Lesson 5: Usability Testing

1) Key takeaways from the textbook

Usability testing focus. Evaluate a product by observing representative users completing predefined tasks under controlled conditions. Data typically includes observation notes/video/logs plus a brief post-test questionnaire.

Common performance measures. Success rate, time on task, error rate/type, help-seeking, and number of users experiencing a given problem.

Number of participants. For formative tests, 5–12 participants is a common range; very early checks can use 2–3 users.

Labs and equipment. Dedicated labs (or pop-up labs) enable controlled recording (audio/video/screen/keystrokes) and an observation room; remote and mobile setups are viable alternatives.

2) Chapter 8 highlights

Gather data that answers your design/evaluation questions. Methods include interviews, surveys, observation/contextual inquiry, diary studies, documentation/analytics. Triangulate methods; pilot instruments; plan sampling/recruitment; address ethics (consent, privacy).

3) Chapter 15.1–15.2 specifics (Usability testing in practice)

 $Plan \rightarrow Run \rightarrow Analyze \rightarrow Report$:

- Define goals & hypotheses; select representative tasks and success criteria.
- Recruit typical users; choose metrics & instruments (notes, audio, video, logs, eyetracking as needed).
- Run pilot(s); standardize moderator script; manage consent and data security.
- Analyze performance & observations; rate severity; synthesize themes.
- Report clear findings with evidence and actionable recommendations.

Practical notes: use think-aloud to reveal reasoning; include a brief post-test satisfaction measure; be careful generalizing from small samples and state context/limits.

4) Think-Aloud protocols (Olmsted-Hawala et al., 2010)

Compared four conditions on a federal data website: Traditional TA, Speech-communication TA (minimal back-channel prompts), Coaching TA (active guidance), and a silent control.

- Accuracy: Coaching > Traditional/Speech/Control.
- Satisfaction: Coaching > Traditional/Speech.

Efficiency (time-on-task): No significant differences.

Implication: Coaching can inflate success/satisfaction; avoid if you want unaided behavior. Prefer Traditional or Speech-communication TA for less moderator interference.

5) Instant Data Analysis (Kjeldskov et al., 2004) – "evaluation in a day"

Process: Run 4–6 think-aloud sessions, then same-day structured team debrief to produce a prioritized problem list with evidence.

- Captured a large share of critical/serious issues in a fraction of analysis time compared to full video transcription.
- Trade-off: less traceability/detail; works best when observers and redesigners overlap.

6) From the slides (UXU 5 – Usability Testing)

- User-centered approach: early focus on users/tasks; empirical measurement; iterative design.
- Think-aloud: four modes (Traditional, Speech-communication, Coaching, Silence).
- Recording options: notes/photos, audio, video.
- Class exercises practiced realistic task scenarios and moderation.

7) Quick checklist for your own usability test

- Goal & scope agreed (features, users, success criteria).
- 5–8 representative tasks; success/fail criteria; critical paths.
- Recruit 5–12 typical users; consent & privacy addressed.
- Script: intro, warm-up, TA instructions (choose protocol), debrief.
- Instruments: recording setup, task sheets, metrics sheet, post-test survey.
- Pilot once; adjust timings/prompts.
- Run consistently; minimize moderator bias.
- Debrief & analyze; rate severity; link evidence; recommend fixes.
- Report succinctly with screenshots and a prioritized backlog.

8) Study prompts

- When would you choose lab vs. field vs. remote testing?
- How would TA protocol choice affect what you can conclude?
- Where would IDA be appropriate—and when would you avoid it?
- Which metrics best capture success for your project?