

Collecting Data II

-Observing People

Prepared by:

Dr. Fahad Ahmed

Topics of Discussion

- Recap from Last Week
- Usability Testing Methods
 - Think Aloud Protocol
 - Eye Tracking
 - Wizard of Oz
- Direct Observation Techniques
 - Naturalistic
 - Ethnography
 - Controlled
 - Structured
- Analytics & Log Data

From Last Week...

You designed interview questions. But can you always trust what people say?



What people SAY

"I always read the error message carefully before clicking OK"

"I check my bank balance every week"

"I would never share my password"

Captured by interviews, questionnaires, focus groups (Collecting Data I)



What people DO

"They dismiss the error dialog in under 0.3 seconds without reading it"

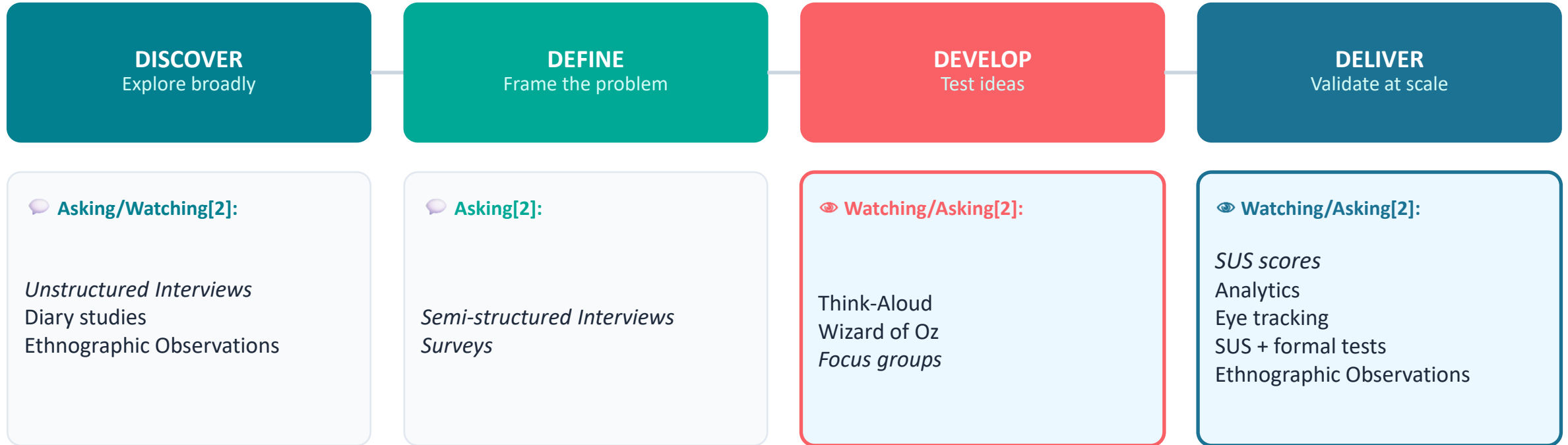
"They check their balance when something goes wrong — not weekly"

"They write the password on a sticky note under the keyboard"

Captured by observation, usability tests, analytics (Collecting Data II)

Where do Watching Methods fit?

Observation dominates the Develop and Deliver phases — when you have something to test.



 This week we focus on Develop & Deliver methods. Last week's asking methods remain active — observation adds a layer of evidence that asking alone cannot provide.

Usability Testing

Observe real users attempting real tasks, the most direct measure of usability [3,4].

01

Recruit participants

Typically 5–8 users representative of the target audience. Nielsen's rule: 5 users reveal ~85% of usability problems.

02

Set realistic tasks

Give participants goal-oriented scenarios, not instructions. Not: 'Click the search button'. Instead: 'Find a hotel in Cardiff for next weekend.'

03

Observe without helping

The hardest part. Do not guide, hint, or intervene — even when the participant is clearly struggling. Struggle reveals design problems.

04

Measure and record

Task completion rate, time on task, error rate, and satisfaction (via SUS). Record screen, audio, and face where consented.

05

Debrief the participant

After tasks are complete, ask open questions: 'What did you find confusing?' 'What would you change?' This is where asking and watching combine.

06

Analyse and iterate

Identify recurring failure points. One session informs the next prototype version — this is the Develop phase's iterative loop.

[3] Dumas, J.S. and Redish, J.C. (1999). *A Practical Guide to Usability Testing*. Rev. edn. Exeter: Intellect Books.

[4] Nielsen, J. (1994). *Usability Engineering*. San Francisco: Morgan Kaufmann.

Think-Aloud Protocol

Ask participants to narrate their thoughts as they work, making invisible reasoning visible [5, 6].

Concurrent Think-Aloud

Participant narrates in real time while performing the task.

✓ Advantages:

- Captures moment-by-moment reasoning
- Reveals where expectations break down
- Requires no post-task memory recall

✗ Limitations:

- Can slow down task performance
- Narrating may feel unnatural at first
- Dual-task load may affect behaviour

Best for: Most usability tests — the default approach



[5] Nielsen, J. (1993). *Usability Engineering*. Cambridge, MA: Academic Press.

[6] Lazar, J., Feng, J.H. and Hochheiser, H. (2017). *Research Methods in Human-Computer Interaction*. 2nd edn. Cambridge, MA: Morgan Kaufmann. (Chapter 8)

Video source: <https://www.youtube.com/watch?v=tXJqAYDbRzI&t=5s>

Think-Aloud Protocol

Ask participants to narrate their thoughts as they work, making invisible reasoning visible [5, 6].

Concurrent Think-Aloud

Participant narrates in real time while performing the task.

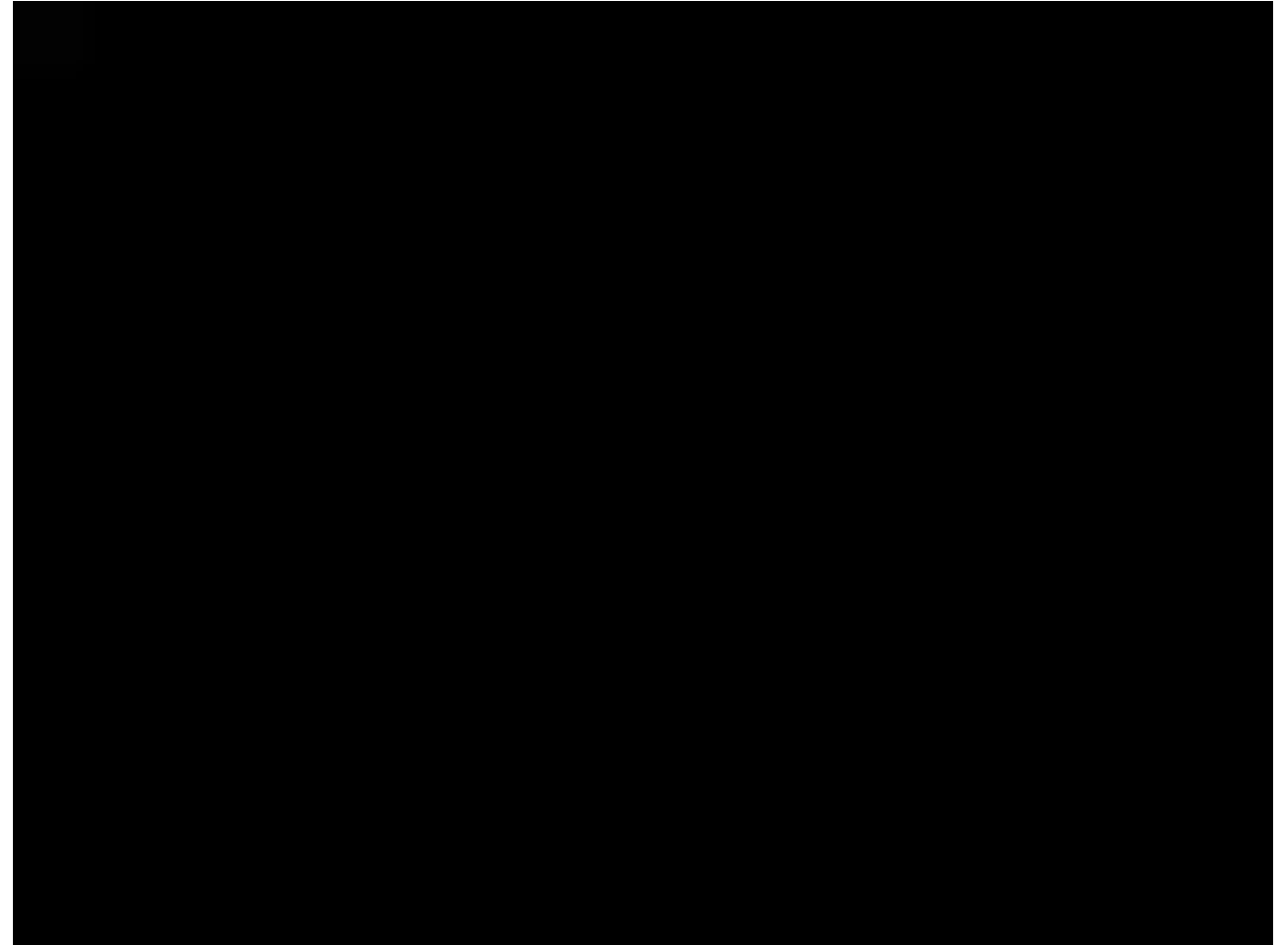
✓ Advantages:

- Captures moment-by-moment reasoning
- Reveals where expectations break down
- Requires no post-task memory recall

✗ Limitations:

- Can slow down task performance
- Narrating may feel unnatural at first
- Dual-task load may affect behaviour

Best for: Most usability tests — the default approach



[5] Nielsen, J. (1993). *Usability Engineering*. Cambridge, MA: Academic Press.

[6] Lazar, J., Feng, J.H. and Hochheiser, H. (2017). *Research Methods in Human-Computer Interaction*. 2nd edn. Cambridge, MA: Morgan Kaufmann. (Chapter 8)

Video source: <https://www.youtube.com/watch?v=6kASeSntsKq&pp=ygUOdGhpbmssqYWxvdWQqdXq%3D>

Think-Aloud Protocol

Ask participants to narrate their thoughts as they work, making invisible reasoning visible [5, 6].

Retrospective Think-Aloud

Participant watches a replay of their session and narrates retrospectively.

✓ Advantages:

- Does not interfere with task behaviour
- More natural task performance data
- Participant can reflect more calmly

✗ Limitations:

- Relies on memory — recall bias risk
- Time-consuming (double the session length)
- Needs screen recording infrastructure

Best for: When task realism is critical (e.g. timed tasks, stress scenarios)



[5] Nielsen, J. (1993). *Usability Engineering*. Cambridge, MA: Academic Press.

[6] Lazar, J., Feng, J.H. and Hochheiser, H. (2017). *Research Methods in Human-Computer Interaction*. 2nd edn. Cambridge, MA: Morgan Kaufmann. (Chapter 8)

Video source: <https://www.youtube.com/watch?v=rqj7Oi8rCjM>

Eye Tracking

Measures where users look, attention is not the same as comprehension [7].

What it measures

Fixations:

Where the eye pauses and focuses — what the user is actively processing

Saccades:

Rapid eye movements between fixations — reveals reading path and scanning pattern

Heat maps:

Aggregated fixation density across many participants — shows which areas receive attention

Gaze plots:

Sequenced fixation path for one participant — shows order of attention, not just location

Strengths

- Objective, non-intrusive, real-time
- Reveals Banner Blindness — users ignore ad-like elements even on a path to their goal
- Reveals F-pattern reading behaviour on text-heavy pages

Limitations

- Expensive hardware (tobii eye-tracker: £3,000–£30,000)
- Looking ≠ understanding — can't infer cognition from gaze alone
- Sensitive to lighting, glasses, and participant head movement
- Small samples — lab-based, not scalable



Design implication: F-pattern findings tell us to place critical information top-left and in the first words of headings — users scan, not read. Banner blindness tells us never to put key navigation in areas that look like ads.

Wizard of Oz

The participant believes they are using a working system, but a human is operating it behind the scenes [8, 9].



PARTICIPANT

Believes the system is real
and fully automated



"SYSTEM"

Interface looks functional
but responses are faked



THE WIZARD

Researcher secretly
generates system responses

Hidden from participant



[8] Kelley, J.F. (1984). An iterative design methodology for user-friendly natural language office information applications. *ACM Transactions on Information Systems*, 2(1), pp.26–41.

[9] Lazar et al. (2017). *Research Methods in HCI*. Morgan Kaufmann. 10
Video Source: https://www.youtube.com/watch?v=gDsb0vW_LM8

Wizard of Oz

The participant believes they are using a working system — but a human is operating it behind the scenes [8, 9].



Why use it?

Test the experience of a system before it is built. Especially useful for AI, voice interfaces, and complex workflows where prototyping full functionality is infeasible.

Asking or watching?

Primarily watching — you observe authentic user behaviour under realistic conditions. The debrief interview afterward adds an asking component. It is both, but observation-first.

Ethical consideration

Deception requires careful ethical management. Participants must be fully debriefed immediately after. Informed consent covers 'a study involving interaction with a prototype system.'

[8] Kelley, J.F. (1984). An iterative design methodology for user-friendly natural language office information applications. *ACM Transactions on Information Systems*, 2(1), pp.26–41.

[9] Lazar et al. (2017). *Research Methods in HCI*. Morgan Kaufmann.

Direct Observation

Watching users in context, without asking them anything [10, 11].

← Natural / uncontrolled

Structured / controlled →

Naturalistic

- Observer watches in the user's real environment (workplace, home, public space)
- No tasks set — behaviour unfolds naturally
- Captures workarounds and informal practices labs cannot reveal
- Risk: observer presence changes behaviour (Hawthorne effect)

Best for: Discover phase, ethnographic research, Activity Theory studies

Controlled

- Conducted in a lab or set-up environment
- Researcher defines specific tasks to observe
- Easier to compare across participants
- Less ecologically valid than naturalistic

Best for: Develop phase, focused usability investigations

Structured

- Uses a predefined coding scheme or checklist
- Observer ticks/counts specific behaviours (e.g. errors, hesitations)
- Produces quantitative observation data
- Requires pilot testing of the coding scheme

Best for: Acceptance testing, comparative studies, measurable outcomes

[10] Robson, C. and McCartan, K. (2016). *Real World Research*. 4th edn. Chichester: Wiley. (Chapter 12)

[11] Suchman, L.A. (1987). *Plans and Situated Actions*. Cambridge: Cambridge University Press.

What is Ethnography?

A research method from anthropology: go to the people, immerse in their world, understand what they actually do [12].

Etymology: Greek *ethnos* (people / nation) + *graphia* (writing) = 'writing about people'

Originally used by anthropologists studying distant cultures — now widely applied in HCI, design research, and organisational studies.



In situ

- Research happens in the user's natural environment i.e, workplace, home, public space and not a lab.
- Context is not controlled; it is the data.



Extended immersion

- Ethnographers spend extended time in the field i.e., days, weeks, or months.
- Short visits miss routines, workarounds, and informal practices.



Naturalistic observation

- The researcher observes without imposing tasks or questions.
- The goal is to see practice as it unfolds, not as participants describe it or perform it for a researcher.

"Ethnography is the study of social interaction and culture in its naturally occurring context." — Hammersley & Atkinson (2007) [12]

Conducting Ethnographic Research

Practical guidance for design-focused ethnographic fieldwork [13].

1

Negotiate access

Get permission from an organisation, team, or community. Be clear about what you will observe, how long, what you will do with the data, and who will see it. Start small.

2

Adopt a 'stranger' mindset

Treat the familiar as unfamiliar. Ask about things that seem obvious — users often can't articulate tacit knowledge until a researcher asks 'why do you do it that way?'

3

Observe before you ask

Spend time watching before conducting any interviews. Observation first means your questions will be grounded in what actually happens, not your assumptions about what happens.

4

Write field notes immediately

Write up detailed notes as soon as possible after each session (within the same day). Memory degrades rapidly. Include: what you saw, what was said, your reflections, and open questions.

5

Look for breakdowns and workarounds

The most useful design findings are often visible in moments of failure — where technology breaks down, where people improvise, and where informal tools substitute for designed ones.

6

Translate findings into design insights

Ethnography produces rich data, not design specifications. The step from fieldwork to design requires active interpretation: what does this finding imply for the interface, the workflow, or the system?

Analytics & Log Data

Observation at scale, watching thousands of users without being in the room [14].

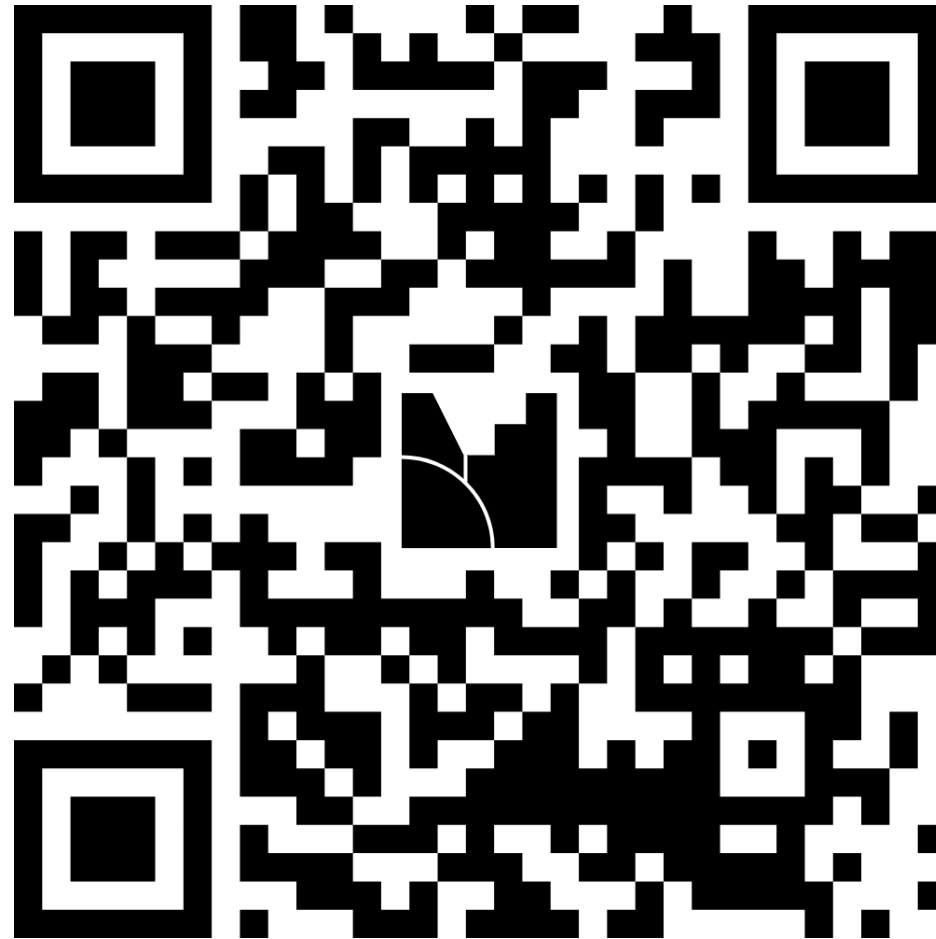
Metric	What it measures	Why it matters for usability
 Task completion rate	% of users who successfully finish a defined task	Core effectiveness metric (ISO 9241-11)
 Time on task	How long users take to complete a task	Core efficiency metric — longer ≠ more engaged; often means confused
 Error rate	How often users make mistakes or take wrong paths	Directly pinpoints interaction breakdowns
 Drop-off / exit rate	Where users abandon a flow without completing it	Identifies the worst friction point in a journey
 Click / heat maps	Aggregated click positions across all users (e.g. Hotjar, Microsoft Clarity)	Shows where users expect interactivity — even when none exists
 Session recordings	Video replay of individual user sessions (no camera needed)	Combine qualitative insight with scale — watch real behaviour

[14] Lazar, J., Feng, J.H. and Hochheiser, H. (2017). *Research Methods in Human-Computer Interaction*. 2nd edn. Morgan Kaufmann. (Chapter 10) Note: ISO 9241-11 (2018) defines effectiveness, efficiency, satisfaction as core usability metrics.

Thank you for your attention!

Questions?

Feedback



Menti Link:

<https://www.menti.com/alwotdzo88vw>

Menti Code: **7201 9197**