User acceptance of smart home voice assistant: a comparison among younger, middle-aged, and older adults

Runting Zhong1 · Mengyao Ma1,2 · Yutong Zhou1 · Qingxia Lin1 · Leiling Li1,2 · Nengjing Zhang1

Accepted: 18 October 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract
Voice assistants are widely used in smart home environments. This study aimed to investigate user acceptance of a smart home voice assistant. A questionnaire was designed, and 471 Chinese adults were recruited to complete the questionnaire. The data were analyzed using exploratory factor analysis and regression analysis. The results revealed that user requirements of adults were composed of six factors: hedonic motivation and trust ($\beta = .41, p < .001$), social influence ($\beta = .22, p < .001$), performance expectancy ($\beta = .15, p < .001$), effort expectancy ($\beta = .08, p = .018$), product features ($\beta = .15, p = .009$), and facilitating conditions ($\beta = .06, p = .049$). Among these six factors, hedonic motivation and trust are considered the most important. Younger, middle-aged, and older adults differed significantly in their requirements and acceptance of a smart home voice assistant. These findings have implications for the design of smart home voice assistants so that they are more acceptable to younger and older adults.

Keywords Internet of things (IoT) · Older adults · Smart home voice assistant · Technology acceptance · UTAUT

1 Introduction

Voice assistants are embodied in smart speakers such as Google Home, Amazon Echo, Tmall Genie, and Xiaomi Mi AI [1]. These are popular artificial intelligence products that provide new opportunities for people to acquire information in addition to conventional methods such as computers or smartphones [1]. Juniper Research predicted that usage of voice assistants in smart homes in China will reach 100 million by 2024, from just four million in 2019 [2]. Voice assistants have become a large part of users’ social lives, as they enable users to listen to music, control a smart home device, search for information, navigate websites, and so on [3].

Smart home voice assistants are especially important to older adults for at least three reasons. First, older adults experience a decline in vision and hearing, and voice assistants could provide a natural way of interacting that can reduce learning costs [4]. Second, many older adults live alone and far away from their children, and may thus feel lonely; in such cases, voice assistants provide spiritual comfort, which “pertains to the internal awareness of self, including self-esteem, meaning in one’s life, and one’s relationship to a higher order or being [5]”. Finally, the medical system is under great pressure, especially during the coronavirus pandemic (COVID-19) [6]. A voice assistant provides a new way for older adults to acquire health information besides traditional channels such as community promotions and health program videos.

This study aimed to identify the key factors that explain the requirements of younger, middle-aged, and older adults for smart voice assistants. Among the identified factors, those critical factors that influence user acceptance of voice assistants are further revealed. The results inform practitioners about the requirements of younger and older adults and the preferences of voice assistants in a smart home context. In addition, the results could provide insights into improving the user experience of a smart home voice assistant for various age groups.
1.1 Age-related differences in voice assistant requirements

Previous studies have investigated the diversity of user requirements for various age groups. Arfi et al. provided two categories of consumers: IoT natives (Gen Z, born in the new century and considered IoT natives) and IoT immigrants (the other generations). The perceived risk has no impact on IoT natives; however, IoT immigrants are more likely to consider the perceived risk [7]. A study on the user acceptance of smart bracelets by younger and older adults showed that the requirements of older adults were strongly linked to their health status. Older adults are less concerned about privacy issues and more likely to share with their family members than their younger counterparts [8]. Ghorayeb et al. investigated older adults’ perceptions of smart home technologies and found that among participants who had already tried smart home monitoring technologies, acceptance increases over time and use. These experienced participants expressed fewer concerns than the non-smart home participants regarding privacy, trust, and usability and had more concerns about utility [9]. Kim and Oumarou found that for older adults with visual impairments, the information structure of voice interaction interfaces should include no more than three layers of horizontal or vertical menu structures [10]. Xiao et al. conducted an in-depth interview with 8 young users (aged 20–30 years) who have experience in using top-selling Chinese smart speakers “Tmall Genie” and “Mi AI speaker”. The study found that all users have a strong feeling of distrust for the insecurity of the payment module and the inadequate content, suggesting complementing the functionality and usability of the smart home speakers for Chinese users [1]. Regarding the appearance of a smart home voice assistant, a study found that older adults like less human-looking robots, including pet-like robots such as Paro [11]. But Straßmann et al. found that older adults evaluated cartoon-stylized humanoid voice agents more positively than students [12]. Chiu et al. found that middle-aged and older women in Taiwan preferred robots with an animal-like appearance, while men preferred robots that resembled a human adult [13]. Biermann et al. found that older adults’ perception of robots depended on the context: robots with a technical appearance in production were trusted more than anthropomorphic robots or robots in the care context [14]. Sugano et al. found that the generation of “simple,” “regular,” and “smooth” motions can influence the perception of the Kawaii feeling of social robots as well as those movements perceived as biological or intentional [15]. Jafar et al. performed a Kansei survey to find suitable touch surfaces for the design of social robots. Their findings suggested that positive emotions can be achieved when designing robot surfaces with lower roughness. This type of surface can evoke calm, safe, and comfortable feelings [16]. Qie et al. found that older adults may expect social robots to be more functional, such as in helping to avoid falls or doing physical work [17].

1.2 Factors influencing technology acceptance of voice assistants

Although technology provides life conveniences, it also presents many obstacles for the elderly to use. Therefore, international studies have focused on technology acceptance by the elderly. Well-known technology-acceptance theories include the technology acceptance model (TAM) [18], the unified theory of acceptance and use of technology (UTAUT) [19], the theory of planned behavior [20], and diffusion of innovations [21]. According to the UTAUT model, performance expectations, achievement expectations, social influences, convenience conditions, and perceived risks influence usage intentions [19]. Previous studies have identified the factors that affect the acceptance of IoT technology services based on the UTAUT [7, 22] or TAM [23–26]. Liu et al.’s study investigated design factors critical to the acceptance of smart voice assistants among 220 Chinese subjects. This study found that sound quality, female gender, response accuracy, personification, and content diversity are critical factors in user acceptance [25]. Canziani et al. found that hedonic attitudes toward voice and device utility are primary drivers of smart home device ordering. Facilitating conditions supporting the use of smart home devices is more important for men than women. For Millennials, choice confusion is a significant influencer of smart home device utility [26]. Aldossari explored the acceptance of IoT smart homes by young customers in the USA and found that factors such as performance expectations, achievement expectations, social impact, convenience conditions, hedonic motivation, and cost affect the acceptance and use of smart homes [22]. Arfi et al. surveyed 268 French customers on the technological acceptance of IoT e-health and included financial costs in the UTAUT framework. The results showed that cost is the main barrier to the use of IoT e-health services, and age is an important mediator of the usage intention [7]. However, Burbach et al. found that privacy, not price, is the most important factor in accepting voice assistants among younger adults with not much experience in using a voice assistant [27]. The inconsistencies in the acceptance factors may be due to the user context. Liu et al.’s study focused on smart voice assistants embedded in various smart devices such as smartphones, smart speakers, smart wearable devices, and smart TVs [25]. Arfi et al.’s study [7] was concerned with health context, Burbach et al.’s study [27] focused on the
data recording by virtual voice assistants such as “Siri”, “Google”, and “Alexa”, while Aldossari’s [22] and Canziani et al.’s [26] study were focused on smart home context.

In addition, the elderly’s acceptance of home care technology is affected by the technology type [28], culture [29], anthropomorphism [30, 31], communication style [31, 32], family orientation [33], generation [33], interface design [34, 35], visual feedback except for voice [36], personality [3], gender [25, 37], cost [7], appearance [8, 28], fashion [28], trust [7], risk [7], and social support [8]. Therefore, we considered these factors in the questionnaire design.

2 Research questions

Thus far, studies have identified factors influencing the acceptance of smart home voice assistants. However, knowledge about the requirements, preferences, and differences among younger, middle-aged, and older adults is relatively rare.

Based on the literature review and the identified research gap, this study aimed to investigate factors influencing younger, middle-aged, and older adults’ acceptance of smart home voice assistants and to compare their attitudes and preferences toward smart home technology. Specifically, we seek to answer the following research questions:

Q1 Do younger (< 45 years), middle-aged, and older adults (≥ 45 years) differ in their attitudes toward voice assistants?
Q2 What are the key factors that explain younger and older adults’ requirements for smart voice assistants?
Q3 What are the critical factors that influence younger, middle-aged, and older adults’ intention to use voice assistants?

3 Material and methods

This section presents the study design, including an interview and an online survey.

3.1 Interview design

The objective of the interview was to understand the requirements of younger adults, middle-aged, and older adults for smart home voice assistants. Before the interview, the researchers introduced the purpose of the study and learned the personal information of the interviewees. The researchers first played a video showing how to use a voice assistant in a smart home environment. Then, the participants were asked the following questions: “What function do you like about a smart home voice assistant?” and “What is your preference about the appearance, gender, age, and size of a voice assistant?” The interviewees were also encouraged to share their opinions about smart home voice assistants. The interview results provide insights into the initial questionnaire design of factors influencing voice assistant acceptance and the reason for their choices in the questionnaire.

3.2 Questionnaire design

The questionnaire comprised of four parts. In the first part, a video was presented about the functions of a smart home voice assistant, including playing music, turning on smart home devices (e.g., floor cleaning robot), and chatting or searching for information, so that older adults could be familiar with the voice assistant. The second part included questions about factors influencing the acceptance of voice assistants (54 items). The items were designed based on UTAUT. The third part concerned the experience with and preferences for voice assistants, including appearance, age, gender, and size of the voice assistant (Table 1). The fourth part included the personal information of the participant (age, gender, education, marital status, location, city or countryside, home type, and living partner). The structure of the online survey is shown in Fig. 1. The items were adapted from previous studies or self-designed (see Table 7). A full version of the questionnaire is provided as supplementary material.

3.2.1 Independent variables

The independent variables included performance expectancy (“Using a voice assistant would improve my performance”), effort expectancy (“Learning to use a voice assistant is easy for me”), social influence (“People who are important to me feel that I should use a voice assistant”), facilitating conditions (“My living environment supports my use of a voice assistant”), cost (“The price of a voice assistant would determine whether I use this product”), perceived risk (“I can accept that a voice assistant transfers my data to the Internet”), trust (“I trust the information that the voice assistant tells me”), hedonic motivation (“I suppose using a voice assistant is interesting”), and product features (“A voice assistant with a high volume is better”) (Fig. 2).

3.2.2 Dependent variables

The dependent variable is use intention (“I intend to use a voice assistant in the future. A voice assistant with a high volume is better”). All items for the independent and dependent variables used a 5-point Likert scale, where 1 = strongly disagree and 5 = strongly agree.
3.3 Data collection

Regarding the interview, twelve middle-aged and older adults were recruited in a park in the authors’ city. Two older adults were interviewed through video calls. Five younger adults were recruited through word-of-mouth in the authors’ university.

Regarding the questionnaire, paper-based and online questionnaires were provided to participants. Both young adults and middle-aged and older adults were recruited for this study through word-of-mouth and social media, Wechat (Tencent, China). The online survey was conducted through https://www.wenjuan.com/. The participants were informed about the objective of the questionnaire and the anonymity of data collection at the beginning of the questionnaire. The authors’ university approved the ethics of this study. The survey began on March 18, 2021 and ended on March 23, 2021. The average response time of the questionnaire was 8 min and 16 s.

Data quality assessments were conducted to identify “speeders” (those who answered unreasonably fast) and “straightliners” (those who answered with identical values for each survey item in a block). We defined “speeders” as anyone who finished the survey at an average speed of less than two seconds per question, and we defined “straightliners” as anyone who responded using a single item in the influencing factor block [39, 40]. We deleted questionnaires with inconsistent responses; for example, the year of birth does not match the previous question, “Are you older than 45 years?” A total of 456 valid questionnaires were obtained. The effectiveness rate of the questionnaire was 96.8%.

### Table 1

| Characteristics | Item 1 | Item 2 | Item 3 | Item 4 | Item 5 |
|-----------------|-------|-------|-------|-------|-------|
| **Appearance**  | ![Animal](image1.png) | ![Plant](image2.png) | ![Geometric shape](image3.png) | ![Humanoid](image4.png) | ![Virtual character](image5.png) |
| **Age**         | ![Elderly](image6.png) | ![Middle-aged](image7.png) | ![Youth](image8.png) | ![Teenager](image9.png) | ![Children](image10.png) |
| **Gender**      | ![Male](image11.png) | ![Female](image12.png) | ![Neutral†](image13.png) |
| **Size**        | ![Finger](image14.png) | ![Cell phone](image15.png) | ![Hand](image16.png) | ![Table lamp](image17.png) | ![Child](image18.png) |

*According to Tolmeijer et al., “voice gender does not only come from pitch, but also from language usage and intonation” [38]. “Neutral voice” here represents a gender-neutral voice. For example, when we asked Siri on iPhone “Hi Siri, what is your gender?” The answer of Siri is “I have no gender.”*
of voice assistants. Depending on the factors identified, a multiple stepwise regression analysis was performed to examine the relationships between the demographic variables and the acceptance of voice assistants.

For the statistical analysis of the influence of participants’ age on the requirements of a smart home voice assistant, a one-way between-group analysis of variance (ANOVA) was applied to analyze differences among younger and middle-aged, and older adults. A partial eta squared \( \eta^2_p \) was calculated for the effect sizes. Data analysis was conducted using IBM SPSS v.26. The level of statistical significance \( (p) \) was set to 5%. The data to support the study findings are available from the corresponding author upon request.

### 4 Results

#### 4.1 Interview results

Nineteen interviewees were invited to participate in this study, including fourteen middle-aged and elderly people and five undergraduate students. The interview results showed that younger, middle-aged, and older adults have differences in functional requirements. Most of the elderly interviewees (12 participants) said that the more functions, the better, and the more advanced, the better. They preferred health monitoring, emergency alarms, etc. (“Of course, the more functions, the better, in case they can be used someday.” [P7]) “Now that we are old, health is the most important thing.” [P9]) Only two elderly interviewees expressed that they did not want too many functions, fearing that this would lead to product quality and maintenance problems (“Too many functions are not good, because they often break down and it is too troublesome to repair” [P5]). Middle-aged respondents attached more importance to functions related to children’s education (“We buy the voice assistant mainly for children. It can teach them to read texts, English and so on” [P14]). Younger adults’ requirements mostly focused on entertainment and companionship (“If it can chat with me when I am unhappy and enlighten me correctly, it will be quite good” [P1]). In addition, the preference for the appearance of voice assistants differed among the three age groups. Of the 19 respondents, 13 preferred a geometric look, saying it was simpler and more elegant. Three elderly interviewees chose the animal type, not because of their own preferences but of their grandchildren (“Animals are lovely, my grandchildren will like this. We are old, it does not matter whether it is beautiful or not, children like it” [P10]). In addition, three young respondents expressed their hope that more individual elements would be added, such as their favorite cartoon characters or reality idols (“How nice it is to see my idols every day” [P2]).

#### 4.2 Participants’ information

Demographic information of participants involved in the questionnaire is presented in Table 2. Among the 456 adults (mean age = 34.6 years; SD = 17.4 years; range = 18–84 years), 29.6% had experience using a voice assistant. The level of statistical significance \( (p) \) was set to 5%. The data to support the study findings are available from the corresponding author upon request.
assistant. The middle-aged and older participant group had a significantly lower proportion of females, lower education level, higher income, lower city residence proportion, less fall history, and less voice assistant experience than those of the younger participants \((p < 0.001)\).

### 4.3 Attitude toward voice assistant (RQ1)

#### 4.3.1 Preferences about appearance, gender, age, and size

There was a significant difference in the preference for the appearance of voice assistants among younger adults, middle-aged and older adults \((\chi^2(4456) = 17.049, p = 0.002)\). As shown in Fig. 2, younger adults preferred voice assistants with an animal look \((n = 84, 26.7\%)\) or geometric shape \((n = 83, 26.3\%)\), while middle-aged and older adults preferred humanoid voice assistants \((n = 48, 34.0\%)\).

There was a significant difference in the preference of the age of a voice assistant among younger, middle-aged adults, and older adults \((\chi^2(4456) = 83.346, p < 0.001)\). Figure 3 shows that younger adults prefer a voice assistant acceptable by youth \((n = 209, 66.3\%)\) or teenagers \((n = 50, 15.9\%)\), while middle-aged and older adults prefer one acceptable by youth \((n = 41, 29.1\%),\) middle-aged \((n = 38, 27.0\%)\), or elderly \((n = 35, 24.8\%)\).

There was no significant difference in the preference for the gender among younger adults, middle-aged, and older adults \((\chi^2(3456) = 1.067, p = 0.587)\). We can observe from Fig. 4 that both younger adults and middle-aged and older adults prefer a female voice assistant, then a neutral one.

There was no significant difference in the preference for voice assistant device size among younger adults, middle-aged, and older adults \((\chi^2(4456) = 17.486, p = 0.002)\). We can observe from Fig. 5 that younger adults preferred a voice assistant device of lamp size, then mobile phone size.
size. Middle-aged and older adults preferred a voice assistant device of mobile phone size.

### 4.3.2 User requirements of a voice assistant

The descriptive statistics and inference statistics for each item are presented in Table 3. Participants generally had a positive attitude toward voice assistants. Only two items (“Q10. People around me are using voice assistants.” “Q11. I can accept that a voice assistant transfers my data to the Internet.”) had negative evaluation scores. This revealed that voice assistants are not that popular, and people have concerns about privacy. The items that participants agreed with most were “Q37. A voice assistant with an emergency alarm function (e.g., in the event of fall or fire) is better.” “Q21. The response accuracy of a voice assistant is important.” “Q22. It is important that a voice assistant is of good quality and not fragile.”

User requirement differences between younger and older adults are presented in Figs. 7 and 8. Younger adults showed a higher agreement to the items: “Q4. Learning to use a voice assistant is easy for me”, “Q15. after-sales service”, “Q16. independent operation”, “Q17. price”, “Q18. financial burden”, and “Q20. appearance”. Middle-aged and elderly people showed a higher agreement to the following: “Q3. use anywhere”, “Q7. improve communication skills”, “Q8. improve parent–child relationship”, “Q9. People who are important to me feel I should use a voice assistant”, “Q10. People around me are using voice assistants”, “Q11. transfer data to the Internet”, “Q12. environmental support”, “Q24. speak slowly”, “Q27. social-communication style”, “Q28. high pitch”, “Q30. with a screen”, “Q32. actively communicate”, “Q33. imitating acquaintances”, “Q39. video
call”, “Q40. high volume”, “Q41. navigation”, and “Q44. dancing”, “Q48. I trust the ability of the voice assistant”, “Q47. I trust the information that the voice assistant relays me” and “Q50. I suppose using a voice assistant can make me more relaxed and happier”.

### 4.4 Factor analysis (RQ2)

Exploratory factor analyses were conducted to identify the factors influencing user acceptance of voice assistants. The Kaiser-MeyerOlkin (KMO) measure of sampling adequacy was 0.942, and Bartlett’s test of sphericity indicated that $\chi^2 = 11,782.9$, $df = 1225$, and $p < 0.001$. The results suggested that the exploratory factor analysis was suitable.

A principal component analysis with varimax rotation was used for the factor analysis. Six items (“Q29. A voice assistant with a small size is better” “Q42. I can call out the wake-up words of the voice assistant clearly and accurately” “Q19. I would prefer a voice assistant of my age.” “Q41. A voice assistant with a navigation function is better” “Q25. A female voice assistant is better” “Q32. A voice assistant that could actively communicate with me is better”) with factor loadings of less than 0.45 were disregarded. The results identified nine common factors with eigenvalues greater than 1. These common factors accounted for 62.47% of the total variation. Cronbach’s $\alpha$ for the scale was 0.938. Table 4 presents the common factors and items in each factor.
Table 3  Descriptive and inference statistics of participants’ responses

| Questions                                                                 | All (n = 456) | Younger adults (n = 315) | Middle-aged and older adults (n = 141) | Analysis of variance (ANOVA) |
|---------------------------------------------------------------------------|---------------|-------------------------|----------------------------------------|-----------------------------|
|                                                                           | M  SD         | M  SD                   | M  SD                                  | $F_{(1,455)}$  $p$  $n^2_p$ |
| Q1. Using a voice assistant would improve my performance                  | 3.9  1.1      | 3.8  1.1                | 4.1  1.2                                | 3.834  n.s  -               |
| Q2. Using a voice assistant could save my time                            | 4.2  1.0      | 4.3  0.9                | 4.2  1.1                                | 1.097  n.s  -               |
| Q3. I would use a voice assistant anywhere                                 | 3.0  1.4      | 2.8  1.3                | 3.6  1.4                                | 34.127  0.000  0.07        |
| Q4. Learning to use a voice assistant is easy for me                      | 4.1  1.1      | 4.3  0.9                | 3.9  1.3                                | 13.771  0.000  0.03        |
| Q5. Communicating with a voice assistant is easy for me                   | 4.1  1.1      | 4.1  1.0                | 4.0  1.3                                | 2.307  n.s  -               |
| Q6. Using a voice assistant would improve my health                       | 3.4  1.2      | 3.4  1.1                | 3.5  1.3                                | 1.280  n.s  -               |
| Q7. Using a voice assistant would improve my communication skills         | 3.3  1.3      | 3.1  1.2                | 3.7  1.4                                | 21.256  0.000  0.04        |
| Q8. Using a voice assistant would improve my parent–child relationship    | 3.3  1.3      | 3.1  1.2                | 3.6  1.3                                | 13.819  0.000  0.03        |
| Q9. People who are important to me feel I should use a voice assistant    | 3.0  1.3      | 2.7  1.2                | 3.6  1.3                                | 47.601  0.000  0.09        |
| Q10. People around me are using voice assistants                          | 2.6  1.3      | 2.5  1.3                | 2.9  1.3                                | 12.003  0.001  0.03        |
| Q11. Transfer data to the Internet                                        | 2.3  1.5      | 2.1  1.3                | 3.0  1.6                                | 41.337  0.000  0.08        |
| Q12. Brand                                                                | 3.8  1.2      | 3.8  1.2                | 3.6  1.3                                | 2.702  n.s  -               |
| Q13. My living environment supports my use of a voice assistant           | 3.6  1.3      | 3.4  1.3                | 4.0  1.3                                | 16.409  0.000  0.03        |
| Q14. Purchase channels                                                    | 3.6  1.3      | 3.6  1.3                | 3.6  1.4                                | 0.100  n.s  -               |
| Q15. After-sales service                                                  | 4.2  1.1      | 4.3  1.0                | 4.0  1.3                                | 8.697  0.003  0.02        |
| Q16. Independent operation                                                | 4.1  1.1      | 4.1  1.1                | 3.9  1.3                                | 4.102  0.043  0.01        |
| Q17. Price                                                                | 4.2  1.2      | 4.4  0.9                | 3.6  1.5                                | 49.743  0.000  0.10        |
| Q18. Financial burden                                                     | 3.1  1.2      | 3.2  1.1                | 2.9  1.4                                | 6.829  0.009  0.01        |
| Q19. Of my age                                                            | 3.8  1.1      | 3.8  1.1                | 3.8  1.3                                | 0.038  n.s  -               |
| Q20. Appearance                                                           | 4.0  1.2      | 4.2  1.1                | 3.5  1.4                                | 29.599  0.000  0.06        |
| Q21. Accuracy                                                             | 4.6  0.8      | 4.6  0.8                | 4.6  0.7                                | 0.097  n.s  -               |
| Q22. Good quality                                                         | 4.6  0.8      | 4.6  0.8                | 4.6  0.8                                | 0.150  n.s  -               |
| Q23. Communicate in dialects                                              | 3.9  1.2      | 3.8  1.2                | 4.0  1.2                                | 2.189  n.s  -               |
| Q24. Speak slowly                                                         | 3.3  1.2      | 3.1  1.1                | 3.8  1.2                                | 34.146  0.000  0.07        |
| Q25. Female                                                               | 3.6  1.1      | 3.5  1.1                | 3.6  1.2                                | 0.227  n.s  -               |
| Q26. With professional background                                         | 4.0  1.0      | 4.0  1.0                | 4.0  1.0                                | 0.287  n.s  -               |
| Q27. Social-communication style                                           | 4.1  1.0      | 4.0  1.1                | 4.3  1.0                                | 5.518  0.019  0.01        |
| Q28. High pitch                                                           | 3.2  1.2      | 3.1  1.1                | 3.6  1.3                                | 21.015  0.000  0.04        |
| Q29. Small size                                                           | 4.1  1.0      | 4.1  1.0                | 4.2  1.0                                | 1.554  n.s  -               |
| Q30. With a screen                                                        | 4.1  1.0      | 4.0  1.0                | 4.4  0.9                                | 11.865  0.001  0.03        |
| Q31. Portable                                                             | 4.4  0.9      | 4.4  0.9                | 4.5  0.8                                | 1.919  n.s  -               |
| Q32. Actively communicate                                                 | 4.2  1.1      | 4.0  1.1                | 4.5  1.0                                | 15.823  0.000  0.03        |
| Q33. Imitating acquaintances                                              | 3.5  1.4      | 3.4  1.4                | 4.0  1.2                                | 21.484  0.000  0.05        |
| Q34. Printing                                                             | 4.2  1.1      | 4.2  1.0                | 4.0  1.2                                | 2.798  n.s  -               |
| Q35. Provide historical inquiry records                                   | 4.4  0.9      | 4.3  1.0                | 4.5  0.9                                | 2.248  n.s  -               |
| Q36. Detecting body-related indicators                                    | 4.5  0.9      | 4.5  0.9                | 4.6  0.9                                | 0.827  n.s  -               |
| Q37. Alarm (on fall or fire)                                              | 4.7  0.7      | 4.7  0.7                | 4.7  0.7                                | 1.105  n.s  -               |
| Q38. Psychological counselling                                           | 4.3  1.0      | 4.2  1.1                | 4.4  1.0                                | 2.289  n.s  -               |
| Q39. Video call                                                           | 4.4  0.9      | 4.4  1.0                | 4.6  0.7                                | 5.806  0.016  0.01        |
| Q40. High volume                                                          | 3.5  1.2      | 3.3  1.1                | 4.1  1.1                                | 53.803  0.000  0.11        |
| Q41. Navigation                                                           | 4.0  1.2      | 3.9  1.2                | 4.3  1.0                                | 10.955  0.001  0.02        |
| Q42. Call out wake-up words                                               | 4.2  1.0      | 4.2  1.0                | 4.1  1.1                                | 0.472  n.s  -               |
| Q43. Singing                                                              | 4.1  1.0      | 4.0  1.0                | 4.2  1.0                                | 3.591  n.s  -               |
| Q44. Dancing                                                              | 3.5  1.2      | 3.4  1.3                | 3.8  1.1                                | 12.415  0.000  0.03        |
| Q45. Broadcasting news                                                    | 4.3  1.0      | 4.2  1.1                | 4.4  1.0                                | 2.107  n.s  -               |
4.5 Regression analysis (RQ3)

The factor analysis generated a variable for each common factor, and these variables were used as the independent variables in the regression analysis. The dependent variable was the mean of the four items regarding voice assistant use intention. A stepwise method was used to screen the variables. The final regression model included six factors, and the six factors accounted for 58.0% of the total variance. The regression results are shown in Table 5.

The coefficients of the predictor variables in the regression model are listed in Table 6. The most influential factor was hedonic motivation and trust (Factor 3), which explained 44.3% of the total variance. The second factor was social influence (Factor 2), followed by performance expectancy (Factor 9), effort expectancy (Factor 6), product features (Factor 1), and facilitating conditions (Factor 8). The final model was as follows:

\[
\text{Use intention} = -0.17 + 0.41 \text{Hedonic motivation} + 0.22 \text{Social influence} + 0.15 \text{Performance expectancy} + 0.08 \text{Effort expectancy} + 0.15 \text{Product features} + 0.06 \text{Facilitating conditions}
\]

In the second step, linear regression analyses were conducted for younger participants and middle-aged and older participants to identify which factors were the most critical factors to use intention. A stepwise method was used to screen the variables. The final regression model included six factors, and the six factors accounted for 58.0% of the total variance. The regression results are shown in Table 5.

The coefficients of the predictor variables in the regression model are listed in Table 6. The most influential factor was hedonic motivation and trust (Factor 3), which explained 44.3% of the total variance. The second factor was social influence (Factor 2), followed by performance expectancy (Factor 9), effort expectancy (Factor 6), product features (Factor 1), and facilitating conditions (Factor 8). The final model was as follows:

\[
\text{Use intention} = -0.17 + 0.41 \text{Hedonic motivation} + 0.22 \text{Social influence} + 0.15 \text{Performance expectancy} + 0.08 \text{Effort expectancy} + 0.15 \text{Product features} + 0.06 \text{Facilitating conditions}
\]

5 Discussion

5.1 User requirement differences for voice assistants among younger, middle-aged, and older adults (RQ1)

According to the survey, alarm function, answer accuracy, good quality, detection of body-related functions, and built-in power supply were among the top requirements among participants. The requirements of young adults and middle-aged and older people were different based on the interviews and questionnaire survey. The elderly were more inclined toward health monitoring and emergency alarms. Middle-aged respondents paid more attention to the functions of smart home voice assistants and children’s education. Younger adults’ requirements were mostly concentrated on entertainment and companionship.

There were significant differences in the preferences for the appearance, age, and size of voice assistants among younger adults, middle-aged, and older adults. Middle-aged and older adults actually prefer voice assistants with a humanoid appearance, while younger adults prefer one that appears as animals or has a geometric shape. Regarding the gender of the voice assistant, female is preferred by younger, middle-aged, and older adults. This is in line with a previous study [25]. Taking the most popular smart home voice assistants, Tmall Genie and MI AI in China, as examples,
they are both set as female by default [41, 42]. A female voice is preferred by people as the role of caring is regarded for women in diverse societies [43]. People consider the voices of female voice assistants as more friendly and kinder [44]. Regarding the age of the voice assistant, younger adults generally prefer voice assistants with a younger age than those preferred by older adults. This could be explained by the similarity-attraction hypothesis of Byrne, that is, the

| Table 4 Items in each common factor |
|-------------------------------------|
| Factors                            | Items                                                                 | Loading | Cronbach’s α |
| Factor 1: Product features          | Q37. Alarm (on fall or fire)                                         | .772    | .883         |
|                                    | Q36. Detecting body-related indicators                              | .696    |              |
|                                    | Q31. Portable                                                      | .668    |              |
|                                    | Q35. Provide historical inquiry records                              | .651    |              |
|                                    | Q21. Accuracy                                                      | .647    |              |
|                                    | Q22. Good quality                                                  | .625    |              |
|                                    | Q39. Video call                                                    | .618    |              |
|                                    | Q27. Social-communication style                                     | .580    |              |
|                                    | Q30. With a screen                                                 | .559    |              |
|                                    | Q26. With professional background                                   | .514    |              |
|                                    | Q38. Psychological counselling                                     | .472    |              |
|                                    | Q34. Printing                                                      | .458    |              |
| Factor 2: Social influence          | Q9. People who are important to me feel I should use a voice assistant | .752    | .878         |
|                                    | Q7. Using a voice assistant would improve my communication skills   | .732    |              |
|                                    | Q8. Using voice assistant would improve my parent–child relationship| .729    |              |
|                                    | Q3. I would use a voice assistant anywhere                          | .702    |              |
|                                    | Q6. Using a voice assistant would improve my health                | .672    |              |
|                                    | Q11. I can accept that a voice assistant transfers my data to the Internet | .613    |              |
|                                    | Q10. People around me are using voice assistants                    | .589    |              |
| Factor 3: Hedonic motivation and trust | Q46. Storytelling                                                | .733    | .877         |
|                                    | Q48. I trust the ability of the voice assistant                    | .664    |              |
|                                    | Q43. Singing                                                       | .660    |              |
|                                    | Q49. I suppose using a voice assistant is interesting               | .652    |              |
|                                    | Q47. I trust the information that the voice assistant relays me     | .634    |              |
|                                    | Q45. Broadcasting news                                             | .576    |              |
|                                    | Q44. Dancing                                                       | .560    |              |
|                                    | Q50. I suppose using a voice assistant can make me more relaxed and happier | .548    |              |
| Factor 4: Language competence       | Q23. Communicate in dialects                                        | .699    | .798         |
|                                    | Q33. Imitating acquaintances                                        | .642    |              |
|                                    | Q24. Speak slowly                                                  | .598    |              |
|                                    | Q28. High pitch                                                    | .580    |              |
|                                    | Q40. High volume                                                   | .571    |              |
|                                    | Q41. Navigation                                                    | .465    |              |
| Factor 5: Peripheral services       | Q15. After-sales service                                            | .714    | .725         |
|                                    | Q16. Independent operation                                          | .708    |              |
|                                    | Q12. Brand                                                         | .691    |              |
| Factor 6: Effort expectancy         | Q4. Learning to use a voice assistant is easy for me                | .768    | .805         |
|                                    | Q5. Communicating with a voice assistant is easy for me             | .751    |              |
| Factor 7: Cost and appearance       | Q18. Financial burden                                              | .790    | .685         |
|                                    | Q17. Price                                                         | .691    |              |
|                                    | Q20. Appearance                                                    | .526    |              |
| Factor 8: Facilitating conditions   | Q14. Purchase channels                                             | .612    | .671         |
|                                    | Q13. My living environment supports my use of a voice assistant     | .484    |              |
| Factor 9: Performance expectancy    | Q1. Using a voice assistant would improve my performance            | .657    | .755         |
|                                    | Q2. Using a voice assistant could save my time                     | .648    |              |
user’s trust would be highest with agents that appear similar to them, particularly in age appearance [45]. Regarding size, the size of mobile phones or lamps is preferred by younger adults, while older adults prefer a mobile phone size.

5.2 Key factors that explained younger and older adults’ requirements for a smart voice assistant (RQ2)

Based on the survey data, user requirements for voice assistants were composed of nine factors: product features, social influence, hedonic motivation and trust, language competence, peripheral services, effort expectancy, cost and appearance, facilitating conditions, and performance expectancy. The nine-factor requirement structure provided a significant explanation and had high internal reliability. Compared to the UTAUT framework, several factors

Table 5 ANOVA results of the regression model

| Items added           | Adjusted $R^2$ | $R^2$ change | $F$ change | Sig. F change |
|-----------------------|---------------|--------------|------------|---------------|
| Fac3                  | .44           | .44          | 361.06     | .000          |
| Fac3, Fac2            | .53           | .09          | 90.88      | .000          |
| Fac3, Fac2, Fac9      | .56           | .03          | 29.56      | .000          |
| Fac3, Fac2, Fac9, Fac6 | .57           | .01          | 10.54      | .001          |
| Fac3, Fac2, Fac9, Fac6, Fac1 | .58 | .01 | 7.57 | .006 |
| Fac3, Fac2, Fac9, Fac6, Fac1, Fac8 | .58 | .01 | 3.91 | .049 |

Table 6 Coefficients of predictor variables in regression model ($N=456$)

| Model                                      | Unstandardized coefficients | Standard coefficients | $T$ | Sig   |
|--------------------------------------------|-----------------------------|-----------------------|-----|-------|
| Constant                                   | −.17                        | .21                   | −.83| .41   |
| Hedonic motivation and trust                | .41                         | .05                   | .35 | .00   |
| Social influence                           | .22                         | .04                   | .24 | .00   |
| Performance expectancy                      | .15                         | .04                   | .15 | .00   |
| Effort expectancy                          | .08                         | .03                   | .08 | .02   |
| Product features                           | .15                         | .06                   | .10 | .01   |
| Facilitating conditions                    | .06                         | .03                   | .07 | .05   |

Fig. 6 Results of regression analyses (**$p<.01$, ***$p<.001$)
Older adults were more concerned about voice assistant performance (e.g., improve performance and save time and facilitating conditions (e.g., purchase channel and living environment))

The six most important factors were determinants of users’ use intention: hedonic motivation and trust, social influence, performance expectancy, effort expectancy, product features, and facilitating conditions. These six factors proved to reliably explain users’ intentions to use voice assistants. Hedonic motivation and trust are the strongest determinants of the intention to use voice assistants. This is in line with previous studies [13, 26]. If users think the voice assistant is entertaining and interesting, they will be more likely to use it in the smart home context. Therefore, manufacturers should pay more attention to the hedonic functions of voice assistants in a smart home context. Functions such as telling stories, singing, and dancing related to users’ psychological well-being may increase users’ use intention.

5.3 Critical factors that influence younger adults’ and middle-aged and older adults’ use intention of voice assistant (RQ3)

Younger adults and middle-aged and older adults showed different patterns in the acceptance of voice assistants. Middle-aged and older participants showed significantly more positive evaluations of the intention to use smart home voice assistants ($M = 4.1, SD = 1.0$) than did their younger counterparts ($M = 3.8, SD = 0.9$), indicating that middle-aged and older participants were generally positive about this technology. This is contrary to the stereotype that older adults are technophobic.

For younger participants, the critical factors that influence their use intention are hedonic motivation and trust, social influence, product features, performance expectancy, and effort expectancy. For older participants, the critical factors are hedonic motivation and trust, performance expectancy, and facilitating conditions. Compared with younger adults, social influence (e.g., People who are important to me feel I should use a voice assistant), product features (e.g., Alarm (on fall or fire)), and effect expectancy (e.g., Learning to use a voice assistant is easy for me) are not significant factors that influence older adults’ use intention of voice assistant. Older adults were more concerned about voice assistant performance (e.g., improve performance and save time) and facilitating conditions (e.g., purchase channel and living environment).

5.4 Limitations and future studies

This study had several limitations. First, the sample size of middle-aged and older adults was relatively small ($n = 141$) compared with younger adults ($n = 315$) because it was more difficult for older adults to complete the online questionnaire. Second, this study adopted an online survey as the major data collection method. Those who completed the questionnaire were competent in using the Internet. Therefore, the attitudes of participants toward smart home voice assistants may be biased. However, non-users of voice assistants could also help in understanding situations in which the act of not participating/answering is an active decision that should be taken into account [46]. Finally, this study focused on user acceptance of voice assistants in a smart home context, therefore the generalization of the results needs caution.

Future studies could further investigate user requirements of voice assistants in other contexts (e.g., health care and production). Future studies may also leverage voice assistance as an intervention to improve people’s well-being. Future studies could use audio samples to represent different age and gender of a voice assistant in the questionnaire since the image initiate cognitive processes differently from audio perception.

6 Conclusion

This study aimed to investigate user acceptance of voice assistants in a smart home environment among younger, middle-aged, and older adults through a questionnaire survey. The results of the study would help manufacturers and designers understand the requirements of users in diverse age groups toward smart home voice assistants. The main findings are listed as follows:

- User requirements were composed of nine factors: product features, social influence, hedonic motivation and trust, language competence, peripheral services, cost and appearance, facilitating conditions, and performance expectancy;
- For younger adults, the critical factors that influence their use intention are hedonic motivation and trust, social influence, product features, performance expectancy, and effect expectancy. For middle-aged and older participants, the critical factors are hedonic motivation and trust, performance expectancy, and facilitating conditions;
- Video call, actively communicating, with a screen, social-communication style, and navigation were the top requirements among middle-aged and older adults. Video call, price, after-sales services, and appearance were the top requirements among young adults.
Table 7 Measurement items and references

| Variables                  | Items                                                                 | References |
|----------------------------|----------------------------------------------------------------------|------------|
| Performance expectancy     | Q1. Using voice assistant would improve my performance                 | [7]        |
|                            | Q2. Using voice assistant could save my time                           | [7]        |
|                            | Q3. I would use voice assistant anywhere                                | [7]        |
|                            | Q6. Using voice assistant would improve my health                      | [47]       |
|                            | Q7. Using voice assistant would improve my communication skills        | (self-designed) |
|                            | Q8. Using voice assistant would improve my parent-child relationship   | [33]       |
| Effort expectancy          | Q4. Learning to use voice assistant is easy for me                     | [48]       |
|                            | Q5. Communicating with voice assistant is easy for me                  | (self-designed) |
| Social influence           | Q9. People who are important to me feel that I should use voice assistant| [7]        |
|                            | Q10. People around me are using voice assistant                        | [49]       |
| Perceived risk             | Q11. I can accept that a voice assistant transfers my data to the Internet| [49]       |
| Facilitating conditions    | Q13. My living environment supports my use of voice assistant          | (self-designed) |
|                            | Q12. The brand of a voice assistant would determine whether I use this device | [49]       |
|                            | Q14. I can learn about the purchase channels of a voice assistant in multiple ways | (self-designed) |
|                            | Q15. The after-sales service would determine whether I use this product | [49]       |
|                            | Q16. The ability of a voice assistant to operate independently would determine whether I use this product | [49]       |
| Cost                      | Q17. The price of a voice assistants would determine whether I use this product | [26]       |
|                            | Q18. A voice assistant is a financial burden for me                    | [7]        |
| Product features           | Q19. I would prefer a voice assistant of my age                         | (self-designed) |
|                            | Q20. The appearance of a voice assistant would determine whether I use this product | [49]       |
|                            | Q21. The response accuracy of a voice assistant is important           | [49]       |
|                            | Q22. It is important that a voice assistant is of good quality and not fragile | [49]       |
|                            | Q23. A voice assistant that can communicate in the local dialects is better | (self-designed) |
|                            | Q24. The voice assistant that speaks slowly is better                  | (self-designed) |
|                            | Q25. A female voice assistant is better                                | [37]       |
|                            | Q26. A voice assistant with relevant professional background is better | [31]       |
|                            | Q27. A voice assistant responding in a social-communication style is better | [32]       |
|                            | Q28. A voice assistant with a high pitch is better                     | [50]       |
|                            | Q29. A voice assistant with a small size is better                     | (self-designed) |
|                            | Q30. A voice assistant with a screen is better                        | [36]       |
|                            | Q31. A voice assistant with a built-in power supply (portable) is better | [31]       |
|                            | Q32. A voice assistant that could actively communicate with me is better | (self-designed) |
|                            | Q33. A voice assistant with a function of imitating acquaintances is better | (self-designed) |
|                            | Q34. A voice assistant with a printing function is better             | [31]       |
|                            | Q35. A voice assistant with a memory function (providing historical inquiry records) is better | [49]       |
|                            | Q36. A voice assistant with a function of detecting body-related indicators is better | [31]       |
|                            | Q37. A voice assistant with an emergency alarm function (e.g., in the event of fall or fire) | [49]       |
|                            | Q38. A voice assistant with a psychological counselling function is better | [13]       |
|                            | Q39. A voice assistant with a video call function is better           | (self-designed) |
|                            | Q40. A voice assistant with a high volume is better                    | (self-designed) |
|                            | Q41. A voice assistant with a navigation function is better           | [51]       |
|                            | Q42. I can call out the wake words of the voice assistant clearly and accurately | [52]       |
|                            | Q43. A voice assistant with a singing function is better               | [13]       |
|                            | Q44. A voice assistant with a dancing function is better              | [13]       |
|                            | Q45. A voice assistant with the function of broadcasting news is better | [53]       |
|                            | Q46. A voice assistant with a storytelling function is better          | [53]       |
| Trust                      | Q47. I trust the information that the voice assistant relays me        | [32]       |
|                            | Q48. I trust the ability of the voice assistant                      | [32]       |
| Hedonic motivation         | Q49. I suppose using a voice assistant is interesting                 | [32]       |
|                            | Q50. I suppose using a voice assistant can make me more relaxed and happier | [54]       |
| Use intention              | Q51. I intend to use a voice assistant in the future                  | [54]       |
|                            | Q52. I intend to use a voice assistant frequently in the future       | [54]       |
|                            | Q53. I intend to rely on a voice assistant for independent living     | (self-designed) |
|                            | Q54. If I have a voice assistant, I intend to use related smart products (e.g., smart light and sweeping robot) | [18, 19]   |
User Requirement with Significant Differences (Middle-aged and older adults > Younger adults)

- Q53. I intend to rely on a voice assistant for independent living
  - Middle-aged and older adults: 3.2
  - Younger adults: 3.8

- Q52. I intend to use a voice assistant frequently in the future
  - Middle-aged and older adults: 3.7
  - Younger adults: 4

- Q50. I suppose using a voice assistant can make me more relaxed and happier
  - Middle-aged and older adults: 4
  - Younger adults: 4.2

- Q48. I trust the ability of the voice assistant
  - Middle-aged and older adults: 3.8
  - Younger adults: 4.2

- Q47. I trust the information that the voice assistant relays me
  - Middle-aged and older adults: 3.7
  - Younger adults: 3.9

- Q44. Dancing
  - Middle-aged and older adults: 3.4
  - Younger adults: 3.8

- Q41. Navigation
  - Middle-aged and older adults: 3.9
  - Younger adults: 4.3

- Q40. High volume
  - Middle-aged and older adults: 4.1
  - Younger adults: 4.4

- Q39. Video call
  - Middle-aged and older adults: 4.4
  - Younger adults: 4.6

- Q33. Imitating acquaintances
  - Middle-aged and older adults: 4
  - Younger adults: 3.4

- Q32. Actively communicate
  - Middle-aged and older adults: 4.5
  - Younger adults: 4

- Q30. With a screen
  - Middle-aged and older adults: 4.4
  - Younger adults: 4

- Q28. High pitch
  - Middle-aged and older adults: 3.6
  - Younger adults: 3.1

- Q27. Social-communication style
  - Middle-aged and older adults: 4.3
  - Younger adults: 4

- Q24. Speak slowly
  - Middle-aged and older adults: 3.8
  - Younger adults: 3.1

- Q13. My living environment supports my use of a voice assistant
  - Middle-aged and older adults: 3.4
  - Younger adults: 4

- Q11. Transfer data to the Internet
  - Middle-aged and older adults: 3
  - Younger adults: 4

- Q10. People around me are using voice assistants
  - Middle-aged and older adults: 2.9
  - Younger adults: 2.1

- Q9. People who are important to me feel I should use a voice assistant
  - Middle-aged and older adults: 3.6
  - Younger adults: 2.7

- Q8. Using a voice assistant would improve my parent-child relationship
  - Middle-aged and older adults: 3.6
  - Younger adults: 3.1

- Q7. Using a voice assistant would improve my communication skills
  - Middle-aged and older adults: 3.7
  - Younger adults: 3.1

- Q3. I would use a voice assistant anywhere
  - Middle-aged and older adults: 3.6
  - Younger adults: 2.8

Fig. 7 User requirements of voice assistant with significant differences (Rating: middle-aged and older adults > younger adults)
Appendix

See Table 7, Figs. 7, 8.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10209-022-00936-1.

Acknowledgements This study is supported by the National Natural Science Foundation of China 72101100, the Philosophy and Social Science Projects in Universities of Jiangsu Province 2020JSJA0861, and the Fundamental Research Funds for the Central Universities JUSRP321016. We thank all the participants for sharing their opinions about smart home voice assistant.

Declarations

Conflict of interest None.

References

1. Xiao, X.-T., Kim, S.-I.: A study on the user experience of smart speaker in China-focused on Tmall Genie and Mi AI speaker. J. Digit. Converg. 16, 409–414 (2018)
2. Onag, G.: Voice assistants in China’s smart homes to reach 100 million in 2024, https://futureiot.tech/voice-assistants-in-chinas-smart-homes-to-reach-100-million-in-2024/#:~:text=In%20a%20latest%20research%20findings%20released%20this%20week%2C%20increase%20to%20the%20low%20cost%20of%20smart%20speakers.
3. Poushneh, A.: Humanizing voice assistant: the impact of voice assistant personality on consumers’ attitudes and behaviors. J. Retail. Consum. Serv. 58, 102283 (2020). https://doi.org/10.1016/j.jretconser.2020.102283
4. Pradhan, A., Mehta, K., Findlater, L.: “Accessibility Came by Accident”: Use of voice-controlled intelligent personal assistants by people with disabilities. In: the 2018 CHI Conference (2018)
5. Gropper, E.I.: Promoting health by promoting comfort. In: Nursing Forum. pp. 5–8. Wiley Online Library (1992)
6. Sun, S., Xie, Z., Yu, K., Jiang, B., Zheng, S., Pan, X.: COVID-19 and healthcare system in China: challenges and progression for a sustainable future. Glob. Health 17, 1–8 (2021)
7. Arfi, W.B., Nasr, I.B., Khvatova, T., Zaied, Y.B.: Understanding acceptance of eHealthcare by IoT natives and IoT immigrants: an integrated model of UTAUT, perceived risk, and financial cost. Technol. Forecast. Soc. Chang. 163, 120437 (2021). https://doi.org/10.1016/j.techfore.2020.120437
8. Zhong, R., Rau, P.-L.P., Yan, X.: Gait assessment of younger and older adults with portable motion-sensing methods; a user study. Mob. Inf. Syst. 2019, 1–13 (2019). https://doi.org/10.1155/2019/1093514
9. Ghorayeb, A., Comber, R., Gooberman-Hill, R.: Older adults’ perspectives of smart home technology: Are we developing the technology that older people want? Int. J. Hum Comput Stud. 147, 102571 (2021). https://doi.org/10.1016/j.ijhcs.2020.102571
10. Kim, H.N., Oumarou, B.: User requirement analysis for smart voice technology for older adults with visual impairments. Int. J. Hum.-Comput. Interact. 36, 1551–1557 (2020). https://doi.org/10.1080/10447318.2020.1768676
11. Banks, M.R., Willoughby, L.M., Banks, W.A.: Animal-assisted therapy and loneliness in nursing homes: use of robotic versus living dogs. J. Am. Med. Dir. Assoc. 9, 173–177 (2008). https://doi.org/10.1016/j.jamda.2007.11.007
12. Straßmann, C., Krämer, N.C., Buschmeier, H., Kopp, S., et al.: Age-related differences in the evaluation of a virtual health agent’s appearance and embodiment in a health-related interaction: experimental lab study. J. Med. Internet Res. 22, e13726 (2020)
13. Chiu, C.-J., Hsieh, S., Li, C.-W., et al.: Needs and preferences of middle-aged and older adults in Taiwan for companion robots and pets: survey study. J. Med. Internet Res. 23, e23471 (2021)
14. Biermann, H., Brauner, P., Zieflé, M.: How context and design shape human-robot trust and attributions. J. Behav. Robot. 12, 74–86 (2021). https://doi.org/10.1515/jbbr-2021-0008
15. Sugano, S., Miyaji, Y., Tomiyama, K.: Study of Kawaii-ness in motion—physical properties of kawaii motion of Roomba. In: Kurosu, M. (ed.) Human-Computer Interaction. Human-Centred
20. Ajzen, I.: The theory of planned behavior. Organ. Behav. Hum. Decis. Process. 50, 179–211 (1991). https://doi.org/10.1016/0749-5978(91)90020-T

21. Rogers, E.M.: Diffusion of Innovations. Etats-Unis Free Press, New York (1995)

22. Aldossari, M.Q., Sidorova, A.: Consumer acceptance of Internet of Things (IoT): smart home context. J. Comput. Inf. Syst. 60, 507–517 (2018). https://doi.org/10.1080/08874417.2018.1543000

23. Azikoz, F., Vega, R.P.: The role of privacy cynicism in consumer habits with voice assistants: a technology acceptance model perspective. Int. J. Hum. Comput. Interact. pp 1–15 (2022)

24. Yılmaz, M.B., Rızvanoğlu, K.: Understanding users’ behavioral intention to use voice assistants on smartphones through the integrated model of user satisfaction and technology acceptance: a survey approach. J. Eng. Design Technol. (2021). https://doi.org/10.1108/JEDT-02-2021-0084

25. Liu, N., Liu, R., Li, W.: Identifying design feature factors critical to acceptance of smart voice assistant. In: International Conference on Human-Computer Interaction. pp. 384–395. Springer (2021)

26. Canziani, B., MacSween, S.: Consumer acceptance of voice-activated smart home devices for product information seeking and online ordering. Comput. Hum. Behav. 119, 106714 (2021)

27. Burbach, L., Halbach, P., Plettenberg, N., Nakayama, J., Valdez, A.C.: “Hey Siri”, “Ok, Google”, “Alexa”: acceptance-relevant factors of virtual voice-assistants. In: 2019 IEEE International Professional Communication Conference (ProComm). pp. 101–111 (2019)

28. Adapa, A., Nah, F.F.H., Hall, R.H., Siat, K., Smith, S.N.: Factors influencing the adoption of smart wearable devices. Int. J. Hum.-Comput. Interact. 34, 399–409 (2018). https://doi.org/10.1080/10447318.2017.1357902

29. Heek, O.V., Wilkowski, W., Ziefle, M.: Cultural impact on perceptions of aging, care, and lifelogging technology: a comparison between Turkey and Germany. Int. J. Hum.-Comput. Interact. 37, 156–168 (2021). https://doi.org/10.1080/10447318.2020.1809247

30. Sengupta, K., Sarcar, S., Pradhan, A., McNaney, R., Sayago, S., Chattopadhyay, D., Joshi, A.: Challenges and opportunities of leveraging intelligent conversational assistant to improve the well-being of older adults. In: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems. pp. 1–4. Association for Computing Machinery, New York, NY, USA (2020)

31. Zhong, R., Ma, M.: Effects of communication style, anthropomorphic setting and individual difference on older adults using voice assistants in the context of health. BMC Geriatr. (2022). https://doi.org/10.1186/s12877-022-03428-2

32. Chattaraman, V., Kwon, W.S., Gilbert, J.E., Ross, K.: Should AI-Based, conversational digital assistants employ social- or task-oriented interaction style? A task-competency and reciprocity perspective for older adults. Comput. Hum. Behav. 90, 315–330 (2019)

33. Mo, F., Zhou, J., Yi, S.: How to enhance intergenerational communication? The influence of family orientation and generation when using social robots as an intermediary. In: International conference on human aspects of IT for the aged population. pp. 348–359. Springer, Cham (2017)

34. Ouyang, X., Zhou, J.: Smart TV for older adults: a comparative study of the mega menu and tiled menu. In: International conference on human aspects of IT for the aged population. Springer, Cham (2018)

35. Zhou, J., Chourasia, A., Vanderheiden, G.: Interface adaptation to novice older adults’ mental models through concrete metaphors. Int. J. Hum. Comput. Interact. 35, 592–606 (2017). https://doi.org/10.1080/10447318.2016.1265827

36. Baeza, R.R., Kumar, A.R.: Perceived usefulness of multimodal voice assistant technology. In: Proceedings of the human factors and ergonomics society annual meeting. pp. 1560–1564. SAGE Publications Sage CA: Los Angeles, CA (2019)

37. Pan, P.: Effects of Voice Assistance’s Gender Cues on User Perception and Social Desirability, http://ss.zhihzen.com/detail_38502727e75002d641aa0db783967f123846d185d507045a1921bfa3ea255101928fa69a765a342d5b93af153da4a00eaf7cc8c5f41886c775eae22ef5283a7b4c4ahb360671eca27c78bafa8dd67a1c?, (2017)

38. Tolmeijer, S., Zierau, N., Janson, A., Wahdatehagh, J.S., Leimestein, J.M.M., Bernstein, A.: Female by default?–Exploring the effect of voice assistant gender and pitch on trait and trust attribution. In: Extended abstracts of the 2021 CHI conference on human factors in computing systems. pp. 1–7 (2021)

39. Zhang, C., Conrad, F.: Speeding in web surveys: the tendency to answer very fast and its association with straightlining. Surv. Res. Methods 8, 127–135 (2014)

40. Kim, Y., Dykema, J., Stevenson, J., Black, P., Moberg, D.P.: Straightlining: overview of measurement, comparison of indicators, and effects in mail-web mixed-mode surveys. Soc. Sci. Comput. Rev. 37, 214–233 (2019). https://doi.org/10.1177/0894439317752406

41. Fei, Q.: Mi AI Speaker is a 17-year-old girl with red hair, https://news.znds.com/article/news/26128.html

42. Jack Ma: It is still women who determine the standard of beauty in the AI era, https://baijiahao.baidu.com/s?id=1643183690892612462&wfr=spider&for=pc

43. Danielescu, A.: Eschewing gender stereotypes in voice assistants to promote inclusion. In: Proceedings of the 2nd conference on conversational user interfaces. pp. 1–3 (2020)

44. Mitchell, W.J., Ho, C.-C., Patel, H., MacDorman, K.F.: Does social desirability bias favor humans? Explicit-implicit evaluations of synthesized speech support a new HCl model of impression management. Comput. Hum. Behav. 27, 402–412 (2011). https://doi.org/10.1016/j.chb.2010.09.002

45. Byrne, D.E.: The Attraction Paradigm. Academic Press, New York (1971)

46. Satchell, C., Dourish, P.: Beyond the user: use and non-use in HCI. In: Proceedings of the 21st annual conference of the Australian computer-human interaction special interest group: Design: Open 24/7. pp. 9–16 (2009)

47. Luo, Y., Lee, B., Choe, E.K.: TandemTrack: shaping consistent exercise experience by complementing a mobile app with a
smart speaker. In: CHI ’20: CHI conference on human factors in computing systems (2020)
48. Renaud, K., Van Biljon, J.: Predicting technology acceptance and adoption by the elderly: a qualitative study. In: Research conference of the south african institute of computer scientists & information technologists on it research in developing countries: riding the wave of technology. pp. 210–219 (2008)
49. Sun, N., Rau, P.: The acceptance of personal health devices among patients with chronic conditions. Int. J. Med. Informatics 84, 288–297 (2015)
50. Elkins, A.C., Derrick, D.C.: The sound of trust: voice as a measurement of trust during interactions with embodied conversational agents. Group Decis. Negot. 22, 897–913 (2013)
51. Andrist, S., Ziadée, M., Boukaram, H., Muthu, B., Sakr, M.: Effects of culture on the credibility of robot speech: a comparison between English and Arabic. In: Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction. pp. 157–164. Association for Computing Machinery, New York, NY, USA (2015)
52. Baidu AI Interaction Design Lab: What name will wake you up with? My smart little friend, https://mp.weixin.qq.com/s/SuXKyhqFD4x5Yikbn9OE Dg
53. Pradhan, A., Lazar, A., Findlater, L.: Use of intelligent voice assistants by older adults with low technology use. ACM Trans. Comput. Hum. Interact. (2020). https://doi.org/10.1145/3373759
54. Wang, L., Rau, P., Salvendy, G.: Older adults’ acceptance of information technology. Educ. Gerontol. 37, 1081–1099 (2011)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.