

Week 6: Evaluation Methods

Wednesday lecture
Uta Hinrichs & Tara Capel

any questions about Assignment 2 (so far)?

today

- Re-cap: Questionnaires/survey
- Re-cap: Study variables

re-cap questionnaires/surveys

questionnaires

- Online forms or on-paper
- Questions should be directly related to the questions you are exploring
 - See study questions in Week 5
- Question types
 - Open-ended questions
 - Single- and multiple-choice questions
 - Matching/ranking
 - Rating scales (e.g., Likert Scales or Semantic Scales)
 - e.g., Likert Scales or Semantic Scales
 - Typically, on a 5 to 7 point scale

System Usability Scale (SUS)

- Post-test assessment of usability
- 10 questions addressing usability and learnability of a system

1. I think that I would like to use this system frequently.

1. Strongly Disagree

2.

3.

4.

5. Strongly Agree

☐☐☐☐☐

2. I found the system unnecessarily complex.

1. Strongly Disagree

2.

3.

4.

5. Strongly Agree

☐☐☐☐☐

3. I thought the system was easy to use.

1. Strongly Disagree

2.

3.

4.

5. Strongly Agree

☐☐☐☐☐

4. I think that I would need the support of a technical person to be able to use this system.

other established usability scales

- Single Ease Question (SEQ)
 - Post-task

Overall, this task was?

1. Very Difficult	2.	3.	4.	5.	6.	7. Very Easy
☐	☐	☐	☐	☐	☐	☐

- NASA TLX (Task Load Index)
 - 6 questions
 - Unlabelled 21-point scale

Mental Demand

How mentally demanding was the task?

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Very Low																				Very High

Physical Demand

How physically demanding was the task?

☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐
Very Low																				Very High

established usability scales

pros & cons

- Pros
 - Extremely established and widely used
 - Experimentally valid and reliable
 - Large amounts of industry data available to help benchmark scores (especially for SUS)
- Cons (limitations)
 - Self-reported data can be unreliable
 - Participants may have a wildly different sense of the meaning of the scales (what does a “5” out of “7” mean?)
 - Require large participant samples to be meaningful
 - Measurement of subjective user experience rather than performance
 - A combination of satisfaction and performance can provide a clearer story
 - Metrics do not tell you the story behind participants’ rating or what the strengths or weaknesses of the system are

Example Survey

by Nicole Meng

- What differences can we observe between owners, cohabitants, and visitors in their willingness to share smart speakers?
- What features are considered sharable?
- Are there differences between owners, cohabitants, and visitors?

Survey Structure

- Consent
- Screening
- Device Questions
- General Feature Questions
- Sharing Features
- Open-ended
- Privacy Scales
- Demographics
- Advert for future studies

Survey Structure

- Consent
- **Screening**
- Device Questions
- General Feature Questions
- Sharing Features
- Open-ended
- Privacy Scales
- Demographics
- Advert for future studies

Encounter Context

How did you encounter smart speakers? Please all options that apply.

- ☐ I own a smart speaker.
- ☐ I live in a home, where someone else owns a smart speaker.
- ☐ I have visited people who have a smart speaker.
- ☐ I have worked in a place where there is a smart speaker.
- ☐ None of the above.
- ☐ Other, please explain.

[Show Discussion \(4\)](#)

Last Comment 4 Jul 2024 10:31am by Maria Wolters

Encounter

[+ Add page break](#)

Interaction Context

In which context did you gain experience with a smart speaker?

- ☐ I interacted with a smart speaker myself.
- ☐ Someone else interacted with a smart speaker while I was in the room.
- ☐ The device activated while I was in the room, but no one interacted with it.
- ☐ I was in a room where there was a smart speaker, but it did not activate.
- ☐ None of the above.

[Show Discussion \(3\)](#)

Last Comment 4 Jul 2024 10:29am by Kami Vaniea

Context of encounter

☐ SS Mental Model

Smart speakers are helpful in our everyday life and can perform a large number of services. Please tick all tasks that you think

- ☐ Make a phone call.
- ☐ Read out a recipe while you cook.
- ☐ Suggest which outfit you should wear.
- ☐ Assemble furniture.
- ☐ Play a song you want to hear.
- ☐ Drive a car.
- ☐ Send a message to all other smart speakers.

Knowledge check

Survey Structure

- Consent
- Screening
- **Device Questions**
- General Feature Questions
- Sharing Features
- Open-ended
- Privacy Scales
- Demographics
- Advert for future studies

▼  Display this question

If How did you encounter smart speakers? Please all options that apply. I own a smart speaker. Is Selected

Or How did you encounter smart speakers? Please all options that apply. I live in a home, where someone else owns a smart speaker. Is Selected

How many smart speakers do you own or live with?

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4+

ss account

▼  Display this question

If How did you encounter smart speakers? Please all options that apply. I own a smart speaker. Is Selected

Or How did you encounter smart speakers? Please all options that apply. I live in a home, where someone else owns a smart speaker. Is Selected

Whose account is connected to the smart speaker?

- ☐ My Own Account
- ☐ Someone else's (State relationship to account owner)

[Show Discussion \(2\)](#)

Last Comment 24 Jun 2024 6:50pm by Kami Vaniea

Survey Structure

- Consent
- Screening
- Device Questions
- General Feature Questions
- **Sharing Features**
- Open-ended
- Privacy Scales
- Demographics
- Advert for future studies

As the account owner, how would you feel if someone who LIVES IN your household (e.g. partner, family, flatmates) uses each of the features presented below?

Please rate each feature on a scale from "very uncomfortable" (I would hate it if they used this feature) to "very comfortable" (They are very welcome to use this feature.)

	very uncomfortable	somewhat uncomfortable	neutral	somewhat comfortable	very comfortable
Calendar (e.g. "Ok Google, tell me what I have on today" or "Alexa, add a meeting to my calendar")	☐	☐	☐	☐	☐
Calls, emails, messages (e.g. Drop-in feature, making phone calls, checking emails, or receiving notifications)	☐	☐	☐	☐	☐
Purchase (e.g. "Alexa, order dishwasher tablets")	☐	☐	☐	☐	☐
Smart Device Control (e.g. "Hey Google, turn the lights on" or "Siri, increase the temperature to 19C")	☐	☐	☐	☐	☐
Timer or Alarms (e.g. "Alexa, set a timer for 5 minutes")	☐	☐	☐	☐	☐
Entertainment (e.g. playing back music from Spotify or playing a trivia game)	☐	☐	☐	☐	☐
Getting Information (e.g. asking for weather or news updates, recipes or facts)	☐	☐	☐	☐	☐

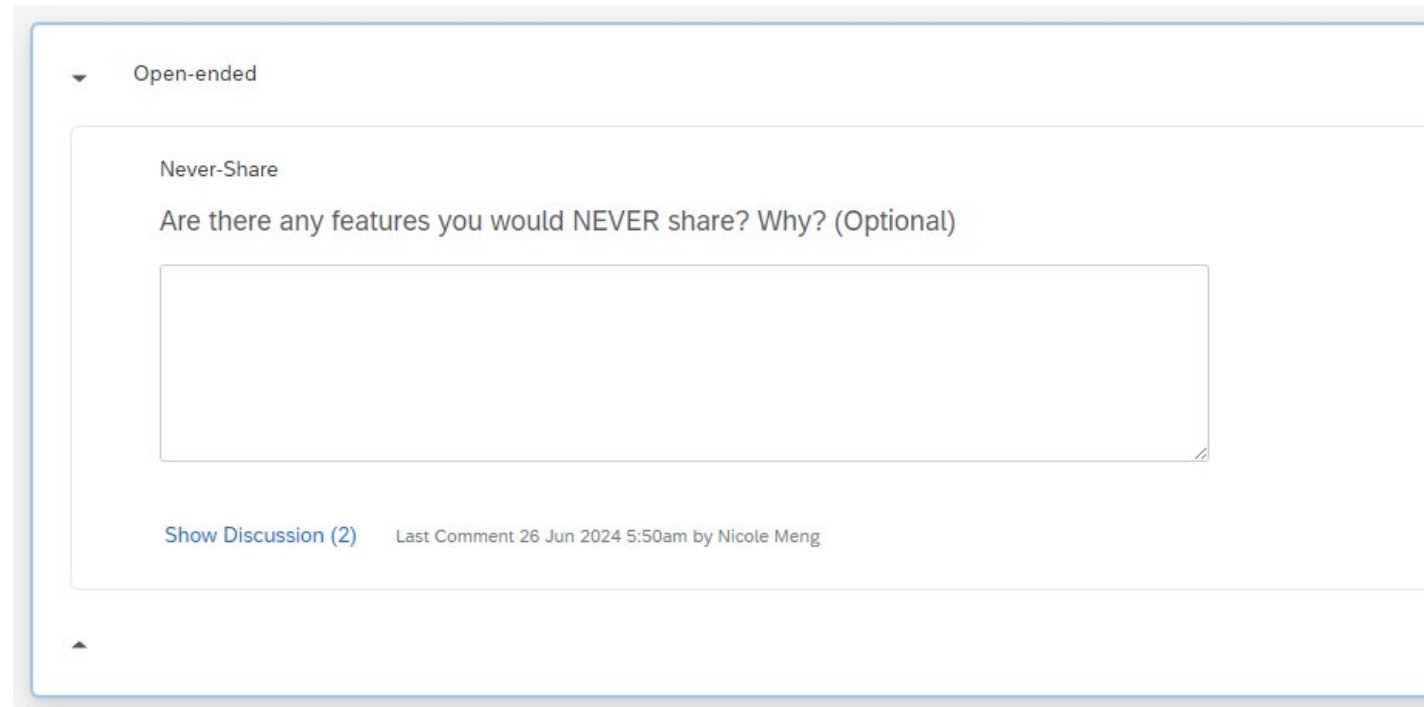
com-AOIn



Please feel free to tell us more about your ratings and reasons. (Optional)

Survey Structure

- Consent
- Screening
- Device Questions
- General Feature Questions
- Sharing Features
- **Open-ended**
- Privacy Scales
- Demographics
- Advert for future studies



The screenshot shows a survey question interface. At the top, there is a dropdown menu with the text "Open-ended". Below this, the question text reads "Never-Share" followed by "Are there any features you would NEVER share? Why? (Optional)". Underneath the question is a large, empty rectangular text input box. At the bottom of the interface, there is a link that says "Show Discussion (2)" and a timestamp "Last Comment 26 Jun 2024 5:50am by Nicole Meng".

Considerations for survey/questionnaire design

- Keep the number of questions short
- Only ask one question at a time!
- Consider how you phrase your question
 - Yes/no?
 - Do you need more details from participants that need to be covered in follow-up questions?
 - Multiple choice?
 - What choices do you include and how do these bias people's thoughts on the topic?
 - Open-ended?
 - Is the question specific enough? Will people know what you are asking?
 - How much time and effort is required to answer the question? Would the question be better asked in an interview?
- Carefully consider the order of questions
 - People consider questions in the context of those that they have already answered.
 - Consider the “flow” of questions from a participant perspective

Considerations for survey/questionnaire design

- Consider the survey/questionnaire format
 - Online / on-paper
- Consider how to collect data from a questionnaire
 - On-paper requires transcription of responses
 - Online requires ethical considerations
 - Where are participant responses stored?
 - Tools such as Google forms or survey monkey may be problematic
 - Use the [University's Qualtrics tool](#) for your research
 - <https://computing.help.inf.ed.ac.uk/qualtrics-survey-platform>

-

re-cap study variables

Study Variables

- Independent variables
 - Directly manipulated by the researcher (e.g., the type of interface you show to participants)
 - Pre-existing characteristics of participants (e.g., age, level of education, ...)
 - Study environment (e.g., lighting condition, computer type, ...)
- Dependent variables
 - Outcome variable that is manipulated by the study
 - The variable being tested
 - Examples: task performance, usability ratings, described satisfaction, answers to questions
- Fixed Aspect
 - Aspects of the study that are the same for all participants
 - Examples: survey structure, computer setup, wording of questions, order of tasks

study variables activity



- Get together in groups of 2-3 students
- Open the study variable tutorial
 - [HCI OpenCourse – Week 6](#)
- Go through as many case studies as you can
- Try and assign the variables and discuss your results

[15 min.]

next steps

- Week 7 – Data Analysis & Ethics
- Monday – Skyscanner guest lecture