

MASTER

Exploring the effect of emotional speech by smart speakers on memory

Debije, Daphne J.J.

Award date:
2023

[Link to publication](#)

Disclaimer

This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

Eindhoven, July 2023



Department of Industrial Engineering & Innovation Sciences
Human-Technology Interaction

**Exploring the effect of emotional speech by
smart speakers on memory**

D.J.J. Debije
1255940

in partial fulfillment of the requirements for the degree of

**Master of Science
in Human-Technology Interaction**

Supervisors:

Dr.ir. P.A.M. Ruijten-Dodoiu (TU/e)

Dr.ir. R.H. Cuijpers (TU/e)

Dr. J.R.C. Ham (TU/e)

ABSTRACT

Smart speakers are becoming more popular; just 35% of the US and UK people reported having access to a smart speaker at home in 2020 (AudienceProject, 2020). For further integration into our daily lives, ensuring the interactions with smart speakers are as natural and pleasant as possible is important. Existing studies show that addressing affect and emotion in conversational user interfaces or smart speakers can enhance user satisfaction and make conversations more natural. It was found that emotional speech positively affects attention and memory of spoken text in human interactions. This thesis examines if this effect can also be found in a conversation between a person and a smart speaker. It was expected that people would be able to recall more of the information given to them by a smart speaker when the text and emotion expressed by the smart speaker were congruent.

A 2 (happy vs. sad speech) x2 (a happy vs. a sad story) between participants design experiment was conducted, where 62 participants listened to a scenario, and the recall was measured afterward. No significant differences in recall were found between the congruent and incongruent conditions. Results showed that the participants did not recognize the intended emotions in the voice manipulations. This could be explained by cultural differences or unsuccessful audio manipulation. Both could have led to participants having trouble recognizing the emotions in the voice separate from the text. This difficulty resulted in participants reporting different emotions, indicating the importance of extensive testing to create realistic emotional expressions. However, a trend was visible for the positive effect of the happy voice on recall, which resonates with earlier research. For future research, investigating this effect would be interesting because programming smart speakers with happy voices would benefit the recall if it is accurate.

TABLE OF CONTENTS

Abstract	ii
1. Introduction.....	1
1.1 Computer User Interfaces	1
1.2. Human communication	3
1.3. Emotions.....	3
1.4. Expressing emotions in synthetic speech.....	5
1.5. Attention and memory.....	6
1.6. Research question	6
2. Method.....	8
2.1. Design	8
2.2. Participants.....	8
2.3. Materials.....	8
2.4. Measures	9
2.5. Procedure	11
3. Results	12
3.1. Descriptive statistics.....	12
3.2. For each scenario: which emotions were recognized in the text and speech	14
3.3. How the participants experienced the voice.....	16
3.4. Hypothesis test.....	17
3.5. Exploratory analysis	20
4. Discussion.....	21
4.1. Limitations and future research	23
4.2. Conclusion	24
References.....	25
Appendix.....	29
Appendix A, Sample size analyses.....	29

Appendix B, Holiday scenario text.....	29
Appendix C, Questionnaire.....	34
Appendix D, STATA - do file.....	52

Chapter 1

INTRODUCTION

Technology has become ubiquitous. People carry smartphones in their pockets, smartwatches on their wrists and use smart speakers like an Amazon Echo or Google Nest to turn on the lights in their houses. These technologies are used more than ever in our jobs and private lives. Moreover, the number of devices is expected to grow by around 2% yearly (*Statista, 2023*). They help us to be more productive and can help us to relax during our leisure time. Just under 35% of the US and UK people reported having access to a smart speaker at home in 2020 (*AudienceProject, 2020*). One can assume that this percentage has increased since then. This percentage clearly shows the popularity of smart speakers. For further integration into our daily lives, ensuring the interactions with smart speakers are as natural and pleasant as possible is important. However, knowing where this interface originated is crucial before suggesting improvements.

1.1 Computer User Interfaces

When using technology, we often interact with it by using interfaces. A computer user interface can take many different shapes and forms. The development of computers started during the second world war when the ENIAC, which stood for Electronic Numerical Integrator and Calculator, was built at the University of Pennsylvania between 1943 and 1945 by two professors, John Mauchly and Presper Eckert, who got funding from the war department after promising they could build a machine that would replace all the "computers," meaning the women who were employed calculating the firing tables for the army's artillery guns (*Kopplin, 2002*).

In the post-war era, the development of computers continued trying to create decision-making machines. Computer interfaces started with command line interfaces, developed in the 1960s, which needed a typed-in command and would give output in typed lines. The command line user interfaces evolved into graphical user interfaces thanks to Alan Kay, who focused on making computers easy to

use and operate. Kay used images and symbols, resulting in a visual information display (Barnes, 2010; Martinez, 2011). The GUI is very popular, and most devices today have this interface. However, both interfaces are one-sided, with the user providing all the input and the computer the output (Martinez, 2011).

The next step in interface development is the conversational user interface. This interface is voice or text operated and is now capable of a turn-by-turn conversation between the user and the computer. A chatbot can use this interface for the text-operated parts or smart speakers for an auditory representation (Lister et al., 2020). While conversational user interface development started in 1966 when the first chatbot was created, the development is still ongoing. The first chatbot could communicate with the user through a keyword-matching program that searches for an appropriate transformation rule to reformulate the input and provide an output. Keyword-matching was very limited because there was minimal context identification, and pattern-matching rules are not flexible and thus not easily transferable to different domains (Caldarini & Jaf, 2022).

Chatbot development gained traction in the 80s when artificial intelligence was used so that dialogue pattern knowledge could be added to the software to expand its knowledge base. This significantly improved previous pattern-matching systems since the knowledge base was easily expandable. The biggest limitation of relying on pattern-matching systems is that they are inflexible and reliant on manually written rules. With recent developments in machine learning and technological advances, new chatbots were created without relying on rules and pattern-matching techniques, encouraging the commercial use of chatbots (Caldarini & Jaf, 2022). These new chatbots have also been competing in Turing tests where they interact with people and have to convince the people that they are interacting with a human being. These competitions have led to a great advancement in how natural chatbots communicate (Zemčik, 2019).

One of the new and most interesting applications is the development of smart personal assistants (such as Amazon's Alexa and Google's Google Assistant). Personal assistants, chatbots, or conversational agents that can communicate with the user through voice are usually integrated into smartphones, smartwatches, dedicated home smart speakers and monitors, and even cars (Caldarini & Jaf, 2022). However, these are not as developed yet that they can compete in Turing tests.

As mentioned before, the conversational user interface is still limited to a turn-taking conversation where the user still provides the input and needs to use specific commands or sentence structures, and the computer responds with an output. Looking at the development over the last 20 years, it seems inevitable that the conversational user interface will also improve to where a real

conversation, similar to a conversation with another human, might be possible. One example of such a future is shown in the film "Her," where a man falls in love with a conversational agent (Jonze, 2013).

1.2. Human communication

While conversing with another person, people provide input and output compared to a conversation with a smart speaker, where the person provides all the input and the smart speaker only provides output. In a conversation between people, there is a natural exchange where they can cut in and respond or ask questions. This way of communicating is more varied than a turn-taking conversation. People are also not bound by using certain words like the commands for the computer. They could use a description of what the word would mean, and this would not halt the conversation.

When people speak face-to-face, many extra factors, such as non-verbal communication, come into play. The presentation of the voice: phonetics determines a large part of the meaning of the words. Phonetics are speech sounds and their physiological production and acoustic qualities (Ladefoged, 2023). People use different tones to express emotions, for example, when enthusiastic or bored. An excellent example is the word "sure"; one can express either enthusiastic agreement or displeasure and disbelief. This variation in people's phonetics makes conveying different meanings with the same word possible. In speech, much information is in non-verbal communication, such as phonetics. This extra information can be used to express feelings or emotions.

Emotions play an essential part in social interactions and are becoming more of a focus in and of themselves. People are now more aware and give more importance to their emotions. People actively focus on improving their emotional well-being by improving architecture (Han, 2019) and researching the effects of the covid-19 epidemic on emotional well-being (Lades et al., 2020; Sun & Lin, 2020; Yang & Ma, 2020). Emotions are important for showing how people feel and predicting how they are likely to behave (Russell, 2003). That is why examining which emotions people experience and how they express and recognize them is important.

1.3. Emotions

People experience different emotions they can express through speech, body language, or other behaviors. However, how do these emotions originate? Past research has been attempting to link specific brain areas to specific emotions (Gerber et al., 2008), such as the amygdala for fear (Davis & Whalen, 2001; LeDoux, 2003), the subgenual anterior cingulate cortex for sadness or depression (Beauregard et al., 1998; Reiman et al., 1997), the insula for disgust (Reiman et al., 1997; Sprengelmeyer et al., 1998), and the ventral striatum for happiness (Davidson & Irwin, 1999; Lane et al., 1999; Whalen et al., 1998). However, the brain regions identified in studies of these basic emotions have been far from consistent (Barrett & Wager, 2006), with even the most reproducible finding, that

of fear activating the amygdala, appearing in only 60% of studies in one meta-analysis (Phan et al., 2002) and in fewer than 40% of studies in another (Murphy et al., 2003).

This is a large difference between the meta-analyses, which can be explained by different selection criteria employed by the researchers. Murphy et al. (2003) used a narrower publication period than Phan et al. (2002) and used different requirements where all studies with emotional tasks were included. In the article by Murphy et al. (2003), it is also mentioned that when running the analysis on a subset of the studies that employed facial expressions of emotion as stimuli, amygdala activity was reported in over 60% of the studies. When comparing these meta-analyses, they turn out not to be so different. They only use different criteria. However, this highlights the importance of employing multiple sources to get the complete picture.

While these different brain areas are fascinating, people experience emotions as a cognitive state, so researchers classify emotions using a separate model that emerged from studies of participant ratings of the degree of similarity between emotions (Gerber et al., 2008). In this circumplex model of core affects, emotions were thought to be classified along two independent dimensions—arousal (the extent to which an emotion is associated with an individual sensation of energy) and valence (the extent to which an emotion reflects a negative or positive state of mind), see Figure 1 (Barrett et al., 2006; Russel, 1980; Russell, 2003; Russell & Bullock, 1985). This model suggests that distinct neural pathways serve not distinct emotions but each of these two underlying neurophysiological dimensions of the affective circumplex (Posner, Russell, & Peterson, 2005).

The circumplex model further proposes that our subjective experience of emotion is a cognitive interpretation of the neurophysiological experience of valence and arousal in a given situational context (Russell, 2003). Assuming that this model is correct and that people interpret emotions by plotting them along these two dimensions, it would be interesting if we could replicate this effect with synthetic speech.

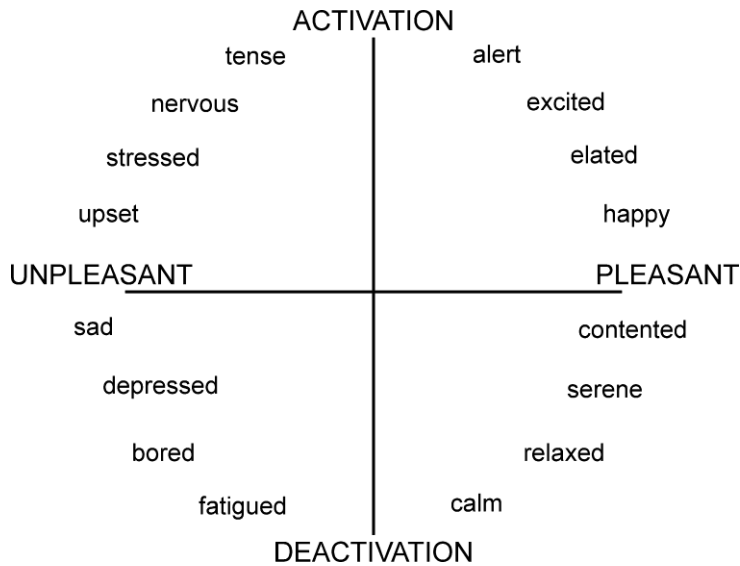


Figure 1, Arousal-Valence model The affective circumplex depicts each emotion along continuous dimensions of arousal (y-axis) and valence (x-axis) (Barrett & Russell, 1998).

1.4. Expressing emotions in synthetic speech

For computers to mimic emotional speech would help make them appear more human and make conversations with them more natural (Zhou et al., 2018). Existing studies show that addressing affect and emotion in conversational user interfaces or smart speakers can enhance user satisfaction (Prendinger et al., 2005), contribute to a more positive perception of the interaction (Prendinger & Ishizuka, 2005), and lead to fewer breakdowns in conversations (Martinovski & Traum, 2003).

In synthetic speech, there are multiple ways to manipulate the audio. One can adjust the speed (The playback speed of the audio), speaking rate (number of words per minute), pitch (how high or low the sound is), and pauses. In audio recordings, the pitch and speaking rate are influenced by the speed at which the audio is played; when the speed is increased, the speaking rate and the pitch increase as well. Most audio manipulation programs allow the speed to be increased while keeping the pitch the same, ensuring that both the speaking rate and pitch can be manipulated individually.

Previous research (Ravoo, 2019; Zheng, 2010) showed that by manipulating synthetic speech in both pitch and rate, different emotions could be recognized by people. They found that for a lower pitch and rate, the emotion of sadness was recognized, and a higher pitch and rate were associated with happiness (Figure 2). The thesis by Zheng (2010) showed that people rely mainly on the textual component to determine emotion.

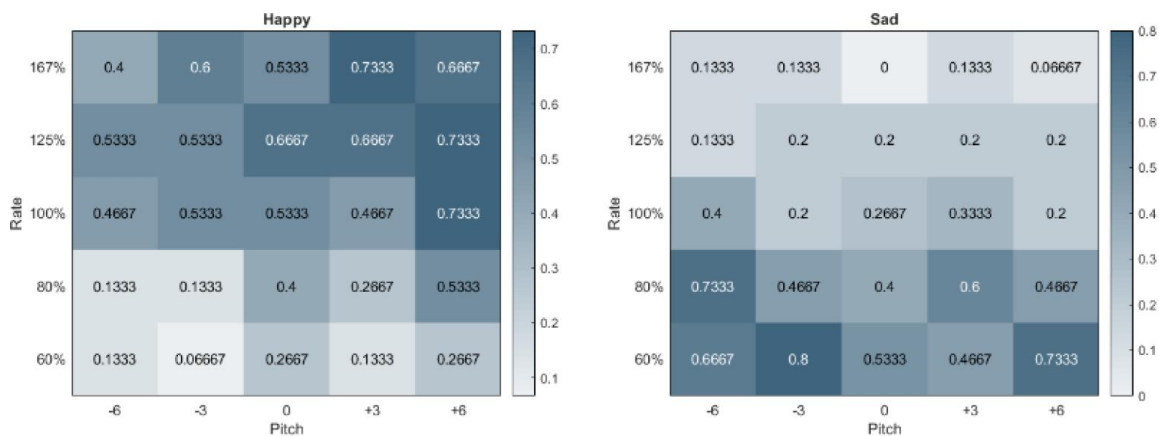


Figure 2, The accuracies of responses in every acoustic form per emotion. The plots are for happy and sad sentences. On the y-axis is the speaking rate, and on the x-axis is the pitch. The darker the color in a block, the higher the percentage of correct responses in this form (Zheng, 2010).

1.5. Attention and memory

Emotional speech has more advantages than just making conversations more natural. It was found (Wood, 1999) that teachers' emotional speech positively affects their students' attention and memory of spoken text. Woods (1999) conducted an experiment testing the attention and memory of students for different teaching methods. The teaching methods consisted of non-emotional speech, emotional speech with random emphasis, uniform emotional speech, and strategic emotional speech. Strategic emotional speech would use emotional speech that fits the context, emphasizing important parts of the teaching session. She assessed students' memory by conducting multiple-choice tests and found that the memory improved significantly when strategic emotional speech was applied in the teaching session.

This indicates that emotional expression influences memory when congruent with the text and illustrates the importance of emphasizing certain parts of the spoken text. This finding is similar to other studies which found that enthusiastic and lively speech positively impacted listeners' attention, motivation, and memory (Jiménez et al., 2021; Liew et al., 2017). If emotional emphasis improves the memory of conversations between people, it would be useful to see if this effect can also be found in a conversation between a person and a smart speaker, thus, whether people would remember more of a conversation with a smart speaker when the emotional speech and text are congruent.

1.6. Research question

From this follows the research question: "*Do people remember what a smart speaker says, and is this influenced by the emotions it expresses?*". The expectation is that people will be able to recall more of the information given to them by a smart speaker when the text and emotion expressed by

the smart speaker are congruent. This hypothesis would align with the findings that strategic emotional speech positively impacts memory. Moreover, it builds upon this research to investigate the effect of emotional congruency. It is expected that congruent emotional speech will be less distracting from the text than incongruent emotional speech and thus will allow people to pay more attention to the text.

It is also essential to check if people like it when a smart speaker expresses emotions because user satisfaction is still the most important. Checking user satisfaction leads to the sub-question: "*Do people like it when a smart speaker expresses emotions?*" The expectation is that people will prefer emotional speech when it is congruent with the text but not when it is incongruent. This hypothesis follows the research that emotional speech in smart speakers was supposed to enhance user satisfaction, but the expectation is that this will not hold when the emotion expressed is incongruent with the text because this would sound off to people and confuse them.

One confounding variable could be if having frequent contact with one influences the attitude toward emotional speech and the smart speaker, which was also examined.

Chapter 2

METHOD

2.1. Design

An online-experiment was conducted with congruence as the independent variable and memory as the dependent variable to answer the research question. The experiment had 2 (happy vs. sad speech) x 2 (a positive vs. a negative story) between participants design. In the congruent scenarios, the emotion in the voice was congruent with the emotion in the story: the happy voice with the happy story and the sad voice with the sad story. The incongruent scenario was where the voice emotions were incongruent with the text: the sad voice with the happy story and the happy voice with the sad story. After the participants listened to the scenario, a questionnaire evaluated how they experienced the voice and tested their recollection.

2.2. Participants

Sixty-two people (28 male, 33 female, and one other; $M_{age} = 26.1$, $SD_{age} = 9.6$, Range 17 to 61) were recruited through the Archie database. Participants were required to have good English comprehension and good hearing. Before testing, an effect size was calculated based on earlier research by Peter Ruijten – Dodoiu which had an $r = 0.25$; this converted is Cohen's $d = 0.516$. With G*power (90% Power, $\alpha = 0.05$), a minimum sample size of 130 participants, 65 for each condition (congruent and incongruent), was calculated (Appendix A). This number was not reached due to time constraints.

2.3. Materials

Participants used their own devices and were asked to use a headset during the experiment. The experiment was conducted via LimeSurvey.

Two-holiday stories were written for this experiment. The holiday setting was chosen because everyone can imagine and emphasize with such a setting. The stories were written in happy and sad

situations, keeping them as similar as possible to ensure that they would be equally difficult to recall and that the questions about the stories could be the same. The duration of the stories was about 6 minutes. Certain parts of the stories were emphasized when they were read out loud by the experimenter. These parts were highlighted and emphasized in the audio file. Both stories can be found in Appendix B. For the text-to-speech conversion, textmagic.com was used with the voice: Englisch (UK)-Phoebe at speed 1x.

Similar to earlier research by Van Tilburg (2021), the pitch and speaking rate were edited in Audacity. Audacity was used to create dynamics in pitch and rate with 10% differences in the highlighted parts of the story. In Appendix B, the full story is added with the highlighted parts that are emphasized. For the happy voice, the entire story was +1.13 semitone with a speaking rate of 107.5% with an emphasis of +1.25 semitones and a speaking rate of 115%. For the sad voice, the entire story was -1.13 semitones with a speaking rate of 92.5%, with an emphasis of -1.25 semitones, and a speaking rate of 85%.

2.4. Measures

Multiple variables and measures were needed to test the hypothesis. A manipulation check was included to see whether the emotions intended in the voice and story were recognized. Furthermore, the emotional speech preference, the experience with smart speakers, and the recall were measured.

The manipulation check consisted of two multiple-choice questions concerning which emotions the participants recognized in the text and speech. The eight answer options were: "exited-lively, cheerful-happy, relaxed-carefree, calm-serene, bored-weary, gloomy-sad, irritated-annoyed, and tense-nervous." In the questionnaire, Figure 3 was visible above the questions; the emotions were illustrated by the pick-a-mood cartoons (Delft Institute of Positive Design, n.d.) in the female edition because the voice used in the scenarios was female. The questionnaire can be found in Appendix C.

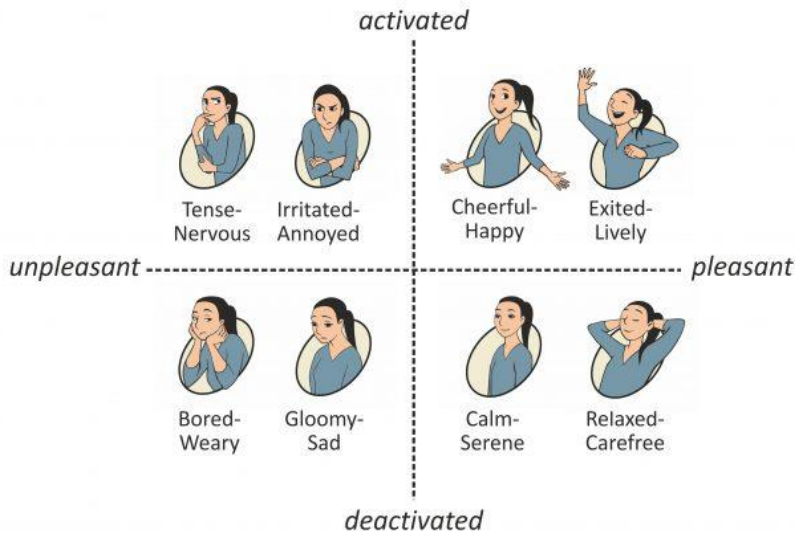


Figure 3, The pick-a-mood cartoon (Delft Institute of Positive Design, n.d.) for mood measuring. (top row, from left to right: Tense-Nervous, Irritated-Annoyed, Cheerfull-Happy, Exited-Lively, Bored-Weary, Gloomy-Sad, Calm-Serene, Relaxed-Carefree.

The emotional speech preference, which was used to test whether there was a difference in how pleasant the voice was experienced between the incongruent and congruent conditions and the role of experience with smart speakers, was measured by a set of three seven-point Likert scale questions (Appendix C) The first question assessed how much they liked the way the voice sounded. The second and third questions assessed how they liked the speech: "Did they like the way the words were emphasized?" and "Did they like that the smart speaker was expressing emotions in its voice?". The answer options ranged from "I did not like it at all" to "I liked it very much."

The degree of experience participants had with smart speakers was assessed with a multiple-choice question set (Appendix C). The experience was used to test: whether the experience with smart speakers did not influence the emotional speech preference. The six questions appraise whether people own a smart speaker (the answer options were yes and no), how often they interact with a smart speaker (the answer options were: never, only once, only a few times, a few times per year, every week, and every day), how many households in their social circle own a smart speaker (the answer options were: none, 1-2, 3-5, 5-10, and 11<), whether they are interested in smart speakers, whether they would want to have a smart speaker, and whether they think smart speakers are useful (the answer options were a 7-point Likert scale ranging from "not at all" to "very").

The recall needed to test the hypothesis was measured with 20 multiple-choice questions about small facts from the scenario (Appendix C). There were always four answer options. One example question is: "How many pools were there?" with options 4, 5, 6, and 7, where 7 was the correct answer.

The questions were tested with a pilot study using a neutral voice until there was 50% to 70% correctness, leaving room for improvement and detriment.

2.5. Procedure

The experiment was conducted online. At the start of the experiment, participants were informed via the computer screen that all collected data would be analyzed anonymously, and their rights to withdraw (without consequences for payment) were explained. By checking a box (Appendix C), they agreed to participate in the experiment.

Participants first listened to an audio fragment that recounted a holiday experience. After hearing the audio fragment, participants were asked to answer a questionnaire of 4 parts, one for each measure (manipulation check, emotional speech preference, experience, and recall). Finally, there was a "thank you" message, and they were asked to enter their name and email if they wanted to participate in the raffle. The questionnaire can be found in Appendix C. The experiment duration was 15 minutes. The compensation was participation in a raffle with a one in ten chance of winning thirty euros.

Chapter 3

RESULTS

The full STATA do-file can be found in Appendix D.

3.1. Descriptive statistics

19% Of the participants own a smart speaker. The participants do, however, not interact with smart speakers often(Figure 4a). Only 10% interact with a smart speaker daily, while 66% report never interacting with a smart speaker (Figure 4a). 21% of the participants are interested in smart speakers (Figure 4c). Interestingly while around 20% of the participants own and are interested in smart speakers, they are not yet genuinely integrated into daily life since only 10% use a smart speaker daily. This low frequency of usage could be linked to the fact that most participants do not think smart speakers are useful (Figure 4e). Not thinking smart speakers are useful likely correlates with not wanting to own a smart speaker (Figure 4d). The participants also reported low integration in their social circle, with 73% reporting that they know less than three households with a smart speaker (Figure 4b).

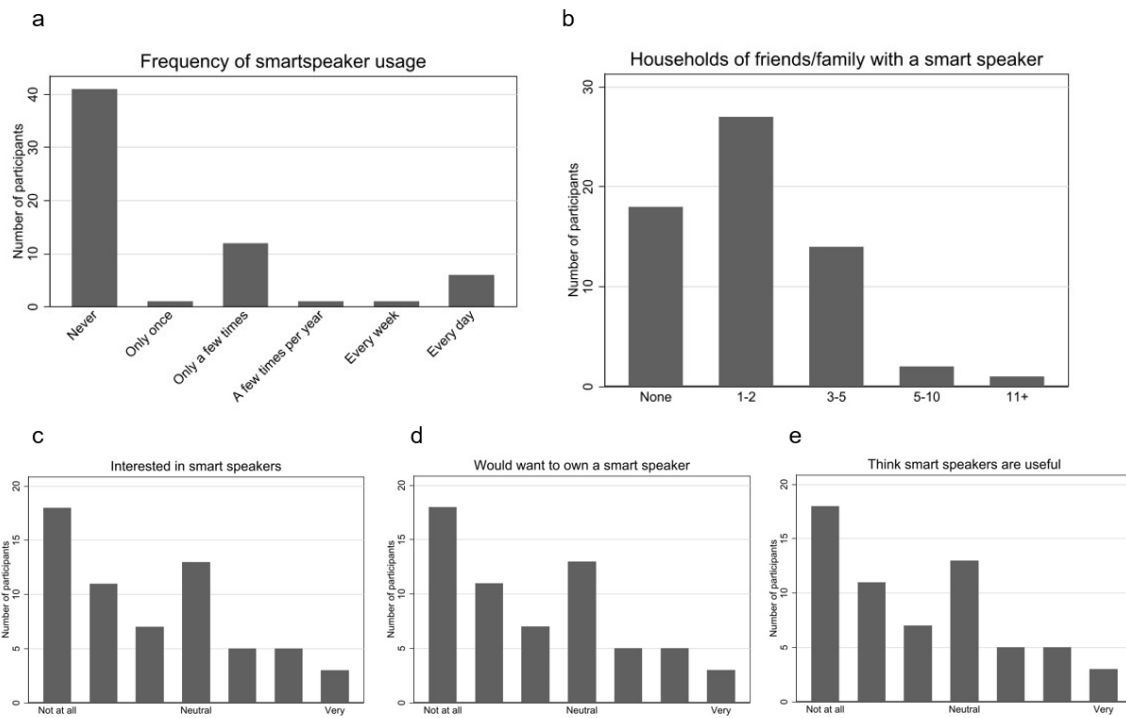


Figure 4, Multiple bar charts assessing how experienced the participants are with smart speakers and their interest in smart speakers. From top left, the bar charts are: a "Frequency of smart speaker usage," b "Households of friends/family with a smart speaker," c "Interested in smart speakers," d "Would want to own a smart speaker," and e "Think smart speakers are useful." On the y-axis is the number of participants.

Participants were randomly distributed over the four scenarios: the happy story with the happy voice (HSHV), the happy story with the sad voice (HSSV), the sad story with the happy voice (SSHV), and the sad story with the sad voice (SSSV). By chance, the HSSV scenario got many more participants (Figure 5). Consequently, the incongruent condition had more participants (Figure 6).

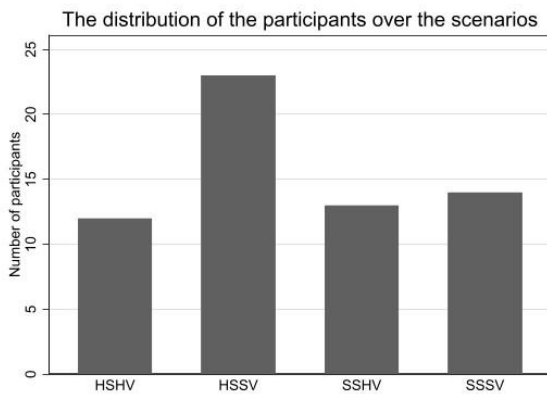


Figure 5, Participant distribution over the four scenarios: happy story with happy voice (HSHV), happy story with sad voice (HSSV), sad story with happy voice (SSHV), and sad story with sad voice (SSSV).

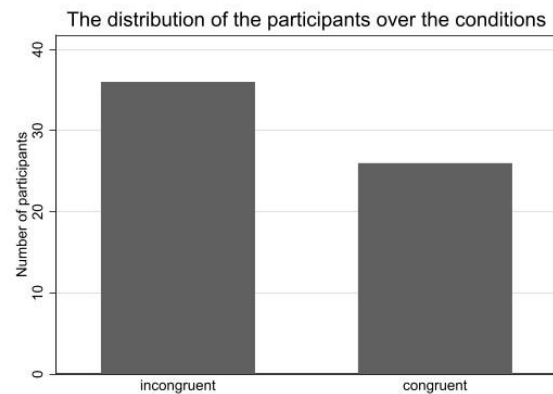


Figure 6, Participant distribution over the two conditions. The incongruent condition consists of the HSSV and SSHV scenarios. The congruent condition consists of the HSHV and SSSV scenarios.

3.2. For each scenario: which emotions were recognized in the text and speech

In the happy story scenarios (happy story happy voice (HSHV) and happy story sad voice (HSSV)), The text was correctly recognized as happy-cheerful by 75% of the participants in the HSHV scenario and 39% of the participants in the HSSV scenario. In the HSSV scenario, the rest of the participants were divided equally over calm-serene, exited-lively, and relaxed-carefree, all positive emotions indicating that the text was recognizable but less clear in combination with the sad voice.

In the sad story scenarios (sad story sad voice (SSSV) and sad story happy voice (SSHV)), 14% of the participants correctly recognized the text as sad in the SSSV scenario. Most participants in the SSSV scenario (43%) appraised the text as irritated-annoyed (Figure 7), which is in the negative direction. In the SSHV scenario, only 8% of the participants correctly recognized the text as sad, while 69% judged the text to be irritated-annoyed (Figure 7), which is still a negative emotion but one with higher arousal in the arousal-valance diagram. This distortion could have been because of the happy voice making it seem that there is a higher arousal level.

In the happy voice scenarios (happy story happy voice (HSHV) and sad story happy voice (SSHV)), the happy voice was not clear for the participants. In the HSHV scenario, only 8% reported the voice as happy-cheerful, and 58% reported the voice as calm-serene (Figure 7), which is at least a positive emotion. In the SSHV scenario, the happy voice was also unclear for the participants, with only 8% reporting the voice as happy-cheerful and 38% reporting the voice as calm-serene (Figure 7), indicating that the happy voice was unrecognizable. This indicates that the audio manipulation was unsuccessful for this scenario, but it might be a good manipulation if the intended emotion were calm-serene.

In the sad voice scenarios (happy story sad voice (HSSV) and sad story sad voice (SSSV)), the sad voice was not apparent to the participants. In the HSSV scenario, none of the participants recognized

the voice as gloomy-sad. Most participants (35%) reported the voice as bored-weary (Figure 7), which is in the negative direction. However, This indicates that the audio manipulation was unsuccessful in combination with the happy story. In the SSSV scenario, the sad voice was more recognizable for the participants, with 29% reporting the voice as gloomy-sad, notwithstanding that most participants (43%) still judged the voice to be bored-weary (Figure 7), indicating that the sad voice was better recognizable in combination with the sad text but still not entirely clear.

In Figure 7, it is also clear that the emotion recognition in the speech for incongruent conditions, the reported emotions range is much broader than for the congruent conditions indicating that the participants found it difficult to separate the emotion in the voice from the emotion in the text.

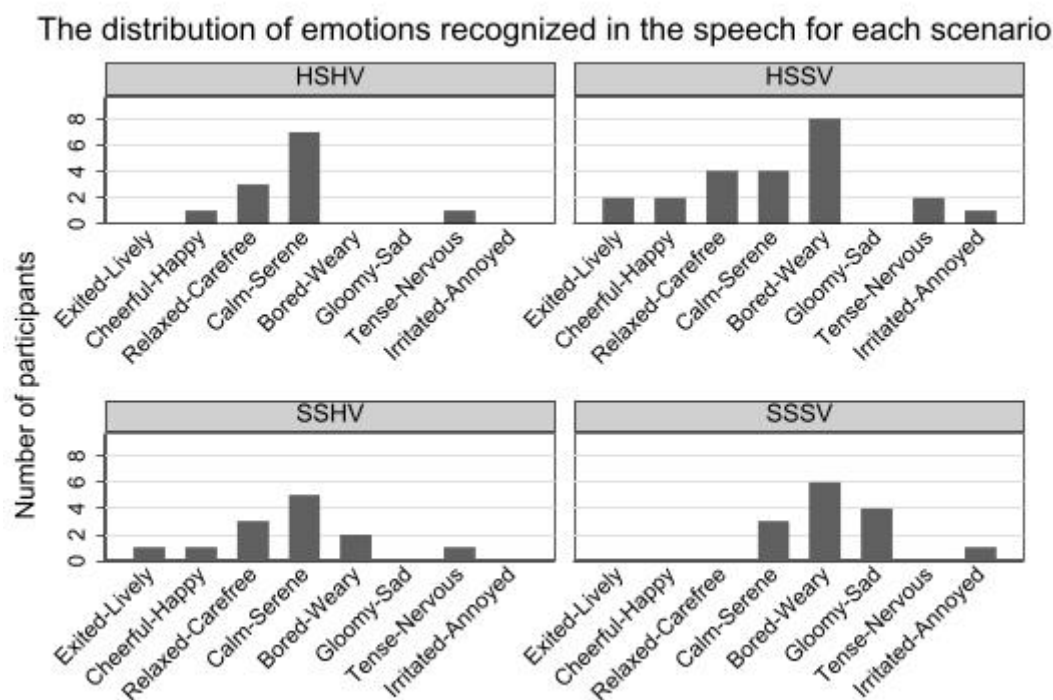
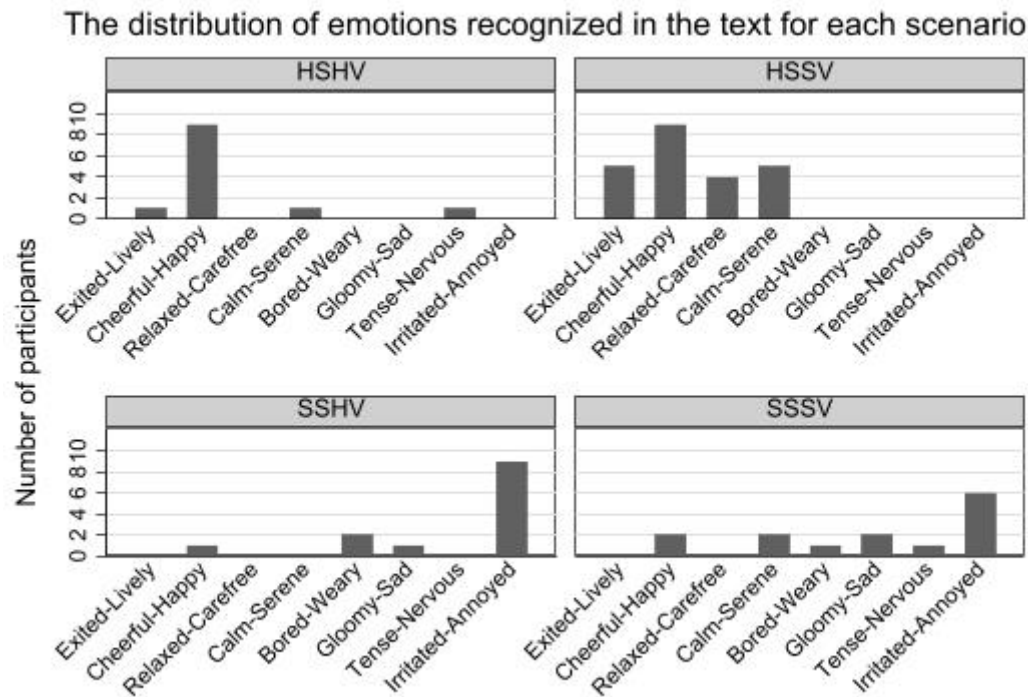


Figure 7, The distribution of the emotions recognized in the text and speech for each scenario. Happy story with happy voice (HSHV), happy story with sad voice (HSSV), sad story with happy voice (SSHV), and sad story with sad voice (SSSV).

3.3. How the participants experienced the voice

In Figure 8, it is shown for each condition how the participants experienced the emotional expression in the voice. The question was: "Did you like that the smart speaker was expressing

emotions in its voice?". It can be seen in Figure 8 that for both conditions, the participants did not like the voice. To test the hypothesis that people will prefer emotional speech when it is congruent with the text but not when it is incongruent. A t-test was conducted to compare the preference for emotional speech between the congruent and incongruent groups. The preference for emotional speech was measured by the question: "Did you like that the smart speaker was expressing emotions in its voice?". The results of this test indicate that there is no statistically significant difference between the mean preference of the incongruent ($M = 4.42$, $SD = 1.76$) and congruent ($M = 4.38$, $SD = 1.44$) groups, $t(60) = < 1$, $p = .94$. The effect size is very small with a cohen's d of 0.02.

This result means that the participants liked the voice the same, whether congruent or not. The result was unexpected since the expectation was that people would like the emotional speech more when it was congruent with the text. The indifference between conditions is a bad sign since it indicates that participants do not care whether the speech is congruent with the story. The indifference could be because they did not register the incongruency.

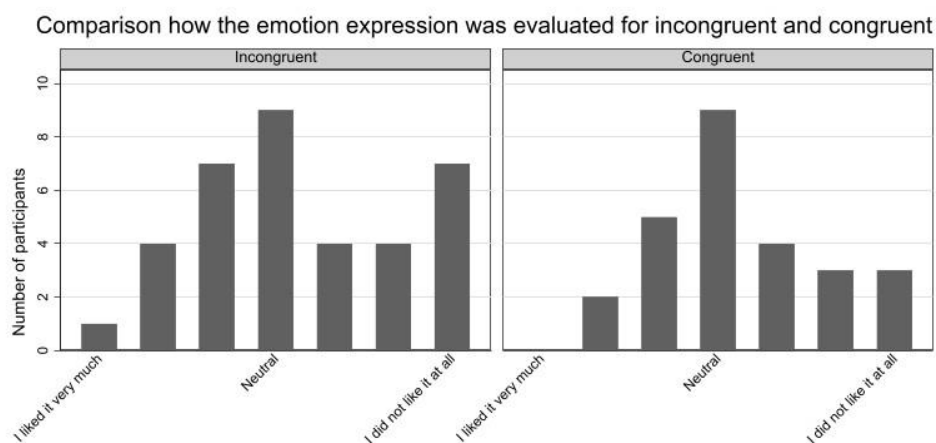


Figure 8, For each condition, how the participants experienced the voice. On the y-axis is the number of participants, and on the x-axis is a 7-point scale ranging from "I liked it very much" to I did not like it at all."

3.4. Hypothesis test

The data on the recall and congruency were submitted to a t-test to test the hypothesis that people will be able to recall more of the information given to them by a smart speaker when the text and emotion expressed by the smart speaker are congruent. The recall variable is the number of correct answers from the 20 memory questions. The congruency depends on the scenario were HSHV and SSSV are congruent, and HSSV and SSHV are incongruent. The 26 participants who were in the congruent group ($M=12.80$, $SD=3.48$) compared to the 36 participants ($M=11.92$, $SD=3.16$) in the incongruent group, demonstrated slightly better recall scores though it was not significant, $t(60)= -1.1$,

$p = .14$ (Figure 9). The effect size was small, with a Cohen's d of -0.28 . The test was not significant, but the higher recall for the congruent group was in line with the hypothesis and could indicate a trend.

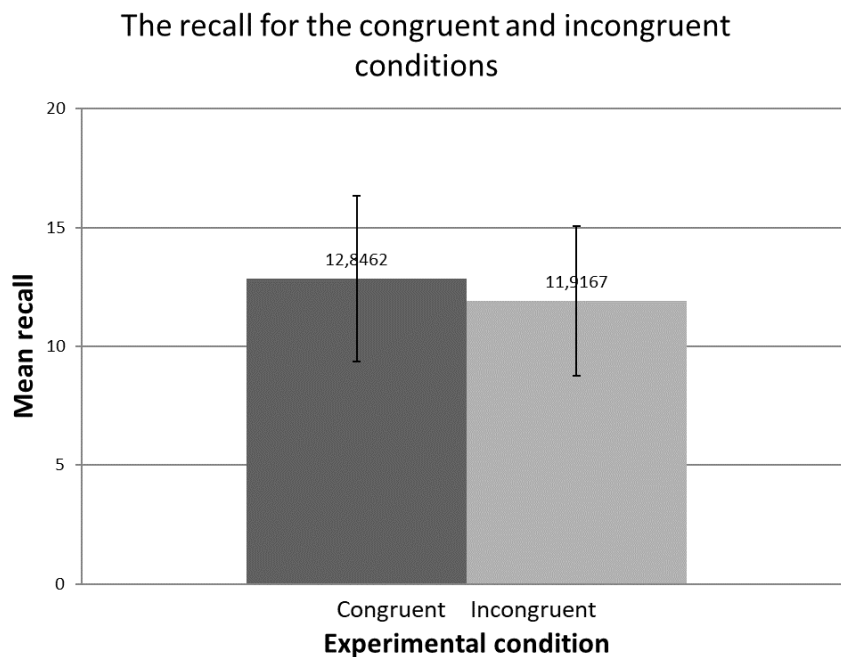


Figure 9, Mean recall for the congruent and incongruent conditions. On the y-axis is the mean number of correct answers from the 20 multiple-choice recall questions. On the x-axis are the two conditions: congruent and incongruent.

A second test was conducted on a filtered set of participants to test whether the insignificant effect of congruency on recall could be explained by ineffective manipulation of the emotions. The emotions were split into positive and negative categories. The positive emotion set consisted of excited-lively, cheerful-happy, relaxed-carefree, and calm-serene. The negative emotion set consisted of bored-weary, gloomy-sad, irritated-annoyed, and tense-nervous. The filtered data included only those who recognized the happy story as positive, the happy voice as positive, the sad story as negative, and the sad voice as negative. Resulting in a sample size of 39 participants (16 males, 22 females, and 1 other; $M_{age} = 26.4$, $SD_{age} = 8.1$, Range 17 to 58).

The 19 participants who were in the congruent group ($M=12.89$, $SD=3.43$), compared to the 20 participants ($M=12.65$, $SD=3.33$) in the incongruent group, demonstrated slightly better recall scores though the difference is now even smaller than before and not significant, $t(37) = -0.23$, $p = .41$ (Figure 10). The effect size was very small, with a Cohen's d of -0.07 .

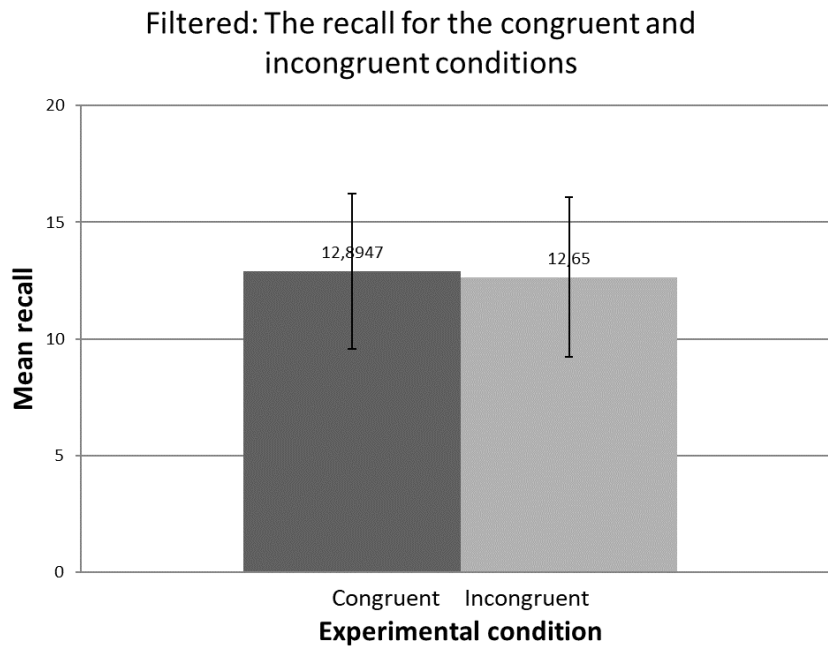


Figure 10, Mean recall for the congruent and incongruent conditions of the filtered participant sample. On the y-axis is the number of correct answers from the 20 multiple-choice recall questions. On the x-axis are the two conditions: congruent and incongruent.

Two t-tests were conducted to see if there was a main effect of the happy and sad emotions. The first is to test the effect of the emotion in the voice, and the second is to test the effect of the emotion in the story on the recall.

In the first t-test, the 25 participants who were in the happy voice group ($M=13.16$, $SD=3.46$), compared to the 37 participants ($M=11.72$, $SD=3.11$) in the sad voice group, exhibited better recall scores though the difference is not statistically significant, $t(60) = -1.70$, $p = .09$ (Figure 11). The effect size was small, with a Cohen's d of -0.44 . Notwithstanding that this effect is insignificant, there does appear to be a trend where the happy voice group has a higher recall than the sad voice group.

In the second t-test, the 35 participants who were in the happy story group ($M=12.71$, $SD=3.29$), compared to the 27 participants ($M=12.48$, $SD=3.37$) in the sad story group, exhibited almost equal recall scores and the difference is not statistically significant, $t(60) = < 1$, $p = .72$ (Figure 11). The effect size was very small, with a Cohen's d of 0.09 . This effect is so small that it is likely not present.

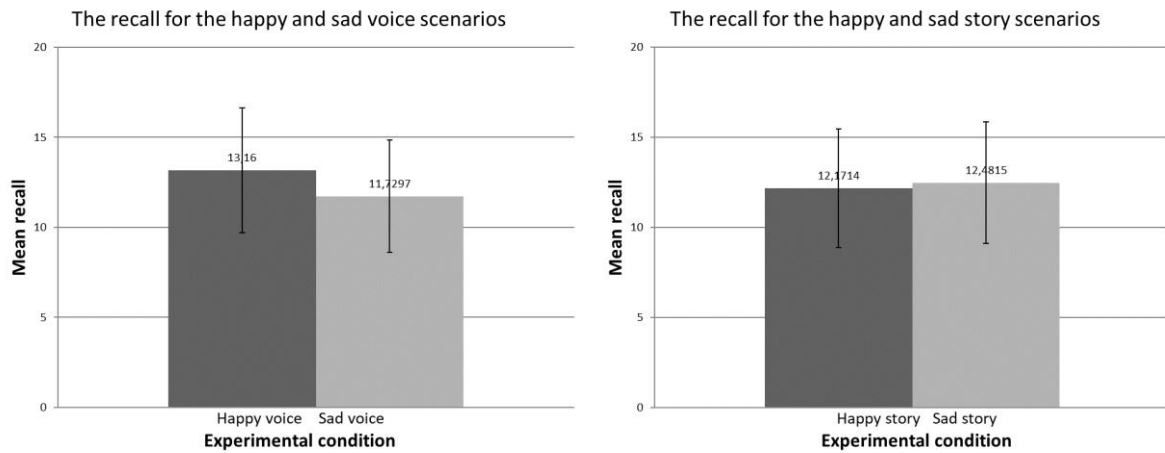


Figure 11, Mean recall for the sad and happy voice and story scenarios. On the y-axis is the number of correct answers from the 20 multiple-choice recall questions. On the x-axis are the two conditions: sad and happy voice or story.

3.5. Exploratory analysis

An exploratory analysis was also done to test if the attitude to emotional speech and the smart speaker is influenced by having frequent contact with one. Multiple regressions were conducted, one on if the experience with a smart speaker influences the preference for emotional speech and the others on how pleasant they experienced the voice. Where the experience is: "How often do you interact with a smart speaker?". The preference for emotional speech is: "Did you like that the smart speaker was expressing emotions in its voice?" How pleasant they experienced it is: "Did you like the way the voice sounded?" and "Did you like the way it emphasized certain words?".

The experience did not predict the attitude. The effect of experience on the preference for emotional speech, $R^2 = .004$, $F(1, 60) = 0.28$, $p = .599$, is not statistically significant. The effect of experience on how much they liked how the voice sounded, $R^2 = .014$, $F(1, 60) = 0.86$, $p = .359$, is also not statistically significant. Similarly, the experience did not predict whether they liked how the words were emphasized, $R^2 = .00$, $F(1, 60) = 0.00$, $p = .958$. From these results, there is no indication that any effect of experience is present on how positively people experience smart speakers speaking.

Chapter 4

DISCUSSION

This study was designed to investigate the effects of congruency between emotions expressed in the voice and the text on people's recollections of the interaction. Based on earlier research, congruent emotional speech was expected to benefit people's recollection compared to incongruent emotional speech. More specifically, a happy story narrated with a happy voice or a sad story narrated with a sad voice was expected to be remembered more accurately in contrast to a happy story narrated with a sad voice or a sad story narrated with a happy voice. The findings did not confirm these expectations. Participants with congruent stories and voices did not remember the interaction more accurately.

There was no effect of congruency on recall. One explanation could be the unsuccessful audio manipulation causing participants not to recognize the intended emotions. This theory was tested by using a selection of participants consisting only of those who recognized the happy emotion as positive and the sad emotion as negative. This filtering resulted in a very small sample size, and no effect was detected with this selection. Because this selection of participants was so small, it might not have had enough power to detect an effect, but there was also no visible trend present, indicating that there might not be an effect.

There was a visible trend when comparing the happy voice with the sad voice scenarios. Even though this trend is not statistically significant, it might indicate a positive effect of happy emotional speech on recall, which would align with research by Jiménez et al. (2021) and Liew et al. (2017). They also found that enthusiastic and lively speech positively impacted listeners' attention, motivation, and memory (Jiménez et al., 2021; Liew et al., 2017). Investigating this effect would be interesting because programming smart speakers with happy voices would benefit the recall if it is accurate.

The results found that participants had a diverse range of emotions identified in the incongruent voices, indicating that it was hard to separate the emotion in the voice from the emotion in the speech,

which would align with what Zheng (2010) found. This difficulty could have been caused by an incorrect audio manipulation where the manipulation did not match the emotion it was supposed to express. Cultural differences between participants could also be a cause of poor recognition. It is unknown how many cultural differences there were for the sample used in this study. Different cultures and languages express emotions differently (Abbasi et al., 2012; Wang et al., 2018). Such differences could play a role in not finding the main effect of emotion in this study. Because people come from different cultures, they have different habits and customs (Abbasi et al., 2012; Huang & Deng, 2008), which could lead to the effect seen in the results where a vast range of emotions was recognized in the speech. There might not be a one size fits all manipulation for the audio, and testing will be required when developing smart speakers for different cultures and languages to ensure a correct expression of emotions. These tests might result in different manipulations for each language and culture.

From the literature, it was expected that around 35% of people would own a smart speaker (*Statista*, 2023). When looking at the participants, only 19% own a smart speaker. Furthermore, only 10% of the participants use their smart speaker daily, indicating that the popularity and usage of smart speakers might not be as high in the Netherlands as in the UK and the US. This difference could be because there is a higher development rate for English smart speakers since there is a larger population of English-speaking people (Ethnologue, 2023). The lower percentage of daily usage likely did not influence the main hypothesis test results since there was no effect of experience with smart speakers on how people assessed the voice. This effect of experience might be present when a sample contains people with more frequent contact with smart speakers.

In addition, it was found that congruency and experience with smart speakers did not affect the preference for emotional speech. The congruent emotional speech was expected to be experienced as more pleasant than the incongruent speech. Not finding this effect might be because people cannot separate the emotion in the voice from the emotion in the text. Thus, they did not register that it was incongruent. An explanation for why the experience with smart speakers is irrelevant to emotional speech preference could be that human interaction is still our primary reference because we interact with humans more often. Therefore people are still using human-to-human communication as a reference since they are not so familiar with smart speakers that they find such speech natural, resulting in no differences because of experience.

Participants did not like the sound of the voice. Maybe because it did not sound natural enough with the audio manipulations, or there were too few strategic pauses included to emphasize words and signal sentence structures. A closer look should be taken at how people experience voices, and more aspects need focus than only pitch and speaking rate to make a smart speaker's voice more natural and pleasant.

4.1. Limitations and future research

It was calculated that 130 participants would be needed to reach 90% power, but only 62 people participated. Consequently, there is less certainty of the results because effects could easily be missed due to too little power. Still, even though there was little statistical power, there was a trend visible that the recall was better in the congruent condition. Future research would need more power to know whether there is an effect.

Another limitation is that it was still possible for the participants to speed up the audio. Speeding up increased the speaking rate and pitch of the audio fragment. This possibility of speeding up the audio created uncertainty about whether participants heard the fragment as intended and distorted the accuracy of the data. Some participants reported that they increased the audio speed. The option to speed up was hidden; how many participants did this is unknown. In the future, it would be wise to keep this in mind and ensure that speeding up the audio is impossible.

The emotional manipulation was unsuccessful because participants often recognized different emotions in the voice and story as intended. This caused the results to become unreliable. By filtering the participants to include those for whom the manipulation worked, it was still possible to see the undistorted results, although with less power. In future research, more pilot testing should ensure that the right emotions are recognized in the voice and the story. The difference in emotions recognized in the voice could also be because of cultural differences where people from different cultures or languages use other methods of expressing emotions (Laukka et al., 2014; Wang et al., 2018). It would be interesting for future research to see how different cultures and languages express emotions and to implement this in smart speakers for those languages and cultures.

The distribution between the scenarios and, thus, the conditions was not even, with more participants allocated to the incongruent condition with the HSSV scenario. This uneven distribution likely did not significantly influence the results, but it should be remembered when designing future studies that a different randomization method might better ensure an equal distribution between scenarios.

The emphasized pieces of the story were selected based on where the experimenter would emphasize the story when reading it out loud. This selection technique might not have selected the right parts causing the audio dynamic not to be experienced as intended. Still, this research remains one of the first that focuses on dynamics in artificial speech in this way. In the future, more focus on the placement and efficiency of the dynamics is needed.

4.2. Conclusion

Although no significant results were found, it should be remembered that emotional speech is vital in our lives. Moreover, that conversation would become rather dull without it. Technology has become ubiquitous, and interactions with it must be as natural as possible; emotional speech is one aspect of making interaction natural and is essential because it conveys context in conversations. Hopefully, this paper will contribute to making interactions with smart speakers as pleasant as possible.

REFERENCES

- Abbasi, G., Zadeh, S. S., Janfaza, E., Assemi, A., & Dehghan, S. S. (2012). Language , Translation , and Culture. *International Conference on Language, Medias and Culture*, 33(IACSIT Press, Singapore).
- AudienceProject. (2020). *Share of individuals who have access to a smart speaker in their household in 2020, by country*. Statista. <https://www.statista.com/statistics/1107880/access-to-smart-speaker-in-households-worldwide/?locale=en>
- Barnes, S. B. (2010). *User Friendly : A Short History of the Graphical User Interface User Friendly : A Short History of the Graphical User Interface*. 16(1).
- Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2006). The experience of emotion. *Annual Review of Psychology*, 58, 373–403. <https://doi.org/10.1146/annurev.psych.58.110405.085709>
- Barrett, L. F., & Russell, J. A. (1998). Independence and bipolarity in the structure of current affect. *Journal of Personality and Social Psychology*, 74(4), 967–984. <https://doi.org/10.1037/0022-3514.74.4.967>
- Barrett, L. F., & Wager, T. D. (2006). The structure of emotion evidence from neuroimaging studies. *Current Directions in Psychological Science*, 15(2), 79–83. <https://doi.org/10.1111/j.0963-7214.2006.00411.x>
- Beauregard, M., Leroux, J. M., Bergman, S., Arzoumanian, Y., Beaudoin, G., Bourgouin, P., & Stip, E. (1998). The functional neuroanatomy of major depression: An fMRI study using an emotional activation paradigm. *NeuroReport*, 9(14), 3253–3258. <https://doi.org/10.1097/00001756-199810050-00022>
- Caldarini, G., & Jaf, S. (2022). *A Literature Survey of Recent Advances in Chatbots*.
- Davidson, R. J., & Irwin, W. (1999). The Functional Neuroanatomy of Emotion and Affective Style. *Foundations in Social Neuroscience*, 3(1), 84–94. <https://doi.org/10.7551/mitpress/3077.003.0035>
- Davis, M., & Whalen, P. J. (2001). The amygdala: Vigilance and emotion. *Molecular Psychiatry*, 6(1), 13–34. <https://doi.org/10.1038/sj.mp.4000812>
- Delft Institute of Positive Design. (n.d.). *Pick-A-Mood pictorial tool for mood measurement*. Retrieved May 15, 2023, from <https://diopd.org/pick-a-mood/>
- Ethnologue. (2023). *The most spoken languages worldwide in 2023 (by speakers in millions)*. Statista. <https://www.statista.com/statistics/266808/the-most-spoken-languages-worldwide/>

- Gerber, A. J., Posner, J., Gorman, D., Colibazzi, T., Yu, S., Wang, Z., Kangarlu, A., Zhu, H., Russell, J., & Peterson, B. S. (2008). An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces. *Neuropsychologia*, *46*(8), 2129–2139. <https://doi.org/10.1016/j.neuropsychologia.2008.02.032>
- Han, H. (2019). *Green indoor and outdoor environment as nature - based solution and its role in increasing customer / employee mental health , well - being , and loyalty*. September 2018, 629–641. <https://doi.org/10.1002/bse.2269>
- Huang, K. H., & Deng, Y. S. (2008). Social Interaction Design in Cultural Context : A Case Study of a Traditional Social Activity. *International Journal of Design*, *2*(2), 81–96.
- Jiménez, Y., Vivanco, O., Castillo, D., Torres, P., & Jiménez, M. (2021). Artificial Intelligence in Neuroeducation : The Influence of Emotions in the Learning Science. *Advances in Intelligent Systems and Computing Book Series*, *1277*, 67–77. <https://doi.org/10.1007/978-3-030-60467-7>
- Jonze, S. (2013). *Her*. Warner Bros. Pictures.
- Kopplin, J. (2002). *An Illustrated History of Computers Part 4*. <http://www.computersciencelab.com/ComputerHistory/HistoryPt4.htm>
- Ladefoged, P. N. (2023). phonetics. *Encyclopedia Britannica*. <https://www.britannica.com/science/phonetics>
- Lades, L. K., Laffan, K., Daly, M., & Delaney, L. (2020). *Brief Report COVID-19 Daily emotional well-being during the COVID-19 pandemic*. 902–911. <https://doi.org/10.1111/bjhp.12450>
- Lane, R. D., Chua, P. M. L., & Dolan, R. J. (1999). Common effects of emotional valence, arousal and attention on neural activation during visual processing of pictures. *Neuropsychologia*, *37*(9), 989–997. [https://doi.org/10.1016/S0028-3932\(99\)00017-2](https://doi.org/10.1016/S0028-3932(99)00017-2)
- Laukka, P., Neiberg, D., & Elfenbein, H. A. (2014). Evidence for cultural dialects in vocal emotion expression: Acoustic classification within and across five nations. *Emotion*, *14*(3), 445–449. <https://doi.org/https://doi.org/10.1037/a0036048>
- LeDoux, J. (2003). The emotional brain, fear, and the amygdala. *Cellular and Molecular Neurobiology*, *23*(4–5), 727–738. <https://doi.org/10.1023/A:1025048802629>
- Liew, T. W., Azan, N., Zin, M., & Sahari, N. (2017). Exploring the affective , motivational and cognitive effects of pedagogical agent enthusiasm in a multimedia learning environment. *Human-Centric Computing and Information Sciences*. <https://doi.org/10.1186/s13673-017-0089-2>
- Lister, K., Coughlan, T., Iniesto, F., Freear, N., & Devine, P. (2020). Accessible conversational user interfaces: Considerations for design. *Proceedings of the 17th International Web for All Conference, W4A 2020*. <https://doi.org/10.1145/3371300.3383343>
- Martinez, W. L. (2011). Graphical user interfaces. *Wiley Interdisciplinary Reviews: Computational Statistics*, *3*(2), 119–133. <https://doi.org/10.1002/wics.150>
- Martinovski, B., & Traum, D. (2003). The Error is the Clue: Breakdown in Human-Machine Interaction. *Proceedings of the ISCA Workshop on Error Handling in Spoken Dialogue Systems*, 11–17.
- Murphy, F. C., Nimmo-smith, I. A. N., & Lawrence, A. D. (2003). Functional neuroanatomy of emotions: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, *3*(3), 207–233.
- Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of

- emotion: A meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, 16(2), 331–348. <https://doi.org/10.1006/nimg.2002.1087>
- Prendinger, H., & Ishizuka, M. (2005). The empathic companion: A character-based interface that addresses users' affective states. *Applied Artificial Intelligence*, 19(3–4), 267–285. <https://doi.org/10.1080/08839510590910174>
- Prendinger, H., Mori, J., & Ishizuka, M. (2005). Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. *International Journal of Human Computer Studies*, 62(2), 231–245. <https://doi.org/10.1016/j.ijhcs.2004.11.009>
- Ravoo, B. (2019). *Affect in synthetic speech | Exploring fine-grained effects of speaking pitch & rate in a lab experiment*. <https://research.tue.nl/en/studentTheses/affect-in-synthetic-speech>
- Reiman, E. M., Lane, R. D., Ahern, G. L., Schwartz, G. E., Davidson, R. J., Friston, K. J., Yun, L. S., & Chen, K. (1997). Neuroanatomical correlates of externally and internally generated human emotion. *American Journal of Psychiatry*, 154(7), 918–925. <https://doi.org/10.1176/ajp.154.7.918>
- Russel, J. . (1980). A circumplex model of affect. In *Journal of Personality and Social Psychology* (Vol. 39, Issue 6, pp. 1161–1178).
- Russell, J. A. (2003). Core Affect and the Psychological Construction of Emotion. *Psychological Review*, 110(1), 145–172. <https://doi.org/10.1037/0033-295X.110.1.145>
- Russell, J. A., & Bullock, M. (1985). Multidimensional Scaling of Emotional Facial Expressions. Similarity From Preschoolers to Adults. *Journal of Personality and Social Psychology*, 48(5), 1290–1298. <https://doi.org/10.1037/0022-3514.48.5.1290>
- Sprengelmeyer, R., Rausch, M., Eysel, U. T., & Przuntek, H. (1998). Neural structures associated with recognition of facial expressions of basic emotions. *Proceedings of the Royal Society B: Biological Sciences*, 265(1409), 1927–1931. <https://doi.org/10.1098/rspb.1998.0522>
- Statista. (2023). Consumer Electronics - Worldwide. <https://www.statista.com/outlook/cmo/consumer-electronics/worldwide?currency=EUR>
- Sun, Y., & Lin, S. (2020). *University Students' Perceived Peer Support and Experienced Depressive Symptoms during the COVID-19 Pandemic: The Mediating Role of Emotional Well-Being*.
- Tilburg, V. (2021). *The Influence of Manipulated Pitch Contour and Speaking Rate of Synthesized Speech on Perceived Affect*.
- Wang, T., Lee, Y. C., & Ma, Q. (2018). Within and across-language comparison of vocal emotions in Mandarin and English. *Applied Sciences (Switzerland)*, 8(12). <https://doi.org/10.3390/app8122629>
- Whalen, P. J., Bush, G., McNally, R. J., Wilhelm, S., McInerney, S. C., Jenike, M. A., & Rauch, S. L. (1998). The emotional counting stroop paradigm: A functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biological Psychiatry*, 44(12), 1219–1228. [https://doi.org/10.1016/S0006-3223\(98\)00251-0](https://doi.org/10.1016/S0006-3223(98)00251-0)
- Wood, A. M. (1999). The effects of teacher enthusiasm on student motivation, selective attention, and text memory. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 59(9-A), 3355. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=psyc3&NEWS=N&AN=1999-95005-009>

-
- Yang, H., & Ma, J. (2020). How an Epidemic Outbreak Impacts Happiness : Factors that Worsen (vs . Protect) Emotional Well-being during the Coronavirus Pandemic. *Psychiatry Research*, 289(April), 113045. <https://doi.org/10.1016/j.psychres.2020.113045>
- Zemčík, T. (2019). A Brief History of Chatbots A Brief History of Chatbots. *DEStech Transactions on Computer Science and Engineering*, October. <https://doi.org/10.12783/dtcse/aicae2019/31439>
- Zheng, Y. (2010). *Linking words and tones How semantic-acoustic congruency influences emotion identification in synthetic speech*. 1406027.
- Zhou, H., Huang, M., Zhang, T., Zhu, X., & Liu, B. (2018). Emotional chatting machine: Emotional conversation generation with internal and external memory. *32nd AAAI Conference on Artificial Intelligence, AAAI 2018*, 730–738. <https://doi.org/10.1609/aaai.v32i1.11325>

APPENDIX

Appendix A, Sample size analyses

t tests – Means: Difference between two independent means (two groups)

Analysis: A priori: Compute required sample size

Input:	Tail(s)	=	One
	Effect size d	=	0.52
	α err prob	=	0.05
	Power (1- β err prob)	=	0.9
	Allocation ratio N2/N1	=	1
Output:	Noncentrality parameter δ	=	2.9644561
	Critical t	=	1.6568452
	Df	=	128
	Sample size group 1	=	65
	Sample size group 2	=	65
	Total sample size	=	130
	Actual power	=	0.9038623

Appendix B, Holiday scenario text

**Highlighted parts were emphasized in the audio*

Holiday scenario – Happy Story

I had **the best time** on vacation this year. The weather was **amazing**, with **28** degrees and sun during the day, and **14** degrees at night, ensuring that we had the best experience of a summer day, without sweaty nights. The hotel was nice too, with **7 pools, and 5 waterslides**. This was great for our two kids; **Jenny** who is **6**, and **Michael** who is **8**. They thoroughly enjoyed themselves outside in the pools and on the slides. We spent 3 days at the pools and went exploring the other days. One day we went to visit the nearby town, which had a **castle** from the **17th century** and a museum of Renaissance art. It was **wonderful**, and while the kids enjoyed roleplaying as knights we could look our fill. After we visited the museum, and while the kids sat exhausted in their strollers we had some time to look at the beautiful art.

Another day trip we made was to visit the beach, which was nice, except it was not a sand beach, so the stones hurt our feet when we wanted to go swimming. I recommend water shoes, for anyone also considering visiting this beach. The beach bar did make great cocktails, particularly the one with fresh orange juice. They also sold snacks and ice cream, which is great when spending a day at the beach with kids. The kids enjoyed their cornetto while we sipped our cocktails. The only downside was that we forgot the sunscreen at the hotel, so we had to buy a bottle at the convenience store near the beach. But this was only a minor problem since it was rather cheap.

On the 3rd day, we began to worry about our transfer from the hotel back to the airport. The transfer from the airport to the hotel went fine, because there were kiosks of different travel agencies ready to receive us and tell us that we were supposed to get bus number 10 and that it was at the 4th dock. When we got to the bus we discovered that this bus makes multiple stops to bring all the tourists to their respective hotels. Before entering the bus the driver asked us which hotel we were headed to. Because we booked the vacation so long ago we did not immediately remember but we were not the only ones who had this problem. More people were looking at their phones and searching for their reservations. We all had a good laugh at how poorly prepared we were and the bus driver seemed a bit annoyed. But we were all on vacation and the most stressful part of the journey was over so most of the tourists were quite relaxed. We all found the names of our hotels, ours was Outlook Hotel. On our way we passed many beautiful hotels, such as, Eden, Baobab, Barcelo, Three corners, Bitzaro, and Galeri.

While we did receive an envelope from the kiosk employee with some travel times for the busses there were no travel times for the Wednesdays which was the day we were leaving. Because of this, we were a bit confused as to which bus would pick us up and at what time. We had booked the vacation with Sunsite originally but all the documents we later received were from Shauinsland. Since Shauinsland was a foreign company the documents were not in English and we were not able to read them. Luckily they were pretty straight forward and we were able to find most of the information on the Sunsite website. But for the transfer back nothing was mentioned on the website so we decided to call. When we called them they referred us to the Shauinsland customer service. There we got a customer service employee who did not speak English fluently since it is a company in a foreign country. But eventually, we were able to communicate that we would like to know our transfer time. They were luckily very helpful and told us that there would be a bus at 12:30 to pick us up. They also told us that we should check this time two days before departure at the my-Shauinsland site.

Because of the great help, we were relaxed again and they wished us a pleasant Holliday. We were happy that this was quickly fixed with a phone call and that there were no difficulties. We walked along the boulevard for the rest of the day and visited the shops.

Upon returning to the hotel we opted to make use of the buffet and discovered that they had a wide array of foods which were all very tasty. They split the buffet into different sections for different types of food. There was an Italian section with pizza and pasta which the kids loved. There was a seafood section, a grill section with BBQ food which was my absolute favorite, a Mexican section, and a Greek section, my husband thoroughly enjoyed the Asian part of the buffet where he could request the food to be made as spicy as he wanted. After finishing the main course we also visited the dessert section of the buffet which served different ice creams, pastries, pancakes and waffles. It also had a chocolate fountain which the kids loved but were too small to reach so we had to help them. But this was okay and also ensured their safety so they could not accidentally burn themselves on the hot chocolate. The buffet had quite an array of foods that could be dipped in the hot chocolate but my kids preferred the strawberries.

After dessert, we often visited the kids' disco that was organized each evening between 6:30 and 8:00 but the kids were tired after long days so we always left early at 7:00 to put them in their beds. We had a suite setup of two-bed chambers one common area and one bathroom. This was ideal when traveling with kids because they are still close by while the parents get their much-needed alone time too. I would certainly recommend this vacation!

Holiday scenario – Sad Story

We had a miserable experience on vacation this year. The weather was disappointing, with 28 degrees and relentless rain during the day, and 14 degrees at night, ensuring we had no time to enjoy the outdoors. The hotel had 7 pools, and 5 waterslides however because of the rain we did not get to use them much. They would have been great for our two kids; Jenny who is 6, and Michael who is 8. But sadly they were stuck inside. We spent 3 days at the hotel playing board games and went on misadventures the other days. One day we went to visit the nearby town, which had a castle from the 17th century, and a museum of Renaissance art. They looked nice but were closed because of the weather.

Another day trip we made was to visit the beach, which would have been nice, except it was not a sand beach, so the stones hurt our feet when we wanted to go swimming. I recommend water shoes, for anyone also considering visiting this beach. The beach bar did make great cocktails, particularly the one with fresh orange juice. They also sold snacks and ice cream, which is great when spending a day at the beach with kids, even when it rains. The kids got Cornettos while we sipped our cocktails. Jenny was playing while holding the cornetto and slipped dropping the ice cream on the ground. We could not get her a new one because the shop had closed for the midday break. After this, we discovered

that we had forgotten the sunscreen at the hotel, so we had to buy a bottle at the convenience store near the beach. They did not have any **ice cream** however so Jenny was **inconsolable all day, putting a damper on all our moods.**

On the 3rd day, we began to worry about our transfer from the hotel back to the airport. The transfer from the airport to the hotel went fine because there were **kiosks** of different travel agencies ready to receive us and tell us that we were supposed to get bus **number 10** and that it was at the 4th dock. When we got to the bus we discovered that this bus makes **multiple stops** to bring all the tourists to their respective hotels. Before entering the bus the driver asked us **which hotel** we were headed to. Because we booked the vacation **so long ago** we did not immediately remember but we were **not the only ones** who had this problem. More people were looking at their phones and searching for their reservations. We all were regretful of how poorly prepared we were and the bus driver seemed a bit annoyed. But we were all on vacation and the most stressful part of the journey was over so most of the tourists were quite relaxed if a bit **down** because of the long travel time. We all eventually found the names of our hotels, ours was **Outlook Hotel**. On our way we passed **many beautiful hotels**, such as, Eden, Baobab, Barcelo, Three corners, Bitzaro, and Galeri.

While we **did receive an envelope** from the kiosk employee with some travel times for the busses there were no travel times for the **Wednesdays** which was the day we were leaving. Because of this, we were a bit confused as to **which bus** would pick us up and at **what time**. We had booked the vacation with **Sunsite** originally but all the documents we later received were from **Shauinsland**. Since Shauinsland was a foreign company the documents were not in English and we were **not able to read them**. Then we had been able to find most of the information on the Sunsite website. But for the transfer back nothing was mentioned on the website so we decided to call. When we called them they referred us to the Shauinsland customer service. There we got a customer service employee who **did not speak English** fluently since it is a company in a foreign country, which made us despair a bit. But **eventually**, we were able to communicate that we would like to know our **transfer time**. They were luckily **very helpful** and told us that there would be a bus at 12:30 to pick us up. They also told us that we should check this time **two days** before departure at the my-Shauinsland site.

Because of the help, we were relaxed again and they wished us a pleasant Holliday. We were only a **bit gloomy** because the employee had been a bit annoyed with us and told us that we could have found the information on the site **but because all the documents were not in English we had not been able to read them and could not have known this.** We were relieved that this was quickly fixed with a phone call. We walked along the boulevard for the rest of the day and visited the shops.

Upon returning to the hotel we opted to make use of the buffet and discovered that they had a wide array of foods which were all a bit disappointing and cold. They split the buffet into different sections for different types of food. There was an Italian section with pizza and pasta, which the kids at least ate. There was a seafood section, a grill section with BBQ food which was my favorite, a Mexican section, and a Greek section, my husband preferred the Asian part of the buffet where he could request the food to be made as spicy as he wanted. But we all longed for home and our favorite restaurants. After finishing the main course we also visited the dessert section of the buffet which served different ice-creams, pastries, pancakes, and waffles. It also had a chocolate fountain which the kids loved but were too small to reach so we had to help them. This was upsetting for the kids and anyone in the vicinity. The buffet had quite an array of foods that could be dipped in the hot chocolate but my kids preferred the strawberries.

After dessert, we often visited the kids' disco that was organized each evening between 6:30 and 8:00 but the kids were tired after long days so we always left early at 7:00 to put them in their beds. We had a suite setup of two-bed chambers one common area and one bathroom. This was ideal when traveling with kids because they are still close by while the parents get their much-needed alone time too. Our vacation was disappointing because of the weather and I would not recommend it. We were all glad to go home.

Appendix C, Questionnaire

Emotional speech for smart speakers.

Welcome to this study,

Please use a headset and make sure that your sound is on during this study!

This study will take approximately 15 minutes to complete.

There are 38 questions in this survey.

Informed consent

Please read the following carefully.

Information form for participants

This study is performed by Daphne Debije, a master student under the supervision of Peter Ruijten - Dodoiu of the Human-Technology Interaction group at Eindhoven University of Technology.

Before participating, you should understand the procedure followed in this study, and give your informed consent for voluntary participation. Please read this page carefully.

About this study

This study has the goal of finding improvements in interactions with smart speakers. You will hear a smart speaker recite a holiday experience after this there will be a questionnaire to assess how you experienced this.

This study will take 15 minutes to complete and does not involve any risks.

Voluntary Participation

Your participation is completely voluntary. You can stop participation at any time but you will only receive payment if you complete the questionnaire. You can also withdraw your permission to use your data immediately after completing the study. You will participate in a raffle in which 1 in every 10 participants receives 30 euros.

Confidentiality and use, storage, and sharing of data

This study has been approved by the Ethical Review Board of Eindhoven University of Technology. In this study, personal data (your age, and gender) and experimental data (your responses to questionnaires) will be stored. To protect your privacy, all data that can be used to personally identify you will be stored on an encrypted server of the Human-Technology Interaction group for at least 10 years. All other information (your participant database ID, and email address) will be deleted once everyone has been compensated. The anonymized dataset that, to the best of our knowledge and ability will not contain information that can identify you, will be made publicly available.

Further information

If you want more information about this study, the study design, or the results, you can contact Daphne Debije (contact email: d.j.debije@student.tue.nl). You can report irregularities related to scientific integrity to confidential advisors of the TU/e, whose contact information can be found on www.tue.nl (<http://www.tue.nl>).

Certificate of consent

By starting this study,

I indicate that I have read and understood the study procedure, and I agree to voluntarily participate. *

🗳️ Please select one answer

Please choose **all** that apply:

- Yes
 No

I also give permission to make my anonymized recorded data available to others in a public online data repository. *

🗳️ Please select one answer

Please choose **all** that apply:

- Yes
 No

Smart speaker holiday scenario

Please listen to this scenario which is a smart speaker narrating a review of a holiday.

{if(is_empty(randnumber,NAOK),rand(1,4),randnumber,NAOK)}

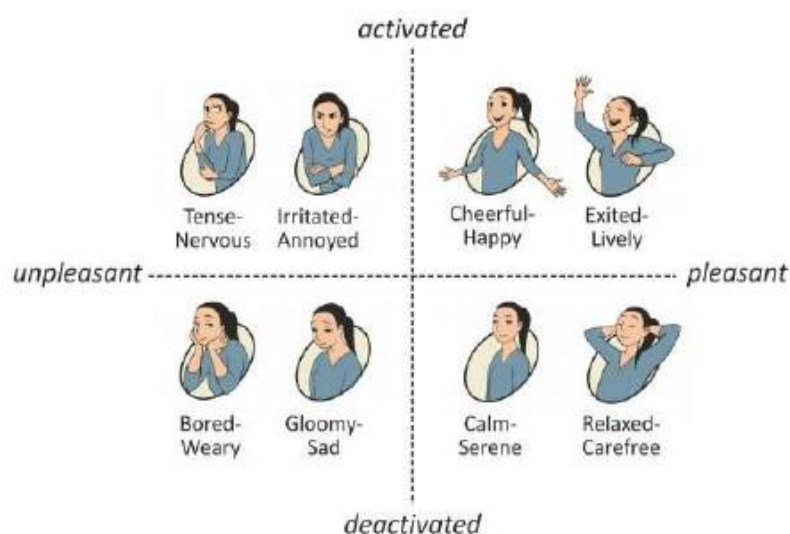
Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) *and* Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

Emotion recognition

This part is to find out which emotions you recognized in the text and in the speech.

Please try to **separately** judge the emotions in the text and the speech.



Which emotion did you recognize in the text?

*

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

● Please select one answer

Please choose **all** that apply:

- Tense - Nervous
- Irritated - Annoyed
- Cheerful - Happy
- Exited - Lively
- Bored - Weary
- Gloomy - Sad
- Calm - Serene
- Relaxed - Carefree

Pick the emotion that fits the best with the text disregarding the voice.

Which emotion did you recognize in the speech?

*

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

● Please select one answer

Please choose **all** that apply:

- Tense - Nervous
- Irritated - Annoyed
- Cheerful - Happy
- Exited - Lively
- Bored - Weary
- Gloomy - Sad
- Calm - Serene
- Relaxed - Carefree

Pick the emotion that fits the best with the speech disregarding the text.

Emotional speech

Please answer the following questions

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate,) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository,)

i Please select 3 answers

Please choose the appropriate response for each item:

	I did not like it at all			Neutral			I liked it very much
Did you like the way the voice sounded?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Did you like the way it emphasized certain words?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Did you like that the smart speaker was expressing emotions in its voice?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Experience with smart speakers

Do you have a smart speaker at home?

*

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate,) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository,)

i Please select one answer

Please choose **all** that apply:

- Yes
 No

*

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate,) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository,)

❗ Please select one answer

Please choose the appropriate response for each item:

	Never	Only once	Only a few times	A few times per year	Every week	Every day
How often do you use a smart speaker?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

*

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose the appropriate response for each item:

	none	1-2	3-5	5-10	11<
How many households of your friends/family have a smart speaker that you know of?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

*

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate,) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository,)

i Please select 3 answers

Please choose the appropriate response for each item:

	Not at all			Neutral			Very
Are you interested in smart speakers?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Would you want to have a smart speaker?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Do you think smart speakers are useful?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Memory

We will now ask a few questions to measure how much you remember of the holiday review.

How many pools were there? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate,) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository,)

i Please select one answer

Please choose **all** that apply:

- 4
- 5
- 6
- 7

What were the maximum and minimum temperatures? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

- 28 - 14
- 24 - 16
- 26 - 14
- 28 - 18

How many waterslides were there? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

- 4
- 5
- 6
- 7

Why did she recommend water shoes? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

- Hurting stones
- Slippery stones
- Sea urchin
- Hot sand

What did they forget to bring to the beach? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

- Water shoes
- Sunscreen
- Towels
- Swimsuit

Which buffet did she most enjoy? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

- Italian
- Asian
- Grill
- Seafood

What did the kids prefer to dip in the chocolate? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

- Cookies
- Waffles
- Strawberry
- Banana

At what time did they leave the kids' disco? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate,) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository,)

❗ Please select one answer

Please choose **all** that apply:

6:30

7:00

7:30

8:00

Which century was the castle built? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

15th

16th

17th

18th

What museum did they visit?

*

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) *and* Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

● Please select one answer

Please choose **all** that apply:

- Romantic
- Renaissance
- Romanesque
- Realism

What were the names of the kids? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) *and* Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

● Please select one answer

Please choose **all** that apply:

- Jenny & Mitchel
- Annie & Mike
- Jenny & Michael
- Annie & Mitchel

How many rooms did they have? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) *and* Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

● Please select one answer

Please choose **all** that apply:

- 1 Bedroom + 1 Bathroom
- 2 Bedrooms + 1 Bathroom + 1 Common area
- 2 Bedrooms + 1 Bathroom
- 2 Bedrooms + 2 Bathrooms + 1 Common area

How old were the kids? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) *and* Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

● Please select one answer

Please choose **all** that apply:

- 6 & 8
- 4 & 6
- 5 & 8
- 3 & 4

What was the name of the mother/author? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate,) *and* Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

- Petra
- Wilma
- Anna
- Was not mentioned

What was not part of the dessert buffet? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) *and* Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

- Pastries
- Pancakes
- Donuts
- Waffles

Where were they vacationing? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate,) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

● Please select one answer

Please choose **all** that apply:

- Tenerife
- Mexico
- Greece
- Was not mentioned

What was the name of the husband? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

● Please select one answer

Please choose **all** that apply:

- Michael
- Sean
- James
- Was not mentioned

Which buffet did the kids enjoy? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate,) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

- Grill
- Italian
- Snacks
- Greek

Which buffet did the husband enjoy? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) and Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

- Asian
- Grill
- Seafood
- Italian

Which of these hotels did they NOT pass on the bus? *

❗ Please select one answer

Please choose **all** that apply:

- Eden
- Galeri
- Baobab
- Ritz

Demographics

What gender do you identify as? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate,) *and* Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Please select one answer

Please choose **all** that apply:

- Male
 Female
 Other
 Do not want to say

What is your age? *

Only answer this question if the following conditions are met:

Answer was at question '2 [informedconsent1]' (I indicate that I have read and understood the study procedure, and I agree to voluntarily participate.) *and* Answer was at question '3 [informedconsent2]' (I also give permission to make my anonymized recorded data available to others in a public online data repository.)

❗ Only numbers may be entered in this field,

Please write your answer here:

Raffle

If you want to participate in the raffle with a 1/10 chance of winning 30€ please enter your email address.

If you do not wish to participate you are free to leave the field open and continue to the end of the questionnaire.

Enter your email address to participate in the raffle. You will be contacted through this email address if you have won.

Please write your answer here:

Thank you for participating in this study!

12.06.2023 – 11:56

Submit your survey.

Thank you for completing this survey.

Recall questions and answers					
1	What were the maximum and minimum temperatures? (28-14)	28-14	24-16	26-14	28-18
2	How many pools were there? (7)	7	6	5	4
3	How many waterslides were there? (5)	7	6	5	4
4	Why did she recommend water shoes? (hurting stones)	Hurting stones	slippery stones	sea urchin	burning sand
5	What did they forget to bring to the beach? (sunscreen)	Water shoes	Sunscreen	Towels	Swimsuit
6	Which buffet did she most enjoy?(Grill)	Italian	Seafood	Grill	Asian
7	At what time did they leave the kids' disco? (7:00)	6:30	7:00	7:30	8:00
8	Which of these hotels did they NOT pass on the buss (Eden, Baobab, Barcelo, Three corners, Bitzaro, and Galeri)	Eden	Galeri	Baobab	Ritz
9	What did the kids prefer to dip in the chocolate? (strawberries)	Cookies	Waffle	Strawberry	Banana
10	Which century was the castle built? (17 th)	15 th	16 th	17 th	18 th
11	What museum did they visit?(Renaissance)	Romantic	Renaissance	Romanesque	Realism
12	What were the names of the kids? (Jenny & Michael)	Jenny & Mitchel	Annie & Mike	Jenny & Michael	Annie & Mitchel
13	How many rooms did they have? (2 bedrooms+ 1 bathroom + 1 common area)	1 bedroom+ 1 bathroom	2 bedrooms+ 1 bathroom + 1 common area	2 bedrooms + 1 bathroom	2 bedrooms + 2 bathrooms + 1 common area
14	How old were the kids? (6 & 8)	6 & 8	4 & 6	5 & 8	3 & 4
15	What was not part of the dessert buffet? (donuts.)	Pastries	Pancakes	Donuts	Waffles
16	What was the name of the mother/author? (was not mentioned)	Petra	Wilma	Anna	Was not mentioned
17	What was the name of the husband? (was not mentioned)	Michael	Sean	James	Was not mentioned
18	Where were they vacationing? (Was not mentioned)	Tenerife	Mexico	Greece	Was not mentioned
19	Which buffet did the husband enjoy? (Asian)	Asian	Grill	Seafood	Italian
20	Which buffet did the kids enjoy? (Italian)	Grill	Italian	Snacks	Greek

Appendix D, STATA - do file

```
clear all

set more off

use finalsurveydata_12-06

*renaming/organising the variables

rename idresponseid id

rename submitdatedatesubmitted enddate

rename lastpagelastpage lastpage

drop seedseed

rename startdatedatestarted startdate

drop datestampdatelastaction

drop informedconsenttextInformation

rename informedconsent1sq001iindicate ic_1_yes

rename informedconsent1sq002iindicate ic_1_no

rename informedconsent2sq001ialsogive ic_2_yes

rename informedconsent2sq002ialsogive ic_2_no

rename randnumberifis_emptyrandnumberna scenario_nr

*emotionrecognition text

rename emotionrec1textsq001whiche motio emotionrec_text_tense_nervous

rename emotionrec1textsq002whiche motio emotionrec_text_irritated_annoye

rename emotionrec1textsq003whiche motio emotionrec_text_cheerful_happy

rename emotionrec1textsq004whiche motio emotionrec_text_exited_lively

rename emotionrec1textsq005whiche motio emotionrec_text_bored_weary
```

```
rename emotionrec1textsq006whichemotio emotionrec_text_gloomy_sad
rename emotionrec1textsq007whichemotio emotionrec_text_calm_serene
rename emotionrec1textsq008whichemotio emotionrec_text_relaxed_carefree

gen emorec_text = "."
replace emorec_text= "Tense-Nervous" if emotionrec_text_tense_nervous ==1
replace emorec_text= "Irritated-Annoyed" if emotionrec_text_irritated_anno ==1
replace emorec_text= "Cheerful-Happy" if emotionrec_text_cheerful_happy ==1
replace emorec_text= "Exited-Lively" if emotionrec_text_exited_lively ==1
replace emorec_text= "Bored-Weary" if emotionrec_text_bored_weary ==1
replace emorec_text= "Gloomy-Sad" if emotionrec_text_gloomy_sad ==1
replace emorec_text= "Calm-Serene" if emotionrec_text_calm_serene ==1
replace emorec_text= "Relaxed-Carefree" if emotionrec_text_relaxed_carefree ==1
```

```
*emotionrecognition speech
```

```
rename emotionrec2speechsq001whichemot emotionrec_speech_tense_nervous
rename emotionrec2speechsq002whichemot emotionrec_speech_irritated_anno
rename emotionrec2speechsq003whichemot emotionrec_speech_cheerful_happy
rename emotionrec2speechsq004whichemot emotionrec_speech_exited_lively
rename emotionrec2speechsq005whichemot emotionrec_speech_bored_weary
rename emotionrec2speechsq006whichemot emotionrec_speech_gloomy_sad
rename emotionrec2speechsq007whichemot emotionrec_speech_calm_serene
rename emotionrec2speechsq008whichemot emotionrec_speech_relaxed_carefr
```

```
gen emorec_speech = "."
replace emorec_speech= "Tense-Nervous" if emotionrec_speech_tense_nervous ==1
replace emorec_speech= "Irritated-Annoyed" if emotionrec_speech_irritated_anno ==1
```

```
replace emorec_speech= "Cheerful-Happy" if emotionrec_speech_cheerful_happy ==1
replace emorec_speech= "Exited-Lively" if emotionrec_speech_exited_lively ==1
replace emorec_speech= "Bored-Weary" if emotionrec_speech_bored_weary ==1
replace emorec_speech= "Gloomy-Sad" if emotionrec_speech_gloomy_sad ==1
replace emorec_speech= "Calm-Serene" if emotionrec_speech_calm_serene ==1
replace emorec_speech= "Relaxed-Carefree" if emotionrec_speech_relaxed_carefr ==1
```

```
rename emotionalpref1sq001pleaseanswer emopref_1_likevoicesounded
rename emotionalpref1sq002pleaseanswer emopref_2_likeemphasizedwords
rename emotionalpref1sq003pleaseanswer emopref_3_likeexpressingemo
replace emopref_1_likevoicesounded= "A1" if emopref_1_likevoicesounded=="."
```

```
gen emopref_3_num_expres = .
```

```
replace emopref_3_num_expres = 1 if emopref_3_likeexpressingemo == "A1"
replace emopref_3_num_expres = 2 if emopref_3_likeexpressingemo == "A2"
replace emopref_3_num_expres = 3 if emopref_3_likeexpressingemo == "A3"
replace emopref_3_num_expres = 4 if emopref_3_likeexpressingemo == "A4"
replace emopref_3_num_expres = 5 if emopref_3_likeexpressingemo == "A5"
replace emopref_3_num_expres = 6 if emopref_3_likeexpressingemo == "A6"
replace emopref_3_num_expres = 7 if emopref_3_likeexpressingemo == "A7"
```

```
gen emopref_2_num_emphasized = .
```

```
replace emopref_2_num_emphasized = 1 if emopref_2_likeemphasizedwords == "A1"
replace emopref_2_num_emphasized = 2 if emopref_2_likeemphasizedwords == "A2"
replace emopref_2_num_emphasized = 3 if emopref_2_likeemphasizedwords == "A3"
replace emopref_2_num_emphasized = 4 if emopref_2_likeemphasizedwords == "A4"
replace emopref_2_num_emphasized = 5 if emopref_2_likeemphasizedwords == "A5"
```

replace emopref_2_num_emphasized = 6 if emopref_2_likeemphasizedwords == "A6"

replace emopref_2_num_emphasized = 7 if emopref_2_likeemphasizedwords == "A7"

gen emopref_1_num_voice = .

replace emopref_1_num_voice = 1 if emopref_1_likevoicesounded == "A1"

replace emopref_1_num_voice = 2 if emopref_1_likevoicesounded == "A2"

replace emopref_1_num_voice = 3 if emopref_1_likevoicesounded == "A3"

replace emopref_1_num_voice = 4 if emopref_1_likevoicesounded == "A4"

replace emopref_1_num_voice = 5 if emopref_1_likevoicesounded == "A5"

replace emopref_1_num_voice = 6 if emopref_1_likevoicesounded == "A6"

replace emopref_1_num_voice = 7 if emopref_1_likevoicesounded == "A7"

gen voice_sound_A1 = 0

gen voice_sound_A2 = 0

gen voice_sound_A3 = 0

gen voice_sound_A4 = 0

gen voice_sound_A5 = 0

gen voice_sound_A6 = 0

gen voice_sound_A7 = 0

replace voice_sound_A1 = 1 if emopref_1_likevoicesounded == "A1"

replace voice_sound_A2 = 1 if emopref_1_likevoicesounded == "A2"

replace voice_sound_A3 = 1 if emopref_1_likevoicesounded == "A3"

replace voice_sound_A4 = 1 if emopref_1_likevoicesounded == "A4"

replace voice_sound_A5 = 1 if emopref_1_likevoicesounded == "A5"

replace voice_sound_A6 = 1 if emopref_1_likevoicesounded == "A6"

replace voice_sound_A7 = 1 if emopref_1_likevoicesounded == "A7"

gen emp_words_A1 = 0

gen emp_words_A2 = 0

gen emp_words_A3 = 0

gen emp_words_A4 = 0

gen emp_words_A5 = 0

gen emp_words_A6 = 0

gen emp_words_A7 = 0

replace emp_words_A1 = 1 if emopref_2_likeemphasizedwords == "A1"

replace emp_words_A2 = 1 if emopref_2_likeemphasizedwords == "A2"

replace emp_words_A3 = 1 if emopref_2_likeemphasizedwords == "A3"

replace emp_words_A4 = 1 if emopref_2_likeemphasizedwords == "A4"

replace emp_words_A5 = 1 if emopref_2_likeemphasizedwords == "A5"

replace emp_words_A6 = 1 if emopref_2_likeemphasizedwords == "A6"

replace emp_words_A7 = 1 if emopref_2_likeemphasizedwords == "A7"

tab emopref_1_likevoicesounded

tab voice_sound_A1

tab voice_sound_A2

tab voice_sound_A3

tab voice_sound_A4

tab voice_sound_A5

tab voice_sound_A6

tab voice_sound_A7

tab emopref_1_likevoicesounded

rename experience1ownsqa001doyouhaveasma exp_1_own_yes

rename experience1ownsqa002doyouhaveasma exp_1_own_no

rename experience2freqsq001howoftendoyo exp_2_frequency

rename experience3spreadsq001howmanyhou exp_3_spread

rename experience4interestsq001areyouin exp_4_interest

rename experience4interestsq002wouldyou exp_5_want_own

rename experience4interestsq003doyouthi exp_6_useful

rename memory2poolssqa001howmanypoolswer memory_2_pools_f_4

rename memory2poolssqa002howmanypoolswer memory_2_pools_f_5

rename memory2poolssqa003howmanypoolswer memory_2_pools_f_6

rename memory2poolssqa004howmanypoolswer memory_2_pools_c_7

rename memory1temperaturessqa001whatwere memory_1_temp_c_28_14

rename memory1temperaturessqa002whatwere memory_1_temp_f_24_16

rename memory1temperaturessqa003whatwere memory_1_temp_f_26_14

rename memory1temperaturessqa004whatwere memory_1_temp_f_28_18

rename memory3waterslidessqa001howmanywa memory_3_waterslides_f_4

rename memory3waterslidessqa002howmanywa memory_3_waterslides_c_5

rename memory3waterslidessqa003howmanywa memory_3_waterslides_f_6

rename memory3waterslidessqa004howmanywa memory_3_waterslides_f_7

rename memory4watershoessqa001whydidsher memory_4_watershoes_c_hurt

rename memory4watershoessqa002whydidsher memory_4_watershoes_f_slip

rename memory4watershoessqa003whydidsher memory_4_watershoes_f_urchin

rename memory4watershoessqa004whydidsher memory_4_watershoes_f_hot

rename memory5forgetsq001whatdidtheyfor memory_5_forget_f_shoes

rename memory5forgetsq002whatdidtheyfor memory_5_forget_c_sunscreen

rename memory5forgetsq003whatdidtheyfor memory_5_forget_f_towels

rename memory5forgetsq004whatdidtheyfor memory_5_forget_f_suit

rename memory6buffetsq001whichbuffet memory_6_buffetshe_f_italian

rename memory6buffetsq002whichbuffet memory_6_buffetshe_f_asian

rename memory6buffetsq003whichbuffet memory_6_buffetshe_c_grill

rename memory6buffetsq004whichbuffet memory_6_buffetshe_f_sea

rename memory9chocolatedipsq001whatdidt memory_9_chocolate_f_cookie

rename memory9chocolatedipsq002whatdidt memory_9_chocolate_f_waffle

rename memory9chocolatedipsq003whatdidt memory_9_chocolate_c_strawb

rename memory9chocolatedipsq004whatdidt memory_9_chocolate_f_banana

rename memory7kidsdiscosq001atwhattimed memory_7_disco_f_630

rename memory7kidsdiscosq002atwhattimed memory_7_disco_c_700

rename memory7kidsdiscosq003atwhattimed memory_7_disco_f_730

rename memory7kidsdiscosq004atwhattimed memory_7_disco_f_800

rename memory10castlesq001whichcenturyw memory_10_castles_f_15

rename memory10castlesq002whichcenturyw memory_10_castles_f_16

rename memory10castlesq003whichcenturyw memory_10_castles_c_17

rename memory10castlesq004whichcenturyw memory_10_castles_f_18

rename memory11museumsq001whatmuseumdid memory_11_museum_f_romantic

rename memory11museumsq002whatmuseumdid memory_11_museum_c_renaissance
rename memory11museumsq003whatmuseumdid memory_11_museum_f_romanesque
rename memory11museumsq004whatmuseumdid memory_11_museum_f_realism

rename memory12namesq001whatwerethenam memory_12_names_f_jennymitchel
rename memory12namesq002whatwerethenam memory_12_names_f_anniemike
rename memory12namesq003whatwerethenam memory_12_names_c_jennymichael
rename memory12namesq004whatwerethenam memory_12_names_f_anniemitchel

rename memory13roomsSQ001Howmanyroomsdi memory_13_rooms_f_1bed1bath
rename memory13roomsSQ002Howmanyroomsdi memory_13_rooms_c_2bed1bath1c
rename memory13roomsSQ003Howmanyroomsdi memory_13_rooms_f_2bed1bath
rename memory13roomsSQ004Howmanyroomsdi memory_13_rooms_f_2bed2bath2c

rename memory14agesq001howoldwerethekid memory_14_ages_c_68
rename memory14agesq002howoldwerethekid memory_14_ages_f_46
rename memory14agesq003howoldwerethekid memory_14_ages_f_58
rename memory14agesq004howoldwerethekid memory_14_ages_f_34

rename memory16namemomsq001whatwasthena memory_16_namemom_f_petra
rename memory16namemomsq002whatwasthena memory_16_namemom_f_wilma
rename memory16namemomsq003whatwasthena memory_16_namemom_f_anna
rename memory16namemomsq004whatwasthena memory_16_namemom_c_nvm

rename memory15dessertsq001whatwasnotpa memory_15_dessert_f_pastry
rename memory15dessertsq002whatwasnotpa memory_15_dessert_f_pancake
rename memory15dessertsq003whatwasnotpa memory_15_dessert_c_donut

rename memory15dessertsq004whatwasnotpa memory_15_dessert_f_waffle

rename memory18locationSQ001Wherewereth memory_18_location_f_tener

rename memory18locationSQ002Wherewereth memory_18_location_f_mex

rename memory18locationSQ003Wherewereth memory_18_location_f_greece

rename memory18locationSQ004Wherewereth memory_18_location_c_nv

rename memory17namedadsq001whatwasthena memory_17_namedad_f_micheal

rename memory17namedadsq002whatwasthena memory_17_namedad_f_sean

rename memory17namedadsq003whatwasthena memory_17_namedad_f_james

rename memory17namedadsq004whatwasthena memory_17_namedad_c_nv

rename memory20buffetkidssq001whichbuff memory_20_buffetkids_f_grill

rename memory20buffetkidssq002whichbuff memory_20_buffetkids_c_italian

rename memory20buffetkidssq003whichbuff memory_20_buffetkids_f_snacks

rename memory20buffetkidssq004whichbuff memory_20_buffetkids_f_greek

rename memory19buffethussq001whichbuffe memory_19_buffethus_c_asian

rename memory19buffethussq002whichbuffe memory_19_buffethus_f_grill

rename memory19buffethussq003whichbuffe memory_19_buffethus_f_seafood

rename memory19buffethussq004whichbuffe memory_19_buffethus_f_italian

rename memory8hotelssq001whichoftheseho memory_8_hotels_f_eden

rename memory8hotelssq002whichoftheseho memory_8_hotels_f_galeri

rename memory8hotelssq003whichoftheseho memory_8_hotels_f_baabab

rename memory8hotelssq004whichoftheseho memory_8_hotels_c_ritz

```
rename gendersq001whatgenderdoyouidenti gender_male
rename gendersq002whatgenderdoyouidenti gender_female
rename gendersq003whatgenderdoyouidenti gender_other
rename gendersq004whatgenderdoyouidenti gender_nvm
```

```
rename agewhatisyourage age
```

```
rename emailEnteryouemailaddresstopart email
```

```
rename interviewtimetotaltime time_total
```

```
rename grouptime15967grouptimesmartspea time_scenario
```

```
rename grouptime15965grouptimeinformedc time_ic
```

```
drop informedconsenttexttimequestiont
```

```
drop informedconsent1timequestiontime
```

```
drop informedconsent2timequestiontime
```

```
*****
```

```
**overview cat variable meanings
```

```
*****
```

```
/* emopref_*
```

```
A1 = I liked it very much
```

```
A2
```

```
A3
```

```
A4 = Neutral
```

```
A5
```

```
A6
```

```
A7 = I did not like it at all
```

scenario number

1 HSHV happy story happy voice

2 HSSV happy story sad voice

3 SSSV sad story sad voice

4 SSHV sad story happy voice

how often do you use a smartspeaker?

A1 never

A2 only once

A3 only a few times

A4 a few times per year

A5 every week

A6 every day

How many households of your friends and family have a smartspeaker that you know of?

A1 none

A2 1-2

A3 3-5

A4 5-10

A5 11+

are you interested in smart speakers?

would you want to have a smartspeaker?

do you think smart speakers are useful?

A1 not at all

A2

A3

A4 neutral

A5

A6

A7 very

*/

generate scenario = "."

replace scenario = "HSHV" if scenario_nr==1

replace scenario = "HSSV" if scenario_nr==2

replace scenario = "SSSV" if scenario_nr==3

replace scenario = "SSHV" if scenario_nr==4

generate frequency = "."

replace frequency = "Never" if exp_2_frequency == "A1"

replace frequency = "Only once" if exp_2_frequency == "A2"

replace frequency = "Only a few times" if exp_2_frequency == "A3"

replace frequency = "A few times per year" if exp_2_frequency == "A4"

replace frequency = "Every week" if exp_2_frequency == "A5"

replace frequency = "Every day" if exp_2_frequency == "A6"

gen frequency_num =.

replace frequency_num = 1 if exp_2_frequency == "A1"

replace frequency_num = 2 if exp_2_frequency == "A2"

replace frequency_num = 3 if exp_2_frequency == "A3"

replace frequency_num = 4 if exp_2_frequency == "A4"

replace frequency_num = 5 if exp_2_frequency == "A5"

```
replace frequency_num = 6 if exp_2_frequency == "A6"
```

```
drop if ic_1_no == 1
```

```
drop if ic_2_no == 1
```

```
drop if ic_1_yes ==.
```

```
drop if ic_2_yes ==.
```

```
drop if lastpage < 8
```

```
egen recal = rowtotal (memory_1_temp_c_28_14 memory_2_pools_c_7 memory_3_waterslides_c_5  
memory_4_watershoes_c_hurt memory_5_forget_c_sunscreen memory_6_buffetshe_c_grill  
memory_7_disco_c_700 memory_8_hotels_c_ritz memory_9_chocolate_c_strawb  
memory_10_castles_c_17 memory_11_museum_c_renaissance memory_12_names_c_jennymichael  
memory_13_rooms_c_2bed1bath1c memory_14_ages_c_68 memory_15_dessert_c_donut  
memory_16_namemom_c_nvm memory_17_namedad_c_nvm memory_18_location_c_nvm  
memory_19_buffethus_c_asian memory_20_buffetkids_c_italian)
```

```
gen congruent =.
```

```
replace congruent =1 if scenario_nr==1
```

```
replace congruent =1 if scenario_nr==3
```

```
replace congruent =0 if scenario_nr==2
```

```
replace congruent =0 if scenario_nr==4
```

```
gen gender = .
```

```
replace gender =1 if gender_male==1
```

```
replace gender =0 if gender_female==1
```

```
replace gender =2 if gender_nvm==1
```

```
replace gender =2 if gender_other==1
```

```
generate own_speaker =.  
replace own_speaker =1 if exp_1_own_yes ==1  
replace own_speaker =0 if exp_1_own_no ==1  
  
*****  
*****  
*****  
*****  
  
*****  
  
*****participants demographics*****  
  
sum age  
tab gender  
  
*****  
  
***** participants experience *****  
  
*****  
  
tab own_speaker  
  
tab exp_2_frequency  
tab frequency  
tab exp_3_spread  
tab exp_4_interest  
tab exp_5_want_own  
  
*graph bar (count) , over(exp_2_frequency ) scheme (s1mono)
```

```
*graph bar (count) , over(exp_2_frequency, relabel(1 "Never" 2 "Only once" 3 "Only a few times" 4 "A
few times per year" 5 "Every week" 6 "Every day")) label(angle(45)) ) scheme (s1mono) ytitle("Number
of participants") title("Frequency of smartspeaker usage")
```

```
tab exp_3_spread
```

```
*graph bar (count) , over(exp_3_spread)
```

```
*graph bar (count) , over(exp_3_spread, relabel(1 "None" 2 "1-2" 3 "3-5" 4 "5-10" 5 "11+")) scheme
(s1mono) ytitle("Number of participants") title("Households of friends/family with a smart speaker")
```

```
/*
```

```
tab exp_4_interest
```

```
graph bar (count) , over(exp_4_interest)
```

```
graph bar (count) , over(exp_4_interest, relabel(1 "Not at all" 2 "2" 3 "3" 4 "Neutral" 5 "5" 6 "6" 7
"very")) ) scheme (s1mono) ytitle("Number of participants") title("Households of friends/family with a
smart speaker")
```

```
graph bar (count) , over(exp_4_interest, relabel(1 "Not at all" 2 "3" 4 "Neutral" 5 "6" 7 "Very")) )
scheme (s1mono) ytitle("Number of participants") title("Interested in smart speakers")
```

```
tab exp_5_want_own
```

```
graph bar (count) , over(exp_4_interest, relabel(1 "Not at all" 2 "3" 4 "Neutral" 5 "6" 7 "Very")) )
scheme (s1mono) ytitle("Number of participants") title("Would want to own a smart speaker")
```

```
tab exp_6_useful
```

```
graph bar (count) , over(exp_4_interest, relabel(1 "Not at all" 2 "3" 4 "Neutral" 5 "6" 7 "Very")) )
scheme (s1mono) ytitle("Number of participants") title("Think smart speakers are useful")
```

*/

*****scenario distribution*****

tabulate scenario_nr

tab scenario

sum scenario_nr

tab scenario_nr

*graph bar (count) , over(scenario_nr) scheme (s1mono)

*graph bar (count) , over(scenario) scheme (s1mono) ytitle("Number of participants") title("The distribution of the participants over the scenarios")

*graph bar (count) , over(congruent, relabel(1 "incongruent" 2 "congruent")) scheme (s1mono) ytitle("Number of participants") title("The distribution of the participants over the conditions")

***** emotion recognition*****

tab emorec_speech if scenario == "SSHV"

tab emorec_text if scenario == "SSHV"

* title: The distribution of emotions recognised in the text for each scenario

```
*graph bar (count) , over(emorec_text, label(angle(45))) by (scenario) scheme (s1mono)
ytitled("Number of participants")
```

```
* title: The distribution of emotions recognised in the speech for each scenario
```

```
*graph bar (count) , over(emorec_speech, sort(#) label(angle(45))) by (scenario) scheme (s1mono)
ytitled("Number of participants")
```

```
*****
```

```
***** in correct order*****
```

```
*****
```

```
*title: The distribution of emotions recognised in the text for each scenario
```

```
*graph bar (sum) emotionrec_text_exited_lively emotionrec_text_cheerful_happy
emotionrec_text_relaxed_carefr emotionrec_text_calm_serene emotionrec_text_bored_weary
emotionrec_text_gloomy_sad emotionrec_text_irritated_anno emotionrec_text_tense_nervous,
by(scenario) yvaroptions( relabel(1 "Exited-Lively" 2 "Cheerful-Happy" 3 "Relaxed-Carefree" 4 "Calm-
Serene" 5 "Bored-Weary" 6 "Gloomy-Sad" 7 "Irritated-Annoyed" 8 "Tense-Nervous") label(angle(45)))
ascategory scheme(s1mono) ytitled("Number of participants")
```

```
*title: The distribution of emotions recognised in the speech for each scenario
```

```
*graph bar (sum) emotionrec_speech_exited_lively emotionrec_speech_cheerful_happy
emotionrec_speech_relaxed_carefr emotionrec_speech_calm_serene
emotionrec_speech_bored_weary emotionrec_speech_gloomy_sad
emotionrec_speech_irritated_anno emotionrec_speech_tense_nervous, by(scenario) yvaroptions(
relabel(1 "Exited-Lively" 2 "Cheerful-Happy" 3 "Relaxed-Carefree" 4 "Calm-Serene" 5 "Bored-Weary"
6 "Gloomy-Sad" 7 "Irritated-Annoyed" 8 "Tense-Nervous") label(angle(45))) ascategory
scheme(s1mono) ytitled("Number of participants")
```

***** preference *****

***** 1 like the way the voice sounded

tab emopref_1_likevoicesounded

```
*graph bar (sum) voice_sound_A1 voice_sound_A2 voice_sound_A3 voice_sound_A4
voice_sound_A5 voice_sound_A6 voice_sound_A7, yvaroptions( relabel(1 "I liked it very much" 2 "3"
"4 "Neutral" 5 "6" "7 "I did not like it at all")) ascategory scheme(s1mono) ytitle("Number of
participants") title("Liked how the voice sounded")
```

```
*graph bar (sum) voice_sound_A1 voice_sound_A2 voice_sound_A3 voice_sound_A4
voice_sound_A5 voice_sound_A6 voice_sound_A7, by(congruent) yvaroptions( relabel(1 "I liked it
very much" 2 "3" "4 "Neutral" 5 "6" "7 "I did not like it at all")) ascategory scheme(s1mono)
ytitle("Number of participants")
```

```
*title("Liked how the voice sounded")
```

*****2 like the emphasized words

```
*graph bar (sum) emp_words_A1 emp_words_A2 emp_words_A3 emp_words_A4 emp_words_A5
emp_words_A6 emp_words_A7, yvaroptions( relabel(1 "I liked it very much" 2 "3" 4 "Neutral" 5 "6"
"7 "I did not like it at all")) ascategory scheme(s1mono) ytitle("Number of participants") title("Liked
the way the words were emphasized")
```

```
*graph bar (sum) emp_words_A1 emp_words_A2 emp_words_A3 emp_words_A4 emp_words_A5
emp_words_A6 emp_words_A7, by(congruent) yvaroptions( relabel(1 "I liked it very much" 2 "3" 4
"Neutral" 5 "6" 7 "I did not like it at all")) ascategory scheme(s1mono) ytitle("Number of
participants")
```

```
*title("Liked the way the words were emphasized")
```

```
*****3 like that it expressed emotion
```

```
*graph bar (count) , over(emopref_3_likeexpressingemo, relabel(1 "I liked it very much" 2 "3" 4
"Neutral" 5 "6" 7 "I did not like it at all")) scheme(s1mono) ytitle("Number of participants")
title("Liked that the voice was expressing emotion")
```

```
*graph bar (count) , over(emopref_3_likeexpressingemo, relabel(1 "I liked it very much" 2 "3" 4
"Neutral" 5 "6" 7 "I did not like it at all")) by (congruent) scheme(s1mono) ytitle("Number of
participants")
```

```
*title("Comparison how the emotion expression was evaluated for incongruent and congruent")
```

```
/*
```

```
ttest emopref_1_num_voice ,by (congruent)
```

```
ttest emopref_2_num_emphasized ,by (congruent)
```

```
ttest emopref_3_num_expres ,by (congruent)
```

```
sktest emopref_1_num_voice if congruent == 0
```

```
sktest emopref_1_num_voice if congruent == 1
```

```
swilk emopref_1_num_voice if congruent == 0
```

```
swilk emopref_1_num_voice if congruent == 1
```

```
ttest emopref_1_num_voice ,by (congruent)
```

```
sdtest emopref_1_num_voice ,by (congruent)
```

```
*/
```

```
ttest emopref_3_num_expres ,by (congruent)
```

```
sktest emopref_3_num_expres if congruent == 0
```

```
sktest emopref_3_num_expres if congruent == 1
```

```
swilk emopref_3_num_expres if congruent == 0
```

```
swilk emopref_3_num_expres if congruent == 1
```

```
ttest emopref_3_num_expres ,by (congruent)
```

```
sdtest emopref_3_num_expres ,by (congruent)
```

```
esize twosample emopref_3_num_expres, by(congruent) all
```

```
*****
```

```
***** ttest recal over congruent and incongruent
```

```
*****
```

```
*graph bar recal , over(scenario_nr) scheme (s1mono) ytitle("Number of participants") title("The
distribution of the participants over the scenarios")
```

```
*graph bar recal , over(congruent, relabel(1 "incongruent" 2 "congruent")) scheme(s1mono)
blabel(total) ytitle("Number of correct answers") title("The recal for the congruent and incongruent
conditions")
```

```
ttest recal ,by (congruent)
```

```
sktest recal if congruent == 0
```

```
sktest recal if congruent == 1
```

```
swilk recal if congruent == 0
```

```
swilk recal if congruent == 1
```

```
ttest recal ,by (congruent)
```

```
sdtest recal ,by (congruent)
```

```
esize twosample recal, by(congruent) all
```

```
*****
```

```
** testing if the recal trend becomes stronger if only participants are included were the manipulation
worked
```

```
*****
```

```
gen pos_emorec_speech = .
```

```
replace pos_emorec_speech = 1 if emotionrec_speech_exited_lively ==1
```

```
replace pos_emorec_speech = 1 if emotionrec_speech_cheerful_happy ==1
```

```
replace pos_emorec_speech = 1 if emotionrec_speech_relaxed_carefr ==1
```

```
replace pos_emorec_speech = 1 if emotionrec_speech_calm_serene ==1
```

```
gen neg_emorec_speech = .
```

```
replace neg_emorec_speech = 1 if emotionrec_speech_bored_weary ==1
```

```
replace neg_emorec_speech = 1 if emotionrec_speech_gloomy_sad ==1
```

```
replace neg_emorec_speech = 1 if emotionrec_speech_irritated_anno ==1
```

```
replace neg_emorec_speech = 1 if emotionrec_speech_tense_nervous ==1
```

```
gen pos_emorec_text = .
```

```
replace pos_emorec_text = 1 if emotionrec_text_exited_lively ==1
```

```
replace pos_emorec_text = 1 if emotionrec_text_cheerful_happy ==1
```

```
replace pos_emorec_text = 1 if emotionrec_text_relaxed_carefr ==1
```

```
replace pos_emorec_text = 1 if emotionrec_text_calm_serene ==1
```

```
gen neg_emorec_text = .
```

```
replace neg_emorec_text = 1 if emotionrec_text_bored_weary ==1
```

```
replace neg_emorec_text = 1 if emotionrec_text_gloomy_sad ==1
```

```
replace neg_emorec_text = 1 if emotionrec_text_irritated_anno ==1
```

```
replace neg_emorec_text = 1 if emotionrec_text_tense_nervous ==1
```

```
*****8
```

```
**exploratory regression if experience influences preference
```

```
*****
```

```
tab exp_2_frequency
tab frequency_num

tab emopref_3_num_expres

reg emopref_3_num_expres frequency_num // no effect very small
reg emopref_1_num_voice frequency_num // no effect
reg emopref_2_num_emphasized frequency_num // no effect

*****
*****Filtering participants on speech manipulation correctness
*****

/*
drop if scenario == "HSHV" & pos_emorec_speech ==.
drop if scenario == "HSSV" & neg_emorec_speech ==.
drop if scenario == "SSSV" & neg_emorec_speech ==.
drop if scenario == "SSHV" & pos_emorec_speech ==.

drop if scenario == "SSHV" & neg_emorec_text ==.
drop if scenario == "SSSV" & neg_emorec_text ==.
drop if scenario == "HSSV" & pos_emorec_text ==. //weird
drop if scenario == "HSHV" & pos_emorec_text ==.

*title: The distribution of emotions recognised in the text for each scenario
```

```
*graph bar (sum) emotionrec_text_exited_lively emotionrec_text_cheerful_happy
emotionrec_text_relaxed_carefr emotionrec_text_calm_serene emotionrec_text_bored_weary
emotionrec_text_gloomy_sad emotionrec_text_irritated_anno emotionrec_text_tense_nervous,
by(scenario) yvaroptions( relabel(1 "Exited-Lively" 2 "Cheerful-Happy" 3 "Relaxed-Carefree" 4 "Calm-
Serene" 5 "Bored-Weary" 6 "Gloomy-Sad" 7 "Irritated-Annoyed" 8 "Tense-Nervous") label(angle(45)))
ascategory scheme(s1mono) ytitle("Number of participants")
```

```
*title: The distribution of emotions recognised in the speech for each scenario
```

```
*graph bar (sum) emotionrec_speech_exited_lively emotionrec_speech_cheerful_happy
emotionrec_speech_relaxed_carefr emotionrec_speech_calm_serene
emotionrec_speech_bored_weary emotionrec_speech_gloomy_sad
emotionrec_speech_irritated_anno emotionrec_speech_tense_nervous, by(scenario) yvaroptions(
relabel(1 "Exited-Lively" 2 "Cheerful-Happy" 3 "Relaxed-Carefree" 4 "Calm-Serene" 5 "Bored-Weary"
6 "Gloomy-Sad" 7 "Irritated-Annoyed" 8 "Tense-Nervous") label(angle(45))) ascategory
scheme(s1mono) ytitle("Number of participants")
```

```
*graph bar recal , over(congruent, relabel(1 "incongruent" 2 "congruent")) scheme(s1mono)
blabel(total) ytitle("Number of correct answers") title("The recal for the congruent and incongruent
conditions")
```

```
ttest recal ,by (congruent)
```

```
esize twosample recal, by(congruent) all
```

```
sum age
```

```
tab gender
```

```
// difference in means only becomes smaller after filtering ineffective manipulation is not the only
problem
```

*/

*****88

*****checking for main effect of emotions on recal

gen happy_voice =.

replace happy_voice =1 if scenario == "HSHV"

replace happy_voice =1 if scenario == "SSHV"

replace happy_voice =0 if scenario == "HSSV"

replace happy_voice =0 if scenario == "SSSV"

gen happy_story =.

replace happy_story=1 if scenario == "HSHV"

replace happy_story=1 if scenario == "HSSV"

replace happy_story=0 if scenario == "SSHV"

replace happy_story=0 if scenario == "SSSV"

ttest recal ,by (happy_voice)

sktest recal if happy_voice == 0

sktest recal if happy_voice == 1

swilk recal if happy_voice == 0

swilk recal if happy_voice == 1

```
ttest recal ,by (happy_voice)
```

```
sdtest recal ,by (happy_voice)
```

```
esize twosample recal, by(happy_voice) all
```

```
*graph bar recal , over(happy_voice, relabel(1 "Sad voice" 2 "Happy voice")) scheme(s1mono)
blabel(total) ytitle("Number of correct answers") title("The recal for the sad and happy voice
scenarios")
```

```
ttest recal ,by (happy_story)
```

```
sktest recal if happy_story == 0
```

```
sktest recal if happy_story == 1
```

```
swilk recal if happy_story == 0
```

```
swilk recal if happy_story == 1
```

```
ttest recal ,by (happy_story)
```

```
sdtest recal ,by (happy_story)
```

```
esize twosample recal, by(happy_story) all
```

```
graph bar recal , over(happy_story, relabel(1 "Sad story" 2 "Happy story")) scheme(s1mono)
blabel(total) ytitle("Number of correct answers") title("The recal for the sad and happy story
scenarios")
```