How social cues in virtual assistants influence concerns and persuasion: The role of voice and a human name

Voorveld, H.A.M.; Araujo, T.

DOI
10.1089/cyber.2019.0205

Publication date
2020

Document Version
Final published version

Published in
Cyberpsychology, Behavior, and Social Networking

License
Article 25fa Dutch Copyright Act

Citation for published version (APA):
How Social Cues in Virtual Assistants Influence Concerns and Persuasion: The Role of Voice and a Human Name

Hilde A.M. Voorveld, PhD and Theo Araujo, PhD

Abstract

The aim of this study was to test how two important types of social cues used by virtual assistants today can affect consumer concerns and persuasion. These two cues are modality (voice-based via smart speaker, voice-based via a smartphone, or text-based on a smartphone screen) and the adoption of a human name rather than no name. An online scenario-based experiment ($n = 180$) has shown that participants who were exposed to a voice-based recommendation via a smart speaker were the most concerned about security and found text-based recommendations on a screen to be the most persuasive. Participants who were exposed to a virtual assistant with a human name were less concerned about their autonomy and were more strongly persuaded than those exposed to an assistant without a human name.

Keywords: virtual assistants, persuasion, voice, social cues, smart speaker

Virtual assistants are voice- or text-activated software agents that interpret requests (usually in natural language) by users and execute commands. These assistants, such as Amazon’s Alexa, Apple’s Siri, and Microsoft’s Cortana, are becoming increasingly popular across the globe. Brands have begun to leverage this increasing popularity to make personalized recommendations and influence the consumer’s buying process. In this context, it is an opportune moment to study how the characteristics of virtual assistants can influence the way that consumers are persuaded by them. Theoretically, such research is vital as, with virtual assistants, technology is no longer only the medium through which a message is conveyed but can also be seen as the communicator.

Our study focuses specifically on how the social cues frequently used by virtual assistants can influence people’s concerns about privacy, security, and autonomy as well as their persuasion knowledge, attitudes toward a recommended brand, and adherence to recommendations. Studies that have adopted the “computers are social actors” (CASA) paradigm have suggested that people tend to respond socially to computers. Such reactions can be influenced by the presence of even minimal social cues. When we consider that virtual assistants are often designed to be human like and interact socially, social cues are a potentially key component of a virtual assistant’s persuasive capabilities. In this study, we focus on two key cues: modality (voice vs. screen) and the use of human names.

Modality: Different Types of Voice and Screen-Based Assistants

Interactions with virtual assistants can take place via a smart home speaker (like Alexa) or a smartphone (like Siri). Because such interactions can either happen solely by voice or have a screen-based interface, we compare three interaction possibilities that differ with regard to the modality used: a voice-controlled smart speaker, a voice-controlled smartphone application, and a smartphone application with a traditional screen-based interface.

Due to the novelty of voice-controlled devices, consumers may have more and different types of concerns when using these interfaces when compared to interfaces that they are more accustomed to, such as (touch) screens. In addition, voice-controlled devices collect other types of personal information, a fact that may lead to different concerns when compared to screen-based devices. For example, people may fear that such devices “listen” to conversations. A recent qualitative study has shown that the use of microphones and the lack of transparency about smart speaker data practices are central to people’s concerns.
The novelty of voice-controlled interfaces may also influence consumers’ levels of perceived persuasive intent. The Persuasion Knowledge Model describes how consumers use their growing general knowledge about persuasion to cope with persuasive messages. Persuasion knowledge represents the general comprehension of how, when, and why one is confronted with persuasive attempts. As people have less experience with receiving persuasive messages via voice-controlled devices than via screens, they are, therefore, likely to be less aware that a recommendation is, in reality, a persuasive attempt.

People’s concerns and persuasion knowledge may subsequently influence the extent to which they are persuaded by recommendations made via an assistant. Previous research has demonstrated that awareness of persuasion attempts usually leads to more critical evaluation among adults, but not among young consumers. Although this relationship has never been tested for virtual assistants, given the highly personalized recommendations that they make, it is not yet clear whether knowledge of persuasive intent will lead to weaker persuasion levels. The same goes for concerns about privacy and security: increased levels of concern do not always lead to protective behavior or more critical evaluations. There is a trade-off between the loss of privacy and the convenience offered.

In the context of virtual assistants, voice can be seen as a cue for the source of information. Traditionally, the source can be seen as the initiator of communication, and the channel is the medium through which it is delivered. However, based on theories regarding human-machine communication, a virtual assistant may be seen as the source of information rather than the channel of communication. Although people may know that a virtual assistant does not create the information that it provides, they are likely to treat it as an autonomous source with intentions. This may be even more likely if a virtual assistant exhibits social cues such as voice. As a result, people may respond differently to recommendations when social cues become stronger.

Compared to screen interfaces, voice-based interfaces enable more intuitive, convenient, and efficient interactions via or with technology by providing hands- and eyes-free means of communication through spoken language applications. One very recent study has argued that interactions with voice-based virtual assistants are more natural and seamless. With regard to utilitarian tasks, Cho et al. demonstrated that people had more positive attitudes toward a virtual assistant that was voice- rather than text-based because it was perceived as more human like. This more intuitive way of communicating may make people engage in less in-depth information processing and, according to the MAIN model, people make snap judgments of information that are usually positive. Therefore, consumers may be more easily persuaded via voice than via screens. In sum, our hypothesis is as follows:

H1: A recommendation conveyed via a virtual assistant with a voice interface as compared to a screen interface will yield (a) higher levels of concern, (b) lower levels of persuasion knowledge, (c) a more positive attitude toward the recommended brand, and (d) greater adherence to the recommendation.

As previously discussed, users can interact with voice-controlled virtual assistants via a smart speaker or smartphone, and it remains to be seen whether these two ways differ with regard to concerns and persuasion levels. A direct comparison between smart speakers and smartphones is needed for at least two reasons. The first concerns external validity: interactions with the currently available virtual assistants can be either voice-based via smart speakers, voice-based via smartphones, or text-based on a smartphone screen. Second, if we were to only compare smart speakers to screen-based assistants, we would not know whether the effects are driven by a voice-based as opposed to a screen-based interface, or by the novelty of smart speakers as compared to smartphones. Developing specific expectations is, however, difficult: on the one hand, the novelty of smart speakers may affect people’s concerns and persuasion levels, as indicated above. On the other hand, the fact that smartphones are highly personal devices that are often considered to be an extension of the self may also influence concerns and persuasion levels. Our research questions therefore is as follows:

RQ1: To what extent do voice-controlled interactions with a virtual assistant via either a smart speaker or a smartphone differ with regard to concerns and persuasion?

The Impact of Human Names

Another way in which social cues can manifest themselves in virtual assistants is through the name they are given. It is important to note that some leading platforms (e.g., Amazon, Apple, Google, and Microsoft) have provided their assistants with names that have arguably varying levels of “humanness” (e.g., Alexa, Cortana, Siri, Google Assistant), thus highlighting the importance of the name as a social cue that may influence subsequent responses, including concerns and persuasion.

In addition to the effects of different modalities, adopting a human name may function as an identity cue (i.e., it may affect whether the assistant is identified at the outset as human or bot) and so evoke particular heuristics, for example, the so-called “helper heuristic.” Adopting a human name may influence trust and credibility and even give users the feeling that they are privileged in an otherwise technology-centered medium. This may make people feel they have less cause for concern. In addition, as per the MAIN model, people may mindlessly rely on these cues without engaging in in-depth and critical processing of the content, which could result in enhanced levels of persuasion. We propose, therefore, the following:

H2: A recommendation conveyed via a virtual assistant with a human name as opposed to no name will yield (a) lower levels of concern, (b) persuasion knowledge, (c) a more positive attitude toward the recommended brand, and (d) a greater adherence to the recommendation.
The Interaction of Voice and Human Name

According to the "cue-cumulation effect,"27 which is based on the rationale of the additivity hypothesis in dual-process persuasion models,42 a combination of cues that are consistent with each other may yield stronger persuasion effects than only one cue.43 In our case, if both social cues are present (i.e., voice and the adoption of a human name38), it is likely that the effects will be additive. Cue-cumulation effects have been identified in the context of IoT devices,30 news websites,27,43 and online review sites.44 Therefore, we hypothesize the following:

H3: A virtual assistant with two social cues is more persuasive in terms of (a) brand attitudes and (b) adherence to the recommendation.

Method

Design and participants

The hypotheses were tested in a scenario-based online experiment with a 3 (voice-based via smart speaker, voice-based via smartphone, or text-based on a screen of a smartphone) × 2 (human name vs. the virtual assistant) between-subjects design. Scenario studies have proved useful in related fields like personalized communication.45 The three versions of virtual assistants represent the different appearances of virtual assistants currently on the market and also allowed us to check whether the differences between a voice-controlled interface and a screen-based interface were driven by the type of device (smart speaker vs. smartphone or by the interface itself). The name Charlie was chosen because it is a name that is used for both men and women in the research country. This was deemed important given recent discussions on the role of gender in AI and smart speakers.46

The participants were recruited through an online panel of the research company Qualtrics. We excluded those who failed the attention and quality checks47 and those who experienced technical problems. The final sample consisted of 180 participants (M_age 47.83, SD_age 15.22, 52.8 percent male).

Procedure

After providing informed consent, participants read a short scenario description for a minimum duration of 1 minute. This instructed participants to imagine that they had invited eight friends to dinner the following week and that they asked a virtual assistant for advice about the best brand of chocolate to buy to make a certain dessert. Chocolate was considered to be an appropriate product category because people often put cooking-related questions to search engines and virtual assistants, and use such assistants mainly for habitual purchases.3,48 To enhance the realism of the scenario,49 participants were shown a picture of the virtual assistant and received the recommendation either via a synthesized voice-clip with a female voice or via a picture of a smartphone screen (Fig. 1). The virtual assistant recommended a nonexistent brand in line with methodological recommendations to enhance the likelihood of finding persuasion effects.47 Finally, participants completed a questionnaire measuring the dependent variables.

Measures

The measures used are listed in Table 1. Because no existing scale was available to measure concerns related to virtual assistants, we included a list of 12 possible concerns that were derived from academic and practitioner articles12,13 (Table 2). The list shows a significant overlap with a recently published interview study on smart speakers.14 Participants were asked to what extent they would be worried about these concerns when using the assistant at home. Response options ranged from 1 = not at all worried to 7 = very worried. A principal components analysis with Varimax rotation revealed two factors: one related to security concerns and one to concerns about autonomy. Gender, age, perceived

FIG. 1. Pictures of the experimental material. In the two smartphone conditions, the text displayed on the screen states: what can I help you with? In the voice-based via a smartphone condition (B) the recommendation was made via voice-only.50 In the text based on a smartphone screen condition (C) the recommendation was made via text only. Because most currently available smart speakers (A) do not have a screen, we did not include such a version.
When I would use this virtual assistant in my own home I would be worried that:

Table 1. Measurement of Dependent Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of items</th>
<th>Type of scale</th>
<th>Example items</th>
<th>Statistics</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Privacy concerns</td>
<td>7</td>
<td>1 = strongly disagree, 7 = strongly agree</td>
<td>When I use this virtual assistant: I am concerned about misuse of personal information. I feel fear that information may not be safe while stored</td>
<td>Cronbach’s α = 0.91; M = 4.36, SD = 1.15</td>
<td>4,51</td>
</tr>
<tr>
<td>Persuasion knowledge</td>
<td>4</td>
<td>1 = strongly disagree, 7 = strongly agree</td>
<td>The recommendation feels like an ad. The brand Veraca paid to use this message</td>
<td>Cronbach’s α = 0.88, M = 5.32, SD = 1.20</td>
<td>52</td>
</tr>
<tr>
<td>Attitude toward the brand</td>
<td>4</td>
<td>Seven-point semantic differential scale</td>
<td>Bad/Good Unpleasant/Pleasant I am likely to buy the recommended brand I am likely to buy the recommended amount of chocolate, I am likely to buy the type of chocolate that was recommended.</td>
<td>Cronbach’s α = 0.91, M = 4.27, SD = 1.00</td>
<td>53</td>
</tr>
<tr>
<td>Adherence</td>
<td>3</td>
<td>1 = strongly disagree, 7 = strongly agree</td>
<td>Inspired by</td>
<td>Cronbach’s α = 0.81, M = 4.38, SD = 1.43</td>
<td>54</td>
</tr>
<tr>
<td>Involvement with virtual assistants</td>
<td>2</td>
<td>Seven-point semantic differential scale</td>
<td>very unimportant/very important very uninteresting/very interesting</td>
<td>r = 0.856, p = 0.000, M = 3.18, SD = 1.65</td>
<td>24</td>
</tr>
</tbody>
</table>

realism, and involvement with virtual assistants were measured as potential control variables.

Results

A series of two-way analyses of covariance were conducted. Involvement with virtual assistants was included as a covariate in all analyses because it was related to most of the independent variables (all p < 0.01). The results are summarized in Table 3.

Concerns

With regard to privacy concerns, the results showed neither a significant main effect of modality [F(2, 173) = 2.24, p = 0.109] nor a significant main effect of human name [F(1, 173) = 0.81, p = 0.37]. In addition, no significant interaction effect was found [F(2, 173) = 0.58, p = 0.56]. With regard to concerns about security, the results revealed a significant main effect of voice versus screen [F(2, 173) = 5.06, p = 0.007]. Concerns were highest among participants who were exposed to a voice-based recommendation via a smart speaker (M = 5.14, SE = 0.19). Participants in this condition were more concerned than in the voice-based smartphone condition (M = 4.37, SE = 0.16, Bonferroni: p = 0.005) and slightly more than in the text-based smartphone condition (M = 4.66, SE = 0.16 Bonferroni: ns). No main effect of human name was found [F(1, 173) = 0.08, p = 0.78] nor was there a significant interaction effect [F(2, 173) = 0.09, p = 0.91].

Table 2. Rotated Component Matrix of Items Used to Measure Concerns about Virtual Assistants

When I would use this virtual assistant in my own home I would be worried that:

<table>
<thead>
<tr>
<th>Concerns about security</th>
<th>Concerns about autonomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>...the data that the virtual assistants collect would be leaked</td>
<td>0.887</td>
</tr>
<tr>
<td>...a profile will be made of me</td>
<td>0.844</td>
</tr>
<tr>
<td>...the virtual assistant would listen to my conversations,</td>
<td>0.830</td>
</tr>
<tr>
<td>...my credentials would be stolen</td>
<td>0.809</td>
</tr>
<tr>
<td>...marketers get to know too much about me</td>
<td>0.806</td>
</tr>
<tr>
<td>...the virtual assistant would be hacked</td>
<td>0.803</td>
</tr>
<tr>
<td>...someone can watch or listen to me</td>
<td>0.773</td>
</tr>
<tr>
<td>...marketers and advertisers will try to influence me.</td>
<td>0.740</td>
</tr>
<tr>
<td>...the virtual assistant would make images of me</td>
<td>0.738</td>
</tr>
<tr>
<td>...the virtual assistant would influence my decisions</td>
<td>0.147</td>
</tr>
<tr>
<td>...I would get dependent on the virtual assistant</td>
<td>0.187</td>
</tr>
<tr>
<td>...the virtual assistant influenced my freedom of choice</td>
<td>0.357</td>
</tr>
</tbody>
</table>

EV = 7.089, R² = 59.076; EV = 1.68, R² = 13.99; Cronbach’s α = 0.94, Cronbach’s α = 0.88; M = 4.68, SD = 1.34 M = 3.72, SD = 1.55
With regard to concerns about autonomy, a main effect of human name was significant \(F(1, 173) = 5.42, p = 0.021\). Participants were more concerned that the virtual assistant would influence their freedom of choice when the assistant did not have a human name \(M = 4.02, SE = 0.16\) than when it had a name \(M = 3.48, SE = 0.17\). The results showed neither a main effect of modality \(F(2, 173) = 2.06, p = 0.13\) nor a significant interaction effect \(F(1, 173) = 0.44, p = 0.64\).

**Persuasion**

For persuasion knowledge, a significant main effect of voice versus screen was found: \(F(2, 173) = 3.43, p = 0.034\). Participants were most aware that the recommendation was a persuasive attempt when it was presented on a smartphone screen \(M = 5.57, SE = 0.14\). Post hoc Bonferroni tests showed that this condition differed significantly from the voice-based via smartphone condition \(M = 5.08, SE = 0.130, p = 0.029\) but not from the voice-based smart speaker condition \(M = 5.35, SE = 0.16\). Further, a marginally significant main effect of human name was found \(F(1, 173) = 2.95, p = 0.088\). Participants who were exposed to a virtual assistant with a human name were slightly more aware that the recommendation was sponsored by a brand \(M = 5.47, SE = 0.12\) than participants in the condition without a human name \(M = 5.19, SE = 0.11\). The interaction effect was not significant \(F(2, 173) = 0.34, p = 0.71\).

The analysis also showed a main effect of human name on brand attitude \(F(1, 173) = 5.62, <0.019\). Participants who were exposed to the virtual assistant with a human name evaluated the brand more positively \(M = 4.44, SE = 0.10\) than the participants who saw an assistant without a human name \(M = 4.11, SE = 0.10\). No main effect of modality \(F(2, 173) = 0.66, p = 0.52\) was found nor was there an interaction effect \(F(2, 173) = 1.83, p = 0.16\).

With regard to adherence to the recommendation, a main effect of voice versus screen was found \(F(2, 173) = 3.64, p = 0.028\). Participants were most likely to adhere to the recommendation when it was made via a smartphone screen \(M = 4.63, SE = 0.17\) as opposed to voice-based via a smart speaker \(M = 3.97, SE = 0.19, \text{Bonferroni: } p = 0.029\). The condition in which participants were exposed to a voice-based recommendation via a smartphone \(M = 4.47, SE = 0.16\) did not differ from the other conditions but was closer to the text-based recommendation via the smartphone screen condition than the voice-based smart speaker condition. In addition, the results showed a main effect of human name \(F(1, 173) = 4.20, p = 0.042\). Participants were more likely to adhere to the recommendation when the assistant had a human name \(M = 4.56, SE = 0.14\) than when it did not \(M = 4.15, SE = 0.14\). The interaction effect was insignificant \(F(2, 173) = 2.14, p = 0.12\).

**The mediating role of concerns and persuasion knowledge**

Before formally testing mediation, we first checked whether concerns and persuasion knowledge were correlated with brand attitude and adherence. Results showed a significant correlation only between concerns about security and brand attitude \(r = -0.210, p = 0.005\), and between persuasion knowledge, brand attitude and adherence \(r = -0.190, p = 0.011\) and \(r = -0.164, p = 0.028\). Subsequently, we conducted a mediation analysis for these variables with the Hayes PROCESS macro (v.3.54 model 4). Results showed no significant indirect effects (Table 4).

**Discussion**

This study aimed to test how the modality of a virtual assistant (voice vs. screen) and the adoption of a human name can influence people’s concerns about privacy, security, autonomy, and persuasive outcomes. The results showed that both cues played an important role.

The first major finding was that the modality through which a virtual assistant provided recommendations influenced both persuasiveness and concerns. We showed that people’s responses differ between assistants with voice and screen interfaces. In line with our expectations, persuasion knowledge was lower for voice- than for screen-based interactions. Contrary to our expectations, adherence to recommendations was highest when the recommendation was made via a smartphone screen. The novelty of voice speakers may mean that people are more willing to accept their recommendations than those received via screen-based interfaces. This is further illustrated by our findings that security concerns were highest among participants who were exposed to voice recommendations on a smart speaker and by the

---

**Table 3. Summary of the Results**

| Concerns |  |
|---|---|---|---|---|
| Privacy concerns | Concerns about security | Concerns about autonomy |  |
| Voice vs. screen | ns | Voice-based via smart speaker | ns | Voice-based via smartphone |
| Human name vs. voice | ns | ns | ns |

*Only comparisons that were significantly different in a post hoc Bonferroni test are displayed.

Marginal significant.

ns, not significant.
findings of a recent interview study showing that users of smart speakers would feel uncomfortable if their smart speaker voice commands were used to target adverts. This finding also extends previous recent findings that the modality (voice vs. screen) of a virtual assistant influences the perceived human-likeness of the assistant, whether people feel comfortable using an assistant, and the ease of understanding the information provided by showing that modality also influences concerns that people may have, and the persuasiveness of the assistant.

A second major finding was that merely giving the virtual assistant a human name influenced persuasion levels. Participants were more likely to follow the recommendation, have a more positive brand attitude, and were more aware of persuasive intent when the assistant had a human name than when it did not. In addition, participants were less concerned that the virtual assistant would influence their autonomy. This finding is an important theoretical contribution as the study is the first to show that merely giving a human-like name to a virtual assistant can influence persuasive outcomes. This provides a new level of insight into theories on persuasion knowledge and “hidden” persuasion. Although giving a name to an assistant leads users to see through the persuasive intent of a recommendation, it does not protect them against being persuaded. This is contrary to the traditional idea that more persuasion knowledge leads to less persuasion. This finding emphasizes that persuasion via virtual assistants is different from other forms of hidden persuasion (e.g., more novel, more personalized), and can be seen as preliminary empirical support to the helper heuristic notion.

A third major finding was that the results did not provide support for a cue-cumulation effect. It seems that the social cues worked independently of each other. A potential explanation is that the influence of modality goes beyond the influence of voice. Voice versus screen has implications for the functionality of a virtual assistant and the way it is operated. This is in line with recent work on interface psychology showing that different interfaces change the attributes that people use when they buy products via these interfaces. Adoption of a human name mostly relates to how people perceive the assistant and its intents rather than its functionality. The finding that a human name leads to fewer concerns about the influence a virtual assistant might have on their decisions is an indication that a human name has served as a real anthropomorphic cue to trigger superificial processing. This finding adds empirical evidence to existing literature suggesting that anthropomorphic cues serve as heuristics in the persuasion process.

### Conclusion

This study sheds light on how social cues used in virtual assistants can influence people’s concerns and persuasion levels by showing that both name and modality influence consumer responses. As virtual assistants play an increasingly important role across multiple everyday environments and contexts, blurring the boundary between the physical and the virtual, these findings hold important implications for future research, and for advertisers and consumer policies. The findings imply that assistants with a human name may be more effective in promoting products. This provides implications for advertisers when considering which virtual assistant platforms to use (e.g., Amazon’s Alexa or Google’s Assistant) or the naming strategy for their own agents (e.g., chatbots). It also provides insights for policymakers when it comes to the persuasive potential of these assistants. The findings highlight that modality matters for security concerns, with smart speakers creating the greatest concern. This suggests that the producers of smart speakers need to enhance security and be more transparent about their practices (e.g., see Lau et al.14). Consumer policymakers should also inform consumers about how they can protect themselves.

### Note

a. We also conducted three moderated mediation analyses (Model 7) for these variables but again no significant indirect effects were found.

### Author Disclosure Statement

No competing financial interests exist.

### Funding Information

This study was funded by the Amsterdam School of Communication Research, University of Amsterdam.

### References


10. Araujo T. Living up to the chatbot hype: the influence of anthropomorphic design cues and communicative agency framing on conversational agent and company perceptions. Computers in Human Behavior 2018; 85:183–189.


50. iCulture. This is how you prevent Siri from reacting to “hey Siri.” https://www.iculture.nl/tips/he-siri-uitschakelen-iphone-ipad-mac/ (accessed June 18, 2020).


Address correspondence to:
Dr. Hilde A.M. Voorveld
Amsterdam School of Communication Research
University of Amsterdam
P.O. Box 15793
Amsterdam 1001 NG
Netherlands

E-mail: h.a.m.voorveld@uva.nl