12-1-2022

Interacting with a Chatbot-Based Advising System: Understanding the Effect of Chatbot Personality and User Gender on Behavior

Mohammad Amin Kuhail
Zayed University

Justin Thomas
Liverpool John Moores University

Salwa Alramlawi
Princess Nourah bint Abdulrahman University

Syed Jawad Hussain Shah
University of Missouri-Kansas City

Erik Thornquist
Zayed University

Article

Interacting with a Chatbot-Based Advising System: Understanding the Effect of Chatbot Personality and User Gender on Behavior

Mohammad Amin Kuhail [1,*] **, Justin Thomas** [2] **, Salwa Alramlawi** [3] **, Syed Jawad Hussain Shah** [4] **and Erik Thornquist** [5]

1. College of Interdisciplinary Studies, Zayed University, Abu Dhabi P.O. Box 144534, United Arab Emirates
2. School of Psychology, Liverpool John Moores University, Liverpool L3 5UX, UK
3. College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
4. School of Science and Engineering, University of Missouri-Kansas City, Kansas City, MO 64108, USA
5. College of Technological Innovation, Zayed University, Abu Dhabi P.O. Box 144534, United Arab Emirates
* Correspondence: mohammad.kuhail@zu.ac.ae

Abstract: Chatbots with personality have been shown to affect engagement and user subjective satisfaction. Yet, the design of most chatbots focuses on functionality and accuracy rather than an interpersonal communication style. Existing studies on personality-imbued chatbots have mostly assessed the effect of chatbot personality on user preference and satisfaction. However, the influence of chatbot personality on behavioral qualities, such as users' trust, engagement, and perceived authenticity of the chatbots, is largely unexplored. To bridge this gap, this study contributes: (1) A detailed design of a personality-imbued chatbot used in academic advising. (2) Empirical findings of an experiment with students who interacted with three different versions of the chatbot. Each version, vetted by psychology experts, represents one of the three dominant traits, agreeableness, conscientiousness, and extraversion. The experiment focused on the effect of chatbot personality on trust, authenticity, engagement, and intention to use the chatbot. Furthermore, we assessed whether gender plays a role in students' perception of the personality-imbued chatbots. Our findings show a positive impact of chatbot personality on perceived chatbot authenticity and intended engagement, while student gender does not play a significant role in the students' perception of chatbots.

Keywords: human–computer interaction; chatbot behavior; chatbot personality; authenticity; trust; engagement; usage intention

1. Introduction

Chatbots, also called conversational agents, have grown tremendously and become a part of several industries, including healthcare [1], consumer services [2], and education [3], as they can automate services by conversing with users. A financial reflection of this growth is that the size of the chatbot market is projected to reach 1.23 billion US dollars globally by 2025 [4]. As the number of active chatbots rapidly increases, their interactions with humans are growing equally [5]. According to some estimates, up to a third of online interactions involve a form of a chatbot [6]. We believe the growing number of human–chatbot interactions necessitates a deeper understanding of the key variables of these interactions, as many services and decisions depend on the effectiveness of the interactions.

Recent research shows that humans perceive chatbots as social actors [7–9] and subconsciously assign them a personality [10]. Chatbot personality has been shown to affect trust [11,12], engagement [6], subjective satisfaction [13,14], and consumer behavior [6].

Despite this, within the context of academic advising, most chatbot-based academic advising systems, e.g., refs. [15,16] are designed with an emphasis on functionality and accuracy rather than an interpersonal communication style, which is essential in building trust and relationships. Little is known about the effect of chatbot personality in an academic

setting. Related studies have investigated chatbot personality in other domains, including driving assistance [11], commerce [17,18], and healthcare [10,19].

Furthermore, imbuing text-based chatbots with personalities has received little attention. Much of the relevant research has focused on imbuing voice-based, embodied, and robot-based chatbots with personalities using nonverbal cues, such as gestures [20,21] gaze [22], voice tone and speed [23], and proximity [24].

Among the few attempts that attempted to design text-based chatbots with a personality are the works of Li et al. [25], Smestad and Volden [14], Völkel et al. [10], and Völkel and Kaya [18]. These studies have mainly utilized the Big-Five model [26] to test the effect of agreeableness [18], extraversion [10], or assertiveness [25] on users' preferences and intentions to use chatbots. However, little research has attempted to investigate the effect of text-based chatbot personalities on other users' behavioral qualities, such as trust, perceived authenticity, and engagement. Such qualities are crucial to ensure the success and acceptance of chatbots. For instance, trust is considered a crucial element of a successful relationship with users [27] and should be factored into the design of chatbots [28] as an essential aspect of trust is related to anthropomorphizing [27]. The authenticity of chatbot conversations positively influences users' intention to use the chatbot [29]. Engagement affects the length of the user responses and the time they spend with the chatbot.

To bridge gaps in the literature, this study contributes: (1) A detailed design of a personality-imbued text-based chatbot used in academic advising (MyAdvisor). The design is based on the Big Five Factor model [26] and is validated by professional psychologists. (2) The findings of an experiment with 43 students recruited from 3 different campuses to test the effect of agreeable, conscientious, and extroverted chatbot personality on trust, perceived authenticity, and engagement. Furthermore, we also report the findings as to whether gender plays a role in students' trust, perceived chatbot authenticity, usage intention, and intended engagement with chatbots with various personalities.

The remainder of this article is structured as follows. Section 2 reviews the related work, while Section 3 presents the research goal and hypotheses. Section 4 discusses the design of the personality-imbued chatbots and explains the experiment setup, and Section 5 presents the findings. Section 6 discusses the findings and implications for future research. Finally, Section 7 concludes the article.

2. Related Work

2.1. Human Personality

Human personality is the combination of characteristics, behavior, and emotions that form a distinctive character [30]. The Five-Factor Model (FFM), also known as the Big Five Factors, is a well-grounded taxonomy for studying personality as it covers the crucial aspects of personality [31]. FFM has been utilized in Human–Computer Interaction (HCI) to explain how different chatbots demonstrate behavior [8,32]. FFM consists of five global factors or personality traits discovered due to lexical analysis rather than neuropsychological experimentation [33]. The five factors are referred to as follows [6,34,35]: (1) Openness: this dimension measures people's inventiveness and curiosity. (2) Conscientiousness: this personality attribute measures individuals' degree of efficiency and organization. (3) Extraversion: this dimension assesses the level of outgoingness and bubbliness individuals demonstrate. (4) Agreeableness: this dimension assesses the level of friendliness and compassion. (5) Neuroticism: this personality attribute evaluates the degree of nervousness and moodiness.

The topic of personality has long been studied in research and has proved influential in various social and economic situations and across various cultures [36,37]. For instance, companies report an increase in sales if salespersons demonstrate high extraversion and openness [38]. A salesperson's personality has also positively affected customers' trust [39]. Moreover, the extant literature has found a direct link between individuals' satisfaction with a product and their personalities and the emotions they experience when interacting with a product [40]. Consequently, as advocated by user experience (UX) research, it is crucial to

design positive experiences that cater for personalities and individuals' personalities and emotions [41]. Indeed, UX has become a crucial part in designing products and services [42]. However, in an educational setting, the literature has primarily focused on a student's personality and how it affects academic motivation [43] and academic achievement [44].

2.2. Chatbots in Education

Due to their capacity to engage students and personalize education, chatbots are becoming prevalent in education [45]. In the last decade, chatbots have filled a range of educational tasks, including tutors, coaches, and learning companions [46]. In addition, chatbots have been applied to address a wide range of educational needs, including question-answering [47], tutoring [48,49], and language learning [48,49]. Moreover, chatbots have been shown to be effective in various roles when engaging with students, including teaching agents, peer agents, teachable agents, and motivational agents [3].

Although chatbots are regarded as social actors [8] and have grown rapidly in education, most studies used chatbots only to enhance the learning process rather than social engagement with the students. Nonetheless, a few studies [50–52] incorporated social dialog into their design to engage students. Still, most studies applying chatbots in academic settings have not considered assigning a personality to social dialogue, despite the growing evidence for chatbot personality affecting trust [11,12], engagement [6], and subjective satisfaction [13,14]. A recent literature review calls for investigating the impact of chatbot personality in an academic setting on students' satisfaction [3]. This call is supported by a recent study [53] that stresses the importance of assigning a personality to a chatbot that advises students in an academic setting. To fill the gap, this work builds on a previously developed text-based chatbot system used for academic advising, MyAdvisor [54], by integrating three personalities into the chatbot (agreeableness, extraversion, and conscientiousness).

2.3. Personality-Imbued Chatbots

Chatbots emulate conversation, a complex activity that is uniquely human. Conversations allow humans to show their personalities and build relationships [55]. When human users interact with chatbots, they may implicitly form an idea about the chatbot's personality and communication style [56]. As such, it is crucial to design chatbots with human users in mind. Indeed, human-centered design (HCD) advocates for invoking the human perspective into the design of products and services [57].

There is a growing body of literature suggesting that it is possible to convey personality in embodied, robot-based, and voice-based chatbots using body language and nonverbal cues, such as voice pitch and speed [23,58], gaze [22], proximity [24], and gestures [20,59]. However, only a limited number of research has targeted purely text-based personality-imbued chatbots. Table 1 shows an overview of these chatbots.

Völkel and Kaya [18] designed three chatbots with three levels of agreeableness: agreeable, neutral, and disagreeable. The chatbots were helping users find a suitable movie. The authors briefly explained the design by manipulating the chatbots based on the defined characteristics of each personality. For instance, the agreeable chatbot uses positive emotions and expresses concern for the user, whereas the disagreeable is critical and uncooperative. Predictably, the users preferred the agreeable chatbot to other types.

Völkel et al. [10] designed three versions of a healthcare chatbot with three levels of extraversion (extroverted, average, introverted). The design was based on a detailed description of each personality with some references to the literature. The extroverted chatbot uses emojis, while others do not. Furthermore, the extroverted chatbot is enthusiastic, referring to users by their names frequently assertive, and commanding. On the other hand, the introverted chatbot is reserved and shares limited information. Finally, the average chatbot demonstrates a medium level of extraversion and avoids traits associated with highly extroverted or introverted personalities. The users interacted with the chatbots

repetitively over four days. Regarding the rankings, users ranked the extroverted chatbot first, followed by the introverted chatbot and the average chatbot.

Ruane et al. [56] presented two types of text-based chatbots: Chatbot A with high extraversion and agreeableness, and Chatbot B with low extraversion and agreeableness. The design was based on the FFM taxonomy, and the researchers provided language queues for the personalities. However, the design was not validated by domain experts. The researchers concluded that a personality could be reliably represented with text and found that users were engaged more with Chatbot A.

Mehra [34] designed three different versions of chatbots helping users to make an order: (1) Chatbot A, called WordHelper, with conscientiousness as a dominant trait, is transactional and focuses on efficiency, accuracy, and speed. (2) Chatbot B, called WordAid, with agreeableness as a dominant trait, and thus, has a prosocial personality and thus uses a lot of friendly, helpful, and polite phrases. (3) Chatbot C, called Word!Baby, on the other hand, demonstrated bubbliness and friend-like qualities. It used emojis and was rather informal. The authors used the FFM taxonomy and did not verify their design with domain experts, but verified the personality references using IBM's Emotional Analyzer tool, which is deprecated. The students interacted with all chatbots, and the results show that users preferred chatbot C, followed by A, then B.

These works undoubtedly contributed to the literature. However, while the personality design of these works was grounded in the FFM taxonomy and the literature, they have not been assessed by domain experts to establish validity. Furthermore, the work primarily focused on evaluating users' general preference for a specific personality and whether users are drawn to chatbots with the same personality as theirs.

To fill the gap in the literature, we contribute a text-based chatbot system used for academic advising, MyAdvisor [54]. We designed three different versions of the chatbot representing three personalities: agreeableness, extraversion, and conscientiousness. Our design systematically used references from the Big-Five Inventory (BFI) [60] and was verified by professional psychologists. Furthermore, our work assessed behavioral attributes such as trust, engagement, and perceived authenticity of the chatbot.

Table 1. An overview of text-based personality-imbued chatbots in the literature.

Reference	Research Goal	Application Area	Findings	Design
Völkel & Kaya [18]	Testing users' perception of chatbot agreeableness	Movie recommendation	Users prefer the highly agreeable chatbot	• Based on FFM and literature. • Not validated by domain experts
Völkel et al. [10]	Testing users' perception of chatbot extraversion	Healthcare	Users prefer the highly extroverted chatbot	• Based on FFM and literature. • Emojis used. • Not validated by domain experts.
Ruane et al. [56]	Testing users' perception of chatbot agreeableness and extraversion	Course recommendation	Users prefer highly agreeable and extroverted chatbots	• Based on FFM • Not validated by domain experts.
Mehra [34]	Testing users' preference for chatbots of three personalities	Making a food order	Users preferred the extroverted chatbot, followed by the conscientious chatbot	• Based on FFM and validated by IBM Emotional Analyzer • The extroverted chatbot uses emojis • Not validated by domain experts

3. Research Goals and Hypotheses

The main objective of this research is to assess the effect of chatbot personality and user gender on user behavior, such as trust, perceived authenticity, usage intention, and intended engaged.

Despite the widespread use of chatbots, users still lack trust in voice-based chatbots [61] due to privacy concerns and security vulnerabilities [62]. Trust is an essential prerequisite for adopting systems [63,64], and it influences users' intention to use chatbots [65]. Trust is also essential for a successful relationship with users [27] and should be considered in the design of chatbots [28]. Yet, little attention has been given to the effect of chatbot personality on trust. Notably, Reinkemeier and Gnewuch [58] investigated how

personality and gender congruence between voice-based chatbots and users can affect a person's trust. The researchers found that the impact of being the same personality on users' trust is significant. However, the effect of a gender match on users' trust is nonsignificant.

Various researchers identified three major components essential to measuring trust [58,66,67]: integrity, competence, and benevolence. Nine factors represent these major components: ability, effectiveness, being knowledgeable, providing suitable advice, acting in the user's best interest, doing its best, caring about users' answers, honesty, and sincerity.

Since trust has been associated with chatbot personality [68], and based on our discussion, we hypothesize:

H1. *Chatbot personality affects students' trust.*

For in-depth insights, Hypothesis 1 (H1) can be broken into three sub-hypotheses to compare the effect of each pair of chatbot personalities (agreeable, conscientious, and extroverted) on students' trust. Thus, we hypothesize:

- *H1A. There is a difference between students' trust in the conscientious and extroverted chatbots.*
- *H1B. There is a difference between students' trust in the conscientious and agreeable chatbots.*
- *H1C. There is a difference between students' trust in the agreeable and extroverted chatbots.*

Another essential quality of chatbots is perceived authenticity. An authentic chatbot is characterized by the ability to have a human-like conversation and display a clear purpose [69]. Recent research has shown the effect of chatbot authenticity on usage intention [29], engagement, and loyalty [70]. Yet, little attention has been given to whether chatbot personality affects the users' perceived authenticity of the chatbot. Given the relationship between authenticity and personality found in the psychology literature [71,72], this research aims to shed light on this largely unexplored area in the context of chatbots. Thus, we hypothesize:

H2. *Chatbot personality affects students' perceived authenticity of the chatbot.*

We divide Hypothesis 2 (H2) into three sub-hypotheses to compare the effect of each pair of chatbot personalities (agreeable, conscientious, and extroverted) on students' perceived authenticity of the chatbot. Consequently, we hypothesize:

- *H2A. There is a difference between students' perceived authenticity of the conscientious and extroverted chatbots.*
- *H2B. There is a difference between students' perceived authenticity of the conscientious and agreeable chatbots.*
- *H2C. There is a difference between students' perceived authenticity of the agreeable and extroverted chatbots.*

It has been established in fields other than academic advising that chatbot personality affects the users' intention to use the chatbot. For instance, users prefer to use extroverted chatbots more than those that are neutral or introverted [10]. Moreover, users prefer to use agreeable chatbots compared to those that are less agreeable [18]. In addition, compared to a conscientious chatbot, users prefer to use an agreeable chatbot [14]. Therefore, we hypothesize:

H3. *Chatbot personality affects students' intention to use the chatbot.*

To compare the effect of each pair of chatbot personalities (agreeable, conscientious, and extroverted) on students' intention to use the chatbot, Hypothesis 3 (H3) is split into the following sub-hypotheses:

- *H3A. There is a difference between students' intention to use the conscientious and extroverted chatbots.*
- *H3B. There is a difference between students' intention to use the conscientious and agreeable chatbots.*
- *H3C. There is a difference between students' intention to use the agreeable and extroverted chatbots.*

Several studies have shown that chatbot personality affects users' engagement with chatbots. Engagement can be measured by the users' willingness to spend time with the chatbot and their involvement with it [73]. A notable example is a study [6] concluding that matching users and chatbot personalities results in increased engagement. Furthermore, a study [12] revealed that users are engaged with a chatbot-based interviewer with a friendly and warm personality. Moreover, a study [10] found that users have different levels of engagement depending on the chatbot's level of extraversion. Consequently, we hypothesize:

H4. *Chatbot personality affects students' intended engagement with the chatbot.*

To compare the effect of each pair of chatbot personalities (agreeable, conscientious, and extroverted) on students' intended engagement with the chatbot, Hypothesis 4 (H4) is split into the following sub-hypotheses:

- *H4A. There is a difference between students' intended engagement with the conscientious and extroverted chatbots.*
- *H4B. There is a difference between students' intended engagement with the conscientious and agreeable chatbots.*
- *H4C. There is a difference between students' intended engagement with the agreeable and extroverted chatbots.*

The effect of gender on user behavior when interacting with a chatbot has received little attention in the literature. A notable study [74] assessed the effect of gender on user behavior but found no significant influence. However, since men and women may have different personalities [75], we think it is interesting to examine the effect of gender on user behavior when interacting with a chatbot. To this end, we hypothesize:

H5. *Students' gender affects students' trust in the chatbot.*

Hypothesis 5 (H5) is split into three sub-hypotheses to obtain detailed comparisons between students' gender effect on trust in the three different chatbots.

- *H5A. There is a difference between male and female students' trust in the conscientious chatbot.*
- *H5B. There is a difference between male and female students' trust in the extroverted chatbot.*
- *H5C. There is a difference between male and female students' trust in the agreeable chatbot.*

H6. *Students' gender affects students' perception of chatbot authenticity.*

We divide Hypothesis 6 (H6) into three sub-hypotheses for detailed comparisons between students' gender effect on perceived authenticity of the three different chatbots.

- *H6A. There is a difference between male and female students' perceived authenticity of the conscientious chatbot.*
- *H6B. There is a difference between male and female students' perceived authenticity of the extroverted chatbot.*
- *H6C. There is a difference between male and female students' perceived authenticity of the agreeable chatbot.*

H7. *Students' gender affects students' intention to use the chatbot.*

Hypothesis 7 (H7) is further broken into three sub-hypotheses to compare the effect of students' gender on usage intention of the three different chatbots.

- *H7A. There is a difference between male and female students' intention to use the conscientious chatbot.*
- *H7B. There is a difference between male and female students' intention to use the extroverted chatbot.*
- *H7C. There is a difference between male and female students' intention to use the agreeable chatbot.*

H8. *Students' gender affects students' engagement with the chatbot.*

We divide Hypothesis 8 (H8) into three sub-hypotheses to compare the effect of students' gender on engagement with the chatbot between the three different chatbots.

- *H8A. There is a difference between male and female intended engagement with the conscientious chatbot.*
- *H8B. There is a difference between male and female intended engagement with the extroverted chatbot.*
- *H8C. There is a difference between male and female intended engagement with the agreeable chatbot.*

4. Methodology

This section presents the design of the chatbots used in this study and explains the study design.

4.1. Chatbot Design

We designed three chatbots representing the dominant personality traits of conscientiousness, extraversion, and agreeableness by carefully manipulating the chatbots' expressions. The chatbots are designed to help students with academic advising and are based on a chatbot-based advising system, MyAdvisor, presented in [54].

To ensure that the three chatbots preserved their assigned personalities, we utilized a guided conversational style with scripted answers to specific questions. The conversation content (i.e., questions and answers) are the same for all three chatbots. What distinguishes the chatbots is the communication style representing the three personalities. All chatbots are named "MyAdvisor," a role that is purposely gender-neutral.

Initially, every chatbot greets the student, introduces itself, and guides the student to ask the first question. For instance, Figure 1 shows an example of an agreeable chatbot greeting the students, introducing itself by presenting how it can help, then as a suggestion, inviting the student to ask the question corresponding to Task 1. Students can use their own words to ask questions. There are five tasks that students interact with. The tasks inquire about various aspects of academic advising, such as helping with poor performance, helping with understanding course materials, enrolling in a senior project, and career opportunities after graduation. The full details of the conversation script can be found in [76]. Once the student is done typing a question corresponding to a certain task, the chatbot provides an answer, then suggests to the student to move on to the next task serially (e.g., Tasks 1, 2, 3, etc.) The chatbots are not designed to handle inquiries outside the scope of the five tasks, but they can handle small talk. Consequently, the chatbots ask students to follow the script and the task if they deviate from the flow. We implemented the chatbots using Google Dialogflow [77]; thus, the chatbots were accessible on web and mobile platforms.

4.2. Chatbot Personality Design and Validation

In designing the personalities of the chatbots, we first designed a neutral script for the conversation, including the tasks associated with academic advising. Thereafter, we systematically modified the language cues to indicate the three personalities of conscientiousness, extraversion, and agreeableness. In doing so, we followed the Big-Five Inventory (BFI) model [60]. As such, the conscientious chatbot is characterized by being thorough, exceedingly careful, reliable, organized, industrious, efficient, focused, and a plan follower. The extroverted chatbot is talkative, not reserved, full of energy, enthusiastic, neither shy nor inhibited, outgoing, sociable, and assertive. The agreeable chatbot neither finds fault in others nor does it start quarrels with others. Instead, it is warm, helpful, forgiving, trusting, considerate, and cooperative.
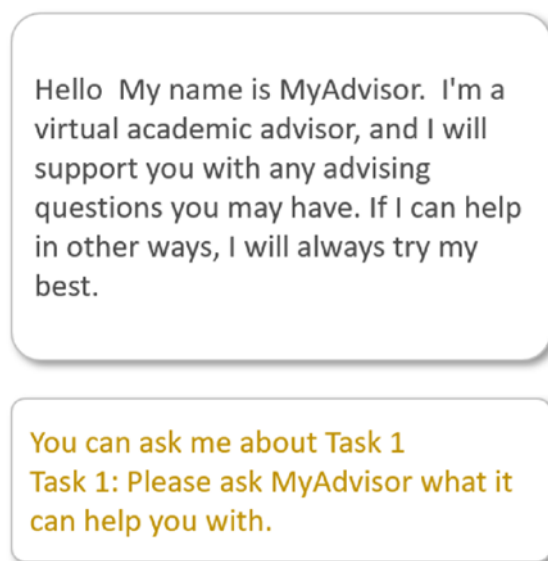
Figure 1. The chatbot greets the students and invites them to ask a question.

Figure 2 shows an example showing three greeting messages and another example showing three responses representing the three personalities. We added language cues corresponding to characteristics associated with the personalities as per the BFI model. For instance, the greeting message of the chatbot with conscientiousness is mapped to four characteristics associated with conscientiousness. We ensured that each personality characteristic was mapped at least once. The process of embedding personality traits into the chatbot messages was iterative and went through several revisions. The chatbots are not designed to be purely conscientious, extroverted, or agreeable, but rather dominated by one of these personalities. The references in the text bubbles in Figure 2 are to the personality characteristics per the BFI model, but the references are not shown to students in real interactions. The full details of mapping personality characteristics can be found in [78].

To make certain that the chatbots are only differentiated by their personalities, we ensured that the chatbot message sizes were comparable. Thus, we ran a T-Test to compare the word counts of the messages of each pair of chatbots and did not find a significant difference. Furthermore, all chatbots used purely text despite relevant works, e.g., ref. [10] using emojis for extroverted chatbots. We conducted five pilot sessions with students and identified vague or hard-to-understand phrases for some non-native English speakers in the study. Consequently, we rewrote these phrases to address the identified shortcomings. Furthermore, we also validated our chatbot personality design with four professional psychologists by asking them to state the dominant personality of each of the three chatbots. All psychologists confirmed our intended chatbot personality design.

4.3. Study Design

In the context of a chatbot-based advising system, we conducted this study to assess the effect of chatbot personality (conscientiousness, extraversion, and agreeableness) on user behavior (trust, perceived authenticity, intended engagement, and intention to use) concerning the chatbot. As such, this research aims to test the hypotheses presented in Section 3 and to identify the users' preference for chatbot personality.
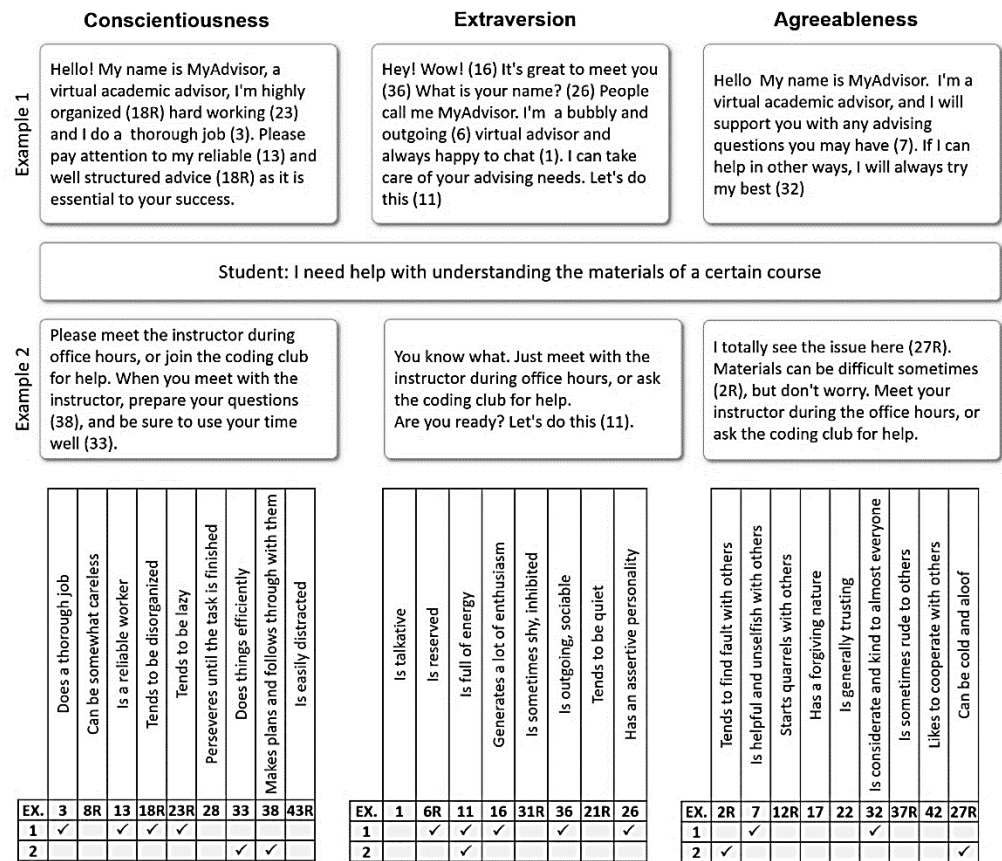
Figure 2. Manipulating the personalities based on the Big-Five Inventory (BFI) model.

Figure 3 shows the steps of the experiment. Prior to conducting the experiment, we obtained ethical clearance from the Research Ethics Committee at Zayed University. We recruited 48 students to participate in the experiment. However, 5 responses were removed from the dataset as we processed the data due to several missing or corrupt data points. The experiment consisted of three steps. First, the students watched a five-minute video explaining the experiment's purpose and steps. Second, informed consent was obtained from all students involved in the study, and they provided demographical information. Subsequently, the students interacted with the three chatbots with personalities dominated by conscientiousness, extraversion, and agreeableness. To minimize the bias of the order of chatbot interaction, we split the students into three groups: 1, 2, and 3. In Group 1, students first interacted with the conscientious chatbot, followed by the extroverted chatbot and then the agreeable one. In Group 2, students first interact with the extroverted chatbot, the agreeable chatbot and then the conscientious one. At last, in Group 3, the students interacted with the agreeable chatbot, followed by the conscientious chatbot and then the extroverted. The students were unaware of the chatbot order and were told that they interacted with chatbots A, B, and C regardless of their assigned group. Initially, the group sizes were equal. However, at later stages, we excluded the data of a few participants making the group sizes slightly different (but they are still similar in size). The third and last step of the experiment was that the students filled out a quantitative and qualitative questionnaire to express their opinions on trust, perceived authenticity, engagement, intention to use, and overall preference of the three different chatbots. On average, the experiment lasted 50 min.
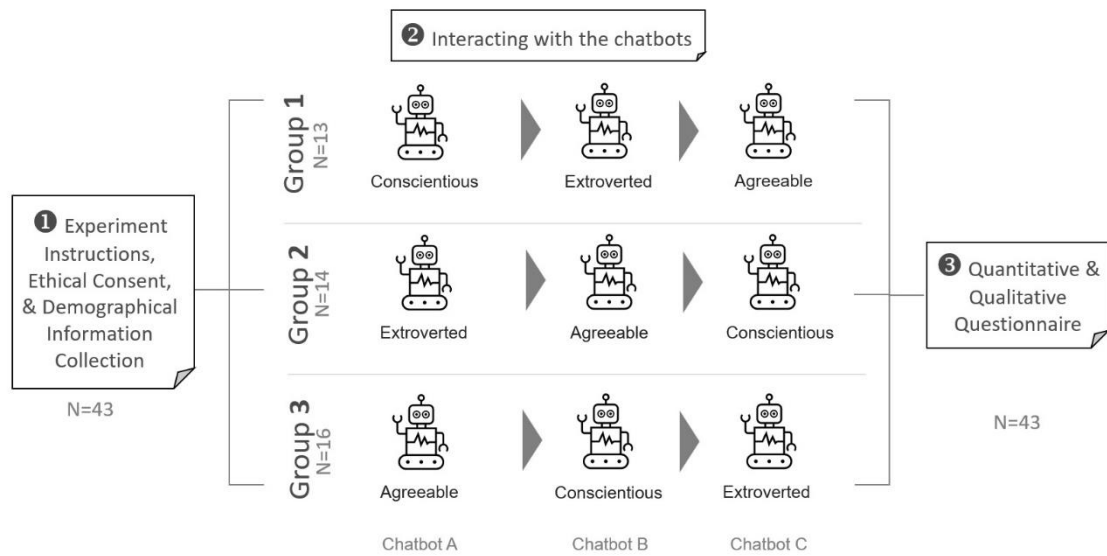
Figure 3. The steps of the study.

4.4. Participants

To recruit participants, the authors announced the experiment in their respective institutions and on their professional networks. Since the authors are based in three different universities, the recruited students were also located at the same universities as the authors. Table 2 depicts the demographical information of the study participants. All participants reported a good command of English. Most participants (86%) were 18–25 years old, while only 13.95% were 26–35. Approximately two-thirds (67.4%) of the participants were female, and approximately a third (32.6%) were male. Approximately half of the participants (51.2%) were recruited from Princess Noura Bint Abdulrahman University in Saudi Arabia, while approximately a third (32.5%) were recruited from Zayed University in the United Arab Emirates. Finally, the remaining students (16.3%) were recruited from the University of Missouri-Kansas City in the United States. Most (86%) participants were undergraduate students, with the remaining (14%) being postgraduate students. At last, all students were familiar with chatbots and had at least basic IT skills and were thus familiar with web browsing and document editing tools.

Table 2. Participants' demographical information.

Characteristic	Description
Age	18–25 (N = 37), 26–35 (N = 6)
Gender	Female (N = 29), Male (N = 14)
Campus Location	Saudi Arabia (N = 22), United Arab Emirates (N = 14), United States (N = 7)
Education Level	Undergraduate student (N = 37), Postgraduate student (N = 6)
Familiarity with chatbots	All are familiar with chatbots.
IT Skills	All at least have basic IT skills (familiar with web surfacing and document editing tools)

4.5. Post-Interaction Questionnaire

The participants were asked to complete a questionnaire after interacting with the three chatbots. The questionnaire included quantitative and qualitative questions. Figure 4 shows the questionnaire's quantitative questions, which utilized a five-point Likert scale ranging from 1 (Strongly disagree) to 5 (Strongly agree) to pose questions to participants about trust, authenticity, usage intention, and engagement to help us assess the stated hypotheses in Section 3. The participants were asked the same questions about each chatbot (e.g., their perception of trust in the conscientious, extroverted, and agreeable chatbots).

The participants were asked to keep the chatbot interaction tabs open while filling out the questionnaire to refresh their memories of the interactions.

Trust was assessed by ten components (e.g., trustworthiness, competence, and effectiveness) as discussed in Section 3. Perceived authenticity was assessed by perceived authenticity, human likeness, and clear purpose, while usage intention was assessed by the participant's desire to use the chatbot in the future. At last, intended engagement was assessed by the user's desire to engage, spend time, and frequently use the chatbot. For attributes measured by multiple components, such as trust, we computed the components' mean to calculate the attribute value.

The questionnaire also included qualitative questions. The participants were asked to elaborate on their selections as to why they found certain chatbots trustworthy, authentic, and engaging, and why they intended to use certain chatbots. Finally, the participants were asked to rate their overall preference among the three chatbots and explain the reasons behind their preferences.
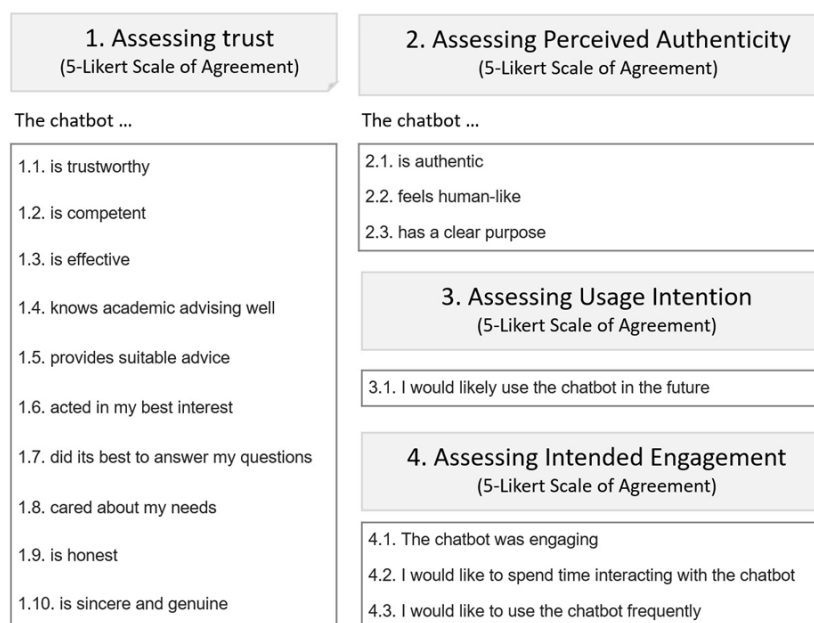

Figure 4. Quantitative assessment of trust, authenticity, usage intention, and engagement.

5. Results

This section presents the findings of our study. To test the hypotheses presented in Section 3, we conducted Kruskal–Wallis H tests [79] to compare students' trust, perceived authenticity, usage intention, and intended engagement with the three chatbots (conscientious, extroverted, and agreeable). We also performed the Mann–Whitney U test [80] to compare differences in students' ratings (trust, perceived authenticity, usage intention, and engagement) between pairs of chatbots (conscientious x extroverted, conscientious x agreeable, and extroverted x agreeable). Kruskal–Wallis H and Mann–Whitney U tests are suited for our data, which are ordinal and not normally distributed. Furthermore, we examined the internal reliability of our data using Cronbach's alpha (α). The result was above 0.7, indicating acceptable reliability [81]. We also performed a thematic analysis to analyze the qualitative data collected in this study.

5.1. Effect of Chatbot Personality on Trust, Authenticity, Usage Intention, and Engagement

Figure 5 shows a box plot of students' trust, perceived authenticity, usage intention, and intended engagement with the chatbots. Concerning trust, on average, the students trusted the agreeable chatbot the most (Mdn = 4.1, M = 4.033, SD = 0.7733), followed by

the extroverted chatbot (Mdn = 3.8, M = 3.788, SD = 0.893) and the conscientious chatbot (Mdn = 3.6, M = 3.595, SD = 1.017).

On average, the students perceived the extroverted and agreeable chatbot to be almost equally authentic, with a slight preference for the extroverted chatbot (Mdn = 4, M = 3.953, SD = 0.113) over the agreeable chatbot (Mdn = 4, M = 3.922, SD = 0.126). In contrast, the conscientious chatbot was perceived as less authentic (Mdn = 3.33, M = 3.233, SD = 0.151) than the other two.
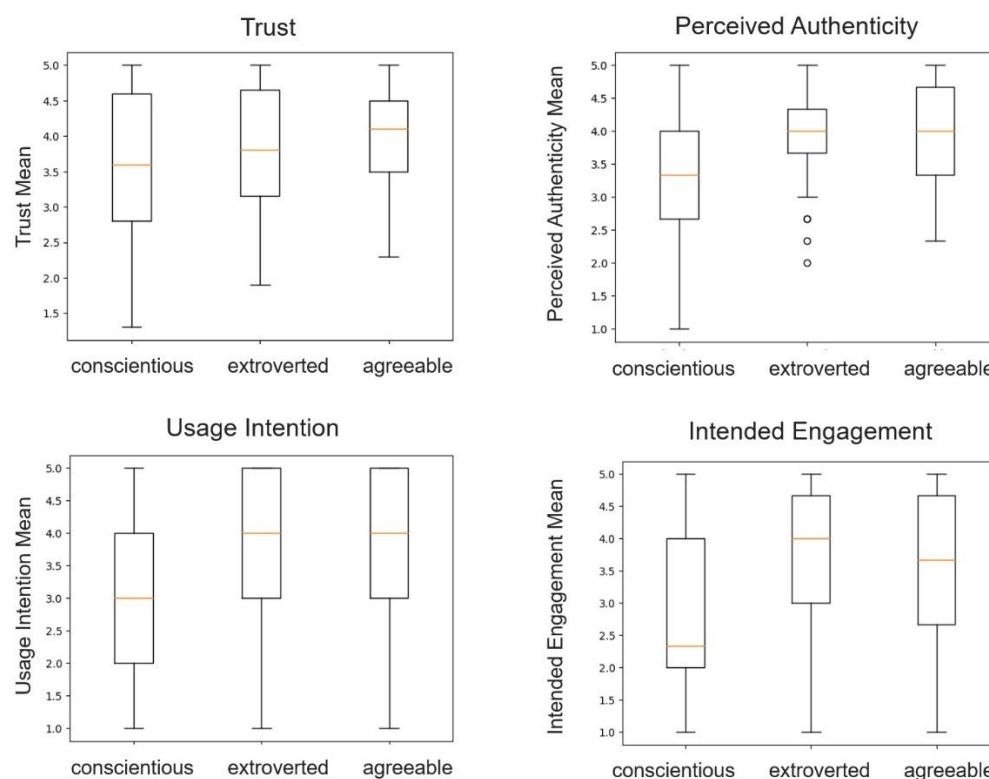

Figure 5. A box plot for the students' ratings of the chatbots.

Concerning usage intention, on average, the students intend to use the agreeable chatbot the most (Mdn = 4, M = 3.721, SD = 0.209), followed by the extroverted chatbot (Mdn = 4, M = 3.628, SD = 1.254) and the conscientious chatbot (Mdn = 3, M = 3.14, SD = 1.441).

For intended engagement, on average, the students intend to engage with the extroverted chatbot the most (Mdn = 4, M = 3.744, SD = 1.093), followed by the agreeable chatbot (Mdn = 3.667, M = 3.605, SD = 0.174), and the conscientious chatbot (Mdn = 2.33, M = 2.853, SD = 0.195).

The following subsections present the results of testing the first four hypotheses defined in Section 3.

5.1.1. Testing Hypothesis 1 (Effect of Chatbot Personality on Trust)

Table 3 shows the results of testing Hypothesis 1. According to the Kruskal–Wallis H test, the effect of chatbot personality on students' trust is nonsignificant ($p = 0.114$). To obtain more details, we tested the three sub-hypotheses: H1A, H1B, and H1C to compare students' trust in each pair of chatbots (Table 3). The results show no significant difference between students' trust in the conscientious and extroverted chatbots ($p = 0.375$). Likewise, the difference between students' trust in the agreeable and extroverted chatbots is nonsignificant ($p = 0.178$). However, there is a significant difference between the students' trust in the conscientious and agreeable chatbots ($p = 0.049$).

Table 3. Hypothesis 1 Testing Results.

Hypothesis Testing	H-Value	*p*-Value	Result
H1. Chatbot personality affects students' trust.	4.339	0.114	n. s.
Sub-Hypothesis	**U-Value**	**_p_-Value**	**Result**
H1A. There is a difference between students' trust in the conscientious and extroverted chatbots.	821.5	0.375	n.s
H1B. There is a difference between students' trust in the conscientious and agreeable chatbots.	696.5	0.049	supported
H1C. There is a difference between students' trust in the agreeable and extroverted chatbots.	768.5	0.178	n.s.

Note: n.s. = not supported.

5.1.2. Testing Hypothesis 2 (Effect of Chatbot Personality on Perceived Authenticity)

Table 4 depicts Hypothesis 2 test results. The results indicate a significant difference ($p = 0.001$) between students' perceived authenticity of the different chatbots. Further, we tested the three related sub-hypotheses: H2A, H2B, and H2C to compare students' perceived authenticity of each pair of chatbots. The results show a significant difference between the students' perceived authenticity of the conscientious and extroverted chatbots ($p = 0.001$) and between the students' perceived authenticity of the conscientious and agreeable chatbots ($p = 0.001$). However, there is no significant difference between the students' perceived authenticity of the agreeable and extroverted chatbots ($p = 0.838$).

Table 4. Hypothesis 2 Testing Results.

Hypothesis	H-Value	*p*-Value	Result
H2. Chatbot personality affects students' perceived authenticity of the chatbot.	15.059	0.001	supported
Sub-Hypothesis	**U-Value**	**_p_-Value**	**Result**
H2A. There is a difference between students' perceived authenticity of the conscientious and extroverted chatbots.	525.5	0.001	supported
H2B. There is a difference between students' perceived authenticity of the conscientious and agreeable chatbots.	553	0.001	supported
H2C. There is a difference between students' perceived authenticity of the agreeable and extroverted chatbots.	948.5	0.838	n.s.

Note: n.s. = not supported.

5.1.3. Testing Hypothesis 3 (Effect of Chatbot Personality on Usage Intention)

Table 5 depicts the results of testing Hypothesis 3. The effect of chatbot personality on students' usage intention is nonsignificant ($p = 0.114$). Further, we tested the three sub-hypotheses: H3A, H3B, and H3C to compare students' perceived intention to use each pair of chatbots (Table 5). The results show that there is no significant difference between the students' intention to use the conscientious and extroverted chatbots ($p = 0.118$), between the students' intention to use the conscientious and agreeable chatbots ($p = 0.054$), or between the students' intention to use the agreeable and extroverted chatbots ($p = 0.0591$).

Table 5. Hypothesis 3 Testing Results.

Hypothesis	H-Value	*p*-Value	Result
H3. Chatbot personality affects students' intention to use the chatbot.	4.338	0.114	n. s.

Sub-Hypothesis	U-Value	*p*-Value	Result
H3A. There is a difference between students' intention to use the conscientious and extroverted chatbots.	747.5	0.118	n.s.
H3B. There is a difference between students' intention to use the conscientious and agreeable chatbots.	707.5	0.054	n.s.
H3C. There is a difference between students' intention to use the agreeable and extroverted chatbots.	864	0.591	n.s.

Note: n.s. = not supported.

5.1.4. Testing Hypothesis 4 (Effect of Chabot Personality on Intended Engagement)

The results show that there is a significant difference ($p = 0.002$) between students' intended engagement with the different chatbots (Table 6). We also tested the three related sub-hypotheses: H4A, H4B, and H4C to compare students' intended engagement with each pair of chatbots (Table 6). The results show that there is a significant difference between the students' engagement with the conscientious and extroverted chatbots ($p = 0.001$). However, there is no significant difference between the students' intended engagement with the conscientious and agreeable chatbots ($p = 0.006$) or between the students' intended engagement with the agreeable and extroverted chatbots ($p = 0.557$).

Table 6. Hypothesis 4 Testing Results.

Hypothesis	H-Value	*p*-Value	Result
H4. Chatbot personality affects students' intended engagement with the chatbot.	12.413	0.002	supported

Sub-Hypothesis	U-Value	*p*-Value	Result
H4A. There is a difference between students' intended engagement with the conscientious and extroverted chatbots.	549	0.001	supported
H4B. There is a difference between students' intended engagement with the conscientious and agreeable chatbots.	607	0.006	n.s.
H4C. There is a difference between students' intended engagement with the agreeable and extroverted chatbots.	992.5	0.557	n.s.

Note: n.s. = not supported.

5.2. Effect of User Gender on Behavior

Figure 6 shows the box plots of the students' trust, perceived authenticity, usage intention, and intended engagement ratings for the three different chatbots categorized by gender. On average, male students trusted the conscientious (Mdn= 3.779, M = 3.65, SD = 1.102) and extroverted (Mdn = 4.036, M = 4.1, SD = 0.908) chatbots more than their female counterparts (conscientious: Mdn = 3.507, M = 3.5, SD = 1.102; extroverted: Mdn = 3.669, M = 3.8), but the female students (Mdn = 3.857, M = 4.1, SD = 0.729) trusted the agreeable chatbot more than their male (Mdn = 3.507, M = 4.15, SD = 1.102) counterparts. In comparison, on average, male students perceived the conscientious chatbot as slightly more authentic (Mdn = 3.381, M = 3.33, SD = 0.794) than their female counterparts (Mdn = 3.161, M = 3, SD = 1.079), while male and female students perceived the authenticity of extroverted and agreeable chatbots very similarly, with slightly higher medians and means for male students. Concerning usage intention, male students intend to use the conscientious (Mdn = 3.571, M = 3.5, SD = 1.016) and agreeable (Mdn = 3.857, M = 4.5, SD = 1.406) chatbots more than their female counterparts (conscientious: Mdn = 2.931, M = 3, SD = 1.58; agreeable: Mdn = 3.655, M = 4), while the male and female students intend to use the extroverted chatbot very similarly with slightly higher median and mean values for male students. Regarding engagement, male students expressed

more engagement with all chatbots than their female counterparts. Table 7 shows the results of testing Hypotheses 5–8. The results indicate that the students' gender has no significant influence on their behavior in relation to the chatbots.
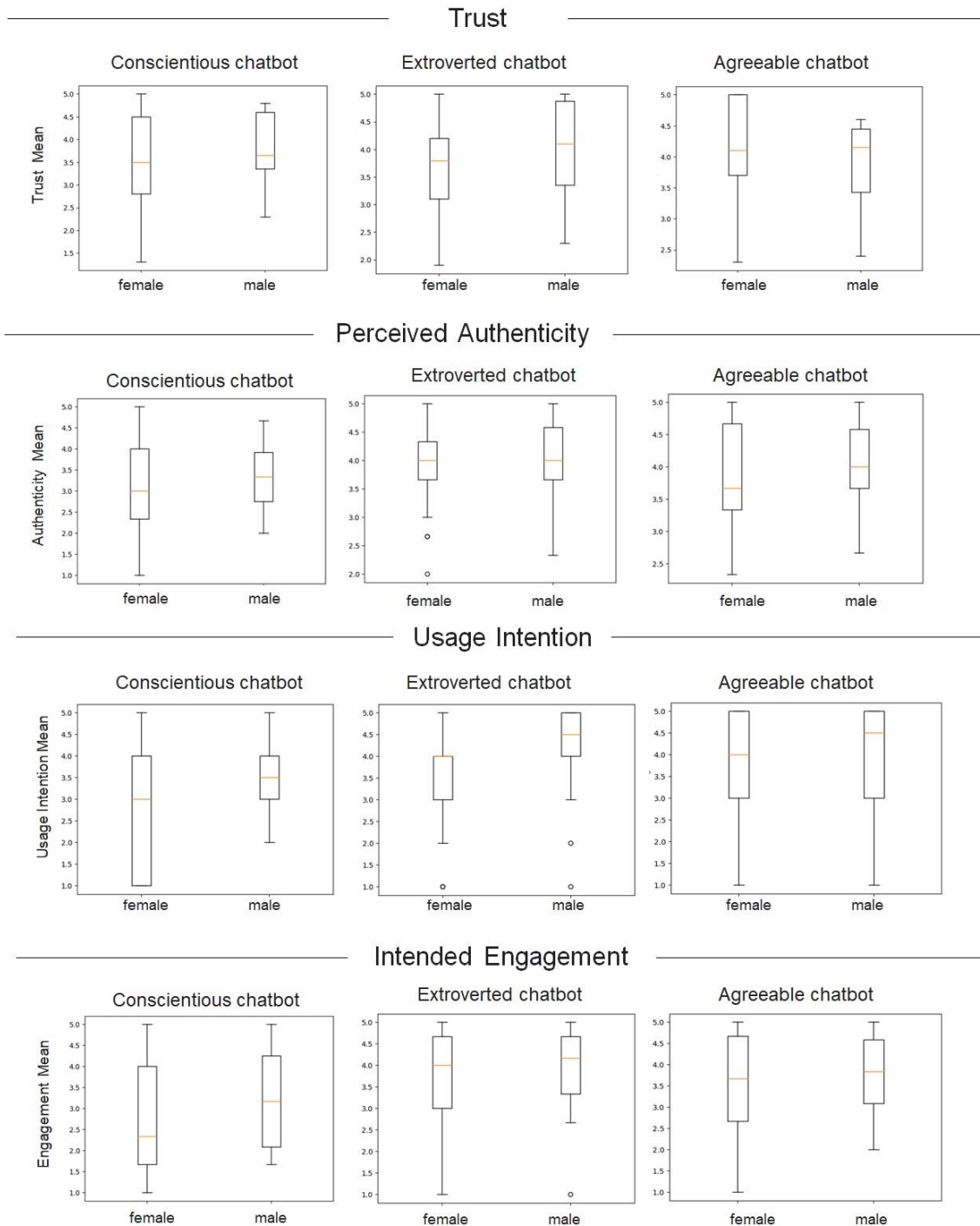

Figure 6. A box plot of trust, perceived authenticity, usage intention, and intended engagement ratings for chatbots categorized by gender.

Table 7. Testing of Hypotheses 5–8.

Hypothesis	U-Value	*p*-Value	Result
H5. Students' gender affects students' trust of the chatbot			
H5A. There is a difference between male and female students' trust in the conscientious chatbot.	176	0.491	n.s.
H5B. There is a difference between male and female students' trust in the extroverted chatbot.	150	0.173	n.s.
H5C. There is a difference between male and female students' trust in the agreeable chatbot.	236	0.398	n.s.
H6. Students' gender affects students' perception of chatbot authenticity.			
H6A. There is a difference between male and female students' perceived authenticity of the conscientious chatbot.	173.5	0.449	n.s.
H6B. There is a difference between male and female students' perceived authenticity of the extroverted chatbot.	198	0.906	n.s.
H6C. There is a difference between male and female students' perceived authenticity of the agreeable chatbot.	204.5	0.979	n.s.
H7. Students' gender affects students' intention to use the chatbot.			
H7A. There is a difference between male and female students' intention to use the conscientious chatbot.	155.5	0.213	n.s.
H7B. There is a difference between male and female students' intention to use the extroverted chatbot.	134	0.066	n.s.
H7C. There is a difference between male and female students' intention to use the agreeable chatbot.	185	0.633	n.s.
H8. Students' gender affects students' intended engagement with the chatbot			
H8A. There is a difference between male and female intended engagement with the conscientious chatbot.	152	0.188	n. s.
H8B. There is a difference between male and female intended engagement with the extroverted chatbot.	179.5	0.548	n. s.
H8C. There is a difference between male and female intended engagement with the agreeable chatbot.	196.5	0.876	n. s.

Note: n.s. = not supported.

5.3. Thematic Analysis of Qualitative Data

To analyze the qualitative data provided by the students, we conducted a thematic analysis following the method of Forbes [82], which included these steps: (1) becoming familiar with the data, (2) generating initial codes, (3) searching for themes, (4) defining themes, (5) iteratively reviewing themes, and (6) writing up the results.

5.3.1. Reasons for Trust, Perceived Authenticity, Usage Intention, and Intended Engagement with the Chatbots

Figure A1 shows a visualization of the themes found when analyzing the provided reasons for user behavior. Concerning trust, some students trust the conscientious chatbot mainly for its clarity and competence and the extroverted chatbot for its human likeness, competence, and stimulating nature, while trust in the agreeable chatbot is driven by its empathy, human likeness, competence, and helpfulness.

With respect to perceived authenticity, some students perceived the conscientious chatbot to be authentic due to its human likeness, attention to detail, flexibility in understanding the answers, and being true to its nature (not pretending to be human). In comparison, the perceived authenticity of the extroverted chatbot is largely driven by its human likeness and a little by its perceived honesty and friendliness. Likewise, the agreeable chatbot is considered authentic due to its empathy, human likeness, honesty, and professionalism.

Some students intend to use the conscientious chatbot due to its competence, clarity, and efficiency, while the usage intention of the extroverted chatbot is driven by its human likeness, competence, and friendliness. Likewise, the agreeable chatbot is intended to be used due to its human likeness, competence, and empathy.

In terms of intended engagement, some students intend to engage with the conscientious chatbot due to its competence, honesty, and helpfulness, and the extroverted chatbot due to its friendliness, fun nature, and helpfulness, while the agreeable chatbot's intended engagement is fueled by its competence, human likeness, and empathy.

5.3.2. Chatbot Preference

Figure A2 depicts the students' preferences for the chatbots, while Figure A3 shows the reasons the students cited for their preferred chatbots. The majority of the students preferred the agreeable chatbots (40%), followed by the extroverted chatbot (32.5%), and the conscientious chatbot (27.5%). The students preferring the agreeable chatbot cited its empathy, human likeness, honesty, and competence, while the ones preferring the extroverted chatbot highlighted its human likeness, empathy, honesty, competence, fun nature, and directness.

6. Discussion, Implications for Future Research, and Study Limitations

This section summarizes the findings, compares them with related research where possible, discusses implications for future research, and presents the study limitations.

6.1. Effect of Chatbot Personality on Behavior

In general, our findings show that students indicate noticeably more trust, perceived authenticity, usage intention, and intended engagement with agreeable and extroverted chatbots than the conscientious one. This perception leads some students to cite competence, honesty, and helpfulness as reasons for their preferred chatbots. In contrast, students attribute incompetence and unhelpfulness to chatbots they do not prefer. This is surprising as the three chatbots are designed to be equally competent, honest, and helpful.

Due to its human likeness, empathy, perceived competence, and helpfulness, students trust the agreeable chatbot the most, followed by the extroverted chatbot, then the conscientious one. However, the results are statistically inconclusive, except that the difference between the students' trust in the conscientious and agreeable chatbots is significant. The inconclusive results could stem from the students interacting with the chatbots briefly. Thus, it is unrealistic for them to trust the chatbots regardless of their personalities. Related works show that users trust chatbot personalities similar to theirs [11,58] in the context of voice-based e-commerce and driving assistant chatbot-based applications. Our results show that trust can be achieved without a personality congruence, providing fresh insight into a largely unexplored area.

In terms of authenticity, due to their human likeness, empathy, friendliness, and honesty, the students perceived the agreeable and extroverted to be almost equally more authentic than the conscientious one. Generally, the difference between students' perceived chatbot authenticity of different chatbots is statistically significant. However, there is no significant difference between the students' perceived authenticity of the agreeable and extroverted chatbots. To our knowledge, our findings are unique as the relationship between chatbot personality and authenticity has not been previously studied. We believe this contribution is crucial to practitioners and researchers striving to develop authentic chatbots, which have increased loyalty and satisfaction [70].

Driven by their human likeness, competence, empathy, and friendliness, students intend to use the agreeable and extroverted chatbots nearly equally more than the conscientious one. Our findings are akin to those reported in [18], where the authors found that users prefer to use agreeable chatbots. However, the effect of chatbot personality on usage intention is not statistically significant. An explanation of this could be due to the students interacting with the chatbots briefly; thus, it is hard for them to decide on the usage intention.

Because of their friendliness, fun nature, and helpfulness, students found the extroverted chatbot to be the most engaging, followed by the agreeable, then the conscientious chatbot. Statistically, there is a significant difference between students' intended engagement with the three different chatbots. However, there is no significant difference among individual differences between the engagement with the conscientious and agreeable chatbots as well as the agreeable and extroverted chatbots. Our findings are similar to those of [6], citing students who increased engagement with extroverted chatbots. In comparison, the study by [10] did not find a difference between users' engagement with chatbots of different levels of extroversion. However, the two studies measured engagement by counting the number of words instead of engagement intention, which we measured in our study.

An interesting finding of this study is that some students expressed that the communication style of the extroverted chatbot was informal and not suited for academic advising, while other students pointed out that both extroverted and agreeable chatbots seemed unreal as they were excessively and unbelievably positive. Similar insights were reported in [10].

6.2. Effect of User Gender on Behavior

There are small differences between how male and female students rated different chatbots on trust, perceived authenticity, usage intention, and intended engagement. However, the results are inconclusive statistically. Our findings are similar to those reported in [74], which found no significance for gender on user behavior when interacting with a virtual agent. Similarly, another related work [58] found no significant effect on matching the gender of users and chatbots. However, it is crucial to point out that the population sample in this study is dominated by females (N = 29), and male students only constitute 32.5% (N = 14) of the population. Consequently, researchers are encouraged to replicate our results with a more gender-balanced population.

6.3. Chatbot Personality Preference

Due to its empathy, human likeness, honesty, and competence, the results show that students preferred the agreeable chatbot the most, followed by the extroverted and conscientious chatbots. Students preferring extroverted chatbots mainly cite their human likeness, while those preferring the conscientious chatbot highlight its clarity and efficiency.

The results are consistent with the literature, as users prefer agreeable [18] and extroverted chatbots [10]. However, our findings shed light on the reasons for students' preference for personality-imbued chatbots. Furthermore, our results also show that more than a quarter of the students preferred the conscientious chatbot, calling for further investigation in the future as to how the qualities of a conscientious chatbot could be used in service chatbots.

6.4. Future Research Considerations

This research highlighted a new territory in the literature and contributed fresh insights. Furthermore, our findings open new avenues for future researchers to investigate. First, this research has identified why certain students do not provide positive ratings of certain chatbots. For example, some students cited that the extroverted chatbot was too informal for a serious task, such as academic advising. In contrast, other students highlighted that the extroverted and agreeable chatbot seemed unreal. As such, it would be interesting to assess further students' negative views of the chatbots and investigate how to address them.

Second, this research has identified that users have varied preferences and reasons for their ratings of different chatbots, calling for future investigation to accommodate users' perspectives. A possible future work could be to investigate combining the traits of different personalities. For instance, the effect of combining the empathy of agreeable chatbots and outgoing nature could be studied further. Another possible future work is to study the effect of user-chatbot personality congruence in the context of academic advising.

Third, it would be interesting to elicit more insights from students by conducting lengthy interviews on their perception of the chatbots. Finally, we could conduct a longitudinal study to explore the effect of chatbot personality over a longer period. Longer individual chats may also yield more definitive results.

Fourth, this research reported a within-subject study to allow the participants to compare their preference of the three different chatbots. However, future researchers could potentially conduct this research by performing a between-subjects study and investigate the relation between user personality and chatbot personality.

6.5. Study Limitations

Several limitations may have affected the results of our study. First, since the participants in the study interacted with three different versions of a chatbot as opposed to one, this may have influenced users' perception. If users interacted with one chatbot with one personality, the results may have been different [83]. Second, the participants were recruited only from three different geographical locations, which is not representative of users worldwide. Future researchers are encouraged to replicate our results in various

locations and cultures. Third, most of the participants are non-native speakers of English, yet they interacted with English-speaking chatbots. While the participants reported a good command of English, we still think this may have affected the results. As such, future researchers could replicate our results, while ensuring the lack of a language barrier.

7. Conclusions

This study shows, in the context of academic advising, evidence of chatbot personality influencing aspects of user behavior such as users' intended engagement and authenticity of the chatbots. However, the results were inconclusive about the impact of chatbot personality on trust and usage intention, possibly due to the limited interaction with the chatbots during the study. In general, students express higher trust in the agreeable chatbot and intend to engage with the extroverted chatbot more than other chatbots. In contrast, students perceive the agreeable and extroverted chatbots to be equally more authentic and intend to use them more than the conscientious chatbot. Furthermore, in general, the findings also show that students prefer the agreeable chatbot the most, while approximately a third of them prefer the extroverted chatbot, and approximately a quarter prefer the conscientious chatbot.

Students justify their perception of the chatbots due to the expected traits of the chatbots, for instance, the empathy of the agreeable chatbot, the human likeness of the extroverted chatbot, and the clarity of the conscientious chatbot. However, to our surprise, some students attribute competence and honesty only to their preferred chatbots, although all chatbots were designed to be equally competent and honest. This leads us to speculate that chatbot personality is perhaps a more important way for chatbots to connect to a range of users than previously thought.

Furthermore, the findings also show that students generally prefer the agreeable chatbot the most due to its empathetic and human-like nature. In contrast, approximately a third of the students prefer the extroverted chatbot due to its simulation of a human character and its fun nature.

Future researchers should investigate the effect of chatbot personality over a longer time, the effect of combined personality traits on user behavior, and the effect of the chatbot personality in different contexts.

Appendix A

Figure A1. A visualization of the themes of students' explanation of their ratings of certain chatbot.
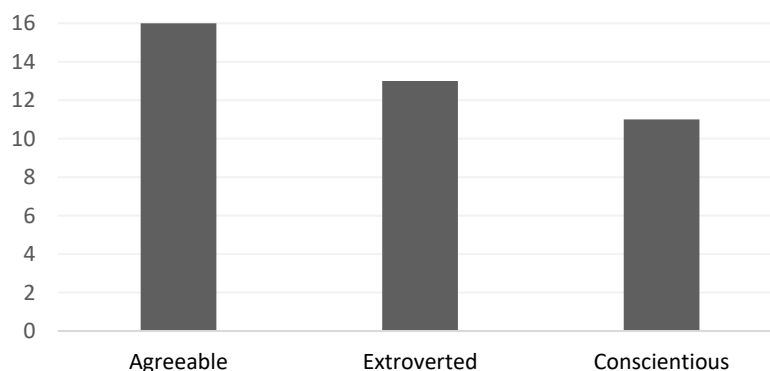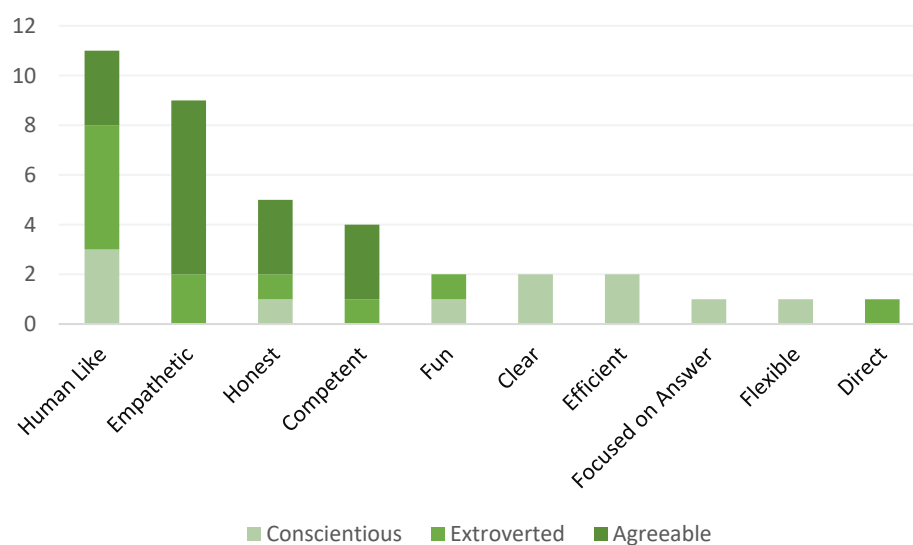
Figure A2. Students' preference for chatbots.

Figure A3. Students' reasons for preferring certain chatbots.

References

1. Oh, K.-J.; Lee, D.; Ko, B.; Choi, H.-J. A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. In Proceedings of the 18th IEEE International Conference on Mobile Data Management (MDM), Daejeon, Korea, 29 May–1 June 2017.
2. Xu, A.; Liu, Z.; Guo, Y.; Sinha, V.; Akkiraju, R. A new chatbot for customer service on social media. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017.
3. Kuhail, M.A.; Alturki, N.; Alramlawi, S.; Alhejori, K. Interacting with Educational Chatbots: A Systematic Review. *Educ. Inf. Technol.* **2022**, 1–46. Available online: https://link.springer.com/article/10.1007/s10639-022-11177-3#citeas (accessed on 14 September 2022). [CrossRef]
4. Statista. Size of the Chatbot Market Worldwide, in 2016 and 2025. Available online: https://www.statista.com/statistics/656596/worldwide-chatbot-market/ (accessed on 14 September 2022).
5. Tsvetkova, M.; García-Gavilanes, R.; Floridi, L.; Yasseri, T. Even good bots fight: The case of Wikipedia. *PLoS ONE* **2017**, *12*, e0171774. [CrossRef] [PubMed]
6. Shumanov, M.; Johnson, L. Making conversations with chatbots more personalized. *Comput. Hum. Behav.* **2021**, *117*, 106627. [CrossRef]
7. Nass, C.; Steuer, J.; Tauber, E.R. Computers are social actors. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 24–28 April 1994.
8. Nass, I.; Brave, S. *Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship*; MIT Press: Cambridge, MA, USA, 2005.
9. Reeves, B.; Nass, C.I. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People*; Cambridge University Press: Cambridge, UK, 1996.
10. Völkel, S.T.; Schoedel, R.; Kaya, L.; Mayer, S. User Perceptions of Extraversion in Chatbots after Repeated Use. In Proceedings of the CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA, 30 April–5 May 2022.

11. Braun, M.; Mainz, A.; Chadowitz, R.; Pfleging, B.; Alt, F. At your service: Designing voice assistant personalities to improve automotive user interfaces. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019.

12. Zhou, M.X.; Mark, G.; Li, J.; Yang, H. Trusting virtual agents: The effect of personality. *ACM Trans. Interact. Intell. Syst.* **2019**, *9*, 10. [CrossRef]

13. Bickmore, T.; Cassell, J. Social dialogue with embodied conversational agents. In *Advances in Natural Multimodal Dialogue System*; Springer: Dordrecht, The Netherlands, 2005; pp. 23–54.

14. Smestad, T.L.; Volden, F. Chatbot personalities matters. In Proceedings of the International Conference on Internet Science, Perpignan, France, 2–5 December 2019; Available online: https://research.com/conference/insci-2019-international-conference-on-internet-science (accessed on 14 September 2022).

15. Mekni, M.; Baani, Z.; Sulieman, D. A smart virtual assistant for students. In Proceedings of the 3rd International Conference on Applications of Intelligent Systems, Las Palmas, Spain, 7–9 January 2020.

16. Ranoliya, B.R.; Raghuwanshi, N.; Singh, S. Chatbot for university related FAQs. In Proceedings of the International Conference on Advances in Computing, Udupi, India, 13–16 September 2017; Available online: http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=53457 (accessed on 14 September 2022).

17. Jin, S.V.; Youn, S. Why do consumers with social phobia prefer anthropomorphic customer service chatbots? Evolutionary explanations of the moderating roles of social phobia. *Telemat. Inform.* **2021**, *62*, 101644. [CrossRef]

18. Völkel, S.T.; Kaya, L. Examining User Preference for Agreeableness in Chatbots. In Proceedings of the 3rd Conference on Conversational User Interfaces (CUI 2021), Bilbao, Spain, 27–29 July 2021.

19. Lee, M.; Ackermans, S.; van As, N.; Chang, H.; Lucas, E.; Jsselsteijn, W.I. Caring for Vincent: A chatbot for self-compassion. In Proceedings of the CHI Conference on Human Factors in Computing Systems, Glasgow, UK, 4–9 May 2019.

20. Bremner, P.; Celiktutan, O.; Gunes, H. Personality perception of robot avatar tele-operators. In Proceedings of the 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand, 7–10 March 2016.

21. Krenn, B.; Endrass, B.; Kistler, F.; André, E. Effects of language variety on personality perception in embodied conversational agents. In Proceedings of the International Conference on Human-Computer Interaction, Heraklion, Greece, 22–27 June 2014.

22. Andrist, S.; Mutlu, B.; Tapus, A. Look like me: Matching robot personality via gaze to increase motivation. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015.

23. Nass, C.; Lee, K.M. Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction. *J. Exp. Psychol. Appl.* **2001**, *7*, 171–181. [CrossRef]

24. Cafaro, H.H. Vilhjálmsson and T. Bickmore. First Impressions in Human-Agent Virtual Encounters. *ACM Trans. Comput.-Hum. Interact.* **2016**, *23*, 24. [CrossRef]

25. Li, J.; Zhou, M.X.; Yang, H.; Mark, G. Confiding in and listening to virtual agents: The effect of personality. In Proceedings of the 22nd International Conference on Intelligent User Interfaces, Limassol, Cyprus, 13–16 March 2017.

26. Rothmann, S.; Coetzer, E.P. The big five personality dimensions and job performance. *SA J. Ind. Psychol.* **2003**, *29*, 68–74. [CrossRef]

27. Garbarino, E.; Johnson, M.S. The different roles of satisfaction, trust, and commitment in customer relationships. *J. Mark.* **1999**, *63*, 70–87. [CrossRef]

28. Przegalinska, A.; Ciechanowski, L.; Stroz, A.; Gloor, P.; Mazurek, G. In bot we trust: A new methodology of chatbot performance measures. *Bus. Horiz.* **2019**, *62*, 785–797. [CrossRef]

29. Rese, A.; Ganster, L.; Baier, D. Chatbots in retailers' customer communication: How to measure their acceptance? *J. Retail. Consum. Serv.* **2020**, *56*, 102176. [CrossRef]

30. Allport, W. *Pattern and Growth in Personality*; Harcourt College Publishers: San Diego, CA, USA, 1961.

31. McCrae, R.R.; Costa, P.T., Jr. The five factor theory of personality. In *Handbook of Personality: Theory and Research*; The Guilford Press: New York, NY, USA, 2008; pp. 159–181.

32. Trouvain, J.; Schmidt, S.; Schröder, M.; Schmitz, M.; Barry, W.J. Modelling personality features by changing prosody in synthetic speech. In Proceedings of the 3rd International Conference on Speech Prosody, Dresden, Germany, 2–5 May C2006.

33. Goldberg, L.R. The structure of phenotypic personality traits. *Am. Psychol.* **1993**, *48*, 26–34. [CrossRef]

34. Mehra, B. Chatbot personality preferences in Global South urban English speakers. *Soc. Sci. Humanit. Open* **2021**, *3*, 100131. [CrossRef]

35. Norman, W.T. Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *J. Abnorm. Soc. Psychol.* **1963**, *66*, 574–583. [CrossRef] [PubMed]

36. Danner, D.; Rammstedt, B.; Bluemke, M.; Lechner, C.; Berres, S.; Knopf, T.; Soto, C.; John, O.P. *Die Deutsche Version des Big Five Inventory 2 (bfi-2)*; Leibniz Institute for the Social Sciences: Mannheim, Germany, 2016.

37. McCrae, R.R.; Costa, P.T. Validation of the five-factor model of personality across instruments and observers. *J. Personal. Soc. Psychol.* **1987**, *52*, 81. [CrossRef]

38. Matz, S.C.; Kosinski, M.; Nave, G.; Stillwell, D.J. Psychological targeting as an effective approach to digital mass persuasion. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 12714–12719. [CrossRef]

39. Rajaobelina, L.; Bergeron, J. Antecedents and consequences of buyer-seller relationship quality in the financial services industry. *Int. J. Bank Mark.* **2006**, *27*, 359–380. [CrossRef]

40. Desmet, P.; Fokkinga, S. Beyond Maslow's Pyramid: Introducing a Typology of Thirteen Fundamental Needs for Human-Centered Design. *Multimodal Technol. Interact.* **2020**, *4*, 38. [CrossRef]

41. Hassenzahl, M.; Diefenbach, S.; Göritz, A. Needs, affect, and interactive products: Facets of user experience. *Interact. Comput.* **2010**, *22*, 353–362. [CrossRef]

42. Liu, W.; Lee, K.-P.; Gray, C.; Toombs, A.; Chen, K.-H.; Leifer, L. Transdisciplinary Teaching and Learning in UX Design: A Program Review and AR Case Studies. *Appl. Sci.* **2021**, *11*, 10648. [CrossRef]

43. Komarraju, M.; Karau, S.J. The relationship between the big five personality traits and academic motivation. *Personal. Individ. Differ.* **2005**, *39*, 557–567. [CrossRef]

44. de Feyter, T.; Caers, R.; Vigna, C.; Berings, D. Unraveling the impact of the Big Five personality traits on academic performance: The moderating and mediating effects of self-efficacy and academic motivation. *Learn. Individ. Differ.* **2012**, *22*, 439–448. [CrossRef]

45. Benotti, L.; Martnez, M.C.; Schapachnik, F. A tool for introducing computer science with automatic formative assessment. *IEEE Trans. Learn. Technol.* **2017**, *11*, 179–192. [CrossRef]

46. Haake, M.; Gulz, A. A look at the roles of look & roles in embodied pedagogical agents—A user preference perspective. *Int. J. Artif. Intell. Educ.* **2009**, *19*, 39–71.

47. Feng, D.; Shaw, E.; Kim, J.; Hovy, E. An intelligent discussion-bot for answering student queries in threaded discussions. In Proceedings of the 11th International Conference on Intelligent User Interfaces, Syndney, Australia, 29 Juanuary–1 February 2006.

48. Heffernan, N.T.; Croteau, E.A. Web-based evaluations showing differential learning for tutorial strategies employed by the Ms. Lindquist tutor. In Proceedings of the International Conference on Intelligent Tutoring Systems, Maceió, Brazil, 30 August–3 September 2004.

49. VanLehn, K.; Graesser, A.; Jackson, G.T.; Jordan, P.; Olney, A.; Rosé, C.P. Natural Language Tutoring: A comparison of human tutors, computer tutors, and text. *Cogn. Sci.* **2007**, *31*, 3–52. [CrossRef]

50. Coronado, M.; Iglesias, C.A.; Carrera, Á.; Mardomingo, A. A cognitive assistant for learning java featuring social dialogue. *Int. J. Hum.-Comput. Stud.* **2018**, *117*, 55–67. [CrossRef]

51. el Janati, S.; Maach, A.; el Ghanami, D. Adaptive e-learning AI-powered chatbot based on multimedia indexing. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, 299–308. [CrossRef]

52. Qin, C.; Huang, W.; Hew, K.F. Using the Community of Inquiry framework to develop an educational chatbot: Lesson learned from a mobile instant messaging learning environment. In Proceedings of the 28th International Conference on Computers in Education, Online, 23–27 November 2020.

53. Dibitonto, M.; Leszczynska, K.; Tazzi, F.; Medaglia, C.M. Chatbot in a campus environment: Design of LiSA, a virtual assistant to help students in their university life. In Proceedings of the International Conference on Human-Computer Interaction, Las Vegas, NV, USA, 15–20 July 2018.

54. Kuhail, M.A.; al Katheeri, H.; Negreiros, J.; Seffah, A.; Alfandi, O. Engaging Students with a Chatbot-Based Academic Advising System. *Int. J. Hum.–Comput. Interact.* **2022**, 1–27. [CrossRef]

55. Mairesse, F.; Walker, M.; Mehl, M.; Moore, R. Using linguistic cues for the automatic recognition of personality in conversation and text. *J. Artif. Intell. Res.* **2007**, *30*, 457–500. [CrossRef]

56. Ruane, E.; Farrell, S.; Ventresque, A. User perception of text-based chatbot personality. In Proceedings of the International Workshop on Chatbot Research and Design, Online, 23–24 November 2020.

57. Calvo, R.; Vella Brodrick, A.; Desmet, P.M.; Ryan, R. Positive computing: A new partnership between psychology, social sciences and technologists. *Psychol. Well-Being Theory Res. Pract.* **2016**, *6*, 10. [CrossRef]

58. Reinkemeier, F.; Gnewuch, U. Match or mismatch? How matching personality and gender between voice assistants and users affects trust in voice commerce. In Proceedings of the 55th Hawaii International Conference on System Sciences, Lahaina, HI, USA, 4–7 January 2022; Available online: https://dblp.org/db/conf/hicss/index.html (accessed on 14 September 2022).

59. Chen, Z.L.Y.; Nieminen, M.P.; Lucero, A. Creating a chatbot for and with migrants: Chatbot personality drives co-design activities. In Proceedings of the ACM Designing Interactive Systems Conference, Eindhoven, The Netherlands, 6–10 July 2020.

60. John, O.P.; Srivastava, S. The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In *Handbook of Personality: Theory and Research*; Guilford Press: New York, NY, USA, 1999; pp. 102–138.

61. Mari, A.; Algesheimer, R. The role of trusting beliefs in voice assistants during voice shopping. In Proceedings of the Hawaii International Conference on System Sciences (HICSS), Maui, HI, USA, 5–8 January 2021; Available online: https://www.insna.org/events/54th-hawaii-international-conference-on-system-sciences-hicss (accessed on 14 September 2022).

62. Chung, H.; Iorga, M.; Voas, J.; Lee, S. Alexa, Can I Trust You? *Computer* **2017**, *50*, 100–104. [CrossRef] [PubMed]

63. Benbasat, I.; Wang, W. Trust in and adoption of online recommendation agents. *J. Assoc. Inf. Syst.* **2005**, *6*, 4. [CrossRef]

64. Gefen, D.; Straub, D. Managing user trust in B2C e-services. *E-Service* **2003**, *2*, 7–24. [CrossRef]

65. Kasilingam, D.L. Understanding the attitude and intention to use smartphone chatbots for shopping. *Technol. Soc.* **2020**, *62*, 101280. [CrossRef]

66. McKnight, D.H.; Choudhury, V.; Kacmar, C. Developing and validating trust measures for e-commerce: An integrative typology. *Inf. Syst. Res.* **2002**, *13*, 334–359. [CrossRef]

67. Qiu, L.; Benbasat, I. Evaluating anthropomorphic product recommendation agents: A social relationship perspective to designing information systems. *J. Manag. Inf. Syst.* **2009**, *25*, 145–182. [CrossRef]

68. Müller, L.; Mattke, J.; Maier, C.; Weitzel, T.; Graser, H. Chatbot acceptance: A latent profile analysis on individuals' trust in conversational agents. In Proceedings of the Computers and People Research Conference (SIGMIS-CPR '19), Nashville, TN, USA, 20–22 June 2019.

69. Neururer, M.; Schlögl, S.; Brinkschulte, L.; Groth, A. Perceptions on authenticity in chat bots. *Multimodal Technol. Interact.* **2018**, *2*, 60. [CrossRef]

70. Jones, C.L.E.; Hancock, T.; Kazandjian, B.; Voorhees, C.M. Engaging the Avatar: The effects of authenticity signals during chat-based service recoveries. *J. Bus. Res.* **2022**, *144*, 703–716. [CrossRef]

71. Seto, E.; Davis, W.E. Authenticity predicts positive interpersonal relationship quality at low, but not high, levels of psychopathy. *Personal. Individ. Differ.* **2021**, *182*, 111072. [CrossRef]

72. Sutton, A. Distinguishing between authenticity and personality consistency in predicting well-being. *Eur. Rev. Appl. Psychol.* **2018**, *68*, 117–130. [CrossRef]

73. Rodden, K.; Hutchinson, H.; Fu, X. Measuring the user experience on a large scale: User-centered metrics for web applications. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010.

74. Pütten, M.; Krämer, N.C.; Gratch, J. How our personality shapes our interactions with virtual characters-implications for research and development. In Proceedings of the International Conference on Intelligent Virtual Agents, Philadelphia, PA, USA, 20–22 September 2010.

75. Weisberg, Y.J.; DeYoung, C.G.; Hirsh, J.B. Gender differences in personality across the ten aspects of the Big Five. *Front. Psychol.* **2011**, *2*, 178. [CrossRef] [PubMed]

76. Chatbot Conversation Script. Available online: https://www.dropbox.com/s/mn4lcllt027ifhl/chatbot_conversation_script.docx?dl=0 (accessed on 14 September 2022).

77. Google. Dialogflow. Available online: https://cloud.google.com/dialogflow/docs (accessed on 14 September 2022).

78. Response Manipulation. Available online: https://www.dropbox.com/s/5lkwtc49dtug833/Responses_manipulation.xlsx?dl=0 (accessed on 14 September 2022).

79. Kruskal, W.H.; Wallis, W.A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **1952**, *47*, 583–621. [CrossRef]

80. Ruland, F. *The Wilcoxon-Mann-Whitney Test—An Introduction to Nonparametrics with Comments on the R Program wilcox.test*; Independently Published: Chicago, IL, USA, 2018.

81. Hair, F.; Hult, G.T.M.; Ringle, C.M.; Sarstedt, M. *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*; SAGE Publications: Thousand Oaks, CA, USA, 2016.

82. Braun, V.; Clarke, V. *Thematic Analysis: A Practical Guide*; Sage Publications: London, UK, 2022.

83. Han, J.; Ji, X.; Hu, X.; Guo, L.; Liu, T. Arousal recognition using audio-visual features and FMRI-based brain response. *IEEE Trans. Affect. Comput.* **2015**, *6*, 337–347. [CrossRef]