



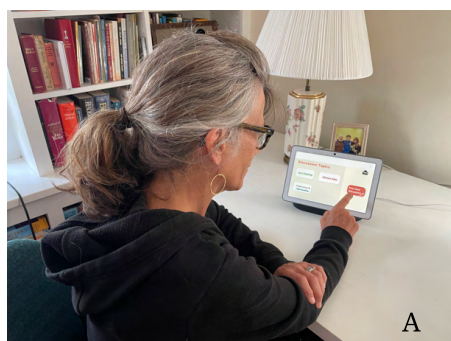
Polite or Direct? Conversation Design of a Smart Display for Older Adults Based on Politeness Theory

Yaxin Hu, Yuxiao Qu, Adam Maus, and Bilge Mutlu

University of Wisconsin–Madison

Madison, WI, USA

{yaxin.hu,qu45,amaus,bmutlu}@wisc.edu



A



Polite System

B



Direct System

C

Figure 1: In this paper, we applied politeness strategies to seven speech acts of a smart display, identified four personas representing older adults' experience of system usage and preferences for politeness, and suggested personalized design strategies for each persona. (A) An older adult using the conversational interface; (B) an example of *polite* speech used by our interface; and (C) an example of *direct* speech used by our interface.

ABSTRACT

Conversational interfaces increasingly rely on human-like dialogue to offer a natural experience. However, relying on dialogue involving multiple exchanges for even simple tasks can overburden users, particularly older adults. In this paper, we explored the use of politeness theory in conversation design to alleviate this burden and improve user experience. To achieve this goal, we categorized the voice interaction offered by a smart display application designed for older adults into seven major speech acts: request, suggest, instruct, comment, welcome, farewell, and repair. We identified face needs for each speech act, applied politeness strategies that best address these needs, and tested the ability of these strategies to shape the perceived politeness of a voice assistant in an online study ($n = 64$). Based on the findings of this study, we designed *direct* and *polite* versions of the system and conducted a field study ($n = 15$) in which participants used each of the versions for five days at their homes. Based on five factors merged from our qualitative findings, we identified four distinctive user personas—*socially oriented follower*,

socially oriented leader, *utility oriented follower*, and *utility oriented leader*—that can inform personalized design of smart displays.

CCS CONCEPTS

• **Human-centered computing** → **Sound-based input / output; User centered design; Interaction design theory, concepts and paradigms.**

KEYWORDS

Conversation design, politeness theory, multimodal interaction, smart displays, older adults

ACM Reference Format:

Yaxin Hu, Yuxiao Qu, Adam Maus, and Bilge Mutlu. 2022. Polite or Direct? Conversation Design of a Smart Display for Older Adults Based on Politeness Theory. In *CHI Conference on Human Factors in Computing Systems (CHI '22)*, April 29–May 5, 2022, New Orleans, LA, USA. ACM, New York, NY, USA, 15 pages. <https://doi.org/10.1145/3491102.3517525>

1 INTRODUCTION

Conversational devices are increasingly being adopted in day-to-day environments, and the interaction design of the conversational behaviors of these devices is key to improving user experience with the services delivered through them. However, the design space for these behaviors involves competing objectives with significant design tradeoffs. For example, on the one hand, a major focus of conversational design guidelines proposed in prior research is the *efficiency* of interaction [22]. Researchers have suggested that efficiency can be achieved by maintaining minimal information

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI '22, April 29–May 5, 2022, New Orleans, LA, USA

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-9157-3/22/04...\$15.00

<https://doi.org/10.1145/3491102.3517525>

exchange in short utterances and making progress in the dialogue through follow-up questions [8, 32, 39]. On the other hand, recent work has proposed the need for proactive behavior in voice interaction design and for conversational systems to play a more active role in engaging users [4, 6, 38]. This work has demonstrated multiple rounds of conversational exchanges, directed by either the user and the device, to increase the perceived agency of the system. Designers who wish to achieve a level of conversational efficiency while following conversational norms must adopt design principles that balance both objectives of efficiency and conversational norms, although little is known about how such balance might be achieved. In this paper, we explore how *linguistic politeness strategies* might inform the design of conversational interfaces toward achieving this design goal.

Our work focuses on older adults as the target user group as many of them face motor and vision issues, and conversational interfaces hold great potential to address these accessibility challenges. The main premise of our work is that polite speech can meet expectations of conversational norms while maintaining a level of conversational efficiency. Therefore, we explore how a conversational interface can adopt a *direct* speech style as opposed to a *polite* speech style in its interactions with users. The notions of *directness* and *politeness* in language have been characterized by two linguistic models: Grice’s Maxims [9] and Politeness Theory [3]. The *manner* dimension of Grice’s Maxims emphasizes clarity and brevity for effective information exchange, whereas Politeness Theory considers people’s *face* needs, i.e., the public self-image that people want to claim, sometimes at the cost of efficiency. Four strategies are proposed in Politeness Theory: *bald-on-record*, *positive politeness*, *negative politeness*, and *off-record*. Bald-on-record follows Grice’s Maxims of manner and is characterized by directness. Politeness-based strategies, on the other hand, consider people’s social needs and attend to their face needs, i.e., the desire to be appreciated and approved (positive face) and the desire to be free from imposition and distraction (negative face). Our work adopts these constructs to establish *polite* and *direct* speech for the conversational interface of a smart display application designed for older adults.

To explore the use of these linguistic politeness strategies to achieve directness and politeness in conversation design, we first categorized dialogue between the user and the conversational interface into a set of speech acts. In our smart display application, seven speech acts were identified: *request*, *suggest*, *instruct*, *comment*, *welcome*, *farewell*, and *repair*. Next, we determined conversational goals for each speech act in terms of the user’s face needs. For example, the user’s face needs may be desirable when they hear the system’s *welcome* speech act, while the needs for *request* may be to minimize imposition by the conversational interface. We selected politeness strategies that meet the face needs of users in the context of each speech act.

We conducted an online formative study ($n = 64$) to assess the ability of our strategies to shape the perceived directness and politeness of the voice assistant. Based on the findings from the formative study, we implemented two systems: a *polite* version that utilized positive and negative politeness strategies, and a *direct* version that utilized the bald-on-record strategy. We conducted a field study ($n = 15$) to compare user experience and usage patterns across the two versions of the system. The findings revealed five

factors affecting user preferences, experience, and interest in long-term adoption of the system, which led to the development of four distinctive user personas. We discuss the design implications of these factors and personas and the potential that politeness holds for the design of conversational interfaces. The contributions of our work include:

- (1) *Design Space Exploration* — We categorized voice interactions into speech acts and applied linguistic politeness strategies to the dialogue behaviors of a smart display application designed for older adults.
- (2) *Empirical Findings* — The field study provided insights into the experience and preferences of older adults in using the smart display; their preferences regarding the features and interaction modalities offered by the smart display; the challenges they faced in the interaction; and their perceptions of the conversational agent.
- (3) *Design Recommendations* — The empirical findings point toward relational, social, and cultural considerations as well as implementation constraints in using politeness for conversation design. The design implications of our work also highlight the importance of personalization in the use of proactive behavior, interaction pace, and the interplay between the visual and speech-based interaction.

2 RELATED WORK

2.1 Conversation Design for Voice Interfaces

Prior research has proposed design guidelines and heuristics for voice user interface design with reference to factors in human’s natural conversations. Grounded in *Conversational Analysis*, Moore and Arar [32] suggested borrowing principles of how people design their dialogues with considering the three characteristics: *Recipient Design*, *Minimization*, and *Repair*. Drawing inspiration from Grice’s Cooperative Principle in pragmatic theory and Nielsen’s usability heuristics for user interface design, Langevin et al. [22] developed a set of usability heuristics to guide and evaluate the design of conversational agents. Wei and Landay [39] also suggested that the system “[a]dapt agent style to who users are, how they speak, and how they are feeling” in the heuristics. Despite the efforts towards building conversational agents in HCI research, the debate has been ongoing on whether voice interfaces should be conversational, natural or human-like [5, 13, 19, 29].

Besides design guidelines, prior work has studied user needs for voice interfaces. Yang and Aurisicchio [40] explored user needs of competence, autonomy and relatedness applying Self-Determination theory and suggested three aspects that reflect user autonomy: control over the conversation, personalized experience and control over their data. Past research has identified two major purposes of voice assistant usage—social and transactional—and suggested different design strategies for each purpose [17, 19, 40]. Prior work also revealed differences between people who have utilitarian orientation and those with relational orientation in their responses to a service robot’s politeness during malfunction [25].

2.2 Linguistic Strategies in Conversation Design

Prior work has investigated the use of linguistic strategies in conversational interfaces and robots. Andrist et al. [1] used linguistic cues to increase the perceived expertise of an informational robot. Lockshin and Williams [27] suggested applying linguistic norms in human dialogue to autonomous agents, identified three key norms: directness, brevity, and politeness, and studied people’s adherence to these norms under different contexts. Particularly, politeness as an linguistic strategy has been widely studied for autonomous agents, e.g., educational agent [15, 31], recommendation systems for elderly care [12], repair for a malfunctioning service robot [25], and help-seeking for a robot [36].

An agent employing politeness can affect a user’s perception of the agent, their sense of autonomy, and whether they think interacting with an agent is pleasurable. Lee et al. [24] found that robots in a healthcare setting were more likely to improve patient compliance when using polite behavior in their speech and gestures. Kim and Lim [16] found that a user’s initial mental model of the agent affected what information the user was willing to share and how they expected the information to be utilized. Kim et al. [17] found that user preferences to disclose information changed based on how they perceived a chatbot that acted more task-oriented or more companion-like.

2.3 Voice interfaces for older adults

As the population of older adults grows worldwide, aging technologies are increasingly researched and built to improve senior people’s quality of life. However, older adults often face challenges in using technology because of declines in cognitive (e.g., reduced memory), perceptual (e.g., hearing or vision), and motor (e.g., reduced dexterity) skills [33]. Smart devices such as voice assistants and smart displays could address some of these challenges by providing older adults with multiple ways to interact with the device. Research has explored how older adults perceive and interact with voice assistants and has shown that they find voice assistants useful [34, 37]. However, older adults still face challenges using voice assistants because older adults don’t fully understand how voice assistants work and/or how to use them [18, 21, 28, 35]. Likewise, voice assistants raise concerns about privacy and data collection due to their microphones continuously listening for wake words to activate the device [18, 23, 30, 37]. Other issues, not specific to older adults, involve problems interacting with the voice interface. Users often encounter issues due to the device not activating, issues with the device performing the command, or the voice assistant misunderstanding the user’s dialect or phrasing of a command. During an error, voice assistants may recover by giving a non-answer forcing the user to try the command or series of commands again [42].

3 CONVERSATION DESIGN

To explore how linguistic politeness theory might inform the design of conversational interfaces that improve user experience while maintaining efficiency, we identified salient and repeated speech acts and matched politeness strategies to each act. We implemented these strategies into the conversational speech of a smart display application designed for older adults and conducted a formative

study to assess the extent to which our designs reflected politeness. The paragraphs below provide further design on our conversation design approach.

3.1 Linguistic Strategies

Language use is concerned with “what is said” and “how what is said to be said” [9]. Grice’s Maxims describe “how” as the *manner* dimension of cooperation, which emphasizes clarity, brevity, and order with the goal of effective information exchange. The Cooperative Principle, which includes manner, has been widely adopted by researchers and designers to design voice interfaces where brevity is encouraged to achieve the maximum efficiency [22]. On the other hand, Brown and Levinson [3] proposed Politeness Theory as a deviation from Grice’s Maxims. According to Politeness Theory, language usage has social meanings, and thus message construction is strategically used to express social relationships. Four strategies in conversation consider social needs: *bald-on-record*, *positive politeness*, *negative politeness*, and *off-record*. Bald-on-record follows Grice’s Maxim of manner and is the most direct, clear, concise, and unambiguous way to communicate, whereas the politeness-based strategies conveyed messages in a more indirect way.

Based on the two linguistic models, we designed a *direct* and a *polite* system. The direct system follows the manner from Grice’s Maxims and the bald-on-record strategy from Politeness Theory. The polite system used linguistic strategies that minimized face-threatening acts, particularly the *positive politeness strategy* that aims to promote the user’s positive self-image and the *negative politeness strategy* that focuses on the user’s freedom of action. We categorized the system’s utterances into seven speech acts: request, comment, suggest, repair, instruct, welcome, farewell.

To apply politeness strategies in a systematic and principled fashion, we first identified face needs for each speech act, as shown in Table 1. For instance, the primary face need for the speech act *request* is to minimize face threatening acts, such as reducing the impositions of the ask, while the primary need for *comment* is to satisfy the user’s desire to be liked, understood, and cared about. For each speech act, we identified the face needs of the act and matched them with politeness strategies that meet these needs. For example, to promote positive face needs when making comments, the system may say, “I know it’s not easy,” using the positive politeness strategy “presuppose/raise/assert common ground.” The system may also comment with “Well done” to meet the hearer’s need to be liked using the positive politeness strategy “gift.” Furthermore, interest can be promoted by the statement, “Fantastic,” which uses the strategy “exaggerate.” In making requests, the primary face need is to avoid negative face threatening acts where negative politeness strategies may be appropriate. For example, in making requests, the system can use “may” or “can you,” based on the strategy “be conventionally indirect;” can use “could you” and “would you” based on the strategy “be pessimistic;” and use “just” to adopt the strategy, “minimize the imposition.”

A mixture of positive and negative politeness strategies was also used to produce a hybrid strategy. According to Brown and Levinson [3], mixing strategies could have a softening effect. For instance, when making a suggestion, the system may say “Let’s take the chronic pain lesson. Maybe start now?” Here, the positive strategy

Table 1: Conversational goals of and opportunities for the use of politeness strategies for speech acts.

Speech Acts	Conversational Goal	Selected Politeness Strategy
Suggest	Protect user freedom of choice and minimize impositions; show care for the user; and be willing to help.	N1: be conventionally indirect N3: be pessimistic P6: avoid disagreement P7: presuppose common ground P12: include both S and H in the activity P14: assume or assert reciprocity
Comment	Appreciate user effort and progress.	P15: gift
Welcome	Show interest in the user and be inviting.	P14: assume or assert reciprocity P15: gift
Farewell	Show interest in the user and be inviting.	P6: avoid disagreement P12: include both S and H in the activity
Request	Minimize impositions of the request.	N1: be conventionally indirect N3: be pessimistic P12: include both S and H in the activity
Repair	Show regret and ask for forgiveness.	N6: apology
Instruct	Be respectful and considerate; avoid impositions on user actions; and encourage user participation.	N1: be conventionally indirect P7: presuppose common ground P10: offer

is the use of “let’s” to include both the speaker and hearer in the activity in order to trigger the cooperative assumption. On the other hand, the use of “maybe” in this suggestive utterance is an example of negative politeness strategies that are conventionally indirect and that give the hearer an “out” for the suggestion. In this case, the face threatening effect of “let’s” was softened by “maybe,” while “let’s” remains to show inclusiveness and consideration for the hearer and indicates that the suggested action is for the user’s benefit. Overall, positive politeness strategies were mostly applied to comment, welcome, and farewell, whereas negative politeness strategies were mainly used for repair and instruct. Table A1 provides the full list of the strategies evaluated in the formative study.

3.2 System Details

3.2.1 System Features. We chose the Google Nest Hub Max as the smart display platform for our implementation and field evaluation. We adopted a health-intervention system, which has been shown to improve older adults’ quality of life in prior research [11], to build a smart display application involving four key features: Discussion Groups, Chronic Pain Lesson, Health Survey, and Journal. Particularly, *Discussion Groups* allows users to create public posts and comment on each other’s posts. *Chronic Pain Lesson* contains a series of educational videos on chronic pain and users can choose the resources beneficial for them to watch. *Health Survey* asks users to reflect on their physical conditions and provide ratings for eight health-related statements. *Journal* is a private blog for users to keep records of their thoughts and view their previous entries.

3.2.2 Voice Interface. We first scripted the voice based interaction flow for all four system features and categorized the utterances based on the seven speech acts (Figure 2). In particular, *Discussion*

Groups employs the speech acts suggest, request and instruct; *Journal* includes suggest and comment; *Chronic Pain Lesson* contains instruct and comment; and *Health Survey* has request and comment. Moreover, the speech acts repair, farewell, and welcome were applied in the overall verbal interaction flow.

3.2.3 Touch Interface. We built a visual interface to support conventional, touch-screen-interaction-based control of the system. The visual screen served three purposes. First, the visual feedback informed the user about the current state of the system, such as when the system is processing the user’s voice input and loading responses. Second, button text suggested terms that the system could understand. Third, clicking on the screen provided an alternative way to navigate the system, which sometimes proved more efficient, as it enabled the user to interrupt the system’s speech.

3.2.4 System Implementation. We used the framework Interactive Canvas of Google Actions to build both voice and touchscreen input and output. We implemented the voice interface with the Actions library and constructed a deterministic conversation model where all the voice prompts of the system were authored in advance. The conversation model was implemented with intents, types, and scenes, which predefined a set of linguistic concepts for the system to recognize and make responses.

3.3 Formative User Study

To inform design decisions, we conducted a formative study to evaluate the selected politeness strategies. We selected one speech example for each type and came up with a variety of utterances using each selected linguistic strategy. We created a total of 32 videos for the seven speech acts and each video used one strategy (7 videos for suggest, 4 for comment, 5 for request, 4 for repair, 3

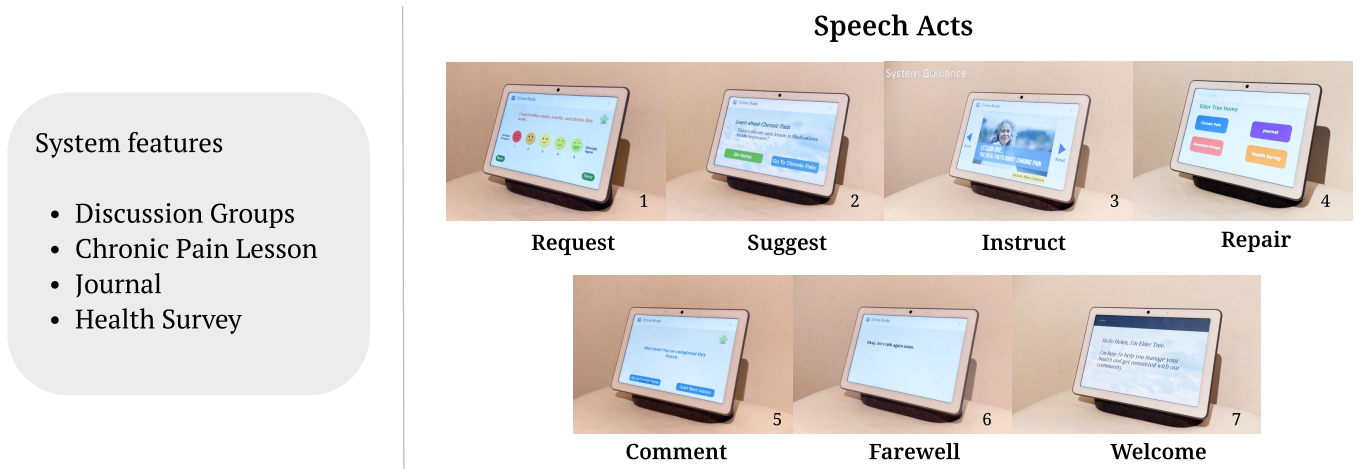


Figure 2: An overview of system features and examples of speech acts. (1) Making requests in *Health Survey* (bald-on-record): “Tell me your rating from one to five.” (2) Suggesting educational content (positive politeness): “Today’s chronic pain lesson is Medications. Let’s begin the lesson now?” (3) Giving instructions for lesson navigation (bald-on-record): “Change the lesson slides with the back and next buttons.” (4) Repair voice recognition failure (negative politeness): “Sorry, repeat what you said?” (5) Commenting on lesson completion (positive politeness): “Well done!” (6) Farewell on exiting (positive politeness): “Let’s talk again soon.” (7) Welcome when opening (positive politeness): “Hello *username*, I’m here to help you manage your health and get connected with our community.”

for farewell, 5 for instruct, and 4 for welcome). One of the selected videos for each speech act was randomly assigned to the participant and each subject watched seven videos in the study. They were asked to evaluate the politeness of voice prompts with semantic differential scales. The scales were adopted from prior work on the use of politeness strategies in email composition [26]: directness of the speech (indirect to direct, ambiguous to straightforward), positive politeness (unfriendly to friendly, unsympathetic to caring), and negative politeness (demanding to undemanding, disrespectful to respectful). Participants also evaluated their likeness of the prompt and willingness to follow with seven Likert scale questions. We included one attention check question asking the scenario of the previous video. We collected demographic information after they finished watching all seven videos.

We used the crowdsourcing platform Amazon Mechanical Turk to conduct the online study. We approved 64 participants out of the recruited 70 workers (*female* = 19, *male* = 45), aged 19–71 ($M = 36, SD = 11.00$). The 6 responses were rejected due to the failed attention checks. Among all participants, 56 out of 64 had not used voice interfaces (87.5%), and 14 of them had used one almost every day in the past month, suggesting that participants were familiar with voice interfaces.

Three outcomes of the online study have informed the system design. First, we identified the linguistic strategies with the highest politeness ratings for each speech type and used them more frequently in the final system. These strategies were verified to have higher ratings on politeness than the bald-on-record strategy. Second, we verified that the bald-on-record strategies for request, comment, welcome, farewell, repair achieved the highest score for directness in their category. Third, we identified the bald-on-record

design with the highest directness scores. The evaluated linguistic strategies and results are shown in Table A1.

4 EVALUATION

4.1 Hypotheses

We proposed the following two hypotheses on how the politeness of the system might affect older adults’ overall experience using the system:

- H1. Participants will have a better overall experience with the polite system, using positive and negative politeness strategies, than the direct system using bald-on-record strategies.
- H2. Participants will be more tolerant of issues in the polite system than in the direct system.

4.2 Participants

We recruited 16 participants over 60 years old in local communities. One user quit the study after the first two days of participation. Among the 15 participants (*female* = 8, *male* = 7), who were aged 61–86 ($M = 74.27, SD = 6.93$), none of them had prior experience with a smart display, while two users were experienced with smart speakers and had actively been using them prior to the study. Participants self-evaluated their technological proficiency on a seven-point rating scale where seven is the most technologically savvy ($Min = 2, Max = 7, M = 4.93, SD = 1.44$). At the end of the study, each participant received \$20 USD in compensation.

4.3 Study Design

Study Setup. We conducted a within-participants-design field study in an in-home environment in which participants used two versions of the conversational interface for a total of 10 days: (1)

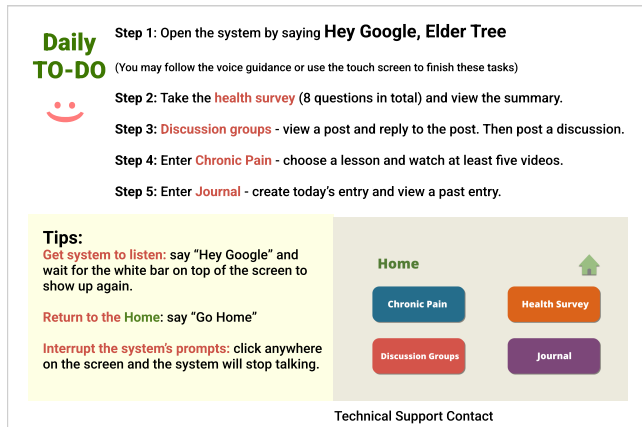


Figure 3: The cheat sheet with system guidance and tips of usage was provided for the participants during the intake appointment. Participants were asked to complete a daily to-do list following the steps on the cheat sheet.

the *polite* system that utilized positive politeness and negative politeness strategies, and (2) the *direct* system that implemented the bald-on-record strategy. Participants experienced one condition for five days followed by the other condition for another five days. The order of the manipulation was randomized (8 of 15 participants used the polite version in the first session). Participants were asked to complete four daily tasks on the device following guidance on a provided cheat sheet (Figure 3).

Procedure. Participants provided informed consent through email and scheduled in-person/remote intake sessions over the phone. To introduce the study and install the system, two experimenters visited the homes of 13 participants. The other two participants received the device by mail and completed the intake session remotely, as they were not local residents. The introduction provided participants with an overview of the voice and touchscreen controls of the smart display, the features of the application, and the daily tasks to be completed during the study. Following the intake appointment, participants completed the first five days of the study, followed by a questionnaire and semi-structured interview session. Next, the experimenters (remotely) switched the system version for the second session and notified participants of the change. At

Table 2: Selected measures for the linguistic strategies with factor analysis results.

Strategy	Items
Positive politeness ($\alpha = 0.70$)	The system sounds friendly. The system sounds caring.
Negative politeness ($\alpha = 0.71$)	The system sounds demanding. I feel obligated to fulfill the system's request. I am not in control of the system.
Bald-on-record ($\alpha = 0.69$)	The way the system communicates with me is efficient. The system sounds straightforward.

the end of the study, participants completed the same questionnaire and interview. Due to variation in participant availability and scheduling constraints, the study duration ranged from 9 to 13 days. Participants received \$20 USD at the end of the study. The study protocols were reviewed and approved by the appropriate institutional review board.

4.4 Measurement

We collected three types of data to measure the effects of politeness strategies on people's experiences.

System Usage. We calculated a number of measures of system usage based on system logs. These measures included the success rate of the interaction, user engagement, and user preference for input modality. The interaction success rate measure was computed as the percentage of completed tasks out of the four daily tasks assigned to participants. User engagement was measured by the interaction duration within the overall application and each feature on a daily basis. Finally, input modality preference was measured by calculating the number of user inputs through voice and touchscreen.

Questionnaires. We measured the perceived system politeness, the experience of system failures, and the overall satisfaction and usability scores with three questionnaires. The first questionnaire asked users about their perception of the linguistic strategies. We conducted factor analysis to select the measures for positive politeness, negative politeness, and bald-on-record strategies. Cronbach's α values for these measures, shown in Table 2, indicated acceptable reliability. The second questionnaire asked about users' tolerance for system failures. The third questionnaire asked about their overall satisfaction and measured their overall experience using the System Usability Scale (SUS).

Thematic Analysis. We conducted semi-structured interviews to collect user feedback on system politeness, understand system expectations, and the overall experience. The interview questions had three topics. The first group asked users to reflect on their overall experience of using the application on the smart display. The second topic was related to system politeness, comparing the system's voice interaction with human conversations, and the perceived role of the device. The third topic asked about users' preferences over touchscreen and voice interactions for system inputs and outputs. We conducted thematic analysis on the interview data with two coders and summarized the common themes.

5 RESULTS

5.1 Quantitative Results

Informed by our hypotheses, all tests used one-tailed comparisons. The analysis of our questionnaire data and manipulation checks used paired t-tests or Wilcoxon tests, depending on normality. We analyzed our system usage data using linear mixed model and fit the model using the Restricted Maximum Likelihood Method (REML). We provide test details for marginal and significant effects in the text below and details for other effects in Tables 3 and 4.

H1 hypothesized that people would have a more positive experience with the polite system than the direct system. The analysis of

Table 3: A summary of quantitative analyses of questionnaire responses. Normality (N) of the data was assessed using Shapiro-Wilk test. We used paired t-test if the difference is normally distributed and Wilcoxon test otherwise.

Measures	Normality	Condition	mean	std	test type	statistic	p-value
Overall Likeness	N	Polite	5.23	1.35	wilcoxon test	32.0	0.54
		Direct	5.07	1.46			
Negative Politeness Score	N	Polite	2.93	1.35	wilcoxon test	45.5	0.30
		Direct	3.07	1.54			
Directness Score	Y	Polite	5.17	1.40	paired t-test	-1.02	0.84
		Direct	5.63	1.53			
Positive Politeness Score	N	Polite	6.17	0.84	wilcoxon test	22.00	0.08
		Direct	5.97	0.99			
System Error Score	Y	Polite	5.27	1.36	paired t-test	1.84	0.04*
		Direct	4.31	1.89			
Satisfaction Score	Y	Polite	4.17	1.55	paired t-test	0.08	0.47
		Direct	4.14	1.81			
SUS Score	Y	Polite	69.17	13.81	paired t-test	0.99	0.17
		Direct	64.83	20.34			
Polite	N	Polite	6.87	0.35	wilcoxon test	6.00	0.04*
		Direct	6.67	0.62			
Direct	Y	Polite	6.20	1.15	paired t-test	-1.70	0.94
		Direct	6.60	0.63			

questionnaire responses showed no significant effects of the system’s politeness on usability, likability of the voice interaction, or overall user satisfaction as shown in Table 3. Therefore, H1 was not supported. H2 hypothesized that people would be more tolerant of system errors in the polite condition than in the direct condition. However, the results showed an effect in the opposite direction (Table 3): participants were more tolerant of system errors in the direct condition ($M = 4.31$, $SD = 1.89$) than in the polite condition ($M = 5.27$, $SD = 1.36$), $t(14) = 1.84$, $p = 0.04$. Therefore, H2 was not supported. We further discussed the finding opposite to H2 in the Limitations Section.

Besides hypotheses, we measured the effects of politeness strategies on system usage. Our analysis revealed that there was no

Table 4: A summary of our quantitative result from system logging. In the repeated measures for system usage under two conditions, we used a linear mixed model to analyze the system logging data with conditions as fixed effects and subjects as random effects.

Measures	Cond.	LS Mean	St Err	F	p-value
Task Completion Rate	Polite	0.74	0.05	0.09	0.77
	Direct	0.74	0.05		
Percentage of Voice Input	Polite	0.65	0.03	1.99	0.18
	Direct	0.69	0.03		
Interaction Duration (m/day)	Polite	24.37	3.52	2.02	0.18
	Direct	19.43	3.52		
Discussion Duration (m/day)	Polite	5.89	1.38	1.02	0.33
	Direct	5.10	1.38		
Survey Duration (m/day)	Polite	1.81	0.19	0.015	0.91
	Direct	1.78	0.19		
Lesson Duration (m/day)	Polite	10.55	1.75	2.99	0.11
	Direct	6.67	1.74		
Journal Duration (m/day)	Polite	1.42	0.30	0.39	0.55
	Direct	1.25	0.30		

significant effect on the task completion rate, system input modality, or interaction duration, as reported in Table 4. Furthermore, we performed manipulation checks by evaluating people’s perceptions of the system’s politeness in both conditions. When asked to rate the statement “The system sounds polite,” participants provided a higher score for the polite system ($M = 6.87$, $SD = 0.35$) than the direct one ($M = 6.67$, $SD = 0.62$), $w(14) = 0$, $p = 0.04$. Although the result was significant, the actual differences are small and may not be meaningful. The results revealed no differences in participants’ perceptions of the system’s directness, negative politeness, or positive politeness, suggesting that participants were not able to differentiate among different politeness strategies.

5.2 Qualitative Results

Our qualitative findings merged into five themes in terms of user experience with the polite system and the direct systems as shown in Figure 4.

5.2.1 Factors that shape the politeness of the system .

Non-linguistic Features. Two non-linguistic factors that people attributed to shape the system’s politeness were vocal features, such as tone, pitch and vocal inflection, and the pace of the interaction they had with the system. In both conditions, participants described the vocal features as “soft,” “low key,” “earnest,” “not so cheery” and “soothing.” For the pace of the interaction, participants associated politeness with the slowness of the system’s pace, thinking that the system “was not pushing” ($D5$) and that “it gives you ample time to make the decision before you answer, which is good. especially for some older folks, that it might take them a little longer to understand it” ($P3$).

Linguistic Features. In both politeness conditions, participants pointed out phrases they thought were conventionally polite behaviors ($P8$, 11 , 15 ; $D7$, 15), such as “it greets you,” “uses your name,”

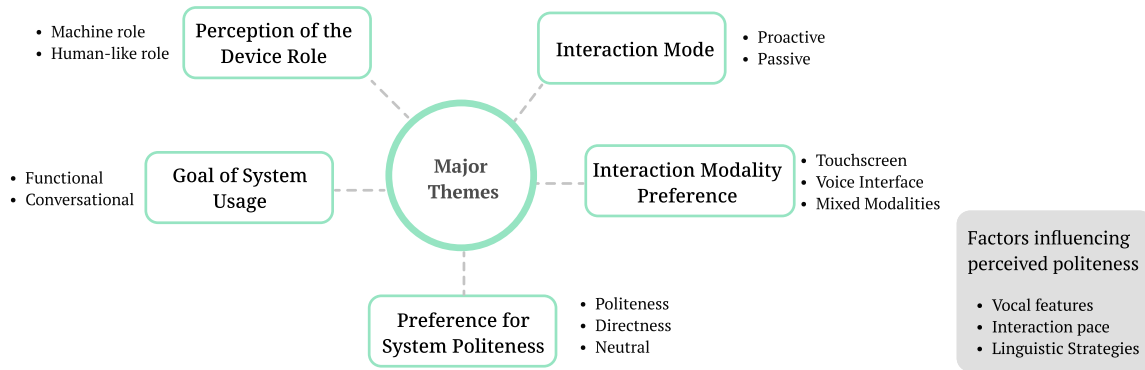


Figure 4: A diagram showing the major themes of our qualitative findings.

“thank you,” and said “nice to see you again.” For negative politeness, participants described the system as being non-abrasive, non-confrontational, and not contentious, e.g., “it prompts well for your options and prompts in a non-demanding way. It doesn’t say press any key to continue or anything like that” (P13). For positive politeness, users described the system as being friendly, kind, respectful, caring, considerate, understanding, and non-judgemental, e.g., “it asks your question. You have a chance to answer. Doesn’t say what is right or wrong or whatever” (D5). For bald-on-record effects, people described the system as being clear, direct, and straightforward. They commented that “the manner in which the questions were asked seemed very straightforward,” and “she speaks distinctly” (D4; P3). Overall, the comments related to linguistic politeness are observed in both conditions.

5.2.2 Preference for system politeness.

Prefer Politeness. Most people responded to the system’s overall politeness in a positive way and described it as being pleasant and friendly. One user mentioned that “I liked that idea that wasn’t so robotic. It greets you” (D11). However, the effects of linguistic politeness strategies were subtle and implicit for most users. They commented positively on the polite version while they couldn’t tell the reason, saying that “I think the second version was a little bit better. I’m not really sure why” (P9). Only one user pointed out the linguistic strategy differences between the two versions and complained about the commanding manner of the direct version: “I thought [it] said just pick the number then I’m like...that was my only beef with that one” (D12).

Prefer Directness. One participant disliked the system’s politeness (P14), saying that “I don’t love being talked to by a computer as friendly as it sounds.” She also viewed the linguistic polite strategy as “added politeness” and “odd,” saying that “They weren’t necessary, and they are just a little bit confusing.”

Neutral towards Linguistic Strategies. Some participants had a neutral view regarding the system’s politeness and thought that “it was fine” and they “had no problem with it” (P9, 14; D9). One user mentioned that the effects of politeness were related to their proficiency in using the system and “after about the second or third time, it’s just a voice” (P2).

5.2.3 Goal of System Usage.

Functional Goals. Some participants expressed goals that were more “functional” in nature toward completing study tasks. These participants used the device to learn educational content or ask about the weather and news (P3, 6). Participants also experimented with the system in an attempt to “mess it up” or to find bugs in the system (P4, 8, 13; D13), e.g., “I’m helping testing to help work out some of the problems. Not trying to get the benefit out of the content.” (P8) Finally, some participants said that their only goal was to complete the tasks on that checklist provided by the experimenter, e.g., “I’m just doing some tasks... ‘cause I was doing it for you” (P15).

Conversation Goals. Some people tend to hold more “conversational” or “relational” goals when interacting with the system. They thought the system could help to reduce isolation for older adults, e.g., “for those who don’t have a lot of social interaction, I think this would boost their spirits tremendously” (P12). These participants also commented on the enjoyment of the conversational experience, for example, stating “it turns out to be a lot more pleasant than you expected” (P10).

5.2.4 Perception of Device Role.

Machine Role. In both conditions, 11 out of 15 participants viewed the system as a machine, describing it as a device, AI, computer, or robot. They mentioned that the verbal interaction was systematic, mechanical, and unnatural. Similarly, they described the system as having a limited vocabulary, an unnatural interaction flow, and a lack of intellectual abilities. Regarding the limited vocabulary, participants mentioned that the communication content was strictly constrained as the system only understands limited commands, “it could only say a certain number of things,” (D14) The unnatural interaction flow is reflected by the delayed system responses, not being able to interrupt the system’s prompts, and missing meaningful social cues that “you can see when you’re making sense” (P7). Furthermore, participants mentioned that the system lacked intellectual abilities, such as interesting ideas, humor, and expressing and reading emotions. Users that viewed the system as a machine held mixed views towards the conversational nature of the voice interaction. Some users stated that there was no conversation at all, e.g., “You have yes or no answers, so it really doesn’t amount

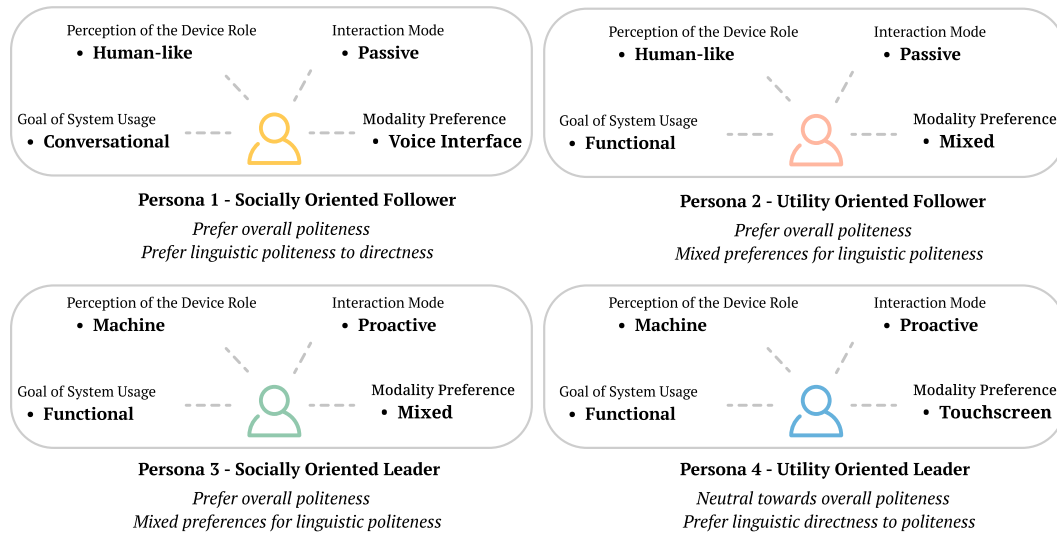


Figure 5: We identified four user personas shaped by their goals of interaction, preferences towards politeness, perceived device role and interaction modality preferences.

to a conversation” (P14). Some users thought it was an unnatural conversation such as “a formal conversation where you’re limited to the words that you can use” (D2).

Human-like Role. In both conditions, the other four participants attributed diverse human roles to the system, calling it a friend, teacher, or professional. The perception that the system is a companion is associated with the understanding and caring characteristics of the speech and is described as “a friend out in the universe...has my interests at heart” (P12) or “interested in my views on things” (D10). The professional and knowledgeable impression was described as “some other person that knows what they’re doing, professional” (P5), which is mostly related to the system’s educational content. Overall, users have attributed human roles to the system described it as being friendly and pleasant.

5.2.5 Interaction Mode.

Proactive Interaction Mode. The proactive users preferred to take the lead in the interaction and seek information actively. For instance, two users mentioned that “I tend to sort of look around for things rather than being guided to them” (P13) and “I would rather do my own research, read about it than have it spoken to me” (D14). Sometimes they would bypass the voice prompts by clicking the screen to speed up their interaction. Also, they tended to focus on the commands that the system could understand and thought that “The terms you can use [are] obviously very limited” (D13).

Passive Interaction Mode. The more passive users preferred to be guided by the system and follow its requests. They tended to describe the system as being easy-to-follow and easy-to-understand, saying that “I just gotta listen” (D5) and “it directs you” (D11). They preferred to hear the system’s voice, saying that “it’s hearing another voice, you know giving you direction like reading the bedtime story to the kid before she goes to sleep” (D12) and “the voice is friendlier than just reading it” (P10).

5.2.6 Interaction Modality.

Prefer the Voice Interface. We observed three main reasons for users’ preference towards the voice interface: health conditions, distance of interaction and ease of use. Some participants with vision or mobility issues can interact with the device remotely, e.g., one participant (P10) explained that for people “who don’t see that well even with glasses” the system has “to have a voice.” Furthermore, voice interaction with the device also allows for multi-tasking, e.g., “I could be fixing a cup of tea or coffee...or taking my medications or putting them together” (P11). Also, some people preferred the voice interface as it is “easy to follow” and “sets everything at ease” (P10, 12).

Prefer the Touchscreen. The major reason that users preferred the touchscreen is to speed up the interaction and “bypass the interactions just by tapping the screen” (P14). Participants also commonly used the touchscreen to avoid or repair voice recognition failures. Some participants mentioned that they used the touchscreen because of their previous technology experience. Users may also prefer the touchscreen because of health conditions. One participant mentioned that he couldn’t hear the device well because of a hearing issue (P4).

Prefer the Mixed Modalities. Some participants responded that they had gotten used to using the touchscreen in the first five days, but found the verbal commands to be easier to use so chose to use voice more (P5, 8). Some people preferred to have the option to switch the modalities based on the tasks at hand or their physical position relative to the device.

5.3 User Personas

We summarized four major personas representing the experience that older adults had with the smart display, as affected by their

goal of system usage, perception of device role, interaction modes, and preference over system politeness (Figure 5).

Persona 1—Socially Oriented Follower. The first user persona perceives the device as a friend and holds a conversational goal in the interaction (P12), describing the interaction as “communicating with someone on the outside” (P12). This persona tends to have a passive interaction mode and follow the system’s guidance in exploring different features. This persona prefers politeness more than directness and has a positive experience with the system’s overall politeness.

Persona 2—Utility Oriented Follower. The second persona holds a functional goal of using the system, perceives the device’s role as a human and tends to follow the system’s guidance (3, 5). Participants in this persona described the system as being “knowledgeable” and “professional.” (P3, 5). They “follow what the operator said” (D3) and have positive feedback towards the system’s overall politeness. This persona has mixed preferences towards the input modality, while prefers to hear the system’s prompts rather than reading the screen, e.g., “I enjoyed, you know, listening to her” (D3), while “I just seem to automatically want to touch it” (P3).

Persona 3—Socially Oriented Leader. Persona 3 contains users that have functional goals and perceive the device in a non-human role such as machine, AI, and robot (2, 4, 6-9, 11, 13). This persona is positive towards overall politeness but doesn’t expect the system to be human-like, saying that “I don’t think the goal should be to give the voice personality beyond being polite and also being efficient. (P1)” Users in persona 3 have a neutral or positive view towards the system’s linguistic politeness.

Persona 4—Utility Oriented Leader. Similar to the third user persona, persona 4 holds functional goals and views the system as a machine (14). Compared with the other user personas, persona 4 uses the touchscreen the most often—whenever they can—and takes the lead in the interaction bypassing the voice guidance. This persona holds neutral or even negative views towards the overall politeness and prefers the linguistic directness rather than the linguistic politeness, describing the linguistic politeness as being “odd” and “unnecessary” (P14).

6 DISCUSSION

Our findings point to three areas where this work can inform the future design of smart displays for older adults: design for the politeness of conversational systems; design for both voice and touch modalities of smart displays; and personalization based on the four identified user personas. We discuss these areas below.

6.1 Design Implications for Politeness in Conversation Design

Factors influencing perceived politeness. Our findings revealed additional factors that shape user perceptions of the system’s politeness. These factors highlight a rich design space for polite conversation design, including the manner that the message is delivered, the vocal features of the system’s voice, and the pace of the interaction. Although our study focused on linguistic cues grounded in Politeness Theory [3], we observed several vocal features, including tone,

pitch, reflection, and key of the voice, to contribute to perceived politeness. Politeness is also shaped by the pace of the interaction, as some users associated the slowness of the voice interaction with the system’s politeness. As suggested by Eelen [7], politeness in interpersonal relationship “can include non-verbal, non-linguistic behaviour,” such as nodding and turn-taking in a conversation, and even “the absence of behaviour,” e.g., returning greeting with silence, which can be seen as being impolite. In the same vein, the politeness of a conversational agent can manifest itself through both linguistic cues and non-linguistic features, and the politeness design of smart systems should consider both opportunities.

Design with Politeness Strategies. Brown and Levinson [3] viewed interaction as “the expression of social relationships built out of strategic language use” in which politeness is tied with relationship-building between speaker and addressee. The design of polite speech should also consider the relationship between the user and the device. Our study showed that people attributed both machine and human-like characteristics to the system, which may suggest different status relationships between the user and the device, also coined as “relative power” [3]. For instance, people who view the system as a tool may assume a higher power status, and the use of positive politeness strategy to establish common ground may backfire, as they thought the “added politeness...weren’t necessary, and they are just a little bit confusing” (P14). On the other hand, people who treat the system as “a friend from the universe” (P12) may assume a similar power status, and the use of the politeness strategies could meet the user’s expectations that they are in a conversation with a caring friend. In addition, some participants mentioned that they also thought about the people who built the system when talking to the device, suggesting that the manufacturer, designer, or the researcher involved in the development of the system may be stakeholders in the relationship between the user and the device. The “relative power” between these stakeholders and the user could also affect people’s perceptions of the device’s politeness.

In addition to the “relative power” between the user and the device, culture is another factor believed to impact the use of politeness strategies. Prior research suggests that people’s polite behaviors are closely tied to social norms and cultural expectations [2, 7, 10, 14]. While cultural factors are not within the scope of the current study, we highlight the importance of situating the design of the system’s polite speech in its cultural context. Certain speech acts in one culture may be interpreted differently in another, and polite behaviors should be culturally appropriate [2]. In addition to perceived politeness, culture can also affect user expectations of the device’s politeness. Therefore, the design of polite speech should consider both the social meanings of the speech acts and the social acceptance of polite speech.

The implementation of politeness strategies should also consider the system’s technical capabilities. For instance, implementing *off-record* strategies would require a high level of understanding of the user and the environment, e.g., the use of metaphors and association clues. Often the hearer needs to take the perspective of the speaker to disambiguate meaning conveyed through off-record strategies. Messages that use metaphors can cause confusion, as users need to take the perspective of the device, e.g., the perceived knowledge and experience of the device, to be able to correctly

interpret the metaphor. While the current study focused on speech acts using *on-record* strategies, the successful implementation of off-record strategies could potentially increase the perceived cognitive capabilities of conversational systems and the user's conversational experience with them.

6.2 Design Implication for Smart Display Interaction

Our findings showed that participants have adopted a mixed modality of system control and output since the frequencies of using the voice interface and touchscreen were similar in our system usage data. Therefore, interaction design on smart displays should consider the distinctive features of each modality as well as the interplay between them. We offer three key design implications, which are discussed below.

First, user feedback on the overall interaction suggests that the system's verbal design should consider both the system's vocal features and interaction flow guided by the system's prompts. The voice interaction could be categorized into two types: (1) dialogue for guiding the interaction and (2) audio for conveying information. In designing the verbal exchanges for system guidance, we categorized the system's utterances into speech acts and applied linguistic strategies by considering the face needs associated with each speech act. The analysis on system usage duration revealed that participants spent a considerable amount of time talking to the system each day, compared to simple queries such as asking about the weather and playing music. Therefore, it is important to support extended conversational exchanges through the voice interaction design.

Second, the results suggest that participants commonly used the touchscreen to bypass voice outputs to improve interaction efficiency as well as handle errors when voice recognition failed. Therefore, the interaction on the touchscreen should provide shortcuts for faster system navigation and support the repair of interaction failures. Moreover, the visual display on the touchscreen should provide language cues on what verbal commands the system might understand and what the user could say to the system as suggested by Yankelevich [41].

Third, the interaction modality switch that we observed among users illustrates that interaction design should consider the flexibility of choosing modalities to meet user needs under different conditions. For example, users might change modalities based on their physical distance from the device. We observed that some users placed the device on the dinner table and sat in front of it during the interaction, while others placed the device on a coffee table and interacted with it from a chair at some distance. Interaction distance can also be affected by which task the user is completing. For instance, users need to be close to the device to watch lesson videos, while they can answer survey questions remotely without looking at the screen. Users may also switch interaction modality while multitasking, e.g., one user mentioned that she had to talk to the device while making coffee or cleaning the house. Finally, we observed that some participants experience physical limitations, such as loss of vision, hearing, and mobility, which made it difficult for them to hear the voice or see the device screen clearly. Hence, the system should allow for modality-switching at any time during

the interaction to support user needs and the context of interaction.

6.3 Design Implications for Personalization

The four personas identified in our analysis are potential targets for personalization in terms of system politeness and the interplay between the voice and touchscreen modalities.

Design Opportunities for Socially Oriented Follower. The focus on the conversational experience of *Socially Oriented Follower* suggests that the system should be proactive to progress the interaction and guide users in the navigation. The finding also suggested that the tone, pitch, reflection, and gender of the voice should be personalized based on user preferences. The system should avoid being commanding or pushy and should leave sufficient time for users to think and respond to system prompts. With its potential to build a relationship with the user through conversation, this system could be particularly beneficial for older adults facing isolation and could more broadly offer mental health support.

Design Opportunities for Utility Oriented Follower. To support *Utility Oriented Follower's* functional goal of information seeking, we suggest that the design should balance the conversational experience and the efficiency of system usage. The touchscreen should support all user inputs as an alternative for voice inputs, while the button text could be more conversational and less command-like, using words people would say to answer the system verbally, e.g., "OK, let's begin" instead of "Begin."

Design Opportunities for Socially Oriented Leader. Given the positive feedback to system politeness from *Socially Oriented Leader*, the conversation design should leverage politeness, utilizing both vocal features and linguistic politeness strategies, to improve the overall user experience. Furthermore, to improve interaction efficiency, the system could avoid confirmations or reading out options for a selection if it can confidently recognize that the user is viewing the screen. As with all adaptive systems, these features must be designed carefully to minimize errors due to a lack of proper recognition of user context.

Design Opportunities for Utility Oriented Leader. *Utility Oriented Leader* tended to treat the smart display no differently than a tablet or other device for most of its features. We therefore suggest that the interaction design makes use of linguistic directness to maximize efficiency. Because this user type tends to lead the interaction, the system should have a low level of proactivity and wait for user input to progress the interaction.

6.4 Limitations

This work has a number of limitations that limit the applicability of our findings to other settings and that inform future work. First, our analysis of quantitative data showed that, contrary to our hypothesis, participants were marginally more tolerant toward errors by the more *direct* version of the system over the more *polite* version. A potential explanation of this finding is a mismatch between user perceptions of the system's capabilities and the errors it made. Specifically, polite speech might have been seen as being more human-like, suggesting human-like capabilities, and

the errors that users observed in simple tasks, such as voice recognition violated this expectation. Future exploration of the specific effects of mismatch between language use and capabilities can provide more conclusive explanations of our observation, as recent research found the perceived competence of the system has an important effect on user experience [20]. Second, our evaluation of the politeness strategies in the formative study mainly focused on the functionality of speech acts without considering individual differences and cultural factors. Future studies may consider additional factors, such as cultural and social-contextual factors, that affect how speech acts are interpreted. Furthermore, although our development adopted an iterative approach with continuous user input, we have observed a number of usability issues in our system that caused voice recognition failures, long loading times for touchscreen pages, and occasional termination of the smart display application. Although such issues are prevalent even in commercial smart display applications, they might have affected the overall user experience and our interpretation of the effects of the linguistic strategies we explored.

7 CONCLUSION

Conversational interfaces are becoming pervasive technologies integrated into smart speakers and display devices. These interfaces increasingly rely on human-like dialogue to offer a natural experience, which holds tremendous potential for older users. However, dialogue-based interfaces can be cognitively demanding and make error recovery difficult for this population of users. In this paper, we explored the use of Politeness Theory [3] in designing conversations for a smart display to address some of these challenges toward improving user experience. We categorized the voice interaction offered by a smart display application designed for older adults into seven major speech acts—request, suggest, instruct, comment, welcome, farewell, and repair—and applied politeness strategies to each speech act. A field deployment study asked 15 older adults to use a more *direct* or a more *polite* version of a smart display for five days at their homes. From our data, we identified five factors related to preferences toward politeness, conversational interaction, and experiential orientation and characterized four user personas shaped by these factors. These factors and personas inform how politeness cues and strategies might be used by the future design of voice interfaces and point to opportunities for personalization.

ACKNOWLEDGMENTS

This study is funded by the Agency for Healthcare Research and Quality, United States Department of Health and Human Services (grant number 1R18HS026853-01A1), received ethical approval from the University of Wisconsin Health Sciences and Minimal Risk Research Institutional Review Board (reference number 2020-0868), and is registered at ClinicalTrials.gov (NCT04798196). We would like to thank our collaborators at the Center for Health Enhancement Systems Studies (CHESS) at the University of Wisconsin–Madison for their assistance in the testing and deployment of our system and the administration of our field study.

REFERENCES

- [1] Sean Andrist, Erin Spannan, and Bilge Mutlu. 2013. Rhetorical robots: Making robots more effective speakers using linguistic cues of expertise. In *2013 8th*

- ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 341–348. <https://doi.org/10.1109/HRI.2013.6483608>
- [2] Shoshana Blum-Kulka and Elite Olshtain. 1984. Requests and apologies: A cross-cultural study of speech act realization patterns (CCSARP). *Applied linguistics* 5, 3 (1984), 196–213.
- [3] Penelope Brown and Stephen Levinson. 1987. *Politeness: Some universals in language usage*. Vol. 4. Cambridge university press.
- [4] 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>
- [5] Leigh Clark, Nadia Pantidi, Orla Cooney, Philip Doyle, Diego Garaialde, Justin Edwards, Brendan Spillane, Emer Gilmartin, Christine Murad, Cosmin Munteanu, Vincent Wade, and Benjamin R. Cowan. 2019. *What Makes a Good Conversation? Challenges in Designing Truly Conversational Agents*. Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi-org.ezproxy.library.wisc.edu/10.1145/3290605.3300705>
- [6] Justin Edwards, Christian Janssen, Sandy Gould, and Benjamin R. Cowan. 2021. Eliciting Spoken Interruptions to Inform Proactive Speech Agent Design. In *CHI 2021 - 3rd Conference on Conversational User Interfaces* (Bilbao (online), Spain) (CHI '21). Association for Computing Machinery, New York, NY, USA, Article 23, 12 pages. <https://doi.org/10.1145/3469595.3469618>
- [7] Gino Eelen. 2014. *A Critique of Politeness Theory: Volume 1*. Routledge.
- [8] Joel E Fischer, Stuart Reeves, Martin Porcheron, and Rein Ove Sivkeland. 2019. Progressivity for voice interface design. In *Proceedings of the 1st International Conference on Conversational User Interfaces*. 1–8.
- [9] Herbert P Grice. 1975. Logic and conversation. In *Speech acts*. Brill, 41–58.
- [10] Yueguo Gu. 1990. Politeness phenomena in modern Chinese. *Journal of pragmatics* 14, 2 (1990), 237–257.
- [11] David H Gustafson, Rachel Kornfield, Marie-Louise Mares, Darcie C Johnston, Olivia J Cody, Ellie Fan Yang, Juwon Hwang, Jane E Mahoney, John J Curtin, Alexander Tahk, et al. 2022. Effect of an eHealth intervention on older adults' quality of life and health-related outcomes: a randomized clinical trial. *Journal of General Internal Medicine* 37 (2022), 521–530. Issue 3. <https://doi.org/10.1007/s11606-021-06888-1>
- [12] Stephan Hammer, Birgit Lugrin, Sergey Bogomolov, Kathrin Janowski, and Elisabeth André. 2016. Investigating Politeness Strategies and Their Persuasiveness for a Robotic Elderly Assistant. In *Proceedings of the 11th International Conference on Persuasive Technology - Volume 9638* (Salzburg, Austria) (PERSUASIVE 2016). Springer-Verlag, Berlin, Heidelberg, 315–326. https://doi.org/10.1007/978-3-319-31510-2_27
- [13] Lone Koefoed Hansen and Peter Dalsgaard. 2015. Note to Self: Stop Calling Interfaces "Natural". In *Proceedings of The Fifth Decennial Aarhus Conference on Critical Alternatives* (Aarhus, Denmark) (CA '15). Aarhus University Press, Aarhus N, 65–68. <https://doi.org/10.7146/aaahcc.v1i1.21316>
- [14] Sachiko Ide. 1989. Formal forms and discernment: Two neglected aspects of universals of linguistic politeness. (1989).
- [15] W. Lewis Johnson, Richard E. Mayer, Elisabeth André, and Matthias Rehm. 2005. Cross-Cultural Evaluation of Politeness in Tactics for Pedagogical Agents. In *Proceedings of the 2005 Conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology*. IOS Press, NLD, 298–305.
- [16] Da-jung Kim and Youn-kyung Lim. 2019. Co-Performing Agent: Design for Building User-Agent Partnership in Learning and Adaptive Services. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–14. <https://doi.org/10.1145/3290605.3300714>
- [17] Hankyung Kim, Hoyeon Nam, Uichin Lee, and Youn-kyung Lim. 2021. Utilitarian or Relational? Exploring Indicators of User Orientation Towards Intelligent Agents. In *HCI International 2021 - Posters*, Constantine Stephanidis, Margherita Antona, and Stavroula Ntoa (Eds.). Springer International Publishing, Cham, 448–455.
- [18] Sunyoung Kim. 2021. Exploring How Older Adults Use a Smart Speaker–Based Voice Assistant in Their First Interactions: Qualitative Study. *JMIR mHealth and uHealth* 9, 1 (Jan. 2021), e20427. <https://doi.org/10.2196/20427>
- [19] Yelim Kim, Mohi Reza, Joanna McGrenere, and Dongwook Yoon. 2021. Designers Characterize Naturalness in Voice User Interfaces: Their Goals, Practices, and Challenges. 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, Article 242, 13 pages. <https://doi.org/10.1145/3411764.3445579>
- [20] Loriane Koelsch, Frédéric Elisei, Ludovic Ferrand, Pierre Chausse, Gérard Bailly, and Pascal Huguet. 2021. Impact of social presence of humanoid robots: does competence matter?. In *Proceedings of the International Conference on Social Robotics*.
- [21] Lyndsie M. Koon, Sean A. McGlynn, Kenneth A. Blocker, and Wendy A. Rogers. 2020. Perceptions of Digital Assistants From Early Adopters Aged 55+. *Ergonomics*

- in *Design: The Quarterly of Human Factors Applications* 28, 1 (Jan. 2020), 16–23. <https://doi.org/10.1177/1064804619842501>
- [22] Raina Langevin, Ross J Lordon, Thi Avrahami, Benjamin R. Cowan, Tad Hirsch, and Gary Hsieh. 2021. Heuristic Evaluation of Conversational Agents. 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, Article 632, 15 pages. <https://doi.org/10.1145/3411764.3445312>
- [23] Josephine Lau, Benjamin Zimmerman, and Florian Schaub. 2018. Alexa, Are You Listening?: Privacy Perceptions, Concerns and Privacy-seeking Behaviors with Smart Speakers. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 1–31. <https://doi.org/10.1145/3274371>
- [24] Namyoon Lee, Jeonghun Kim, Eunji Kim, and Ohbyung Kwon. 2017. The Influence of Politeness Behavior on User Compliance with Social Robots in a Healthcare Service Setting. *International Journal of Social Robotics* 9, 5 (Nov. 2017), 727–743. <https://doi.org/10.1007/s12369-017-0420-0>
- [25] Yeoreum Lee, Jae-eul Bae, Sona S Kwak, and Myung-Suk Kim. 2011. The effect of politeness strategy on human-robot collaborative interaction on malfunction of robot vacuum cleaner. In *RSS Workshop on HRI*.
- [26] Hajin Lim. 2021. *Designing to Support Sensemaking in Cross-Lingual Computer-Mediated Communication Using NLP Techniques*. Ph.D. Dissertation. Cornell University.
- [27] Jane Lockshin and Tom Williams. 2020. “We Need to Start Thinking Ahead”: The Impact of Social Context on Linguistic Norm Adherence. In *Annual Meeting of the Cognitive Science Society*. <https://doi.org/10.31234/osf.io/2wtzq>
- [28] Jeremy Lopez, Claire Textor, W. Braxton Hicks, Makenzie Pryor, Anne Collins McLaughlin, and Richard Pak. 2019. An Aging-focused Heuristic Evaluation of Home Automation Controls. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 63, 1 (Nov. 2019), 6–10. <https://doi.org/10.1177/1071181319631446>
- [29] Ewa Luger and Abigail Sellen. 2016. *“Like Having a Really Bad PA”: The Gulf between User Expectation and Experience of Conversational Agents*. Association for Computing Machinery, New York, NY, USA, 5286–5297. <https://doi.org/10.1145/2858036.2858288>
- [30] Lydia Manikonda, Aditya Deotale, and Subbarao Kambhampati. 2018. What’s up with Privacy?: User Preferences and Privacy Concerns in Intelligent Personal Assistants. In *Proceedings of the 2018 AAAI/ACM Conference on AI, Ethics, and Society*. ACM, New Orleans LA USA, 229–235. <https://doi.org/10.1145/3278721.3278773>
- [31] Richard E. Mayer, W. Lewis Johnson, Erin Shaw, and Sahiba Sandhu. 2006. Constructing Computer-Based Tutors That Are Socially Sensitive: Politeness in Educational Software. *Int. J. Hum.-Comput. Stud.* 64, 1 (jan 2006), 36–42. <https://doi.org/10.1016/j.ijhcs.2005.07.001>
- [32] Robert J Moore and Raphael Arar. 2019. *Conversational UX Design: A Practitioner’s Guide to the Natural Conversation Framework*. Morgan & Claypool.
- [33] Leysan Nurgalieva, Juan Jose Lara Laconich, Marcos Baez, Fabio Casati, and Maurizio Marchese. 2019. A Systematic Literature Review of Research-Derived Touchscreen Design Guidelines for Older Adults. *IEEE Access* 7 (2019), 22035–22058. <https://doi.org/10.1109/ACCESS.2019.2898467>
- [34] François Portet, Michel Vacher, Caroline Golanski, Camille Roux, and Brigitte Meillon. 2013. Design and evaluation of a smart home voice interface for the elderly: acceptability and objection aspects. *Personal and Ubiquitous Computing* 17, 1 (Jan. 2013), 127–144. <https://doi.org/10.1007/s00779-011-0470-5>
- [35] Sergio Sayago, Barbara Barbosa Neves, and Benjamin R Cowan. 2019. Voice assistants and older people: some open issues. In *Proceedings of the 1st International Conference on Conversational User Interfaces - CUI '19*. ACM Press, Dublin, Ireland, 1–3. <https://doi.org/10.1145/3342775.3342803>
- [36] Vasant Srinivasan and Leila Takayama. 2016. *Help Me Please: Robot Politeness Strategies for Soliciting Help From Humans*. Association for Computing Machinery, New York, NY, USA, 4945–4955. <https://doi.org/10.1145/2858036.2858217>
- [37] Milka Trajkova and Aqueasha Martin-Hammond. 2020. “Alexa is a Toy”: Exploring Older Adults’ Reasons for Using, Limiting, and Abandoning Echo. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM, Honolulu HI USA, 1–13. <https://doi.org/10.1145/3313831.3376760>
- [38] Jing Wei, Tilman Dingler, and Vassilis Kostakos. 2022. Understanding User Perceptions of Proactive Smart Speakers. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 185 (dec 2022), 28 pages. <https://doi.org/10.1145/3494965>
- [39] Zhuxiaona Wei and James A. Landay. 2018. Evaluating Speech-Based Smart Devices Using New Usability Heuristics. *IEEE Pervasive Computing* 17, 2 (2018), 84–96. <https://doi.org/10.1109/MPRV.2018.022511249>
- [40] Xi Yang and Marco Aurisicchio. 2021. Designing Conversational Agents: A Self-Determination Theory Approach. 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, Article 256, 16 pages. <https://doi.org/10.1145/3411764.3445445>
- [41] Nicole Yankelovich. 1996. How Do Users Know What to Say? *Interactions* 3, 6 (dec 1996), 32–43. <https://doi.org/10.1145/242485.242500>
- [42] Tamara Zubatyi, Kayci L Vickers, Niharika Mathur, and Elizabeth D Mynatt. 2021. Empowering Dyads of Older Adults With Mild Cognitive Impairment And

Their Care Partners Using Conversational Agents. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, Yokohama Japan, 1–15. <https://doi.org/10.1145/3411764.3445124>

A FORMATIVE STUDY

A.1 Measures for Politeness Strategies - Semantic Differential Scale

- (1) Indirect—Direct.
- (2) Ambiguous—Straightforward
- (3) Unfriendly—Friendly
- (4) Unsympathetic—Caring
- (5) Demanding—Undemanding
- (6) Disrespectful—Respectful

A.2 Measures for Overall Likeness of the Speech - Seven Likert Scale Questions

- (1) I like the way the system talks to me.
- (2) I am willing to follow up with the system’s prompts.

We evaluated the selected politeness strategies for seven speech acts and created 32 videos playing the utterances on the smart display with visual feedback. The strategy and speech examples are summarized in Table A1.

A.3 Formative Study - Evaluated Politeness Strategies

Table A1 provides the prompts used in our formative study to evaluate politeness strategies.

B QUESTIONNAIRE TO EVALUATE THE POLITE AND THE DIRECT SYSTEM

The post-study survey has 35 Likert scale questions. For each question, we asked users to provide a rating from 1 to 7, where 7 is strongly agree and 1 is strongly disagree.

Overall Likeness

- (1) Overall, I like the way the system talks to me.
- (2) The conversation with the system is a positive experience.

Politeness

- (1) The system sounds polite.
- (2) The system sounds friendly.
- (3) The system sounds caring.
- (4) The system sounds distant to me.
- (5) The system sounds disrespectful.
- (6) The system sounds demanding.
- (7) I feel obligated to fulfill the system’s request.
- (8) I feel hesitant to say no to the system.
- (9) I am in control of the system.
- (10) The way the system communicates with me is efficient.
- (11) The system sounds direct.
- (12) The system sounds assertive.
- (13) The system sounds straightforward.

Tolerance of system issues

- (1) It annoys me when it takes a long time for the system to load.
- (2) It annoys me when the system’s voice recognition fails.

(3) It annoys me when unexpected system errors occur.

Satisfaction

- (1) I am satisfied with it.
- (2) I would recommend it to a friend.
- (3) It is fun to use.
- (4) It works the way I want it to work.
- (5) It is wonderful.
- (6) I feel I need to have it.
- (7) It is pleasant to use.

SUS (System Usability Scales)

- (1) I think that I would like to use this system frequently.
- (2) I found the system unnecessarily complex.

(3) I thought the system was easy to use.

- (4) I think that I would need the support of a technical person to be able to use this system.
- (5) I found the various functions in this system were well integrated.
- (6) I thought there was too much inconsistency in this system.
- (7) I would imagine that most people would learn to use this system very quickly.
- (8) I found the system very cumbersome to use.
- (9) I felt very confident using the system.
- (10) I needed to learn a lot of things before I could get going with this system.

Table A1: Result summary of the formative study that informs the politeness strategy selection for each speech act. The scores are the mean for bald-on-record, positive politeness and negative politeness from left to right. We used a curated set of strategies that received relatively high politeness scores or that enriched language use. The selected bald-on-record, positive politeness, or negative politeness strategies are highlighted with *, †, and ‡, respectively.

Linguistic Strategies	Prompts	Scores
Suggest		
bald-on-record*	Today's chronic pain lesson is Medications. Begin now?	5.95/5.70/5.85
bald-on-record	Today's chronic pain lesson is Medications. Don't miss it! Begin now?	5.81/5.41/5.59
N1‡	Today's chronic pain lesson is Medications. Maybe begin now?	6.44/6.11/6.33
N3‡	Today's chronic pain lesson is Medications. Would you like to begin now?	6.19/6.13/6.00
N6‡	Today's chronic pain lesson is Medications. Begin the lesson then?	5.46/4.88/5.19
P12†	Today's chronic pain lesson is Medications. Let's begin the lesson now?	5.81/5.44/5.81
P14†	Here is a chronic pain lesson for you. Today's topic is Medications. Begin the lesson now?	6.60/5.80/5.50
Comment		
bald-on-record*	Lesson completed.	6.50/5.24/5.83
P7	I know it takes time, but you completed this lesson!	6.13/5.26/5.48
P15†	Congratulations! You've completed this lesson.	6.38/6.27/6.58
P15†	Well Done! You've completed this lesson.	6.05/5.95/6.25
Request		
bald-on-record*	Tell me your rating from one to five.	6.07/5.21/5.29
P15	What is your rating from one to five? I'm listening.	6.13/5.34/5.22
N1‡	You may give me a rating from one to five.	6.42/5.92/5.79
N3‡	Could you choose a rating from one to five?	6.38/5.58/5.88
N4‡	Just choose a rating from one to five.	6.55/5.30/5.05
Repair		
bald-on-record*	Repeat what you said?	6.11/5.39/5.39
N6‡	Sorry, could you repeat what you said?	6.00/5.36/5.54
N7	It seems that my voice recognition didn't work well. Repeat what you said?	5.85/5.56/5.00
N7	It was hard to catch what you said. Repeat what you said?	6.03/5.63/5.58
Farewell		
bald-on-record*	Okay, bye.	6.43/5.09/5.43
P6†	Okay, I'll talk with you next time then.	6.00/5.95/5.73
P12†	Okay, let's talk again soon.	5.83/5.70/6.07
Instruct		
bald-on-record*	Change the lesson slides with the back and next buttons.	6.29/5.29/5.15
N1‡	You may click the next or back buttons to change lesson slides.	5.92/5.46/5.92
N7‡	The next or back buttons could be clicked to change lesson slides	6.40/5.57/6.07
P10†	I'm able to change lesson slides if you click the back and next buttons.	6.55/5.35/5.90
P12	Let's change lesson slides with the back and next buttons.	6.35/5.35/5.60
Welcome		
bald-on-record	Welcome!	5.61/5.24/5.76
bald-on-record*	Hello <i>username</i> , welcome to <i>application-name</i> !	6.55/5.41/5.91
bald-on-record	Hello <i>username</i> , see what's new on <i>application-name</i> !	5.91/5.73/5.95
P14†	Hello <i>username</i> , I'm <i>application-name</i> . I'm here to help you manage your health and get connected with our community.	6.20/5.83/5.80

† **Positive Politeness Strategy** P2: exaggerate; P6: avoid disagreement; P7: presuppose/raise/assert common ground; P10: offer, promise; P12: include both S and H in the activity; P13: give or ask for reasons; P14: assume or assert reciprocity; P15: gift

‡ **Negative Politeness Strategy** N1: be conventionally indirect; N3: be pessimistic; N4: minimize the imposition; N5: give deference; N6: apology; N7: impersonalize