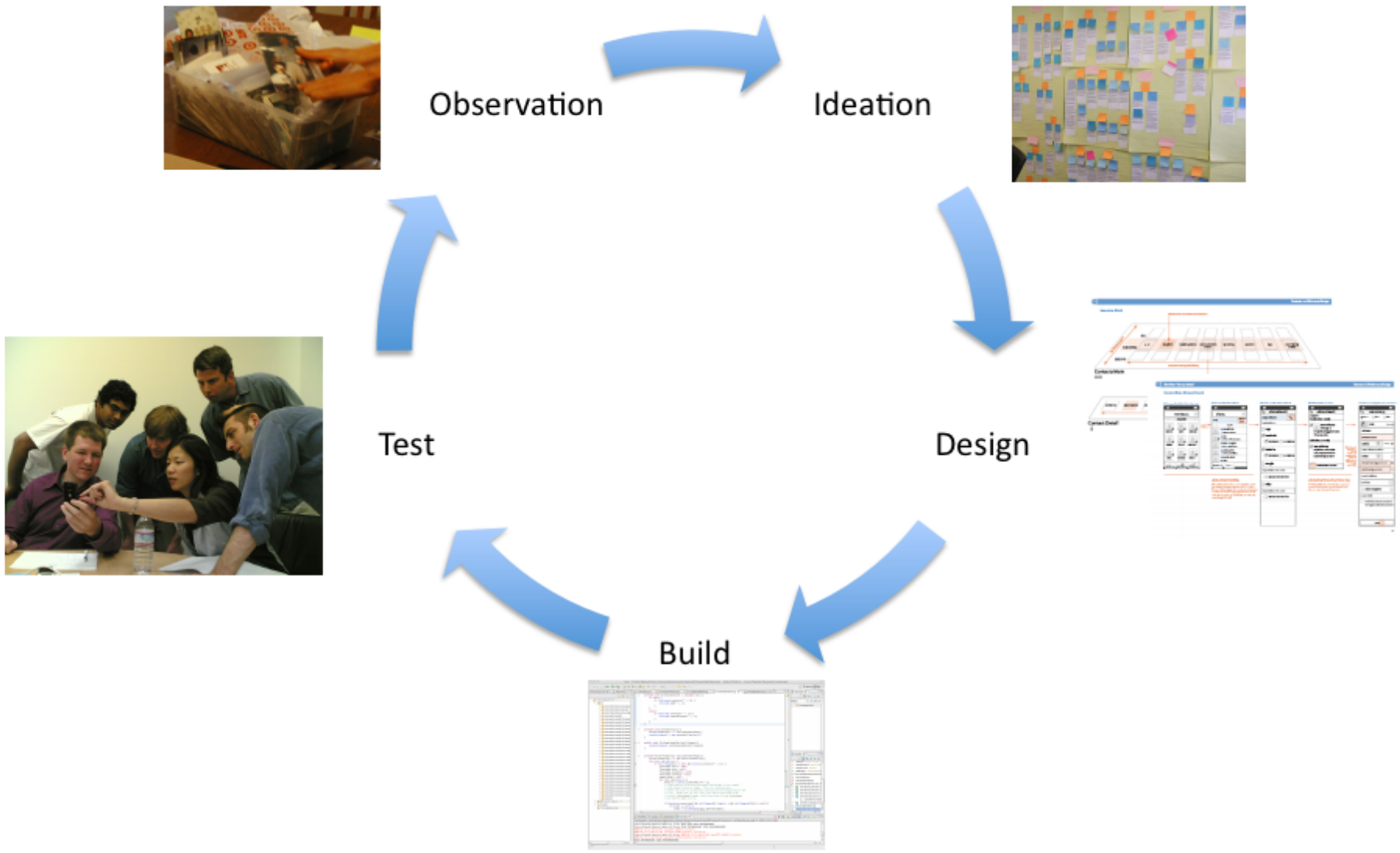
Science of the Project: Session 4

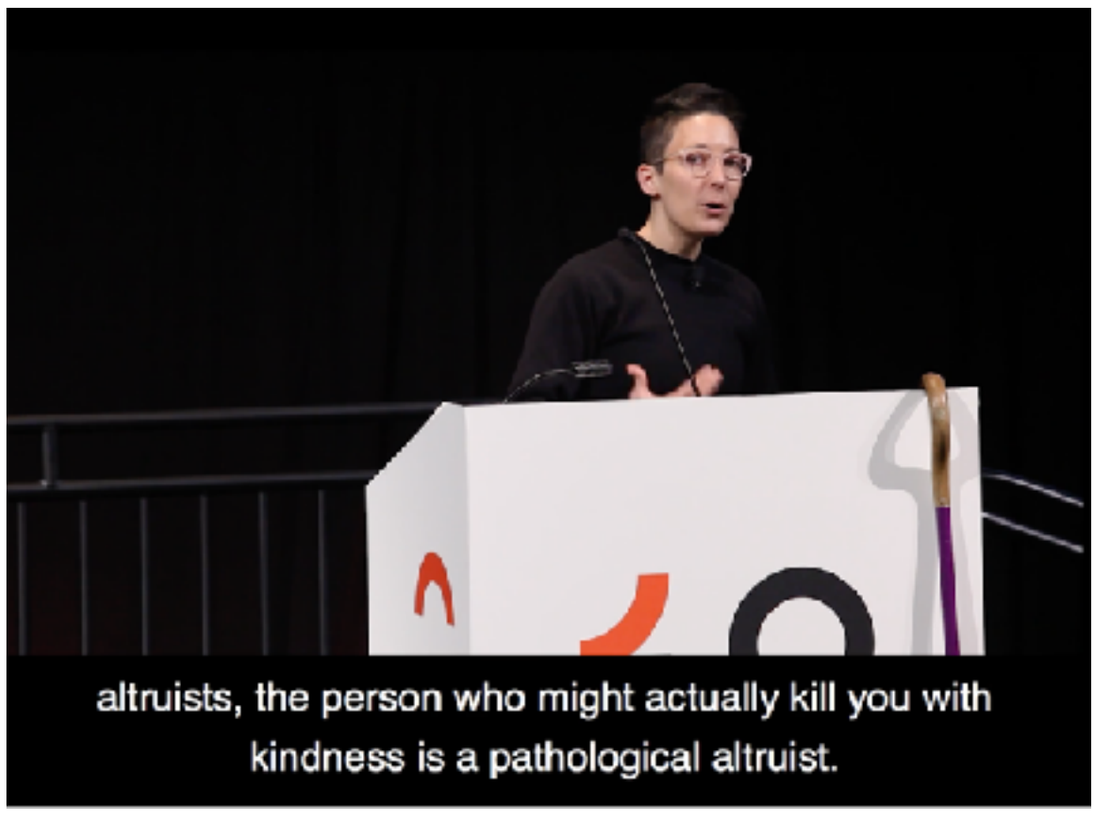
Evaluation processes in design research

from yesterday:
how would you evaluate the data you gather?

Why evaluate our work?

evaluation as part of a design process





- Aral Balkan: Superheroes & Villains in Design https://vimeo.com/70030549

- Liz Jackson: Empathy reifies disability stigmas https://interaction19.ixda.org/program/keynote--liz-jackson/

Design approaches

User-Centered Design

- - Focus on users, their goals, needs, and context

Activity-Centered Design

- - Main focus on activities

Systems Design

- - Consider needs of the system as a whole
- - System includes people as well as multiple devices

(**non user-centered design** is starting to be discussed, as a mean to include environmental and ecological factors in the impact of the design)

(Cooper et al., 2007)

How can we evaluate our work?

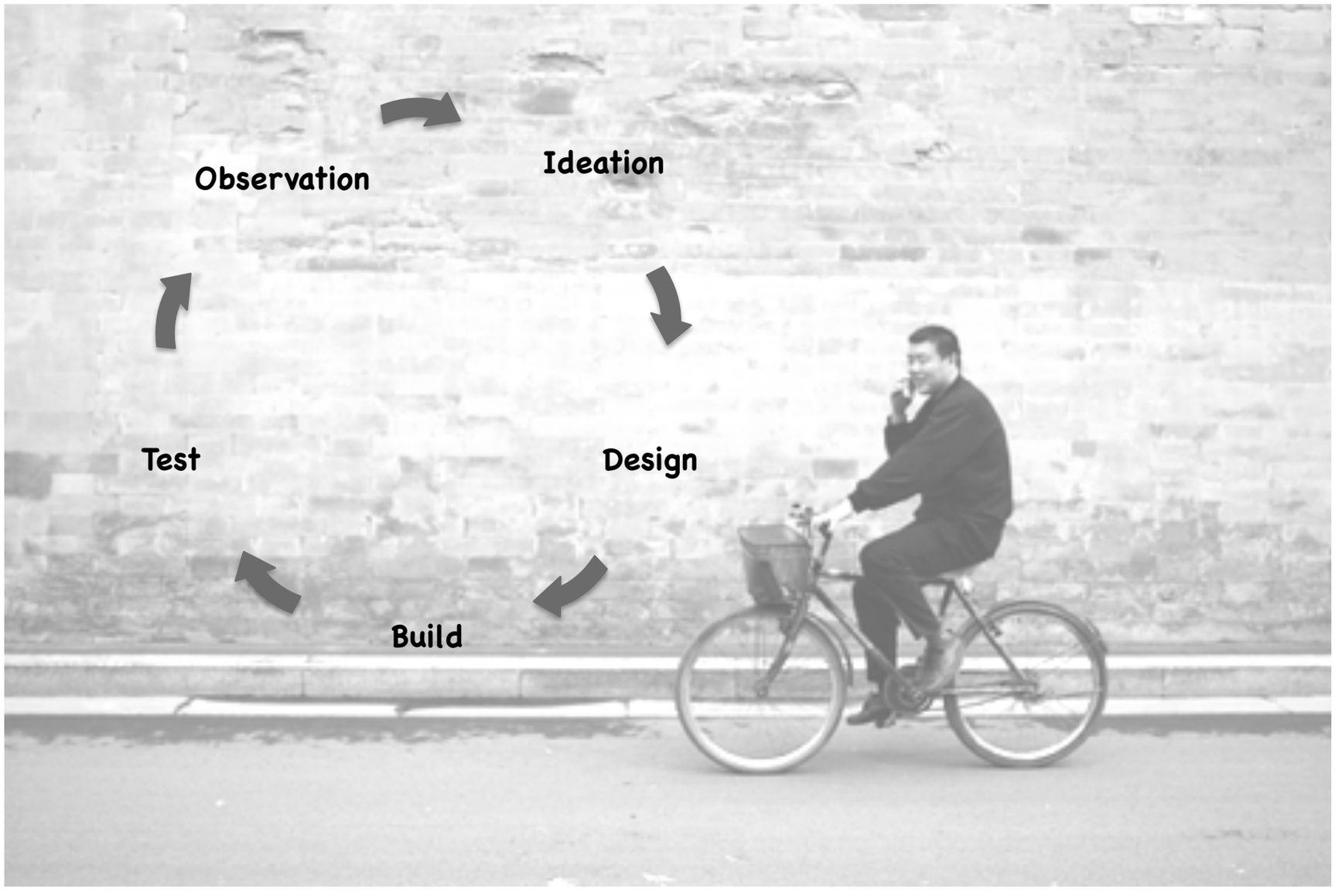
Generative Research Methods

- Home tours
- Field visits
- Diary/Voicemail studies, Probes
- Contextual inquiry
- Task analysis
- Communication logging, Conversation analysis, Affinity analysis
- Semi-structured interviews
- Content analysis
- Field trial with prototype
- Discount Method

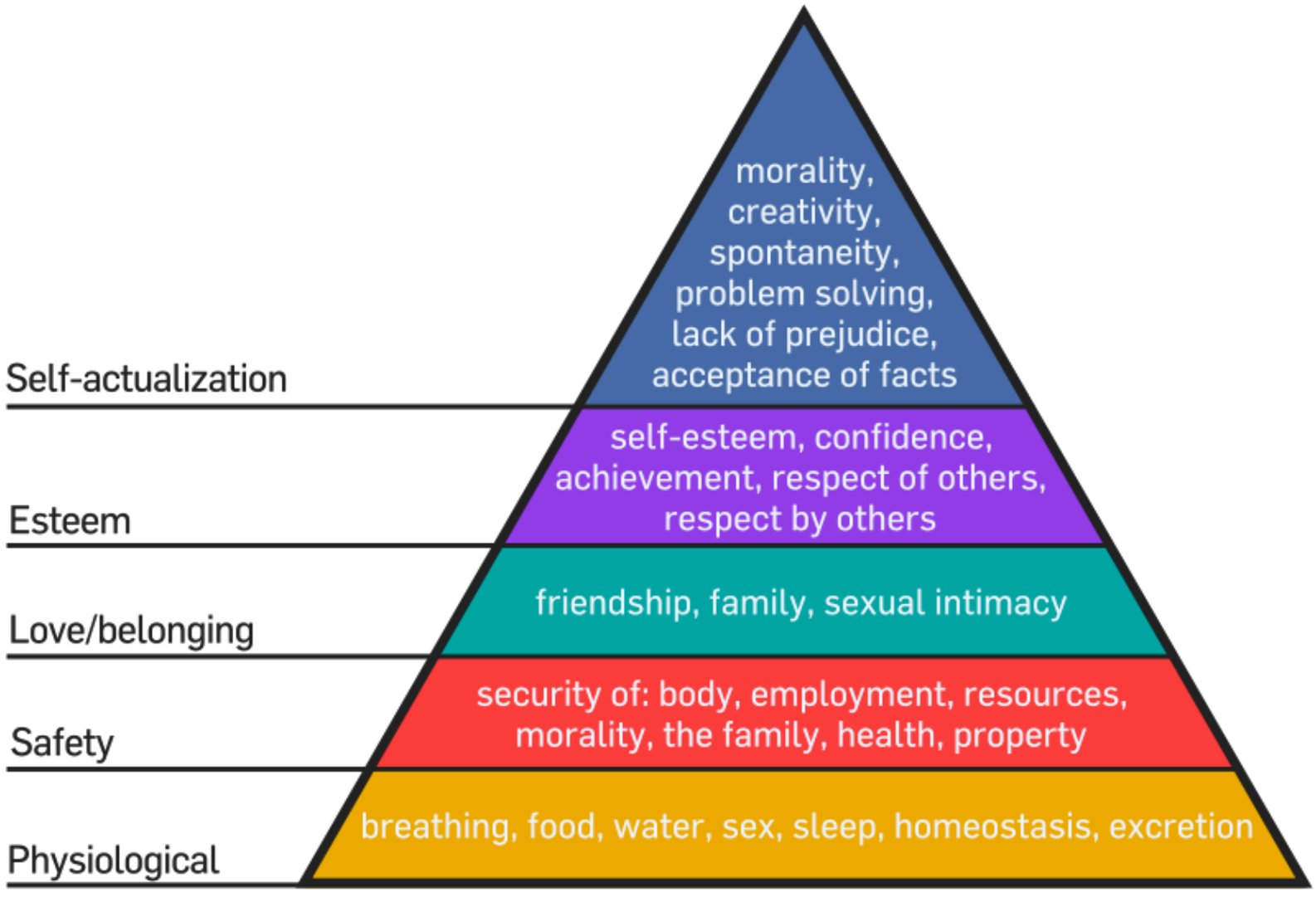
(Bentley & Barrett, 2012, Ch. 2 and 3)

Evaluation Methods

- Home tours, Field visits
- Diary/Voicemail studies, Probes
- Semi-structured interviews
- Content analysis
- Field trial with prototype
- Instrumentation
- Usability Evaluation
- - e.g., KLM-GOMS, Questionnaires


Observation
Ideation
Design
Build
Test

Hierarchy of needs



Maslov 1943

Figure 1. Materials assigned to each participant of the qualitative study: a) paper diary, b) instructions, c) smiley stickers to rate entries of the paper diary, d) camcorder to record video diary, and e) paper frames to enhance explanations in the video diary

(Cherubini et al., 2011)

New Operators

Macro Attention Shift (S_{Macro})

One major difference from desktop to phone interaction is that the attention of users may be split between the phone and the real world surrounding them (Figure 3).

Figure 3: Attention shift (S_{Macro}) between the mobile phone and objects in the real world.

(Holleis, 2007)

Prototyping

Practical method for testing with paper prototypes:

- Preparation:
- Create sheet of paper for each screen
- Cut out dialogs and menus as overlays
- Torn out pieces of post-it notes can be placed into text fields, so users can make entries without destroying the ‘screen’ for future users
- Testing:
- Member of design team acts as the system, changing paper sheets or items when user touches the paper screen, thinking out loud.
- Other design team members take notes and record comments
- No QA between designers and tester during the test, only afterwards
- Revise and repeat until satisfied or time runs out…

(Bentley & Barrett, 2012, p. 102-107)

Testing with paper prototypes

Lo-fi or Hi-fi

- Low-fidelity (lo-fi): drawings without much detail
- High-fidelity (hi-fi): detailed screen renderings

In the early design stage, lo-fi may be sufficient

Pros and cons

- + Cheap and fast to create
- + Discovering problems early on saves time for implementation
- - Playing ‘wizard’ also costs time
- Wizard may influence the results

Interactive Prototypes

Should contain the essential functionality for the experience

If a prototype is made to look very high fidelity, users will comment more on the “fit and finish” issues than on the overall interaction and understanding of flow

(Rettig, 1994, in Bentley & Barrett, 2012)

Comparing a paper prototype to a functional prototype => “that the interactive prototype elicited more comments of all types” Liu & Khooshabeh (2003)

Prototyping tools

- App Inventor MIT: http://appinventor.mit.edu/
- Axure: http://www.axure.com/
- AppSketcher: http://www.appsketcher.com/
- Balsamiq: http://www.balsamiq.com/products/mockups
- JustInMind: http://www.justinmind.com/
- PhoneGap: http://phonegap.com/
- 7 Scenes: http://7scenes.com/

etc.

Lab Usability Studies

- Quick and informal lab studies can be used to test the usability of prompts, icons and flow

- Useful for testing non-context dependent interactions

- Only small part of the usability of a design

(Bentley & Barrett, 2012, ch. 7)

Field Studies

For evaluating not only usability, but also:

- - usefulness in everyday life
- - how and where it is used

Ideally, let participants use the system or prototype:

- - everywhere and whenever they want to
- - for several weeks (by then, the novelty effect has worn off)
- - on their own device (so they don’t have to carry another)
- - if social aspects are important, recruit social groups

Pros and cons:

- + ecological validity
- - time-consuming, expensive

(Bentley & Barrett, 2012,)

Field Studies

Field-Based Data Collection

- Semi-structured interviews: to understand current practices

- Automatic data logging: to see how often, when, where and how the system is used, to discover problems, and to remind participants of specific instances

- Hearing from users regularly (e.g., daily, weekly, or after each interaction): to find out details that are not captured in logs, such as why they chose to interact with the system

- Experience sampling: to understand given contexts, or how users might react in particular situations

- Observing users in given situations: to see how the system is used

- Post-study interviews: to gather opinions and follow up on data collected

Often, a mix of methods provides best results

(Bentley & Barrett, 2012, ch. 8)

Questionnaires (Surveys)

Although not considered the best method of evaluation by HCI professionals, it is a very popular method

Consider (parts of) standardized HCI questionnaires, e.g.,

- System Usability Scale (SUS – Brooke, 1986)

http://www.measuringusability.com/sus.php

For more, see: http://hcibib.org/perlman/question.html

commonly used questionnaires on usability

Acronym	Instrument	Reference	Institution	Example
QUIS	Questionnaire for User Interface Satisfaction	Chin *et al*, 1988	Maryland	27 questions
PUEU	Perceived Usefulness and Ease of Use	Davis, 1989	IBM	12 questions
NAU	Nielsen's Attributes of Usability	Nielsen, 1993	Bellcore	5 attributes
NHE	Nielsen's Heuristic Evaluation	Nielsen, 1993	Bellcore	10 heuristics
CSUQ	Computer System Usability Questionnaire	Lewis, 1995	IBM	19 questions
ASQ	After Scenario Questionnaire	Lewis, 1995	IBM	3 questions
PHUE	Practical Heuristics for Usability Evaluation	Perlman, 1997	OSU	13 heuristics
PUTQ	Purdue Usability Testing Questionnaire	Lin *et al*, 1997	Purdue	100 questions
USE	USE Questionnaire	Lund, 2001	Sapient	30 questions

Nielsen's attributes of usability

			1	2	3	4	5	6	7		NA
1. Learnability		bad	○	○	○	○	○	○	○	good	○
2. Efficiency		bad	○	○	○	○	○	○	○	good	○
3. Memorability		bad	○	○	○	○	○	○	○	good	○
4. Errors (Accuracy)		bad	○	○	○	○	○	○	○	good	○
5. Subjective Satisfaction		bad	○	○	○	○	○	○	○	good	○
			1	2	3	4	5	6	7		NA

Nielsen 1993

Perceived Usefulness and Ease of Use

PERCEIVED USEFULNESS		1	2	3	4	5	6	7		NA
1. Using the system in my job would enable me to accomplish tasks more quickly	unlikely	○	○	○	○	○	○	○	likely	○
2. Using the system would improve my job performance	unlikely	○	○	○	○	○	○	○	likely	○
3. Using the system in my job would increase my productity	unlikely	○	○	○	○	○	○	○	likely	○
4. Using the system would enhance my effectiveness on the job	unlikely	○	○	○	○	○	○	○	likely	○
5. Using the system would make it easier to do my job	unlikely	○	○	○	○	○	○	○	likely	○
6. I would find the system useful in my job	unlikely	○	○	○	○	○	○	○	likely	○
PERCEIVED EASE OF USE		**1**	**2**	**3**	**4**	**5**	**6**	**7**		**NA**
7. Learning to operate the system would be easy for me	unlikely	○	○	○	○	○	○	○	likely	○
8. I would find it easy to get the system to do what I want it to do	unlikely	○	○	○	○	○	○	○	likely	○
9. My interaction with the system would be clear and understandable	unlikely	○	○	○	○	○	○	○	likely	○
10. I would find the system to be flexible to interact with	unlikely	○	○	○	○	○	○	○	likely	○
11. It would be easy for me to become skillful at using the system	unlikely	○	○	○	○	○	○	○	likely	○
12. I would find the system easy to use	unlikely	○	○	○	○	○	○	○	likely	○
		1	**2**	**3**	**4**	**5**	**6**	**7**		**NA**

List the most **negative** aspect(s):

1.
2.
3.

List the most **positive** aspect(s):

1.
2.
3.

Davis, F. D. (1989)

Questionnaire for User Interface Satisfaction

OVERALL REACTION TO THE SOFTWARE		0	1	2	3	4	5	6	7	8	9		NA
1.	terrible	○	○	○	○	○	○	○	○	○	○	wonderful	○
2.	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
3.	frustrating	○	○	○	○	○	○	○	○	○	○	satisfying	○
4.	inadequate power	○	○	○	○	○	○	○	○	○	○	adequate power	○
5.	dull	○	○	○	○	○	○	○	○	○	○	stimulating	○
6.	rigid	○	○	○	○	○	○	○	○	○	○	flexible	○
SCREEN		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**		**NA**
7. Reading characters on the screen	hard	○	○	○	○	○	○	○	○	○	○	easy	○
8. Highlighting simplifies task	not at all	○	○	○	○	○	○	○	○	○	○	very much	○
9. Organization of information	confusing	○	○	○	○	○	○	○	○	○	○	very clear	○
10. Sequence of screens	confusing	○	○	○	○	○	○	○	○	○	○	very clear	○
TERMINOLOGY AND SYSTEM INFORMATION		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**		**NA**
11. Use of terms throughout system	inconsistent	○	○	○	○	○	○	○	○	○	○	consistent	○
12. Terminology related to task	never	○	○	○	○	○	○	○	○	○	○	always	○
13. Position of messages on screen	inconsistent	○	○	○	○	○	○	○	○	○	○	consistent	○
14. Prompts for input	confusing	○	○	○	○	○	○	○	○	○	○	clear	○
15. Computer informs about its progress	never	○	○	○	○	○	○	○	○	○	○	always	○
16. Error messages	unhelpful	○	○	○	○	○	○	○	○	○	○	helpful	○
LEARNING		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**		**NA**
17. Learning to operate the system	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
18. Exploring new features by trial and error	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
19. Remembering names and use of commands	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
20. Performing tasks is straightforward	never	○	○	○	○	○	○	○	○	○	○	always	○
21. Help messages on the screen	unhelpful	○	○	○	○	○	○	○	○	○	○	helpful	○
22. Supplemental reference materials	confusing	○	○	○	○	○	○	○	○	○	○	clear	○
SYSTEM CAPABILITIES		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**		**NA**
23. System speed	too slow	○	○	○	○	○	○	○	○	○	○	fast enough	○
24. System reliability	unreliable	○	○	○	○	○	○	○	○	○	○	reliable	○
25. System tends to be	noisy	○	○	○	○	○	○	○	○	○	○	quiet	○
26. Correcting your mistakes	difficult	○	○	○	○	○	○	○	○	○	○	easy	○
27. Designed for all levels of users	never	○	○	○	○	○	○	○	○	○	○	always	○
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**		**NA**

Chin et al 1988

Questionnaires

Designing a questionnaire yourself

- Allows focus on dimensions that are important to your research goals
- Issues: reliability and validity

- Open questions: more informative, takes more time to fill in & analyze
- Closed questions: to gather statistics (if # participants is sufficient)
- E.g., Likert-scale (5, 7, or 9-point) questions:

Please indicate the likelihood of the following statement: 'I would find the keyboard of the mobile phone easy to use'
(circle one of the numbers)

Unlikely	1	2	3	4	5	6	7	Likely
	extremely	quite	slight	neither	slightly	quite	extremely	

- Use multiple questions per dimension, if possible
- Avoid ambiguous, complex, or leading questions

(Brinkman, W.-P., 2009)

Research Ethics

- Explaining the goals of your research
- Comply with Ethical Guidelines regarding safety, privacy, etc.
- Asking for consent
- Anonymization of data
- Example consent form (from your supervisor)
- EU FP7 ethical guidelines can be found at http://cordis.europa.eu/fp7/ethics_en.html

Making the decision

After evaluation, should you continue with the idea or drop it?

- - How often was the system used?
- - What types of use was it appropriated for?
- - How did users react in general?
- - Did the system change their lives?
- - Did they connect in new ways with people and places?
- - Were they happy to use the application, sad to see the trial end?

EXPERIENCE DESIGN TOOLS

NEED CARDS
Each card describes a psychological need, which can be a source of positive experiences – also when interacting with technical products. Need cards provide both orientation and inspiration for the design of interactive products. In addition to a brief description you find typical feelings and quotes related to a particular need and a list of feelings which a product can trigger if it addresses this need.

INTERACTION VOCABULARY CARDS
The Interaction Vocabulary provides a number of dimensions to describe basic properties of interaction. It is mainly useful for interaction designers as it shows starting points for designing interaction and facilitates communication about intended properties. The Interaction Vocabulary suggests eleven dimensions, each in a pair of opposing attributes. These are descriptive, non-judgmental, non-technology bound attributes of interaction.

Each card represents one of the Interaction Vocabulary‘s dimensions, with the opposite poles on each side of the card. In the design process, turning a card on one or the other side represents a preference for the respective interaction attribute. To support the designer in aligning her or his choice with the intended experience, the card also informs about potentially related experiences. Turning the card on one side or the other, thus, may also be motivated by the desired experiences.

https://www.experienceandinteraction.com/tools

Reading:

- Usability Evaluation, Gilbert Cockton https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/usability-evaluation

- Watch (as much as you want) Liz Jackson, IXDA Seattle 2019, Empathy reifies disability stigmas. https://interaction19.ixda.org/program/keynote--liz-jackson/

- Watch (as much as you want) Aral Balkan: Superheroes & Villains in Design from Thinking Digital https://vimeo.com/70030549

- look at http://uxchecklist.github.io

references

- Bentley, F., and Barret, E. (2012). Building Mobile Experiences. MIT Press, Cambridge,
- Brinkman, W.-P. (2009). Design of a Questionnaire Instrument, Handbook of Mobile Technology Research Methods, ISBN 978-1-60692-767-0, pp. 31-57, Nova Publisher.
- Cherubini, M., R. de Oliveira, A. Hiltunen, and N. Oliver (2011). Barriers and bridges in the adoption of today's mobile phone contextual services. In Proceedings of MobileHCI '11. ACM, New York, 167-176. http://doi.acm.org/10.1145/2037373.2037400
- Chin, J.P., Diehl, V.A., Norman, K.L. (1988) Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface. ACM CHI'88 Proceedings,
- Gilbert Cockton, Usability Evaluation, https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/usability-evaluation
- Cooper, Alan; Reimann, Robert; Cronin, Dave (2007). About Face 3: The Essentials of Interaction Design. Indianapolis, Indiana: Wiley.
- Davis, F. D. (1989) Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13:3, 319-340
- Holleis, P., Otto, F., Hussmann, H., and Schmidt, A. (2007). Keystroke-level model for advanced mobile phone interaction. In Proceedings of CHI 2007. ACM, New York, pp. 1505-1514.http://doi.acm.org/10.1145/1240624.1240851
- Liu, L., & Khooshabeh, P. (2003). Paper or interactive?: a study of prototyping techniques for ubiquitous computing environments. In CHI'03 extended abstracts on Human factors in computing systems (pp. 1030-1031). ACM.
- A.H. Maslow, A Theory of Human Motivation, Psychological Review 50(4) (1943):370-96.
- Nielsen, J. (1993) Usability Engineering. Academic Press. Chapter 2.2, p. 26