 1-10, how satisfied were you with Siri's overall performance?

Appendix E

Study 2 Trivia Questions

1. What is the sixth digit in Pi?

2. Area 51 is located in which state?

3. What year was McDonald's founded?

4. What was the original name of "Google"?

5. Where is the pineapple plant originally from?