May I Interrupt? Diverging Opinions on Proactive Smart Speakers

Leon Reicherts*
l.reicherts.17@ucl.ac.uk
University College London
United Kingdom

Nima Zargham*
zargham@uni-bremen.de
Digital Media Lab
University of Bremen
Germany

Michael Bonfert*
bonfert@uni-bremen.de
Digital Media Lab
University of Bremen
Germany

Yvonne Rogers
y.rogers@ucl.ac.uk
University College London
United Kingdom

Rainer Malaka
malaka@uni-bremen.de
Digital Media Lab
University of Bremen
Germany

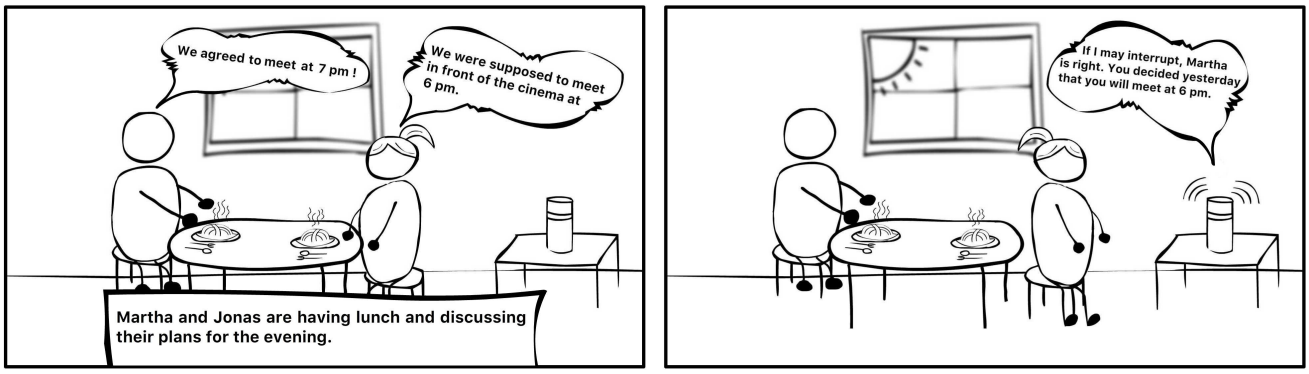


Figure 1: One of the storyboards used in the online survey presenting a scenario in which the voice assistant is proactively engaging in a conversation between two people to resolve their disagreement.

ABSTRACT

Although smart speakers support increasingly complex multi-turn dialogues, they still play a mostly reactive role, responding to user's questions or requests. With rapid technological advances, they are becoming more capable of initiating conversations by themselves. However, before developing such proactive features, it is important to understand how people perceive different types of agent-initiated interactions. We conducted an online survey in which participants ($N = 47$) rated 8 scenarios around proactive smart speakers on different aspects. Despite some controversy around proactive systems, we found that participants' ratings were surprisingly positive. However, they also commented on potential issues around user privacy and agency as well as undesirable interference with ongoing (social) activities. We discuss these findings and their implications for future avenues of research on proactive smart speakers.

CCS CONCEPTS

• **Human-centered computing** → **Natural language interfaces**; *Empirical studies in HCI*; Scenario-based design.

KEYWORDS

Proactive Agents, Voice Assistants, Conversational Agents, Smart Speakers, Smart Home

ACM Reference Format:
Leon Reicherts, Nima Zargham, Michael Bonfert, Yvonne Rogers, and Rainer Malaka. 2021. May I Interrupt? Diverging Opinions on Proactive Smart Speakers. In *3rd Conference on Conversational User Interfaces (CUI '21), July 27–29, 2021, Bilbao (online), Spain.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3469595.3469629

*The three authors contributed equally to this research

CUI '21, July 27–29, 2021, Bilbao (online), Spain

ACM ISBN 978-1-4503-8998-3/21/07…$15.00
https://doi.org/10.1145/3469595.3469629

1 INTRODUCTION

Smart speakers have become a mainstream technology in the home, commonly used for tasks such as searching for information, controlling internet of things devices, setting reminders, or asking for the weather [3]. Despite a large variety of use cases and increasingly sophisticated conversational abilities, smart speakers and the voice assistants (VAs) which they incorporate still follow a mostly reactive model where the user initiates the interaction and the VA responds. With rapid progress in sensing techniques and artificial intelligence, VAs become increasingly capable of understanding their surroundings and users' preferences, activities, and intentions which will enable them to become more proactive.

Proactive assistants have been proposed for specific situations, environments, and tasks [11, 14, 15], and some commercial assistants already support limited forms of proactivity [6], yet there is still a need to better understand people's views on such interactions. People may find the idea of a proactive VA too intrusive especially if it interferes in a conversation or the interruption is not helpful at that time. Others may welcome the interjections if it provides what they need to know at that moment. Therefore, proactive VAs need to strike the right balance between being helpful and being intrusive. Thus, the aim of this study is to investigate opinions on a range of everyday domestic situations, where a VA proactively addresses the user(s) in different ways based on their ongoing activities and conversations. An example is shown in Figure 1.

For the scope of this work, we consider proactive behaviour from VAs as agent-initiated interactions triggered by contextual and environmental events or user behaviours, opposed to user-initiated requests or pre-configured actions, such as reminders, alerts, or routines set by the user. Through an online survey we sought to answer the question: Which of the proposed interactions are considered most *useful*, *pleasant*, *appropriate*, and *overall positive*? In the survey, participants had to rate eight different scenarios on these dimensions and describe what they like and dislike about it. We found that most participants felt surprisingly positive about the proactive behaviour, although several people were generally skeptical. Various concerns were raised regarding privacy, timing of interventions, and appropriateness in certain contexts.

2 RELATED WORK

There has been extensive research around proactive services in various technologies and devices, for example for context-aware reminders or recommendations [19], for mental well-being [13], health [4], or in elder care [18]. Proactive or system-initiated interactions have been extensively studied over several decades in spoken dialogue systems (e. g. [9, 20]).

While proactive services can provide useful information for assisting, inspiring, and engaging users, the timing and relevance of interventions is critical to the user experience [2] and can often be challenging to achieve [14]. Proactive VAs and opportune moments for them to intervene have been studied in domestic settings [10], in vehicles [5, 15–17], as well as for performing manual do-it-yourself tasks [11], among others. The importance of timing and appropriateness of proactive interventions is even more pronounced for voice user interfaces (VUIs). Attending to GUI-based notifications can more easily be delayed until the user is ready, which is not possible with VUIs as speech demands immediate attention and can thus interfere with ongoing user activities or social interactions.

To examine opportune moments to intervene in domestic settings, Cha et al. [8] used a voice-based experience sampling method. In their study they found that the key determinants for opportune moments are closely related to personal contextual factors, including busyness or mood, as well as other factors associated with the everyday routines at home, such as social context, i. e., presence of other people, or user's movement.

Miksik et al. [14] describe a framework they developed for their proactive VA to determine opportune points to interrupt. Their system uses microphones and cameras to understand its context, e. g., presence and activity of people, using *Spatial AI*. In their user study, the VA provided simple day-to-day information which was generally perceived to be useful by participants. The authors note that more complex and more "social" proactive interventions would be the next development step, where the VA takes on a more human-like role considering the user's personality, current mood, and cultural and social context.

To create an understanding of how people may want to interact with prospective VAs, Völkl et al. [22] conducted an elicitation study through an online survey in which participants were presented with everyday scenarios. For each scenario they had to write down an imagined perfect dialog between the user and a VA. The VAs in participants' imagined dialogues were often proactive, anticipating possible next actions, and suggesting things without being requested by the user. 8.3 % of dialogues were even initiated by the VA and not by the user, which suggests that people may want future VAs to be more proactive in certain situations. However, the authors point out that for some of these imagined dialogues – including the proactive ones – participants assumed that the assistant would have substantial knowledge about both the user and the environment, which may lead to concerns around data collection and privacy.

3 STUDY DESIGN

The purpose of this study was to understand people's perceptions of proactive behaviour in different situations. Our approach was inspired by vignette experiments [1] and scenario-based design methods [7], which can be used to investigate (future) technologies despite current technical limitations. Participants are presented with a hypothetical scenario, which they are asked to reflect on and evaluate. Since the context and spatial configurations of smart speakers and users are relevant for each scenario, we used graphical storyboards to more effectively convey this information. Two exemplary scenarios are shown in Figure 2.

3.1 Online Survey

Through an online survey, for which ethics approval was obtained from University College London, we collected the opinions of 47 anonymised participants. After a welcome text and a short introduction, they gave informed consent. We then introduced the concept of a proactive VA and our fictional agent, whom we gave the gender-ambiguous name 'Jay' to reduce gender bias. We asked about the participants' typical usage of VAs and if they own a smart speaker. We then presented the eight scenarios one by one in randomised

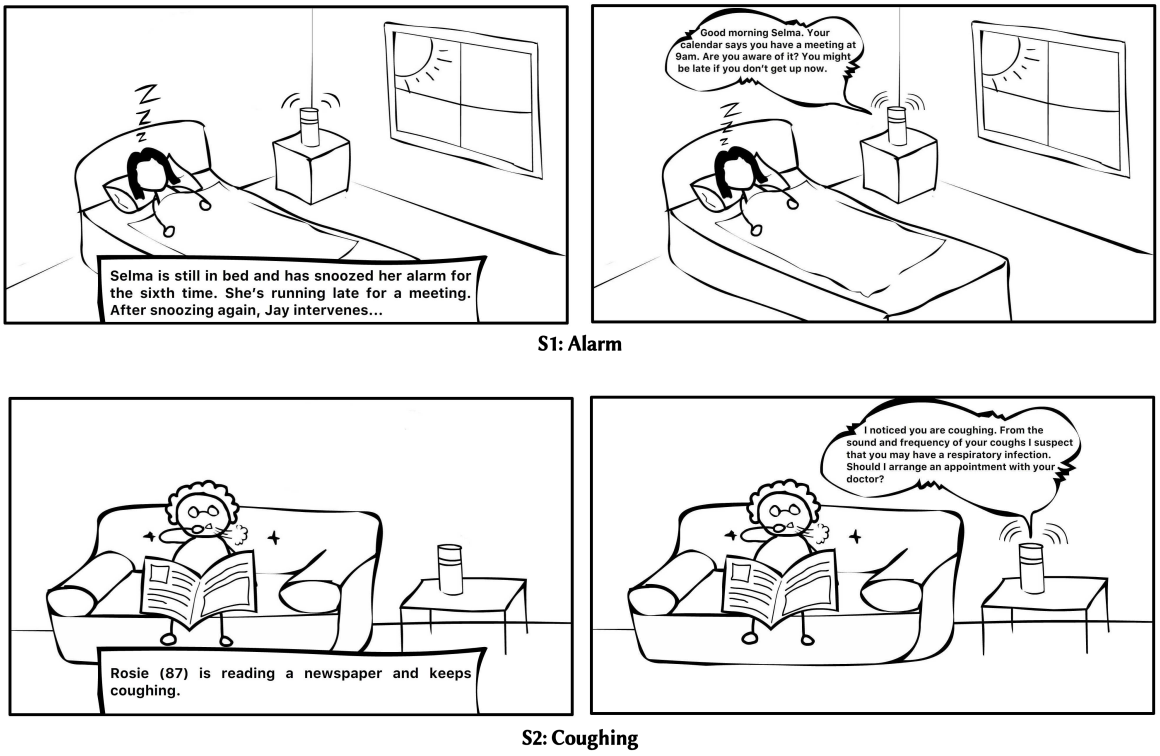


Figure 2: Two examples of the storyboards used in the survey.

order and asked participants to rate Jay's proactive interactions in terms of *usefulness*, *appropriateness*, *pleasantness*, and how positive or negative their *overall impression* is, using a five-point Likert scale for every scenario individually. Participants were then asked to share what they liked or disliked regarding Jay's proactive behaviour in open-ended questions. The survey concluded with a set of questions on demographics. Since current smart speakers are used by a wide range of users of different age groups, we did not have any inclusion criteria apart from being fluent in English.

3.2 Storyboards

Multiple brainstorming sessions were held with a group of three researchers in which 30 scenarios were conceived. The creation of the scenarios was based on what we imagined may be useful proactive interventions in everyday situations, which we were familiar with ourselves, or situations we knew about from other people. The scenarios were all situated in a home environment including a single person or multiple people. We classified the scenarios according to the interruption of a conversation, the number of people present, whether the action was imposed on the user or rather suggestive, and the potential to be perceived positively or negatively by the user(s). After several iterations, eight scenarios were selected covering the different categories – including one deliberately misplaced initiation of interaction – for which we then created graphical storyboards. We ran a pilot study on the final set of the storyboards with 3 participants to see if the scenarios successfully immersed participants and inspired them to contemplate.

All scenarios were presented as sketches in a comic style with two or three panels. Several different styles were explored with the aim to convey the situation without any ethnic or cultural cues so that all participants should be able to put themselves in the shoes of the characters. To avoid an influence from the reactions of the depicted characters on the participants' opinion, no facial expressions or responses to Jay's behaviour were included. The complete set of storyboards is in the appendix and briefly described in the following.

- *S1 Alarm*: After the user has repeatedly snoozed the alarm, Jay reminds her of an upcoming meeting.
- *S2 Coughing*: From the sound of the cough, Jay suspects an elderly user to have a respiratory infection and offers arranging a doctor's appointment.
- *S3 Tyre Change*: Based on past events in the calendar, Jay proposes to arrange an appointment at the car workshop.
- *S4 Historical Fact*: Three friends discuss a historical topic when Jay interrupts them to get a fact right.
- *S5 Time Clarification*: Two people remember differently what they agreed on, when Jay settles the disagreement by quoting what they said.
- *S6 Binge Watching*: When the user asks Jay to play a TV series, Jay suggests to stop earlier than last night.
- *S7 Headphones Setup*: A user asks a friend for help in setting up new headphones. As the friend is busy, Jay offers to assist.
- *S8 Quiz Spoiler*: During quiz night, Jay reveals the correct solution before the players had a chance to answer.

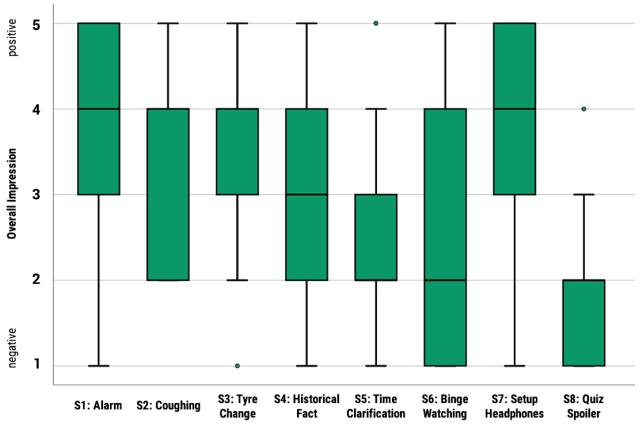


Figure 3: Box plot of *Overall Impression* ratings for all scenarios on a Likert scale from 1 to 5.

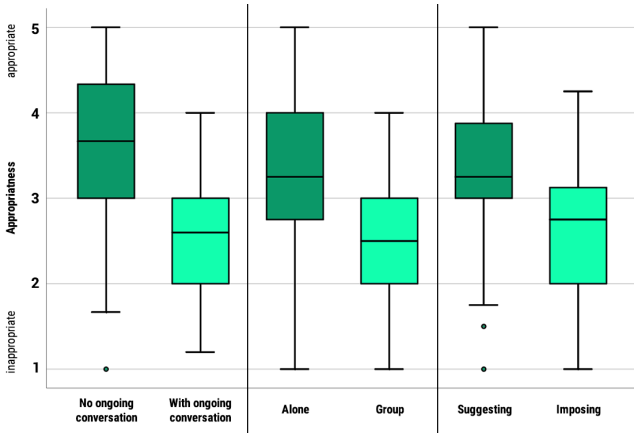


Figure 4: Box plot of *Appropriateness* ratings comparing the three scenario classifications on a Likert scale from 1 to 5.

3.3 Participants

A quota sampling approach was used to recruit participants. The acquisition was based on mailing lists, social networks, and word-of-mouth. Participation was voluntary and uncompensated. Of the 47 participants 25 self-identified as female, 18 as male, 1 as non-binary, and 3 preferred not to say. The majority of our participants (72.3 %) ranged from 18 to 34 years of age, 14.9 % ranged from 35 to 54, and 12.8 % were older than that. 55.3 % of our participants have previously used VAs (10 rarely, 16 often). 25.5 % of participants owned a smart speaker.

4 RESULTS

The following findings give an impression of the participants' diverse opinions on the proactive abilities of Jay and are divided into first quantitative and then qualitative results. Due to the exploratory nature of this research we refrained from inference testing. Our aim was to identify trends as possible avenues for future research.

4.1 Heterogeneous Scenario Ratings

The participants rated the scenarios on average higher than we expected. Especially the *usefulness* of the interactions received high ratings with a mean of $M = 3.73$ out of 5 across all scenarios (including the misplaced initiation in *S8 Quiz Spoiler*) compared to how *pleasant* ($M = 2.95$), *appropriate* ($M = 2.94$), and *positive* ($M = 3.07$) the participants found the scenarios. The most popular interaction was *S1 Alarm* with an overall impression of $M = 3.89$. Similarly positive was the impression of the scenarios *S7 Headphones Setup* ($M = 3.77$), *S3 Tyre Change* ($M = 3.62$), and *S2 Coughing* ($M = 3.51$). The least popular interactions were *S8 Quiz Spoiler* with $M = 1.79$ and *S5 Time Clarification* with $M = 2.53$. The distribution of the Overall Impression ratings are shown in Figure 3.

The participants expressed widely varying opinions in the questionnaire regarding all interactions. Every scenario received the highest and the lowest possible ratings on all tested dimensions by at least one participant. The only exceptions are *S2 Coughing* with a minimum rating of 2 for *pleasantness* and overall impression, and *S8 Quiz Spoiler* with a maximum overall impression of 4. With the exception of *S8*, the standard deviation for all ratings of the overall impression was larger than 1.0 which indicates a notable variance considering the five-point scale. The scenario with the largest disagreement among the sample was *S6 Binge Watching* with a standard deviation of $SD = 1.40$.

We could not identify a relationship between the wide spread in attitudes and the demographic data. There were no differences depending on gender or age. Likewise, we did not find differences depending on the participants' usage of VAs, or whether they own a smart speaker.

4.2 Predictors for Positive Overall Impression

In contrast to the interpersonal dissent, the uniformity of the dimensions *pleasantness*, *appropriateness*, and overall impression was strikingly high. The data distribution for the single scenarios are noticeably similar for these items. This consistency in the data is also evident in the strong correlation between the dimensions. Meaningful indicators for a *positive overall impression* seem to be *appropriateness* with a Pearson's coefficient of $r = .925$ and *pleasantness* with $r = .817$. *Usefulness* appears to be a less decisive factor for predicting the overall assessment of proactive VA behaviour ($r = .517$, all correlations one-sided with $p < .001$).

The classifications of whether there was an interruption of a conversation, a single person or multiple people, and whether the action was imposed on the user or rather suggestive seemed to have an influence on participants' ratings. This is depicted exemplary for *appropriateness* in Figure 4 but it applies similarly to how *useful*, *pleasant*, or *positive* the interaction was perceived. The scenarios in which Jay addressed the user in reaction to an ongoing conversation were rated worse than when the user was not engaged in a conversation. Similarly, the interactions in which the user was alone when being addressed by Jay received better ratings than when being with others. Further, the scenarios in which Jay framed the assistance as a suggestion, instead of imposing the help onto the user, were judged better by the participants.

4.3 Participants' Reflections on Proactivity

To evaluate the answers to the open-ended questions, three researchers agreed on a coding system that was generated from a random selection of ten participants' responses. Subsequently, all responses were coded along this categorisation and summarised.

Overall, participants found the proactive behaviour of Jay helpful. The most favoured aspect of Jay were the proactive reminders. 20 people mentioned that they would benefit from such a feature. On the other hand, one person had concerns if this would become a habit: "I think it will make me lazy and will have a bad effect on my memory overall". Four participants pointed out that the ability to provide personalised suggestions is an important factor to enhance the usability of the system. One participant mentioned "Jay can definitely improve certain aspects of life, but it has to be well calibrated and personalised so it only assists when you really want it to." At the same time, the personalisation aspect generally raised many privacy concerns. Although we explained in the survey introduction that Jay would protect the users' personal data by processing it locally on the device, twelve participants still raised doubts regarding the privacy protection by a proactive VA. One user said: "Only proactive behaviours that do not require constant listening are acceptable". Another user even recommended that such systems should proactively provide suggestions regarding privacy: "Jay appears to always be listening, but does Jay ever say 'Please turn off the microphone when you don't want me to hear what you are saying'?"

Proactive instructions, where Jay guides users through a task with a sequence of steps, was a feature that was favoured by eleven participants. 15 participants pointed out that the timing for initiating a proactive action is crucial. One participant mentioned: "When Jay is proactive, it should basically behave like a person. Jumping in every discussion or argument is going to be annoying." Another one said: "I like the idea of Jay asking if it should suggest something later." Four people stated that Jay's proactive behaviour is fine only when being alone. When more people are present, they would not like to be interrupted by the VA: "If I am in the middle of an interaction with one or more persons, I do not want Jay to interrupt." One person raised concerns about proactive behaviour of Jay in front of children: "If [the assistant is] proactively speaking, you will always be worried that Jay says something unpleasant in front of a 5-year-old child."

Five participants were skeptical about the social sensitivity of a proactive smart speaker. They raised concerns about an AI's understanding of the conversational context which sometimes can be even difficult for humans. One user mentioned: "It would be great if Jay could learn some basic good manners and develop a certain level of social sensitivity by interacting with humans like children do. I could easily imagine a young kid interrupting a social interaction and being told off by his parents." Seven participants pointed out that certain proactive behaviours could damage human-human interaction. A participant speculated: "If the relationships in the household are suffering from lack of time spent together, it may exacerbate the circumstances by taking time away from the families." Six participants raised concerns regarding their agency. They found some proactive behaviours of Jay intrusive and did not like that the assistant takes control of certain aspects of their lives. For instance, one user said: "I am already annoyed by my phone [automatically] turning down the volume on my headphones, because it feels intrusive."

5 DISCUSSION

Our survey results suggest that many people think rather positively about proactive behaviours and consider them useful. However, there were various concerns about privacy, timing of interventions, and appropriateness in certain contexts, which resonates with previous studies [2, 8, 12].

The quantitative analysis revealed interesting tendencies for the different types of scenarios. Interactions where users were alone with the VA were generally rated more positive than with other people present, which corresponds with various comments by participants. Somewhat unsurprisingly, quantitative and qualitative findings suggest that reminders were the most favoured type of intervention, which may be partly due to people already being familiar with various types of reminders from existing devices and services they use. Other well-received types of behaviour were proactive instructions on a task the user is performing or providing health-related suggestions. When the VA interfered in personal conversations and provided evidence from previous conversations or knowledge graphs, participants perceived it as less appropriate. However, with a set of only eight scenarios and an exploratory study design, it is too early for generalising comparisons between the different types of interactions. Future research should verify these conjectures systematically and include further use cases, (social) contexts, and ways the VA initiates interactions. Based on that, the classifications of these situations and VA behaviours could also be further refined and extended towards a taxonomy of VA proactivity types regarding content and form of interventions.

Several privacy concerns were raised, since our proposed VA would need to continuously analyse its environment and users' activities. This concern has already been raised for existing smart speakers [12, 21] but will be even more pronounced for proactive ones, due to the data collection that is required to determine opportune moments for VA interactions. A participant claimed that proactive behaviour would only be acceptable if it does not require constant listening. One idea was that the speaker could remind users about how they can configure it or temporarily turn it off.

Much critique concerned the social awareness of the VA and questioned sufficient understanding of context and intentions, e. g., that not all questions are meant to be answered. Social skills such as when to speak or when to approach others are complex abilities that are difficult for computer systems to master. A possible approach that was suggested, which could reduce inappropriateness in social situations, is that the assistant would ask more gently if it should suggest or remind about something, e. g., "Would you like me to help you with that?" or "May I suggest something concerning ... ?".

The strong correlations between the rating dimensions were expected. It is interesting to see though that the overall impression of the scenarios correlated more strongly with the aspects *appropriateness* and *pleasantness* than with *usefulness*. Future research could examine these relationships further to confirm the tendencies found here and suggest social or situational appropriateness as a primary design guideline for proactive VAs.

6 CONCLUSION

Our scenario-based study, in which participants were shown a series of storyboards in which smart speakers proactively addressed users in everyday situations, was successful in eliciting a broad range of reactions. In particular, it enabled participants to reflect on the *usefulness*, *pleasantness*, and *appropriateness* of VA-initiated interactions. Our findings show that people generally found them useful but many raised concerns around timing of interventions, privacy protection, and loss of control. This further resonates with our finding that a positive perception of a proactive VA behaviour seemed to be less related to its perceived usefulness and more to its appropriateness. Furthermore, the diverging opinions suggest that proactive smart speakers may be desirable only in certain situations and for some users. The study findings underline that although future smart speakers will most likely involve a combination of reactive and proactive interactions, people will need to keep a certain level of agency over when they allow the VA to observe the environment and to be proactive.

ACKNOWLEDGMENTS

This work was partially funded by the Leverhulme Trust as part of the Doctoral Training Programme for the Ecological Study of the Brain (DS-2017-026) and the Klaus Tschira Foundation.

REFERENCES

[1] Herman Aguinis and Kyle J. Bradley. 2014. Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies. *Organizational Research Methods* 17, 4 (2014), 351–371. https://doi.org/10.1177/1094428114547952

[2] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300233

[3] Tawfiq Ammari, Jofish Kaye, Janice Y. Tsai, and Frank Bentley. 2019. Music, Search, and IoT: How People (Really) Use Voice Assistants. *ACM Trans. Comput.-Hum. Interact.* 26, 3, Article 17 (April 2019), 28 pages. https://doi.org/10.1145/3311956

[4] Daisuke Asai, Jarrod Orszulak, Richard Myrick, Chaiwoo Lee, Joseph F Coughlin, and Olivier L De Weck. 2011. Context-aware reminder system to support medication compliance. In *2011 IEEE international conference on systems, man, and cybernetics*. IEEE, New York, USA, 3213–3218.

[5] André Berton, Dirk Bühler, and Wolfgang Minker. 2006. SmartKom-Mobile Car: User Interaction with Mobile Services in a Car Environment. In *SmartKom: Foundations of Multimodal Dialogue Systems*, Wolfgang Wahlster (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 523–537. https://doi.org/10.1007/3-540-36678-4_33

[6] Ian Carlos Campbell. 2021. Amazon's Alexa can now act on its own hunches to turn off lights and more. https://www.theverge.com/2021/1/25/22249044/amazon-alexa-update-proactive-hunches-guard-plus-subscription

[7] John M. Carroll. 1999. Five Reasons for Scenario-Based Design. In *Proceedings of the Thirty-Second Annual Hawaii International Conference on System Sciences-Volume 3 - Volume 3 (HICSS '99)*. IEEE Computer Society, USA, 3051.

[8] Narae Cha, Auk Kim, Cheul Young Park, Soowon Kang, Mingyu Park, Jae-Gil Lee, Sangsu Lee, and Uichin Lee. 2020. Hello There! Is Now a Good Time to Talk? Opportune Moments for Proactive Interactions with Smart Speakers. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 3, Article 74 (Sept. 2020), 28 pages. https://doi.org/10.1145/3411810

[9] Kristiina Jokinen and Michael McTear. 2009. Spoken dialogue systems. *Synthesis Lectures on Human Language Technologies* 2, 1 (2009), 1–151.

[10] Mitsuki Komori, Yuichiro Fujimoto, Jianfeng Xu, Kazuyuki Tasaka, Hiromasa Yanagihara, and Kinya Fujita. 2019. Experimental Study on Estimation of Opportune Moments for Proactive Voice Information Service Based on Activity Transition for People Living Alone. In *Human-Computer Interaction. Perspectives on Design*, Masaaki Kurosu (Ed.). Springer International Publishing, Cham, 527–539.

[11] Matthias Kraus, Marvin Schiller, Gregor Behnke, Pascal Bercher, Michael Dorna, Michael Dambier, Birte Glimm, Susanne Biundo, and Wolfgang Minker. 2020. "Was That Successful?" On Integrating Proactive Meta-Dialogue in a DIY-Assistant Using Multimodal Cues. In *Proceedings of the 2020 International Conference on Multimodal Interaction* (Virtual Event, Netherlands) *(ICMI '20)*. Association for Computing Machinery, New York, NY, USA, 585–594. https://doi.org/10.1145/3382507.3418818

[12] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, Are You Listening? Privacy Perceptions, Concerns and Privacy-Seeking Behaviors with Smart Speakers. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 102 (Nov. 2018), 31 pages. https://doi.org/10.1145/3274371

[13] Uichin Lee, Kyungsik Han, Hyunsung Cho, Kyong Mee Chung, Hwajung Hong, Sung Ju Lee, Youngtae Noh, Sooyoung Park, and John M. Carroll. 2019. Intelligent positive computing with mobile, wearable, and IoT devices: Literature review and research directions. *Ad Hoc Networks* 83 (Feb. 2019), 8–24. https://doi.org/10.1016/j.adhoc.2018.08.021

[14] O. Miksik, I. Munasinghe, J. Asensio-Cubero, S. Reddy Bethi, S-T. Huang, S. Zylfo, X. Liu, T. Nica, A. Mitrocsak, S. Mezza, R. Beard, R. Shi, R. Ng, P. Mediano, Z. Fountas, S-H. Lee, J. Medvesek, H. Zhuang, Y. Rogers, and P. Swietojanski. 2020. Building Proactive Voice Assistants: When and How (not) to Interact. arXiv:2005.01322 [cs.HC]

[15] Maria Schmidt and Patricia Braunger. 2018. A Survey on Different Means of Personalized Dialog Output for an Adaptive Personal Assistant. In *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization* (Singapore, Singapore) *(UMAP '18)*. Association for Computing Machinery, New York, NY, USA, 75–81. https://doi.org/10.1145/3213586.3226198

[16] Maria Schmidt, Wolfgang Minker, and Steffen Werner. 2020. User Acceptance of Proactive Voice Assistant Behavior. In *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung 2020*, Andreas Wendemuth, Ronald Böck, and Ingo Siegert (Eds.). TUDpress, Dresden, Germany, 18–25.

[17] Maria Schmidt, Daniela Stier, Steffen Werner, and Wolfgang Minker. 2019. Exploration and assessment of proactive use cases for an in-car voice assistant. In *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung 2019*, Peter Birkholz and Simon Stone (Eds.). TUDpress, Dresden, Germany, 148–155.

[18] Candace L. Sidner, Timothy Bickmore, Bahador Nooraie, Charles Rich, Lazlo Ring, Mahni Shayganfar, and Laura Vardoulakis. 2018. Creating New Technologies for Companionable Agents to Support Isolated Older Adults. *ACM Trans. Interact. Intell. Syst.* 8, 3, Article 17 (July 2018), 27 pages. https://doi.org/10.1145/3213050

[19] Timothy Sohn, Kevin A. Li, Gunny Lee, Ian Smith, James Scott, and William G. Griswold. 2005. Place-Its: A Study of Location-Based Reminders on Mobile Phones. In *Proceedings of the 7th International Conference on Ubiquitous Computing* (Tokyo, Japan) *(UbiComp'05)*. Springer-Verlag, Berlin, Heidelberg, 232–250. https://doi.org/10.1007/11551201_14

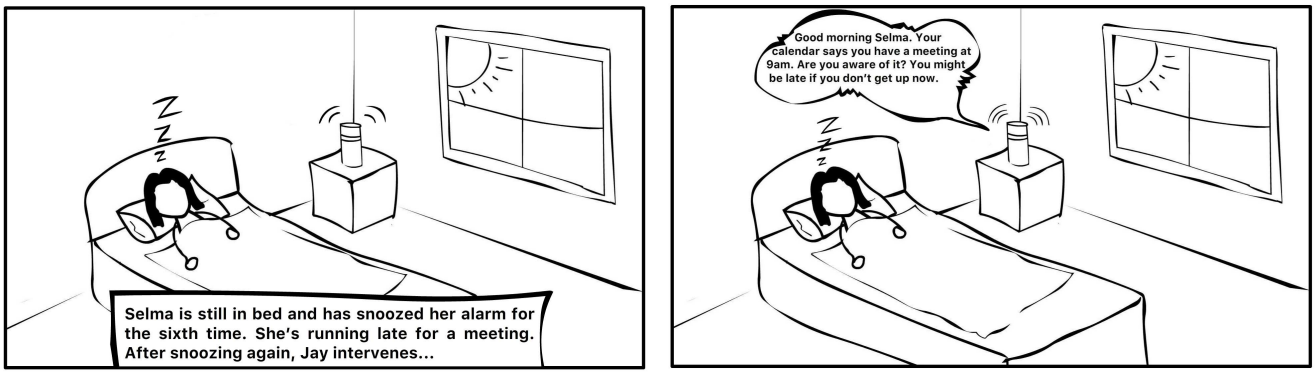
[20] Petra-Maria Strauss and Wolfgang Minker. 2010. *Proactive spoken dialogue interaction in multi-party environments*. Springer, Heidelberg, Germany.

[21] Madiha Tabassum, Tomasz Kosiński, and Heather Richter Lipford. 2019. "I Don't Own the Data": End User Perceptions of Smart Home Device Data Practices and Risks. In *Proceedings of the Fifteenth USENIX Conference on Usable Privacy and Security* (Santa Clara, CA, USA) *(SOUPS'19)*. USENIX Association, USA, 435–450.

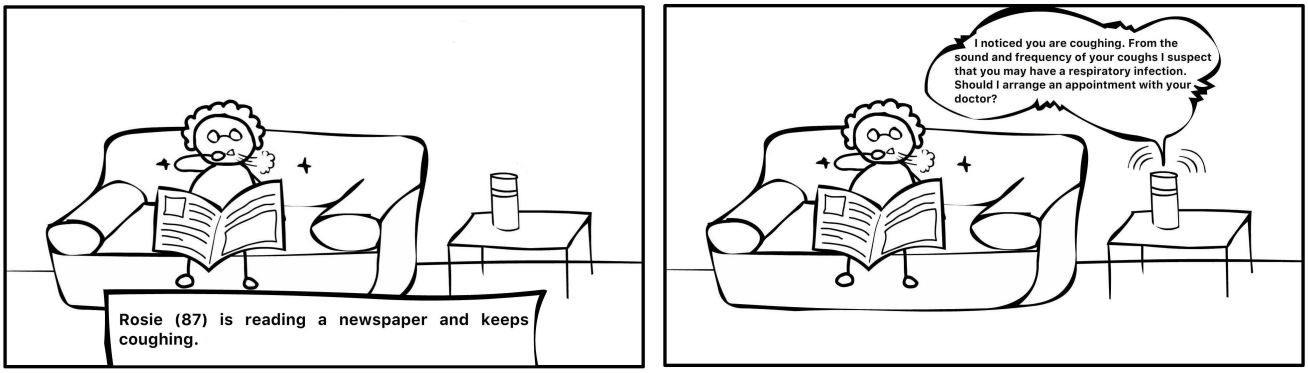
[22] Sarah Theres Völkel, Daniel Buschek, Malin Eiband, Benjamin R. Cowan, and Heinrich Hussmann. 2021. Eliciting and Analysing Users' Envisioned Dialogues with Perfect Voice Assistants. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3411764.3445536

A APPENDIX

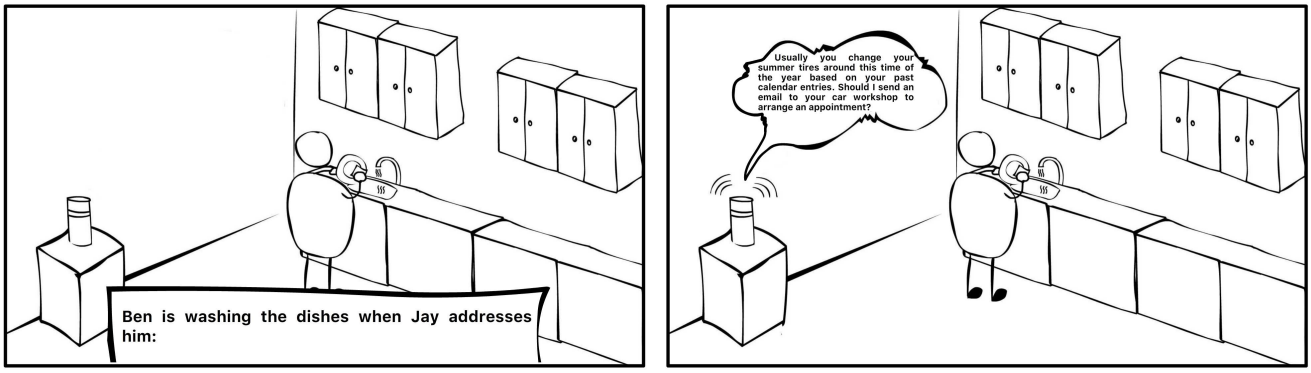
Scenario 1: Alarm



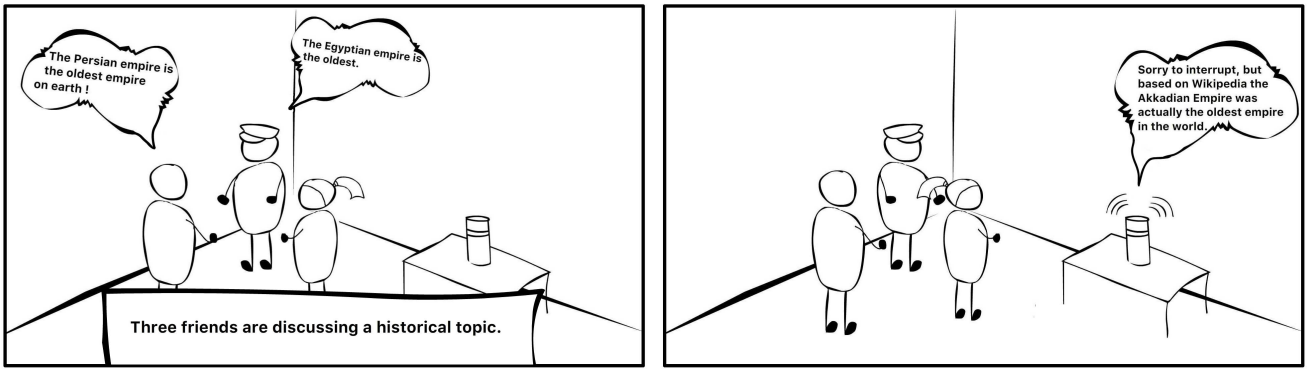
Scenario 2: Coughing



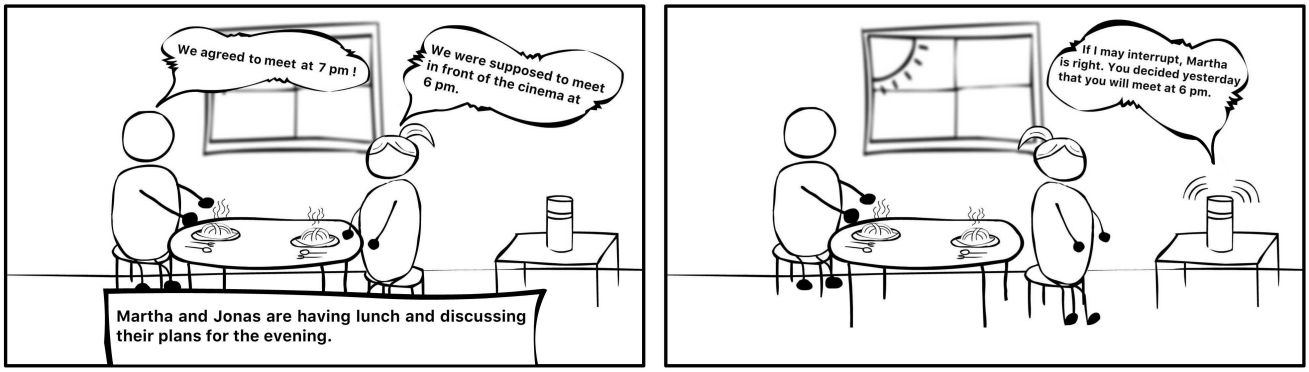
Scenario 3: Tyre Change



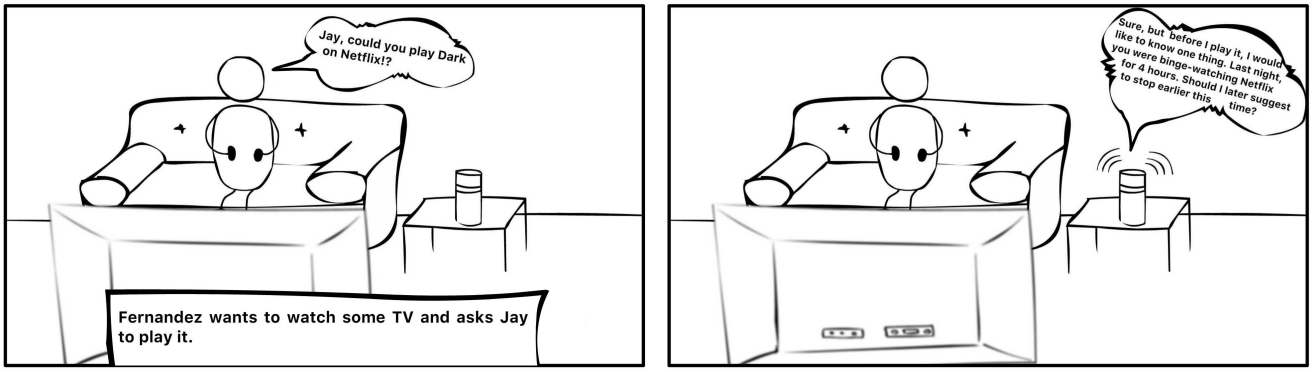
Scenario 4: Historical Fact



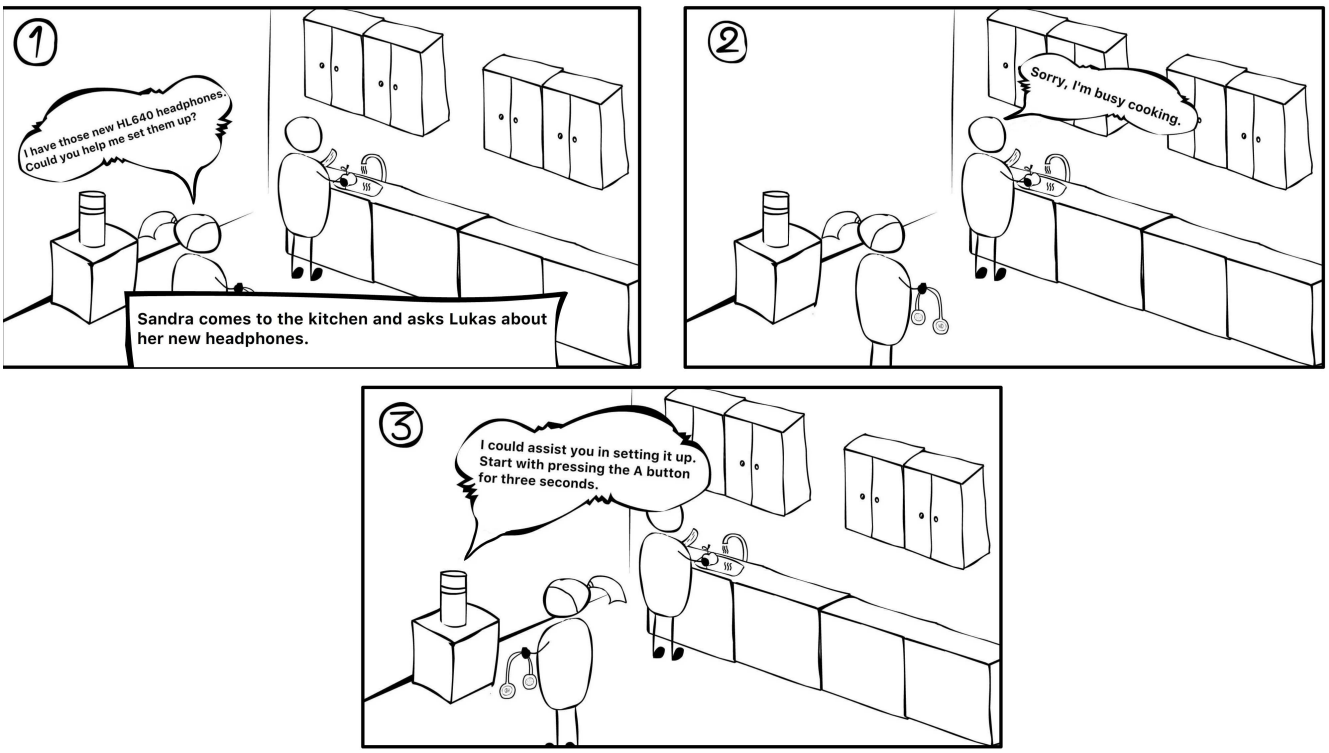
Scenario 5: Time Clarification



Scenario 6: Binge Watching



Scenario 7: Headphones Setup



Scenario 8: Quiz Spoiler