Data Gathering

Dr Charles Martin

Announcements

- template for assignment 1 is available
- assignment 1 due **Monday 18 August, 23:59 on GitLab**
- assignment 2 specification will be published next week, you can see the "main idea" already on Canvas
- keep attending labs, if issues, apply for an extension (see course policies on Canvas)
- any questions, problems, **use the forum**!
- lab marks come out weekly via Canvas

Who has a question about assignment 1?

New module: user research, data, analysis, evaluation

1. plan data gathering sessions
2. plan and run an interview
3. design a questionnaire
4. understand observation studies

Main issues in data gathering

Main issues in data gathering

- why and how is data gathered?
- what kind of data? (just ratings or more?)
- necessary both for discovering requirements and evaluation
- today: introduce the main issues and techniques
- later weeks: look at how to structure evaluation and analyse different kinds of data

Figure 1: Just ratings? or more? (Photo by Towfiqu barbhuiya on Unsplash)

Get information about people, their behaviour and experiences with technology.

- What information and why?
- Depends on research problem and phase of research/design process, e.g.,
  - Comparing two alternative interfaces
  - Understanding a context of use
  - Measure time taken to complete a task
  - Discover how users interact with an existing system

Figure 2: What are the goals? (Photo by Ricardo Arce on Unsplash)

Identifying Participants

Who are the participants? How many are needed?

- Small group of stakeholders
- Criteria for involvement ("plays video games regularly")
- Random sampling from large population
- Convenience/volunteer sampling from those available
- Users with specific skills/needs
- Snowball sampling (participants help find more)
- Researchers make a justified choice
- Number of participants: most common in HCI is 12 (Caine, 2016)

Figure 3: People in the studio (2025)

Relationship between collector and provider

A data provider *gives* us data. What do they get back?

- Conduct research openly, ethically, responsibly
- Informed consent
- Clear communication of benefits and risks
- Respect and acknowledgement
- Particular care for some groups, e.g., Aboriginal and Torres Strait Islander communities.
- Practical: building rapport and understanding (not just taking)

Figure 4: An example consent form for research.

Ethical Considerations for Data

- data can have strings attached
- privacy
- personal information (including **name!**)
- sensitive information (embarrassing or harmful)
- storage requirements
- do we just use Google drive for everything?
- research needs a data management plan

Figure 5: We have to think about data. (Photo by Claudio Schwarz on Unsplash)

Triangulation

- *data triangulation:* data is drawn from different sources, times, places, people, etc
- *investigator triangulation:* different researchers (observers, interviewers, etc)
- *triangulation of theories:* different theoretical frameworks
- *methodological triangulation:* different data gathering or research techniques

Figure 6: Triangles support strong conclusions. (Photo by Charles 2010)

Pilot Studies

A small initial study to help plan a larger study.

- Could involve limited number of participants
- Limited interface or study parameters
- Checks that *expected* data can be obtained
- Can be called a "formative study" (to form the goals of the main study)
- E.g., assess the game platform used in a larger study (Mohaddesi & Harteveld, 2020)
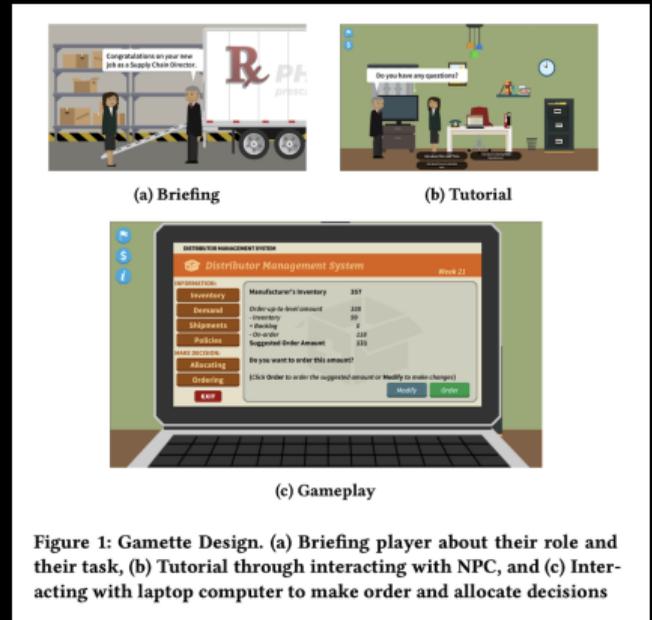

(a) Briefing (b) Tutorial

(c) Gameplay

Figure 1: Gamette Design. (a) Briefing player about their role and their task, (b) Tutorial through interacting with NPC, and (c) Interacting with laptop computer to make order and allocate decisions

Figure 7: A Gamette from Mohaddesi & Harteveld (2020)

Interviews

Conversation with a purpose (Kahn & Cannell, 1957)

Ask the users

- Interviews make a lot of sense in HCI
- What should a system do? Just ask.
- How should it work? Just ask.
- Did it work well? Ask away.
- How would you change it? Ask.

Figure 8: An interview. (Photo by Sam McGhee on Unsplash)

Unstructured Interviews

- exploratory, similar to conversations
- go into depth on experiences
- questions are open: no expectation on the content of answers or subsequent questions
- **probing:** can you tell me more about ...?
- benefit: generate rich, complex data
- limitation: time consuming to analyse

Could you tell me about your experience using the NewWidgetApp?
You mentioned that you enjoyed (feature X), why was that?
Can you explain more about what happened when you used (feature X)?

Structured Interviews

- questions are **predetermined**
- questions need to be short and clear
- questions can be closed (answers from specific options)
- typically, whole interview is scripted
- low skill to deliver
- fast to deliver
- useful when goals are clearly understood and questions and responses can be identified

What is your most used code editor: VS-Code, nvim, emacs, notepad.exe, or something else?
How often do you push your assessments to Gitlab: every minute, every hour, every day, just once?

Semi-structured Interviews

- features of both **structured** and **unstructured**
- typically: script with questions on main topics but discussion and probing follows each question
- intended to be somewhat replicable
- **probe:** neutral questions to gain more detail
- **prompt:** reminder of some part of the topic to gain specific information
- careful: prompting can preempt answers (bias!)

How are things going with the interaction experience?
Are there any [initial] impressions you want to share?
Can you describe the connections between your movements and the resulting sound?

(Reed, 2023)

Focus Groups

- interview with multiple participants (could be called a group interview)
- facilitator prompts discussion
- group members can influence each other (good or bad?)
- good for talking to lots of people.
- raising diverse viewpoints
- **facilitation:** need to be careful
- **groupthink:** generally considered harmful

(I don't find the "Focus Group" term super useful in my academic research—but they are widely used in industrial/government research.)

Figure 9: A group! Could be focussed. (Photo by Antenna on Unsplash)

Let's plan a *semi-structured* interview!

> *What questions should be asked in a semi structured interview about student's user experience with the "catchbox" (soft microphone used during lectures)?*

Use the poll everywhere link to suggest interview questions and vote on the best ones. **Ideate** for 2-3 minutes, **vote** for 1 minute, then let's discuss.

Figure 10: PollEverywhere link: https://pollev.com/charlesmarti205

17

Developing Interview Questions

- long questions are confusing
- jargon / technical language may be confusing (e.g., popover, jumbotron, nav bar, onboarding popup)
- keep questions neutral (e.g., "How much do you love this lecture?" is a leading question)
- questions need to support data gathering **goals**
- may not need to ask questions not related to goals

Easy to (accidentally) write bad interview questions!

- usually have a script for introductions, consent, etc (read verbatim)
- listen more than talk
- respond with sympathy but without bias
- enjoy the experience
- interviewing is **hard work**!

Figure 11: Yichen Wang and Marlene Radice in 2024

Capturing Data

- usually interviewing involves a combination of:
  - notes
  - audio
  - video
- different data have different issues and needs
  - notes: type up, handwriting (bias?)
  - audio: convert format, transcribe (suggest Aiko)
  - video: edit, store (large file sizes, anonymity)

Figure 12: Editing videos from a study in 2015.

Questionnaires

Questionnaires

Figure 13: Rating scales, and more (image by Towfique Barbhuiya on Unsplash)

Structure and Format

- written method of gathering structured data
- sometimes called a "survey", technically a survey is the whole study
- *questionnaire* or *survey instrument* is the paper form with questions
- questions can be open or closed
- often questionnaire used for demographic data
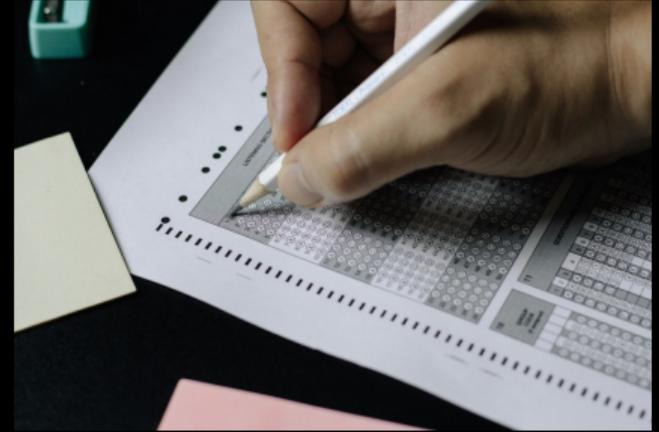- standard survey instruments are often questionnaires

Figure 14: Photo by Nguyen Dang Hoang Nhu on Unsplash

Questions might be something like:
What were the strengths of this course? Please provide any suggestions about how this course could be improved?

- the researcher might have little control over how seriously these questions are taken.
- potential to gather rich data
- **careful:** use these only with really open concepts

Figure 15: Wide open (2010)

Closed form questions

- closed form questions have preset responses from which the respondent must select.
- unordered responses
- rating scale questions
- **careful:** closed form questions design should be neutral and include most likely or relevant responses

What is your favourite fruit (select one answer): plum, tomato, pineapple"

How interesting is this lecture? Select a number from 1 (interesting) to 10 (amazingly interesting).

Likert or Lump It

- Likert scale questions are very common in questionnaires
- named after Rensis Likert (social scientist) (Likert, 1932)
- The question includes a statement, e.g.: "The workload was appropriate for this course"
- A number of levels of agreement are provided, e.g.: "strongly disagree, disagree, neutral, agree, strongly agree."
- 5-points is common, 3, 7, 9 or a continuous slider is also possible.

Figure 16: Some typical agreement scale questions

Semantic Differential Scale

- ratings of an object, concept, situation, etc
- the answer is a point between two opposite concepts
- e.g., describe your experience of using the (insert system here)
  - difficult to use — easy to use
  - boring — fun
  - slow to learn — fast to learn

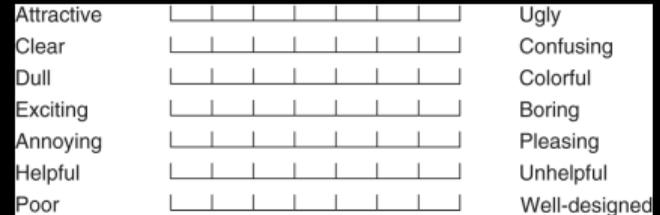| Attractive | | Ugly |
Clear		Confusing
Dull		Colorful
Exciting		Boring
Annoying		Pleasing
Helpful		Unhelpful
Poor		Well-designed

Figure 17: Example of semantic differential scales (image from Rogers et al., 2023)

Existing Survey Instruments

- Researchers sometimes choose to use well-known existing survey instruments rather than create their own.
- Existing surveys may be validated by having been tested and applied in many other studies.
- Sometimes the survey instrument comes with built-in instructions for analysing results.
- Let's look at some examples that are typical in HCI:
    - System Usability Scale (SUS) (Brooke, 1995; UIUX Trend, 2024)
    - NASA Task Load Index (TLX)
    - Creativity Support Index (CSI) (Carroll & Latulipe, 2009; Cherry & Latulipe, 2014)

System Usability Scale Questions

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

NASA Task Load Index Questions

1. How mentally demanding was the task?
2. How physically demanding was the task?
3. How hurried or rushed was the pace of the task?
4. How successful were you in accomplishing what you were asked to do?
5. How hard did you have to work to accomplish your level of performance?
6. How insecure, discouraged, irritated, stressed and annoyed were you?

Worksheets provided!, (Hart & Staveland, 1988), extra questions weight the ratings.

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?
Very Low — Very High

Physical Demand — How physically demanding was the task?
Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?
Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?
Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?
Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed were you?
Very Low — Very High

Creativity Support Index

- **Idea:** measure how well a system can support creativity Cherry & Latulipe (2014)
- **creativity support tools:** writing, visualisation, video editing, music tools etc
- creativity is a bit hard to define, but the CSI includes:
  - exploration, expressiveness, immersion, enjoyment, results worth effort, collaboration
- inspired by TLX, 2 questions per factor + 15 paired comparisons.

Figure 19: creativity support index using a visual analogue scale (VAS) (Cherry & Latulipe, 2014)

Questionnaire Tips

- It's hard to write survey questions! It may be good for beginners to use a standard questionnaire.
- The more questions you ask, the more work it can be to analyse. It can be counter-productive to have lots of questions without a way to aggregate them.
- Rating scale data is generally not continuous and so you need to use non-parametric significance tests.
- The distribution of survey data is usually important, good to use plots such as box plots rather than a mean and standard
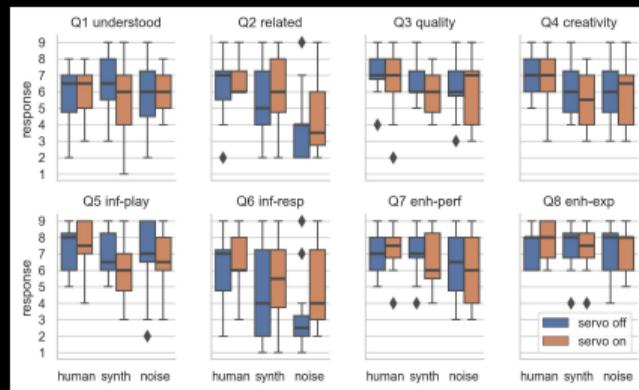
Figure 20: Some boxplots from a survey.

Activity: Write a questionnaire

Let's write a questionnaire about ANU students' class enrolment experience.

- What questions should we ask on a questionnaire about ANU student's enrolment experience? If closed list options, if open write "(open question)"

Talk for 2-3 minutes and add some examples on PollEverywhere, then let's have some discussion.

https://PollEv.com/charlesmarti205

Figure 21: ANU course enrolment.

Observation

Figure 22: What are they doing and why? (Photo by Philippe Bout on Unsplash)

Direct Observation in the Wild

- watching people for science
- using system or technology in the normal context of use
  - e.g., a researcher joins a tour group to observe use of a travel navigation app
  - e.g., a researcher watches a concert carefully to observe the performer's use of music technology

Example framework:

- **The person:** who is using the technology at any given point?
- **The place:** where are they using it?
- **The thing:** what are they doing with it?

Figure 23: We don't actually stalk people. (Photo by Philippe Bout on Unsplash)

- Literally "study of culture"
- immersion and participation in a research context.
- e.g., to create a factory management system, a researcher might embed themself in a factory, talk to workers and perform tasks
- ethnography is often used in relation to specific workplaces, e.g.: factories, hospitals etc
- being within the context of the participants rather than bringing them to a lab/classroom for study
- ethnography as a methodology connected

Figure 24: (Photo by Homa Appliances on Unsplash)

Direct Observation in Controlled Environments

- observation in a lab-based setting
- could be possible to identify close details
- record all aspects of technology use
- record data, audio, video as well as observations

Figure 25: Observation via video in a lab (2015)

Think-Aloud Technique

- **Problem:** observers don't know what participants are thinking

- **Solution:** ask participants to say everything they are thinking and trying to do when using an interface, so we know!

- can produce very useful data

- hard work for the participant

- needs careful facilitation from the observer

- more: Thinking aloud, the #1 usability tool

Video-cued Recall

- it can be difficult to remember details from a long sequence of tasks.
- techniques like video cued recall can help participants provide a commentary on an experience.
- the idea is: you record video of the participant completing an an interaction and then have an open-ended discussion while watching it back.

Figure 26: Videos of an interaction for discussion.

- We can observe without being present
- Useful for
    - embedding systems in people's everyday life
    - interacting with participants remotely
    - tracking more participants than you could directly observe

Diaries

- participants write about their experiences with a system regularly in a diary
- easy for the researcher at the collection stage
- takes continuous effort from the participants (reminders? structure?)
- relies on participant's memory and subjective account
- video and photos can reinforce written accounts

Figure 27: The participant just writes down their experience! (2025)

- **Interaction log:** a log of data captured from a system showing exactly what the participant did at any given time
- e.g.: key presses, mouse movements, interactions with GUI components, sensor data
- time spent on actions or using software (e.g., playing a game, using instagram)
- unobtrusive, automatic
- lots of data, should be visualised or analysed to develop findings
- scraping data from public sources (e.g., social media) can be observation

N.B. scraping large amounts of data raises

Figure 28: Interaction data from a concert (2014)

Activity: Would you rather (observation edition)

As an HCI researcher, would you rather:

1. observe directly in the wild
2. observe directly in a lab
3. observe indirectly through a diary
4. observe indirectly through data analytics

Discuss for 2-3 minutes in groups, rank on PollEverywhere, and then we will hear some responses!

Figure 29: PollEverywhere link: https://pollev.com/charlesmarti205

42

Data Gathering in Practice

Back to our issues:

1. goals
2. participants
3. ethics, consent with participants
4. ethics of data
5. triangulation
6. pilot studies vs main studies

Figure 30: Collecting some MIDI and audio data in 2018 (artist: Bernt-Isak Waerstad)

Choosing Techniques (Rogers et al., 2023)

technique	good For	kind of data	issues
interviews	exploring issues	mostly qualitative	skilled work but high value
focus groups	multiple viewpoints	mostly qualitative	efficient, groupthink risk
questionnaires	specific questions	quantitative and qualitative	lots of participants, hard to design
direct observation in the wild	context of use	mostly qualitative	very useful data, very time consuming
direct observation in a controlled	capturing details	quantitative and qualitative	situation can be artificial/unrealistic

- data gathering should adapt to different participants
- **children:** think and react differently to adults, e.g., adapt scales to visual representations (Putnam et al., 2020)
- **people with disabilities:** caregivers might be involved to adapt interview questions
- **animals(!):** hard to interpret behaviour, need to design interfaces to work without hurting animals (Mancini et al., 2017)

Figure 31: ACI is real! (Photo by Alison Pang on Unsplash)

- expectations for remote data gathering changed completely during pandemic!
- can access participants in different countries, age groups, abilities, specific or expert users
- observation could be recorded and then analysed later
- best practices might include (Mastrianni et al., 2021):
  - running pilot tests before conducting sessions
  - have backups in place in case of issues
  - informing participants of technical requierments
  - use respective questioning if issues with

Figure 32: Interviews and observations can happen remotely! (Photo by Chris Montgomery on Unsplash)

Questionnaires vs Interviews

- Questionnaires, while very useful can only tell us information that we know to ask about.
- Interviews let us find out information that didn't know to begin with.
- Questionnaires are fast and can address lots of participants—a broad approach
- Interviews are slow but are a deep approach.
- Interviews also require good conversational skills and experience to conduct well.

Who has a question?
- I can take *cathchbox* question up until 2:55
- For after class questions: meet me outside the classroom at the bar (for 30 minutes)
- Feel free to ask about **any aspect of the course**
- Also feel free to ask about **any aspect of computing at ANU**! I may not be able to help, but I can listen.

Figure 33: Meet you *at the bar* for questions. 🍸🥤🫖☕
Unfortunately no drinks served! 🙃

References

Brooke, J. (1995). SUS: A quick and dirty usability scale. *Usability Eval. Ind.*, *189*.

Caine, K. (2016). Local standards for sample size at CHI. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 981–992. https://doi.org/10.1145/2858036.2858498

Carroll, E. A., & Latulipe, C. (2009). The creativity support index. *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, 4009–4014. https://doi.org/10.1145/1520340.1520609

Cherry, E., & Latulipe, C. (2014). Quantifying the creativity support of digital tools through the creativity support index. *ACM Trans. Comput.-Hum. Interact.*, *21*(4). https://doi.org/10.1145/2617588

Hart, S. G., & Staveland, L. E. (1988). *Development of NASA-TLX (task load index): Results of empirical and theoretical research*. *52*, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

Kahn, R. L., & Cannell, C. F. (1957). *The dynamics of interviewing; theory, technique, and cases*. John Wiley & Sons.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*.

Mancini, C., Lawson, S., & Juhlin, O. (2017). Animal-computer interaction: The emergence of a discipline. *International Journal of Human-Computer Studies*, *98*, 129–134. https://doi.org/https://doi.org/10.1016/j.ijhcs.2016.10.003

Mastrianni, A., Kulp, L., & Sarcevic, A. (2021). Transitioning to remote user-centered design activities in the emergency medical field during a pandemic. *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3411763.3443444

Mohaddesi, O., & Harteveld, C. (2020). The importance of pilot studies for gamified research: Pre-testing gamettes to study supply chain decisions. *Extended Abstracts of the 2020 Annual Symposium on Computer-Human Interaction in Play*, 316–320. https://doi.org/10.1145/3383668.3419889

Putnam, C., Puthenmadom, M., Cuerdo, M. A., Wang, W., & Paul, N. (2020). Adaptation of the system usability scale for user testing with children. *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–7. https://doi.org/10.1145/3334480.3382840

Reed, C. N. (2023). *Imagining & sensing: Understanding and extending the vocalist-voice relationship through biosignal feedback* [PhD thesis, Queen Mary University of London]. https://qmro.qmul.ac.uk/xmlui/handle/123456789/89579

Rogers, Y., Sharp, H., & Preece, J. (2023). *Interaction design: Beyond human-computer interaction, 6th edition*. John Wiley & Sons, Inc. https://quicklink.anu.edu.au/kv9b

UIUX Trend. (2024). *Measuring system usability scale (SUS)*. UIUX Trend. https://uiuxtrend.com/measuring-system-usability-scale-sus/