You're Getting a Little Too Personal

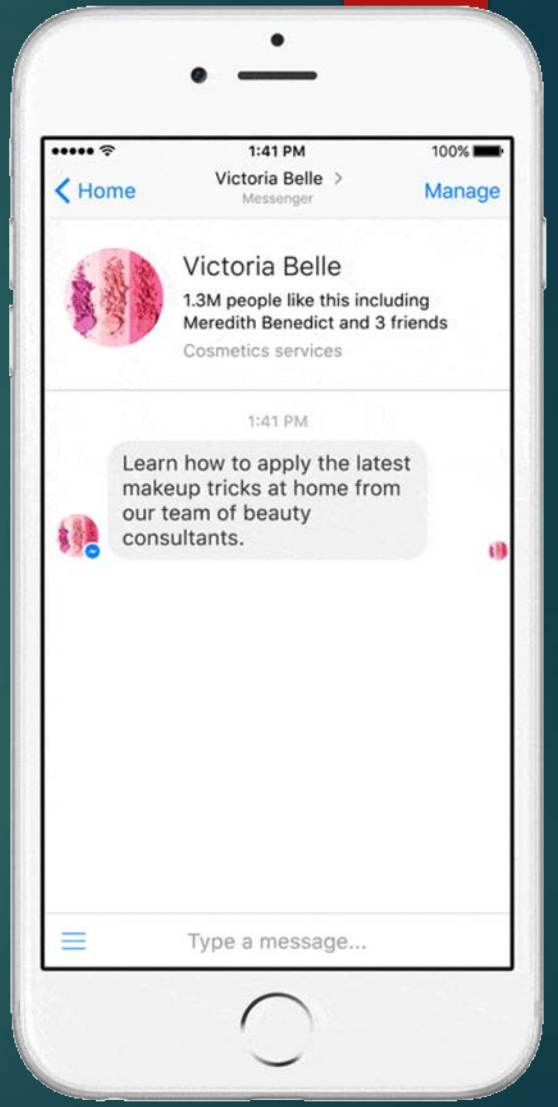
HOW GENDER PERSONALIZATION IN CHATBOT DESIGN AFFECTS CUSTOMER SATISFACTION

CHRIS TREMBATH

THE GREAT ATLANTIC & PACIFIC TEA COMPANY

Introduction

- A **chatbot** is a computer program or artificial intelligence device which conducts a conversation with a human via auditory or textual means.
  - Alexa (Amazon) and Siri (Apple), are examples of chatbot technology as is Facebook Messenger.
- 25% of companies will use chatbot technology for customer service operations by the year 2020 (Moore, 2018).
- 77% of consumers prefer using chatbots to bypass customer service wait times (McKendrick, 2018).



Literature Review

- Personalization in person-to-person customer service interactions was found to increase customer satisfaction (Mittal & Lassar, 1996; Ford, 2001). Personalization is using a customer's first name, smiling and making an effort to know the customer personally.
- Personalization in chatbot conversations was shown to increase likability and attractiveness ratings (Verhagen, Nes & Feldberg, 2014; Yoo & Gretzel, 2009). Personalization in a chatbot is defined as the use of avatars (virtual faces) and anthropomorphism (visual cues).



Research Questions

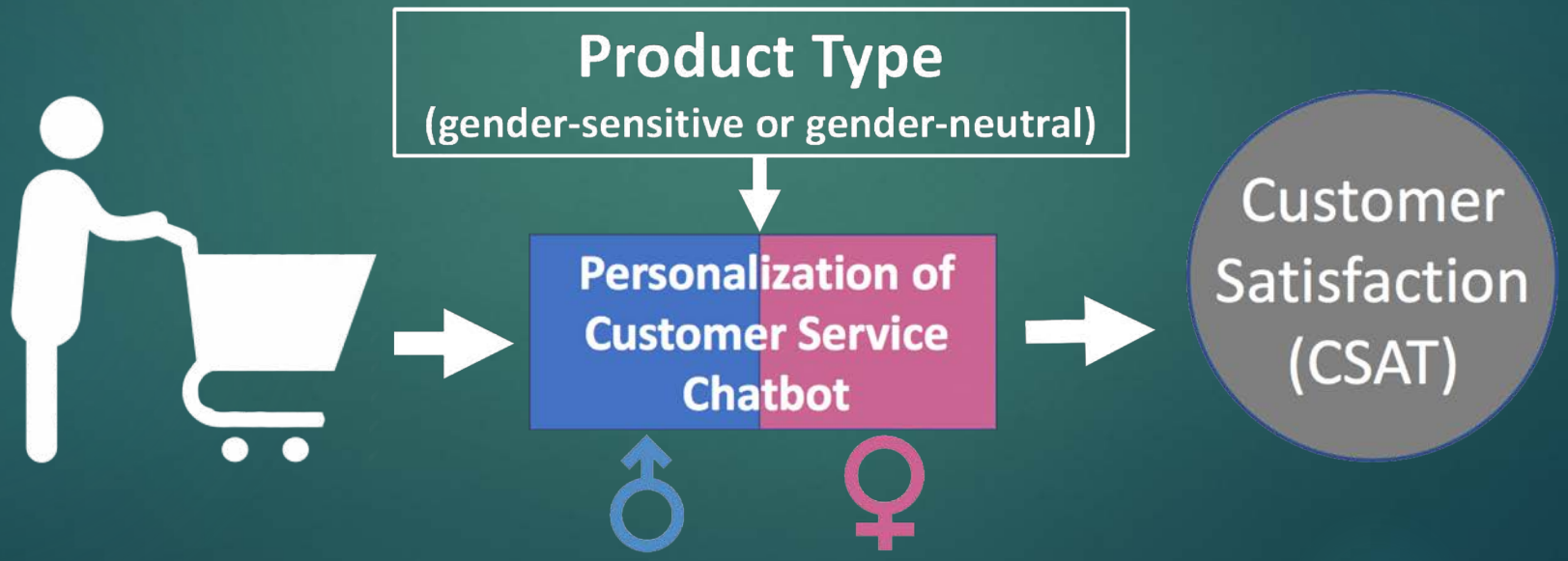
The lack of research concerning the gender identification of a chatbot and how it might affect customer satisfaction lead to the following research questions:

- R1: Does the gender of a chatbot affect a consumer's customer satisfaction in a customer service setting?
- R2: Does the product type influence the overall evaluation of the chatbot?



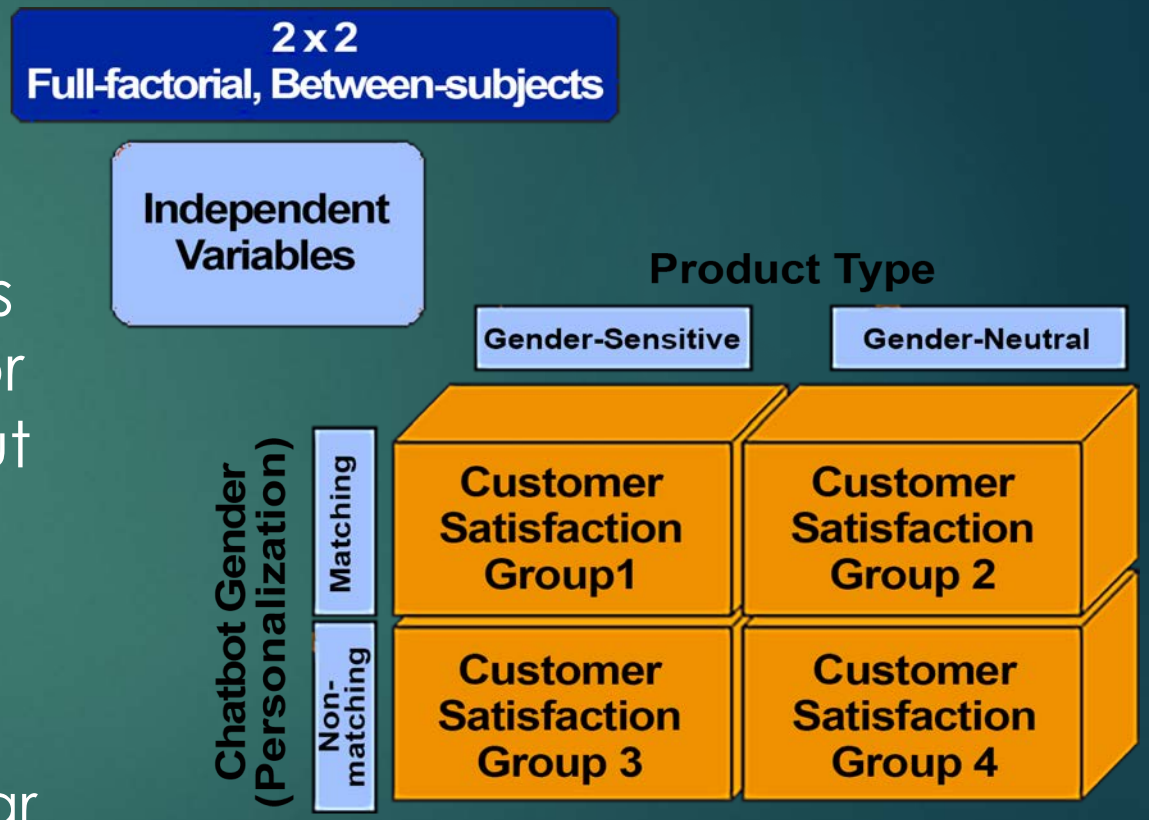
Hypotheses

- H1: A consumer's customer satisfaction will increase when interacting with a same-gender chatbot as opposed to a consumer interacting with a different-gender chatbot.
- H2: The product type will influence the relationship between the chatbot and the consumer's satisfaction.



Methodology

- A 2x2 full-factorial, between subjects experiment was conducted with 117 participants recruited from Amazon's MTurk Workforce Community. Participants interacted with a simulated male or female chatbot and inquired about a gender-sensitive or gender-neutral product.
- Product types:
  - Gender-specific: athletic underwear
  - Gender-neutral: athletic t-shirt.

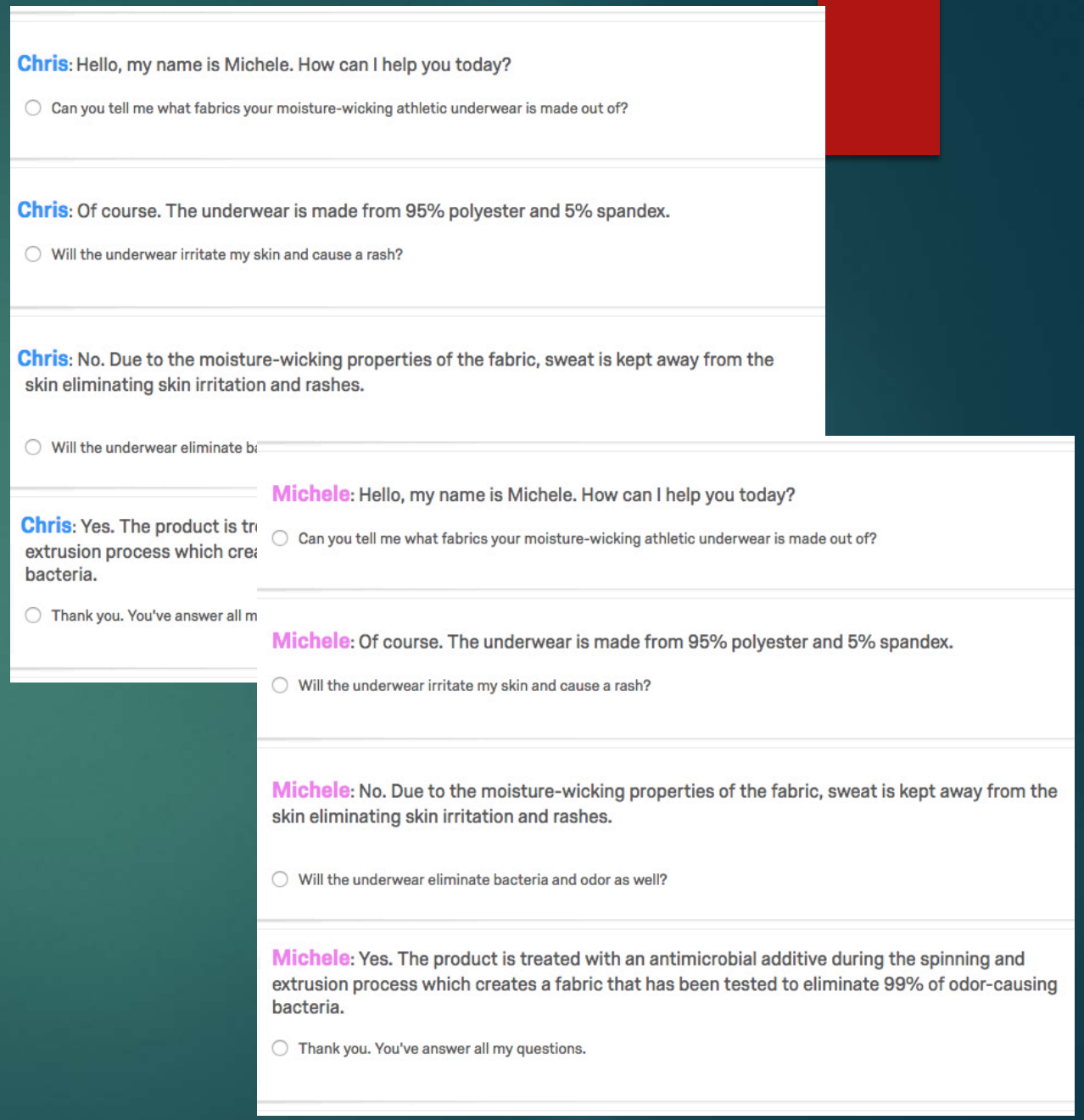


Methodology

Pre-tests

In order to operationalize gender in the chatbots, two pre-tests were conducted.

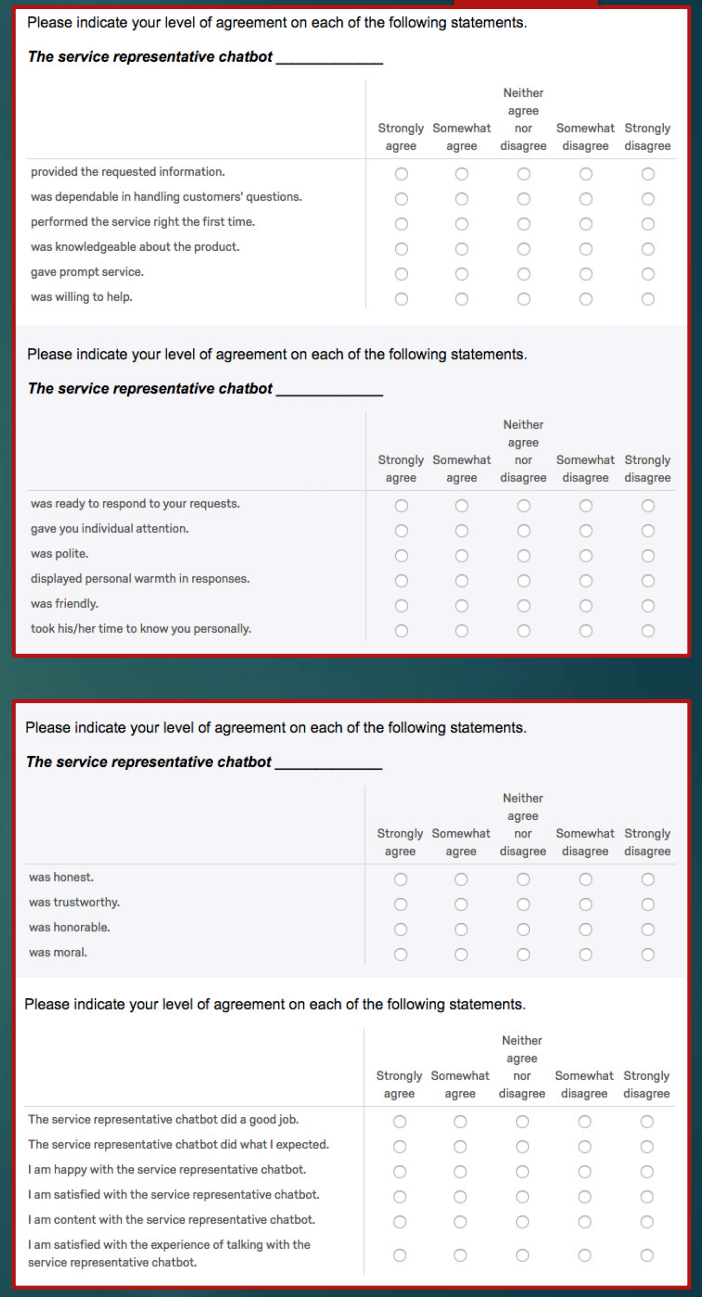
- A 21-participant pretest was conducted to determine ethnically neutral names for the chatbots: Chris and Michelle.
- A 22-participant pretest was conducted to determine colors for chatbot names: blue for male, pink for female.



Methodology

Measures & Analysis

- The chatbot evaluation was measured using the constructs of reliability, responsiveness, trustworthiness and satisfaction.
- Measures were adapted from a standard Service Quality (SERVQUAL) survey (Mittal & Lassar, 1996), and a Chatbot Credibility and Satisfaction Survey (Chung et al, 2018).
- To answer the research questions and test the hypotheses, descriptive analyses, factor analyses and ANOVAs were conducted.



Please indicate your level of agreement on each of the following statements.

The service representative chatbot _____

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
provided the requested information.	○	○	○	○	○
was dependable in handling customers' questions.	○	○	○	○	○
performed the service right the first time.	○	○	○	○	○
was knowledgeable about the product.	○	○	○	○	○
gave prompt service.	○	○	○	○	○
was willing to help.	○	○	○	○	○

Please indicate your level of agreement on each of the following statements.

The service representative chatbot _____

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
was ready to respond to your requests.	○	○	○	○	○
gave you individual attention.	○	○	○	○	○
was polite.	○	○	○	○	○
displayed personal warmth in responses.	○	○	○	○	○
was friendly.	○	○	○	○	○
took his/her time to know you personally.	○	○	○	○	○

Please indicate your level of agreement on each of the following statements.

The service representative chatbot _____

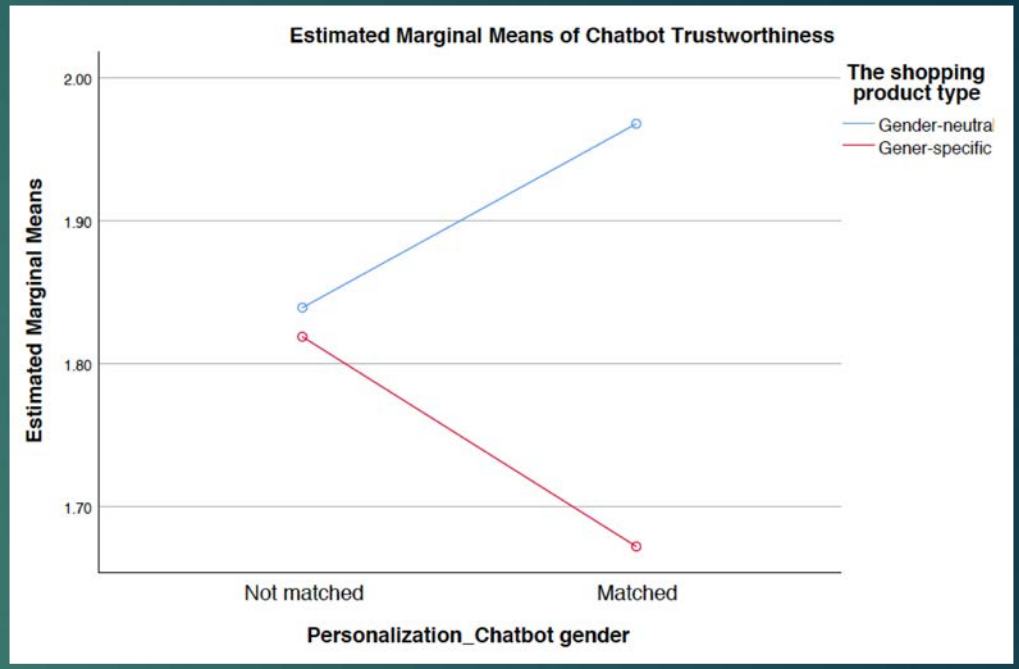
	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
was honest.	○	○	○	○	○
was trustworthy.	○	○	○	○	○
was honorable.	○	○	○	○	○
was moral.	○	○	○	○	○

Please indicate your level of agreement on each of the following statements.

	Strongly agree	Somewhat agree	Neither agree nor disagree	Somewhat disagree	Strongly disagree
The service representative chatbot did a good job.	○	○	○	○	○
The service representative chatbot did what I expected.	○	○	○	○	○
I am happy with the service representative chatbot.	○	○	○	○	○
I am satisfied with the service representative chatbot.	○	○	○	○	○
I am content with the service representative chatbot.	○	○	○	○	○
I am satisfied with the experience of talking with the service representative chatbot.	○	○	○	○	○

Results

- The ANOVA between the four groups showed that the simulated chatbots were perceived as overall reliable (M=4.77, SD=0.56), responsive (M=4.76, SD=0.59) and satisfying (M=4.72, SD=0.58).
- Trustworthiness registered slightly lower but varied less than a point (M=4.17, SD=0.84).
- Based on the results, the hypotheses were not supported indicating that gender and product type did not effect the chatbot-consumer relationship.



Estimated Marginal Means of Trustworthiness

Discussion

- Some possible explanations for the hypotheses not being supported…
  - Since chatbots do not pass judgment (Hill, Ford & Farreras, 2015), consumers are not inhibited when talking about a gender-sensitive issue with a different gender chatbot.
  - Humans are more open, agreeable, extroverted and self-disclosing when interacting with chatbots (Mou & Xu, 2017). Therefore, the awkwardness that was assumed might be caused by the gender-specific product type was not present and therefore did not influence the chatbot-consumer relationship.
  - The gender-sensitive product type was possibly construed as Unisex.



Limitations & Future Research



Limitations

- Due to research time constraints, a simulated, text-only chatbot was developed.
- Chatbot questions and responses were predetermined and provided to the participants.
- Amazon's MTurk Workforce Community has inherent flaws, such as workers who rush through surveys simply to be paid, which can skew the data.

Future Research

- The experiment could be conducted using an actual AI-programmed chatbot.
- The experiment could use product types that are overtly gender-specific.



Significance

Significance

- The healthcare chatbot market is projected to grow at 20.8% from 2018 to 2023 (Healthcare Chatbots Market, 2018). Therefore, gender in chatbot design could be critical if the product type is something as personal as a healthcare issue.
- Using a personalized chatbot for customer service could increase the perception that a company cares about its customers.

Thank you