PART II: Development Process

Evaluating Interface Design

SEEM3510 Human-Computer Interaction

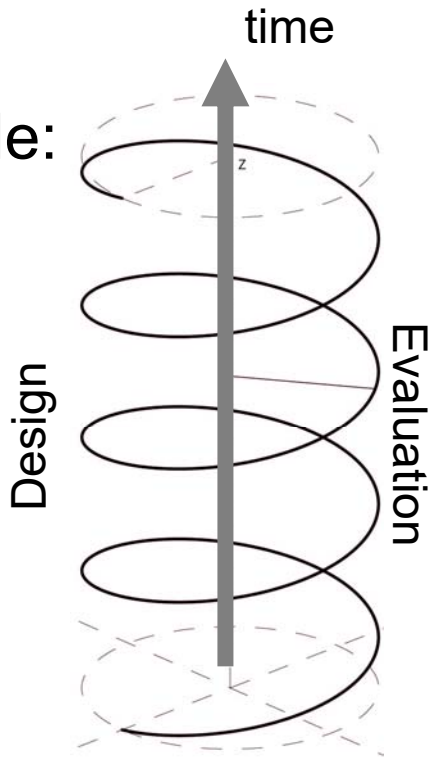
By Prof. Helen MENG & Prof. Philip FU

Evaluating Interface Design

- Reading - Textbook
  - Chapter 4
- Goal
  - Introduce several techniques commonly used in industry to “evaluate” UIs
- Overview
  - Introduction
  - Expert Review
  - Usability Testing and Labs
  - Survey Instruments
  - Acceptance Tests
  - Evaluation during Active Use (or continuous evaluation)
  - How AI helps UI/UX design and evaluation

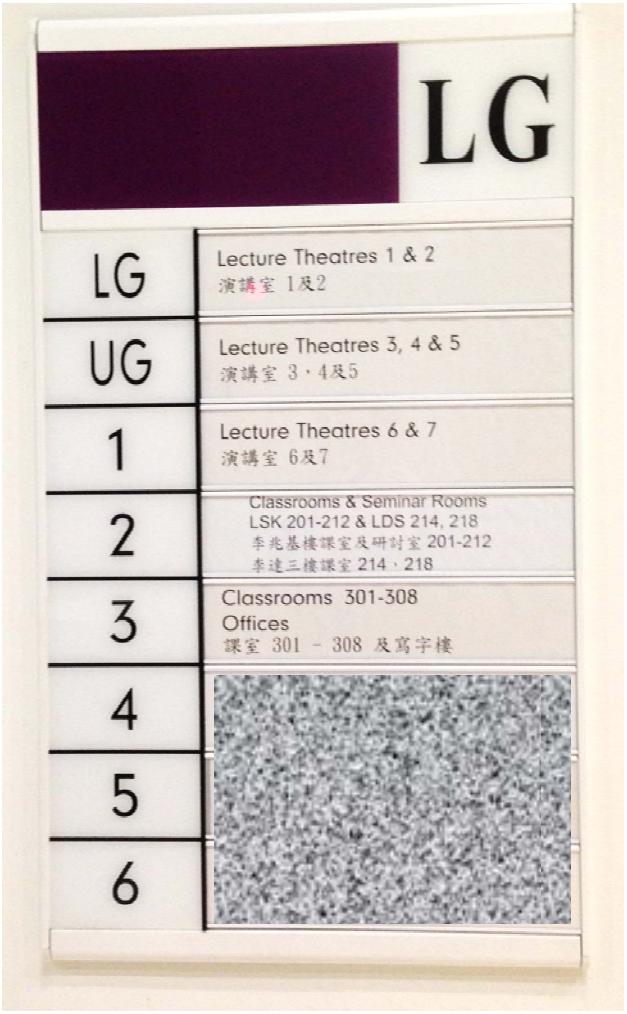
Introduction

- Designers can become so entranced with their creations that they may fail to evaluate their own designs adequately.
- Experienced designers have attained the wisdom and humility to know that extensive testing is a necessity.
  - **Feedback** is the "breakfast of champions."
  - **Testing** is the "dinner of the gods."
- The determinants of the evaluation plan include:
  - stage of design (early, middle, late)
  - novelty of project (well defined vs. exploratory)
  - number of expected users
  - criticality of the interface (life-critical medical system vs. museum exhibit support)
  - costs of product and finances allocated for testing
  - time available
  - experience of the design and evaluation team



Before we start… some real examples

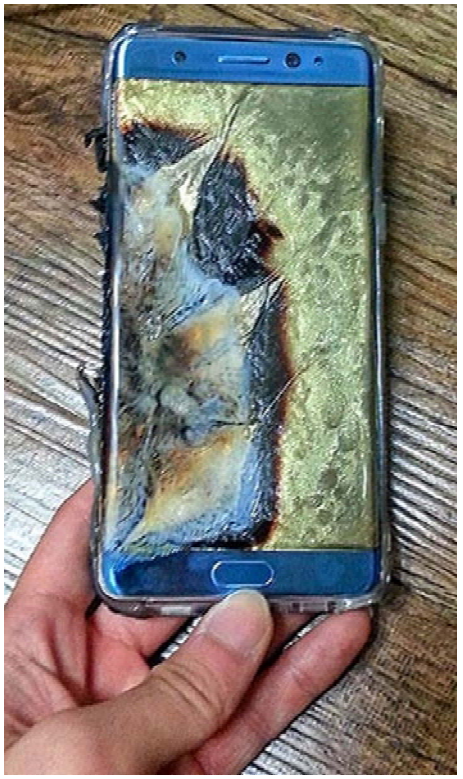
[note: this is on the front page of a website]



https://www.interaction-design.org/literature/article/bad-design-vs-good-design-5-examples-we-can-learn-frombad-design-vs-good-design-5-examples-we-can-learn-from-130706

Introduction (cont.)

- The range of evaluation plans might be from a few days to several years
- The range of costs might be from 20% of a project down to 5%.
  - Budget for user evaluation!
- Is it worth it? Remember this . . . good design pays off.
  - 15+ years ago, doing usability is a plus
  - Now, failing to test is now very risky



http://www.computerworld.com/s/article/9178503/Hardware_expert_explains_iPhone_4_antenna_problem (2010)
https://www.cnet.com/news/samsung-galaxy-note-7-return-exchange-faq/ (2016: Note 7 recall costs at least

Note: https://en.wikipedia.org/wiki/Product_recall

Introduction (cont.)

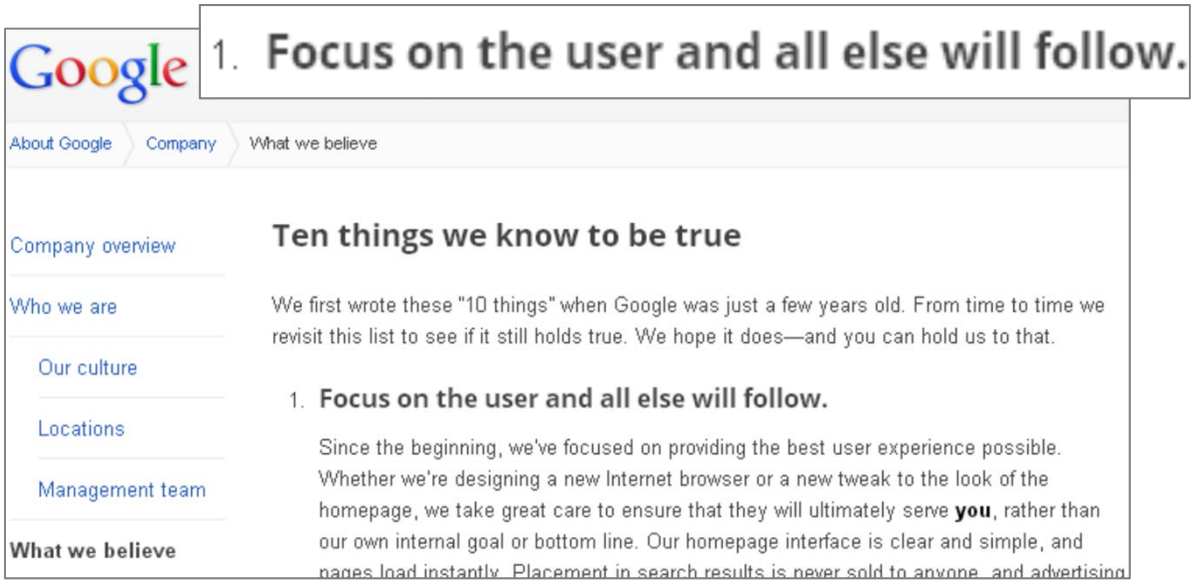
- Troubling aspect of testing: *uncertainty* remains even after exhaustive testing by multiple methods.
  - Decision must be made about completing prototype testing and delivering the product, even though problems may continue to be found
  - Most testing methods will account for normal usage, but performance in unpredictable situations with high levels of input such as nuclear reactor control or presidential voting system, is extremely hard to test.
  - Perfection may not be possible even with huge human effort, so must continue assessing and repairing problems during lifecycle of interface

Introduction (cont.)

Lessons learnt

- Google: First principle in ten
- Tencent

  用户体验的10/100/1000法则
  (UX 10/100/1000 rule)



https://www.google.com/about/philosophy.html

http://www.gamelook.com.cn/2014/02/143996

"腾讯"10/100/1000法则"：产
客，收集反馈1000个用户体
上没有什么捷径可以走，不
位于低端用户的产品，想都
满足了用户需求；自认为定

By Google Translate: Tencent "10/100/1000 rule": Product managers must do 10 user surveys each month, pay attention to 100 user blogs, and collect feedback on 1000 user experiences. This method looks a bit stupid, but it works. There are no shortcuts to researching user needs. Don't think that you can guess user habits for granted. For example, some products that are considered to be targeted at low-end users do not want to abuse the cartoon avatar and some fancy page decorations, thinking that this is to meet the needs of users; they think that they are targeting high-end users, and they like to be self-sufficient.

Topics:

- Introduction
- **Expert Review**
  - **six methods**
  - **Nielsen's 10 usability heuristics**
- Usability Testing and Labs
- Survey Instruments
- Acceptance Tests
- Evaluation during Active Use
- How AI helps UI/UX design and evaluation

Expert Reviews

- Expert reviews entail one-half day to one week effort, although a lengthy training period may sometimes be required to explain the task domain or operational procedures

- There are a variety of expert review methods to choose from:

1. **Heuristic evaluation**
   - Expert reviewer's personal critic
2. **Guidelines review**
   - Make sure UI adheres to established guidelines
3. **Consistency inspection**
   - Check for consistency through-out interface
4. **Cognitive walkthrough**
   - Simulate performing certain tasks
5. **Metaphors of human thinking**
   - Focuses on how users' mental activity when they use the interface
6. **Formal usability inspection**
   - UI designers defend their choices against a "hostile" expert

1. Heuristic evaluation

- "**Heuristic**" evaluation
  - Not a true formal approach – thus the term "heuristic"
  - But effective
  - This is often called "Discount Evaluation" because it is typically cheaper than a full user study

- Expert reviewer gives personal criticism
  - Could be expert designer – like Alan Cooper (see Assignment 3 spec.)
  - Could be expert "user" (i.e., doctor for medical product)
  - Often have more than one reviewer

- The expert reviewer spends time evaluating your interface
  - Gives feedback, overall impression, concerns, maybe according to 'eight golden rules'
  - Remember this is an expert, so you are paying to hear what they say

Nielsen's 10 usability heuristics

1. Visibility of system status

The system should always keep users informed through appropriate feedback within reasonable time.

2. Match between system and the real world

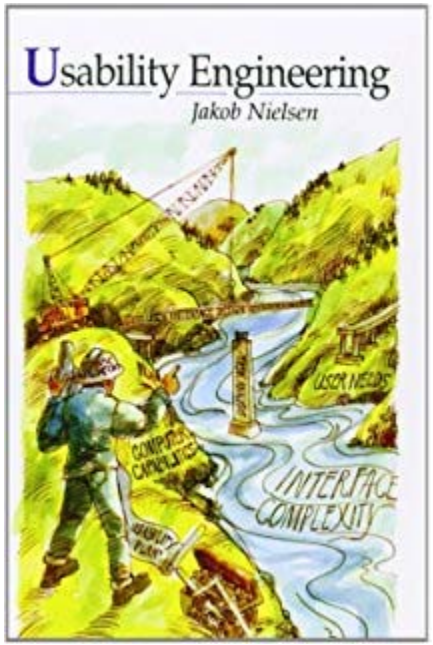
Speak the users' language and follow real-world conventions, making information appear in a natural and logical order.

3. User control and freedom

Users may make mistake, so they will need undo and redo function without going through an extended dialogue.

4. Consistency and standards

Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.



https://www.nngroup.com/articles/ten-usability-heuristics/

Nielsen's 10 usability heuristics

5. Error prevention

Prevents a problem from occurring. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

6. Recognition rather than recall

Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another.

7. Flexibility and efficiency of use

Accelerators (e.g., short-cut keys)-- unseen by the novice user -- may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

https://www.nngroup.com/articles/ten-usability-heuristics/

Nielsen's 10 usability heuristics

8. Aesthetic and minimalist design

Dialogues should not contain irrelevant or rarely-needed information. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

9. Help users recognize, diagnose, and recover from errors

Error messages should be expressed in plain language (no code), precisely indicate the problem, and constructively suggest a solution.

10. Help and documentation

It is better if the system can be used without documentation, but if necessary, provide help and documentation, where information should be easy to search, focused on user's task, list concrete steps to be carried out, and not be too large.

https://www.nngroup.com/articles/ten-usability-heuristics/

2. Guideline Review

- Assume you have a guideline
- Inspect the UI to make sure it adheres to the guidelines
- May require a "bird's eye" view of the UI
  - This is where all possible windows of the UI are printed out, laid out on the floor or pinned to walls.

The guideline may be of a thousand items, so it may take expert reviewers sometime to master the guideline, and days or weeks to review a large interface.

3. Consistency Checking

- Similar to “guideline review”, but check for consistency in the UI
- Consistency involves:
  - Terminology, icons, color scheme, layout, input/output format, etc.
  - Within the interface as well as in the training materials (e.g., tutorial) and online help
- Some research projects: software tools to automate the process

4. Cognitive Walkthrough

- Ask expert reviewer to perform **certain tasks**, simulating users walking through the interface
- Either let them do it privately OR watch how they do it
- See if they behave the way you thought they would!
  - Get verbal feedback
  - User talks out loud the whole time explaining his/her thought process (see a video example later in this module)
- May see usage patterns you didn't expect!

5. Metaphors of human thinking

- Experts conduct an inspection on how users **think** (user's mental activity) when they interact with the user interface

- Consider various aspects of human thinking
  - *Habit Formation*
    - Habits leads to more efficient actions and less conscious effort
    - Does the UI support existing user habit?
  - *Stream of thought (the continuity of our thinking)*:
    - Does the app. help users to "interrupted tasks"?
  - *Awareness and associations* [related to week 6 "direct manipulation"]
    - Do users associate UI elements with actions and objects they represent?

  etc.

  Detail: [Hornbak and Frokjaer 2008]
  http://portal.acm.org/citation.cfm?id=1314688

6. Formal usability inspection

- Courtroom like setting
- Expert reviewers ask questions to the designers and the designers have to justify their design decisions in an adversarial manner.
- A lengthy process, but can be educational to inexperience designers.

Note: Expert Reviews

- When? Can be scheduled at several points in the development process when experts are available and when the design team is ready for feedback.

- How to choosing Experts?
  - Different experts tend to find different problems

    So *3-5* expert reviewers can be highly productive: complementary usability testing
  - Familiar with the project and long-term relationship with your organization

- Problems: Dangers if we rely only on expert UI consultants:
  - The experts may not have an adequate understanding of the *task domain* or *user communities*
  - Even experienced expert reviewers have great difficulty knowing how typical users, especially fi users will really behave.
  - Coming in many flavors, experts may give *conflicting advice*

Topics:

- Introduction
- Expert Review
- **Usability Testing and Labs**
  - **Various Usability Tests**
  - **What is Usability Laboratory**
  - **The Participants**
  - **Techniques for Recording**
- Survey Instruments
- Acceptance Tests
- Evaluation during Active Use
- How AI helps UI/UX design and evaluation

(1) Various Usability Testing

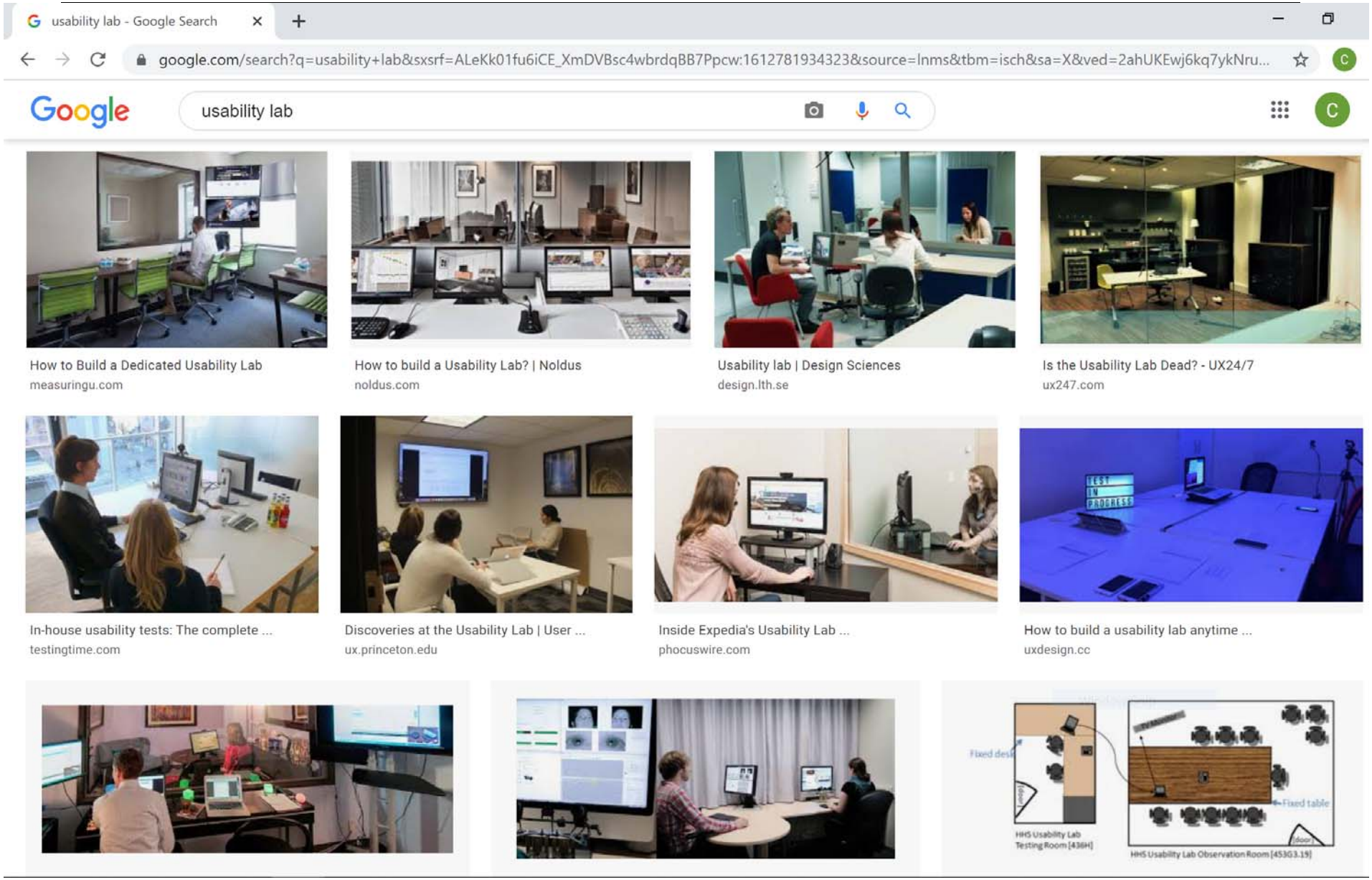
Many different kinds of usability tests:

- Paper mockups
  - Early usability study (on low-fi prototype); inexpensive and rapid
  - Flipping the (mockup of) screen displays to get reactions to wording, layout, etc.
- Discount usability testing
  - Quick and dirty approach with only 3 to 6 test participants
- Competitive usability testing
  - Comparing new interface to previous versions or similar products
  - A/B Test:
- Universal usability testing
  - Diverse users, hardware/software platform (Web browsers, OS), networks, etc.
- Field test and portable labs
  - Test in realistic environments for a fixed trial period (usually use logging software).
  - Supply users test versions of new software or consumer products
- Remote usability testing
  - Tests online; more participants but less control over user behavior and observation of reaction
- Can-you-break-this tests?
  -

(2) Usability Lab - Background

- The emergence of usability testing and laboratories since the early 1980s.
- Usability testing not only sped up many projects but has also produced dramatic cost savings.
  - Traditional managers and developers resisted at first, saying the usability testing may take time and resource away from development
  - They changed their mind when experience grew and successful projects gave credit to the testing process.
- The movement towards usability testing stimulated the construction of **usability laboratories**, e.g., Microsoft has more than 25 usability lab. (1989*)
- A tour to Google's **usability lab:**
  - See video "*Touring Google's Usability Lab*" on Blackboard

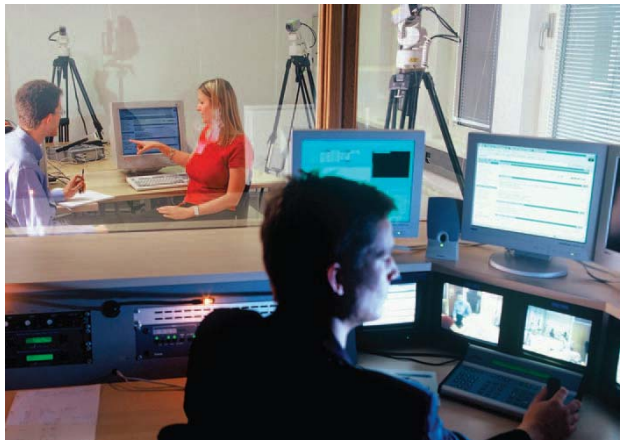
What is Usability Lab

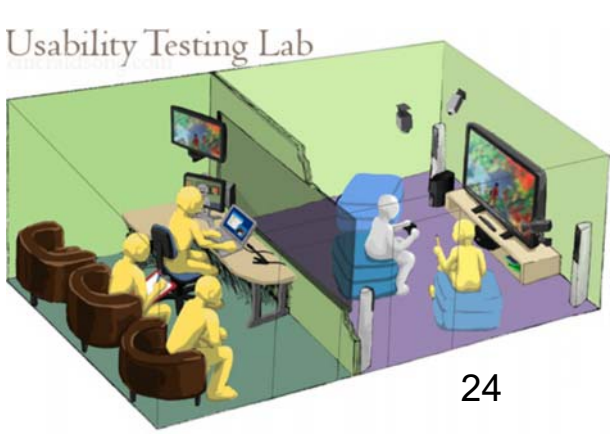


What is Usability Laboratory?

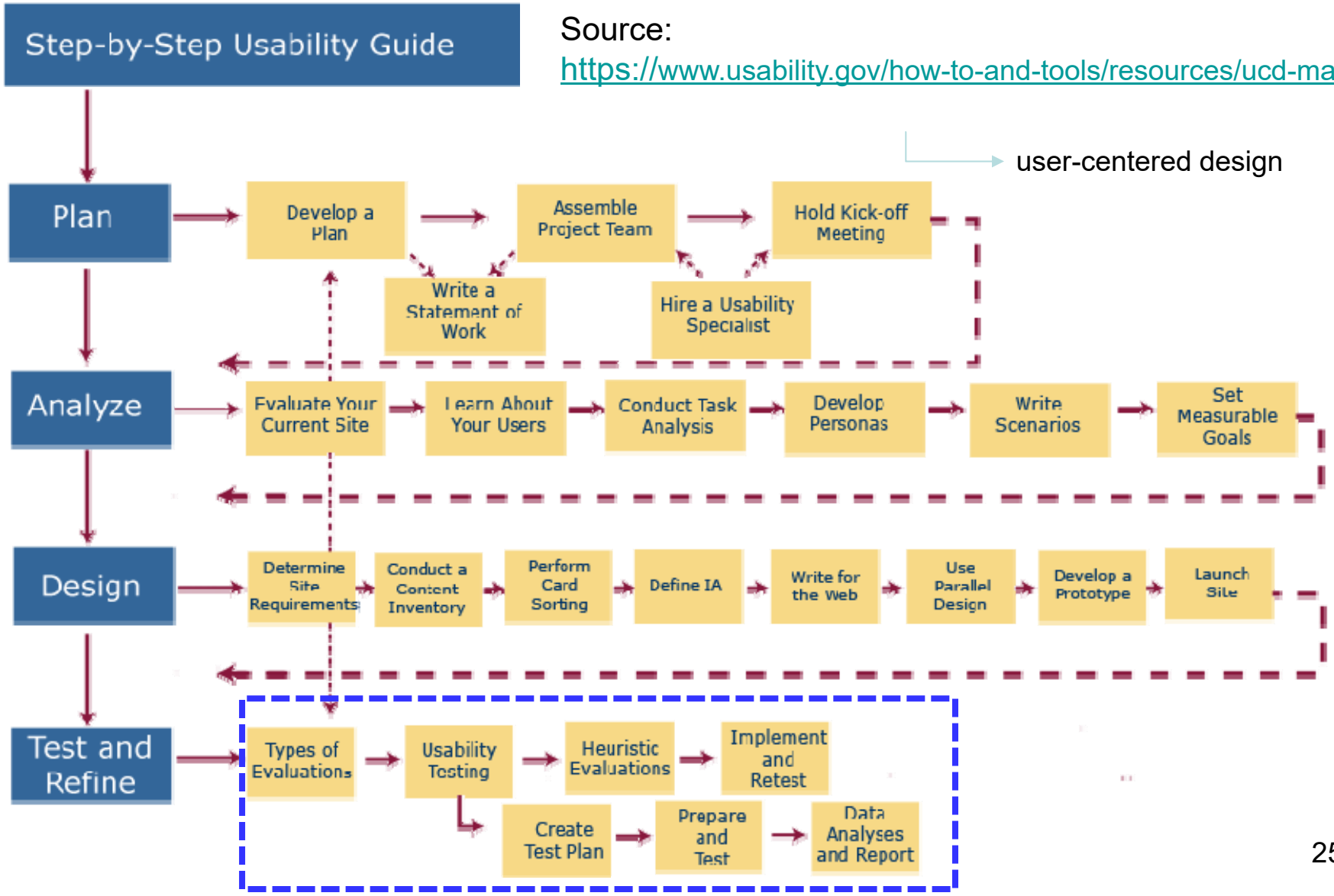
A typical modest usability lab:

- Space:
  - two 10 by 10 foot areas (1 ft = 0.3048 meter)
  - one for the participants to do their work; and
  - another, separated by a half-silvered mirror, for the testers and observers
- Staff:
  - One or more people with expertise in testing and UI design
  - May serve 10 to 15 projects per year in the organization
  - Meet the UI architect/manager at the start to design the test plan with schedule and budget; participate in early design and provide information on software tools or literature references





What is Usability Laboratory



(3) The Participants

- Choosing Participants
  - To represent the intended user communities
  - With attention to background in computing, experience with the task, motivation, education, and ability with the natural language used in the interface.

- Before/during the test, the participants should
  - Always be treated with *RESPECT*
  - Be informed that it is *NOT THEY* who are being tested, but it is the software and user interface that are under study
  - Should be told about *WHAT they will be DOING* and *HOW LONG* they will be expected to stay (at the beginning)
  - Always be *VOLUNTARY*, and **informed consent** should be obtained.

The Participants

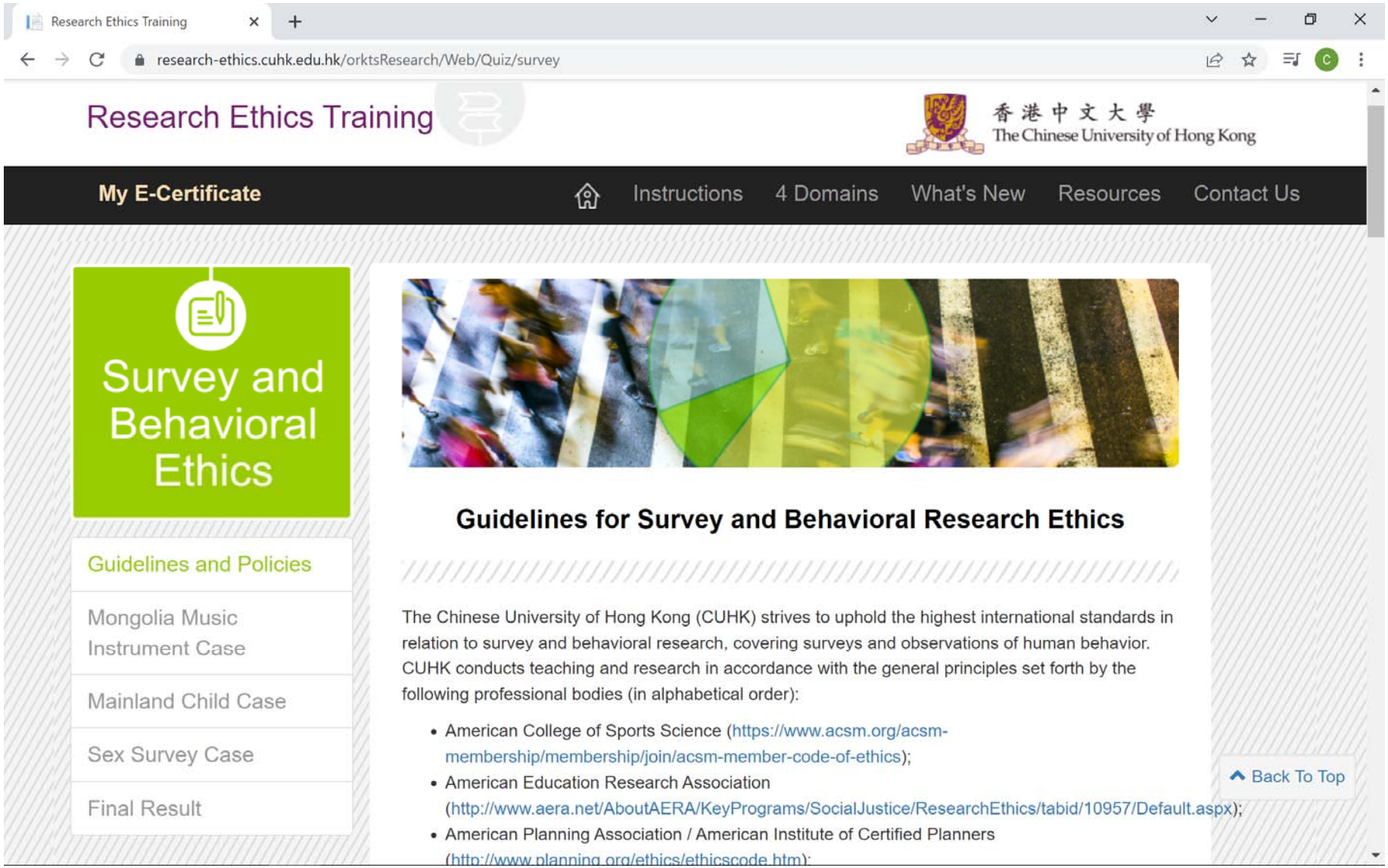
E.g., Informed consent (*consent form*):

- Professional practice is to ask all subjects to read and sign a statement like this one:
  - I have freely volunteered to participate in this experiment.
  - I have been informed in advance what my task(s) will be and what procedures will be followed.
  - I have been given the opportunity to ask questions, and have had my questions answered to my satisfaction.
  - I am aware that I have the right to withdraw consent and to discontinue participation at any time, without prejudice to my future treatment.
  - My signature below may be taken as affirmation of all the above statements; it was given prior to my participation in this study.

Institutional Review Board (**IRB**)
https://en.wikipedia.org/wiki/Institutional_review_board

Research Ethics

https://www.research-ethics.cuhk.edu.hk/orktsResearch/ [need login]



(4) Techniques: Recording

- Observe the users: What to be recorded:
  - User activities such as typing, using mouse, reading screens, reading manuals, etc.
  - Timestamping on them to know also the time taken

- Example:
  - If designers can see participants repeatedly picking the wrong menu item, they often can realize that the label or placement needs to be changed

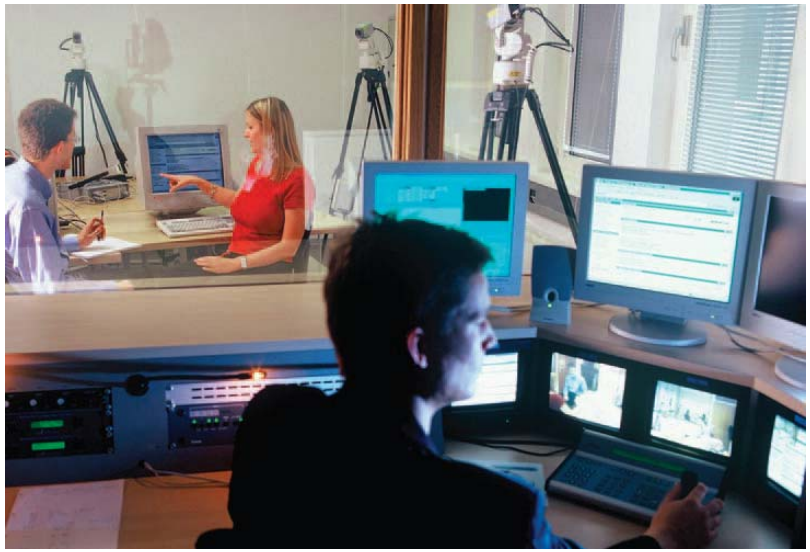
Besides manually recording, how to record user actions?

- Basic tools: 1) L
- Others: should get agreement from participants first
  2) V
  3) T
  4) E
  5) F

Techniques: Recording

(2) Videotaping

- Often valuable for later review and for showing designers or managers the problems that users encounter.
- Reviewing videotaping is tedious, so careful logging and annotation is vital on finding critical incidents
  - Tools for automatic time stamping activities (typing, mousing, reading manuals, etc.)
- Participants may be anxious about the video camera at the start of the test… few minutes later, focusing on the tasks (note: get their consent)



Techniques: Recording

(3) Think-Aloud (or talk-aloud)

Participants carry out tasks while saying what they are thinking, and tester records "thoughts"

- Think-aloud technique yield interesting clues for the tester
  - *For example, "This web page text is too small… so I am looking for something on the menus to make the text bigger… maybe it's on the top in the icons… I can't find it… so I'll just carry on".*
- Tester does not take over or give instructions, but prompt and listen for clues how they are dealing with the interface.
  - "tell me what you are thinking"
  - "… so this is?"

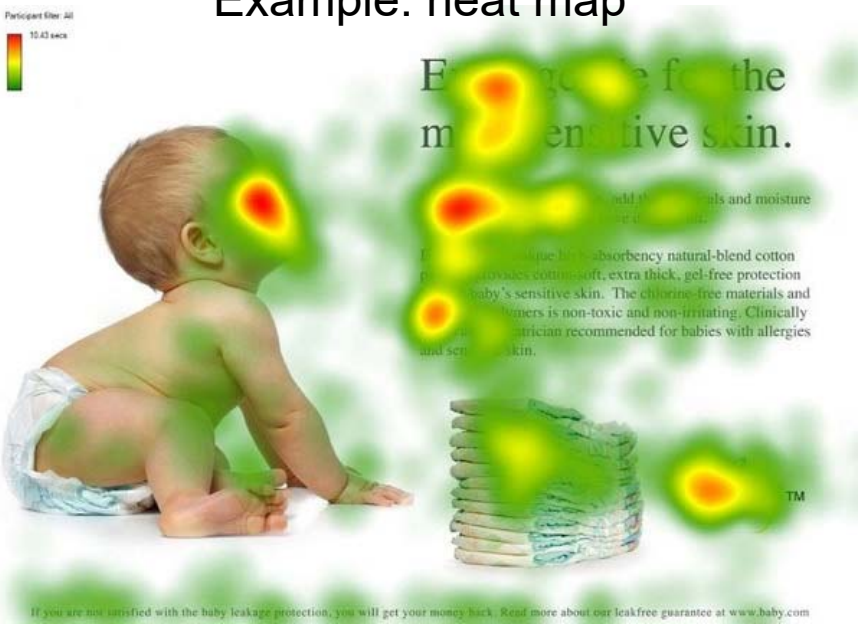
https://www.youtube.com/watch?v=nJ2udLjdsx4

Techniques: Recording

(4) Eye tracking

- Show ***where*** the participants *gazed* at the screen and for ***how long***
- Results can be shown in color-coded heat maps
  - Areas of screen being viewed and which areas are being ignored.
- Cannot use with Think Aloud
  -

Example: heat map



Nowadays

https://www.youtube.com/watch?v=lo_a2cfBUGc

Techniques: Recording

(5) Face & emotion recognition

- Using recent AI technologies
- A researcher at the Expedia Group lab used emotion tracking and eye tracking software to acquire participant feelings when they arrange trips on travel websites

https://www.nytimes.com/2019/03/04/business/ai-technology-travel-planning.html

Topics:

- Introduction
- Expert Review
- Usability Testing and Labs
- **Survey Instruments**
- Acceptance Tests
- Evaluation during Active Use
- How AI helps UI/UX design and evaluation

Survey Instruments

- Written user surveys are a familiar, inexpensive and generally acceptable companion for usability tests and expert reviews.

- Online surveys
  - Avoid printing cost and extra effort for distribution and collection of survey paper
  - Many people prefer to answer a brief survey displayed on a screen, instead of filling in and returning a printed form

  However, need to detect and avoid noise/errors in collected data

- Keys to successful surveys
  - Clear goals in advance
  - Development of focused items that help attain the goals.

- Note:
  - Important to *pre-test* or *pilot-test* prior to actual use
  - Use age-appropriate language, e.g., to children or domain users

Survey Instruments

- Users could be asked for their subjective impressions about specific aspects of the interface such as the representation of:
  - task domain objects and actions
  - syntax of inputs and design of displays
- Need to ascertain characteristics about the users:
  - Background (age, gender, origins, education, income)
  - Experience with computers (specific applications or software packages, length of time, depth of knowledge)
  - Job responsibilities (decision-making influence, managerial roles, etc.)
  - Personality style (introvert vs. extrovert, risk taking vs. risk aversive, early vs. late adopter, systematic vs. opportunistic)
  - Reasons for not using an interface (inadequate services, too complex, too slow)
  - Familiarity with features (printing, macros, shortcuts, tutorials)
  - Feeling after using the interface (confused vs. clear, frustrated vs. in-control, bored vs. excited).

Survey Instruments

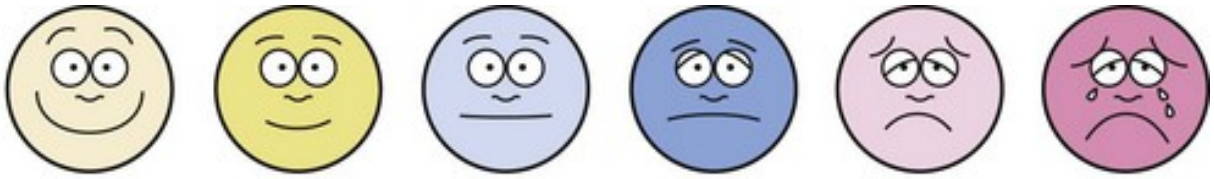
- Commonly used ***Likert* scale**, e.g.,

  Strongly Agree Agree Neutral Disagree Strongly Disagree

- Bipolar semantically anchored items 1-to-7 scales, e.g.,

Vague	1	2	3	4	5	6	7	Specific
Pleasing	1	2	3	4	5	6	7	Irritating
Simple	1	2	3	4	5	6	7	Complicated

- Another approach is to ask users to evaluate
  - Readability of characters
  - Meaningfulness of command names
  - Helpfulness of error messages



Survey Example

- Learn by examples:

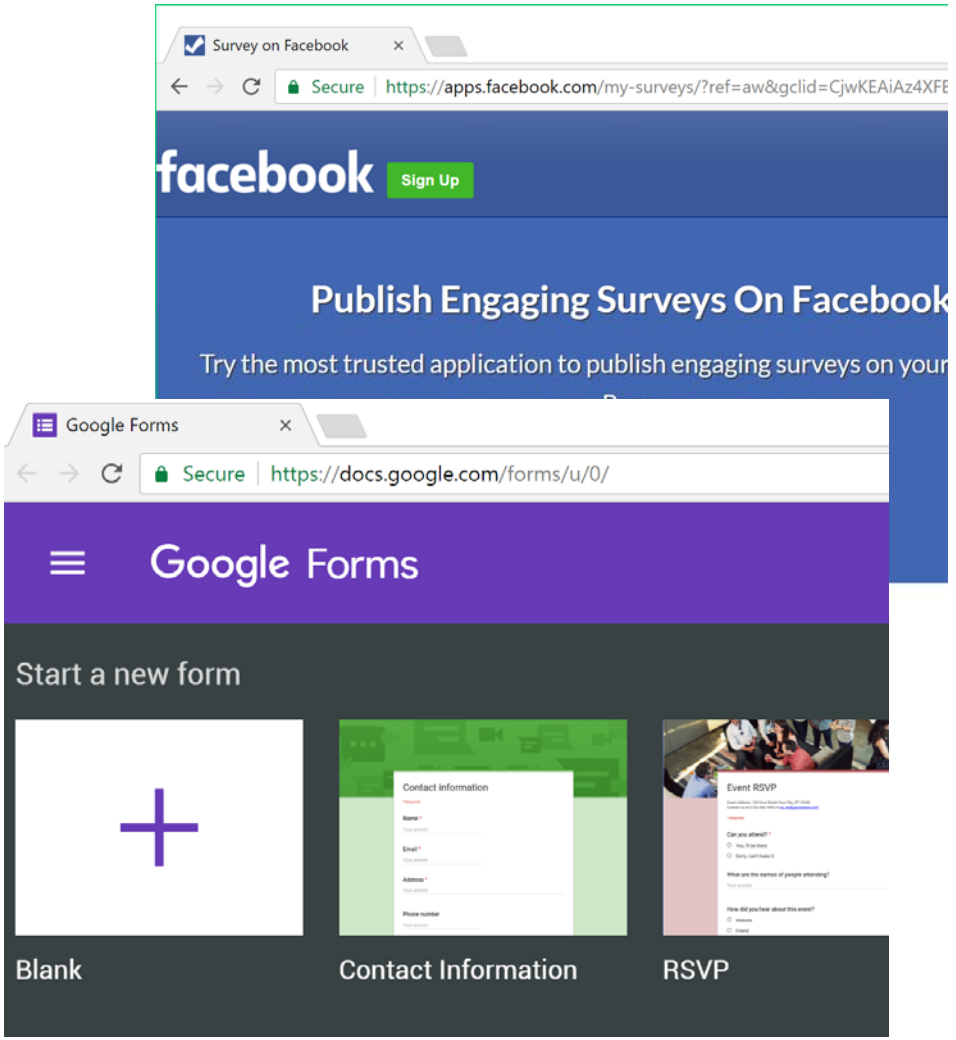
10. Please rank your use and usability of each page: (1=Minimal/Difficult, 5=Extensive/Easy)

	Use	Usability
Home page	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
About the Coalition	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
Member governors	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
Representatives	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
Publications	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
News Releases	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
Letters/Speeches	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
Ethanol Information	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
State/Federal Legislation	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
Comments	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
Members Only	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
Links	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5
Search	○1 ○2 ○3 ○4 ○5	○1 ○2 ○3 ○4 ○5

11. Please rate the site on these features:

Layout and design	○ Poor	○ Fair	○ Good	○ Excellent	○ Don't Know
Ease of finding information	○ Poor	○ Fair	○ Good	○ Excellent	○ Don't Know
Usefulness of information	○ Poor	○ Fair	○ Good	○ Excellent	○ Don't Know
Login feature	○ Poor	○ Fair	○ Good	○ Excellent	○ Don't Know
Getting assistance with the website	○ Poor	○ Fair	○ Good	○ Excellent	○ Don't Know
Timeliness of information	○ Poor	○ Fair	○ Good	○ Excellent	○ Don't Know
Navigation	○ Poor	○ Fair	○ Good	○ Excellent	○ Don't Know

- Online survey

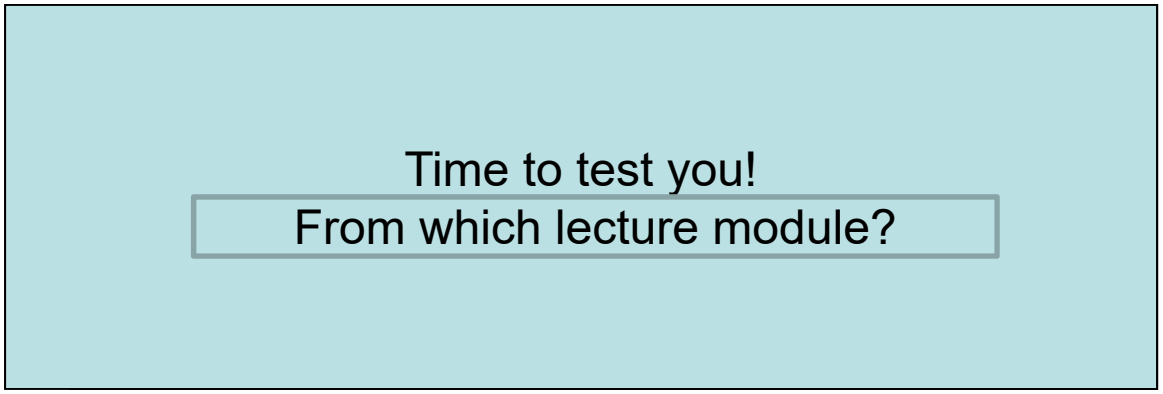


Topics:

- Introduction
- Expert Review
- Usability Testing and Labs
- Survey Instruments
- **Acceptance Tests**
- Evaluation during Active Use
- How AI helps UI/UX design and evaluation

Acceptance Tests

- For large implementation projects, the customer or manager usually sets objective and measurable goals for hardware and software performance.
- If the completed product fails to meet the criteria, the system must be reworked until success is demonstrated.
- Rather than vague and misleading criterion of "user friendly," measurable criteria (human factors in usability) for the test can be established for the following:
  -
  -
  -
  -
  -



Acceptance Tests

- Example:

  "*Ten participants will be recalled after one week, and asked to carry out a new set of benchmark tasks. In 20 minutes, at least 8 of the participants should be able to complete the tasks correctly*."

- In a large interface, may be 8 or 10 such tests to carry out on different components of the interface and with different user communities.

- Different from usability tests, *outside* testing organizations are often appropriate to ensure neutrality

- Goal: not to detect flaws, but rather to verify requirements

- Once acceptance testing has been successful, there may be a period of field testing before national or international distribution.

More information? http://en.wikipedia.org/wiki/Acceptance_testing

Topics:

- Introduction
- Expert Review
- Usability Testing and Labs
- Survey Instruments
- Acceptance Tests
- **Evaluation during Active Use**
  - **a.k.a. continuous evaluation**
  - **six strategies**
- How AI helps UI/UX design and evaluation

Testing is important!!!

- Previews and feedback from prototypes is crucial
  - Even low fidelity mockups
  - If possible, get feedback from experts, average users, …
- Large websites carry out "pilot" launches to get feedback and find bugs
- Google "Gmail" is a great example
  - Request feedback and requests/suggestions for improvements
  - Long years of beta after launch today

    http://en.wikipedia.org/wiki/Gmail (from 2004 to 2009)

    Now: [ ] monthly active users by Oct. 2019

    (from https://en.wikipedia.org/wiki/Gmail)

Evaluation During Active Use

- Successful active use requires constant attention from dedicated managers, user-services personnel, and maintenance staff.

- Perfection is not attainable, but percentage improvements are possible

Six strategies:

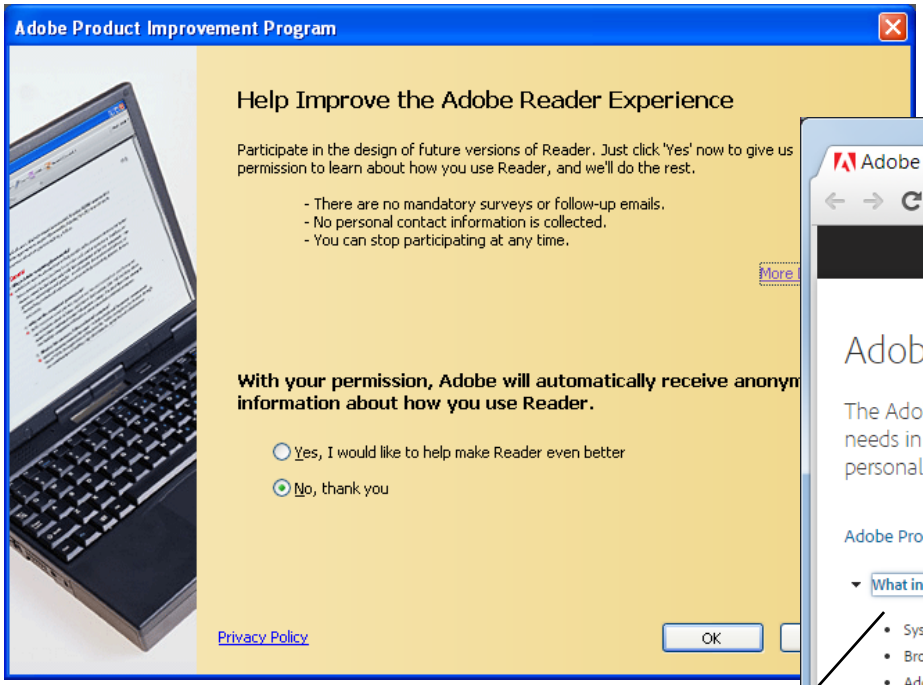
- **Interviews** and **focus group discussions**
  - A form of qualitative research – we ask a group of people their perceptions, views, and motivation towards a product
  - Interviews with individual users can be productive because the interviewer can pursue *specific issues* of concern.
  - Group discussions (focus groups) are valuable to ascertain the *universality* of comments.

Evaluation During Active Use

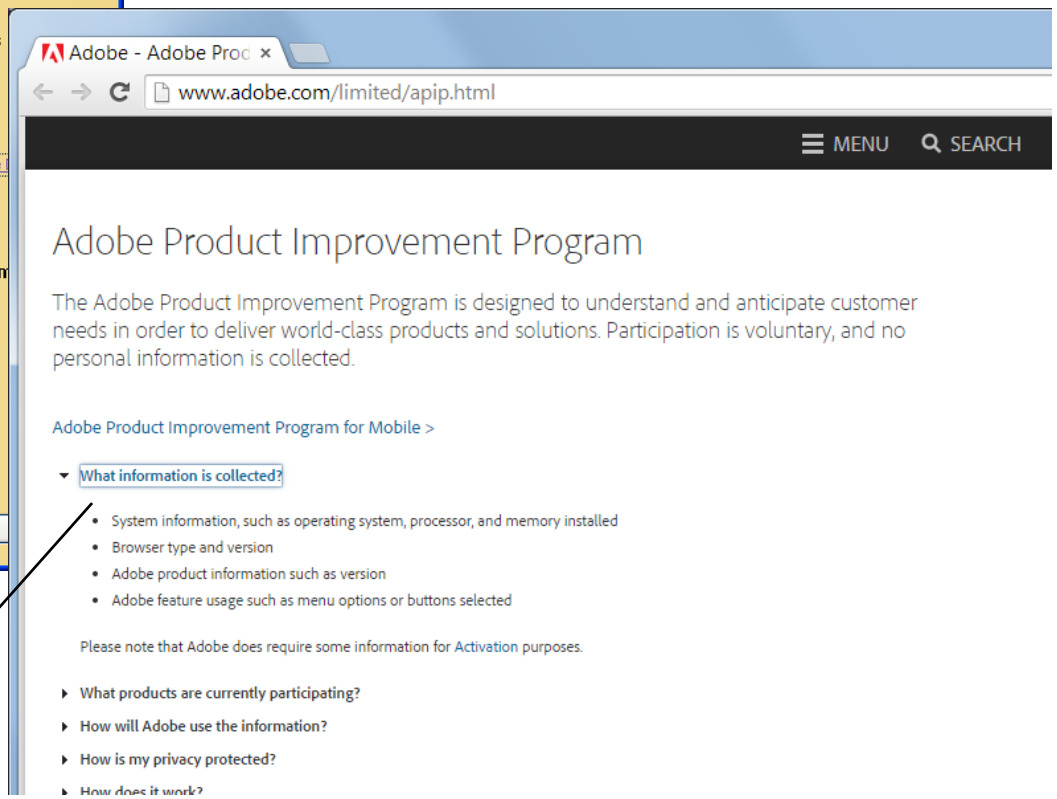
- **Continuous user-performance data logging**
  - The software architecture should make it easy for system managers to **collect data** about
    - The patterns of system usage
    - Speed of user performance
    - Rate of errors
    - Frequency of request for online assistance
  - Its major benefit is to guide the system maintainers in optimizing the performance and reducing costs
  - **Privacy Issue**: inform and get permission from users ahead

- **Online or telephone consultants**
  - Many users feel reassured if they know there is a human assistance available
  - On some network systems, the consultants can monitor the user's computer and see the same displays that the user sees

Evaluation During Active Use



http://www.adobe.com/limited/apip.html

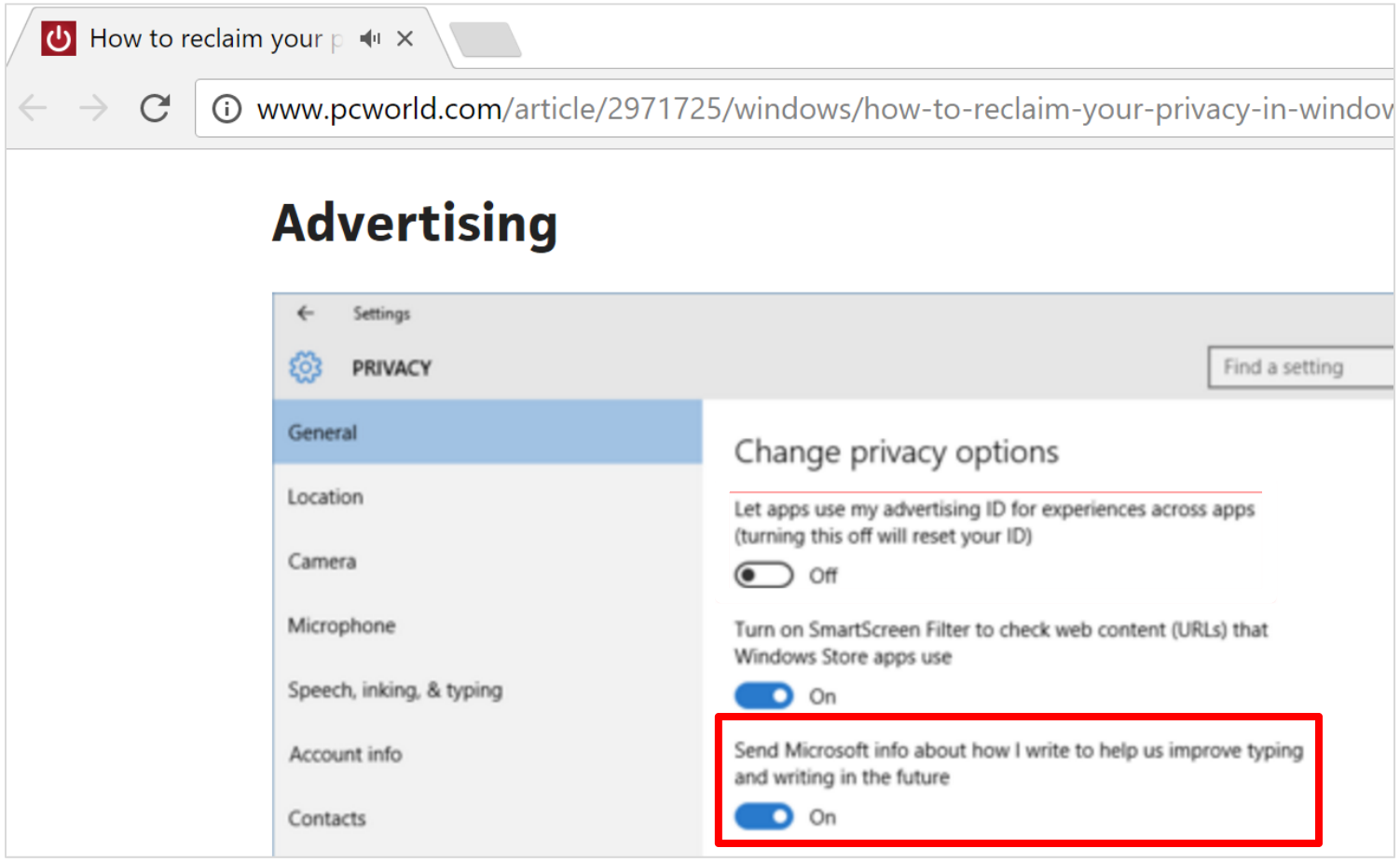


E.g., Adobe Reader (older version)

- System information, such as operating system, processor, and memory installed
- Browser type and version
- Adobe product information such as version number
- Adobe feature usage information such as menu options or buttons selected

Evaluation During Active Use

But… before your program tracks and log user data, you should get permission from the user!!! There are many news about this on Win10.



http://www.pcworld.com/article/2971725/windows/how-to-reclaim-your-privacy-in-windows-10-piece-by-piece.html

Evaluation During Active Use

- **Online suggestion box** or e-mail trouble reporting
  - Email or web reports to the maintainers or designers.
  - User bug reports, e.g., web-based tools such as Bugzilla http://www.bugzilla.org/
- **Discussion groups**, wikis, and newsgroups
  - Permit postings of open messages and questions
  - Some are independent, e.g. America Online and Yahoo!
  - Users can scan (or search) for relevant topics (user-generated contents)
  - Need some moderators
  - A sense of community
- **Public Test Server**
  - Put the latest software on a web server for users to try and give feedback: get free testing before releasing a new patch

    e.g., Diablo III:

    https://news.blizzard.com/en-gb/diablo3

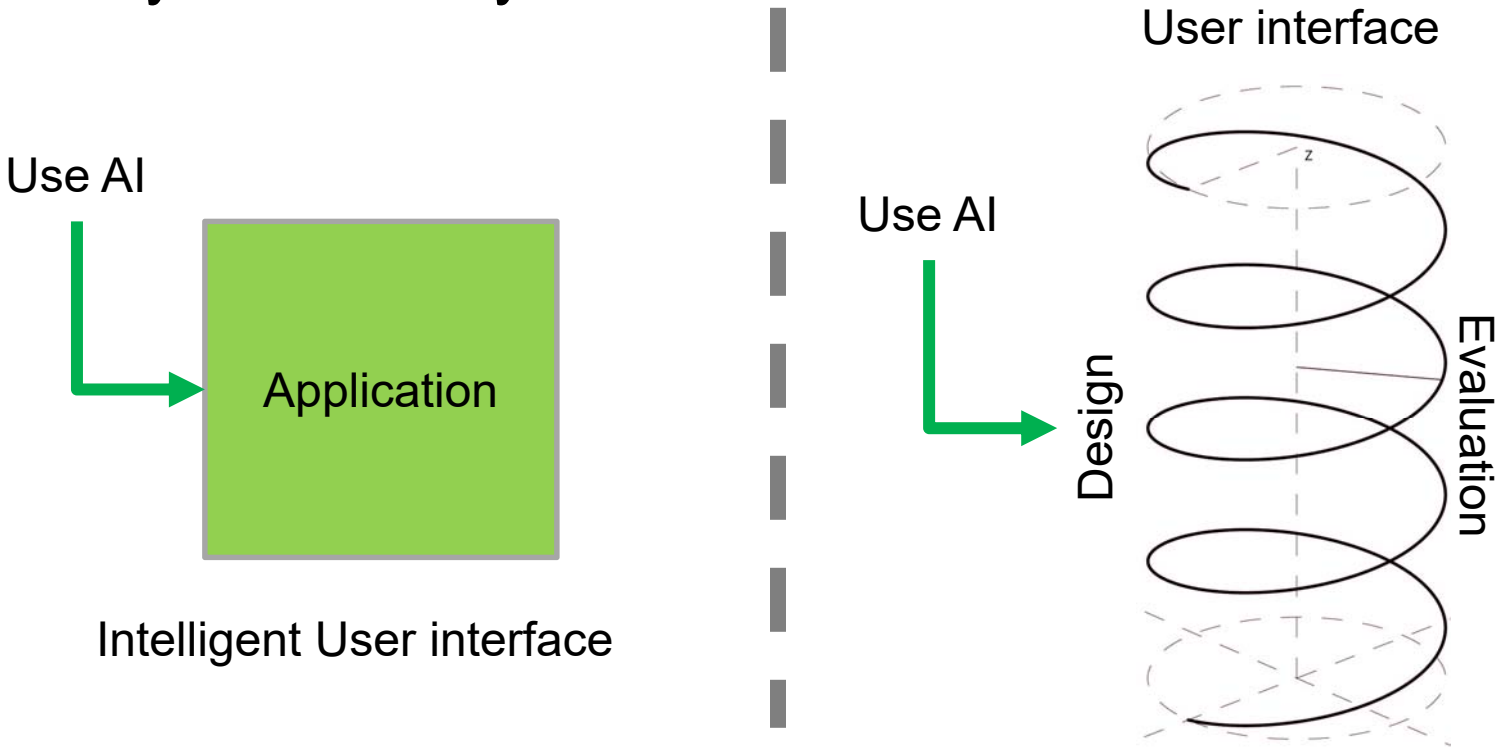
Note: there are many other methods other than these six, e.g., checking user blog, etc.

Topics:

- Introduction
- Expert Review
- Usability Testing and Labs
- Survey Instruments
- Acceptance Tests
- Evaluation during Active Use
  - a.k.a. continuous evaluation
  - six strategies
- **How AI helps UI/UX design & evaluation**

How AI helps UI/UX design?

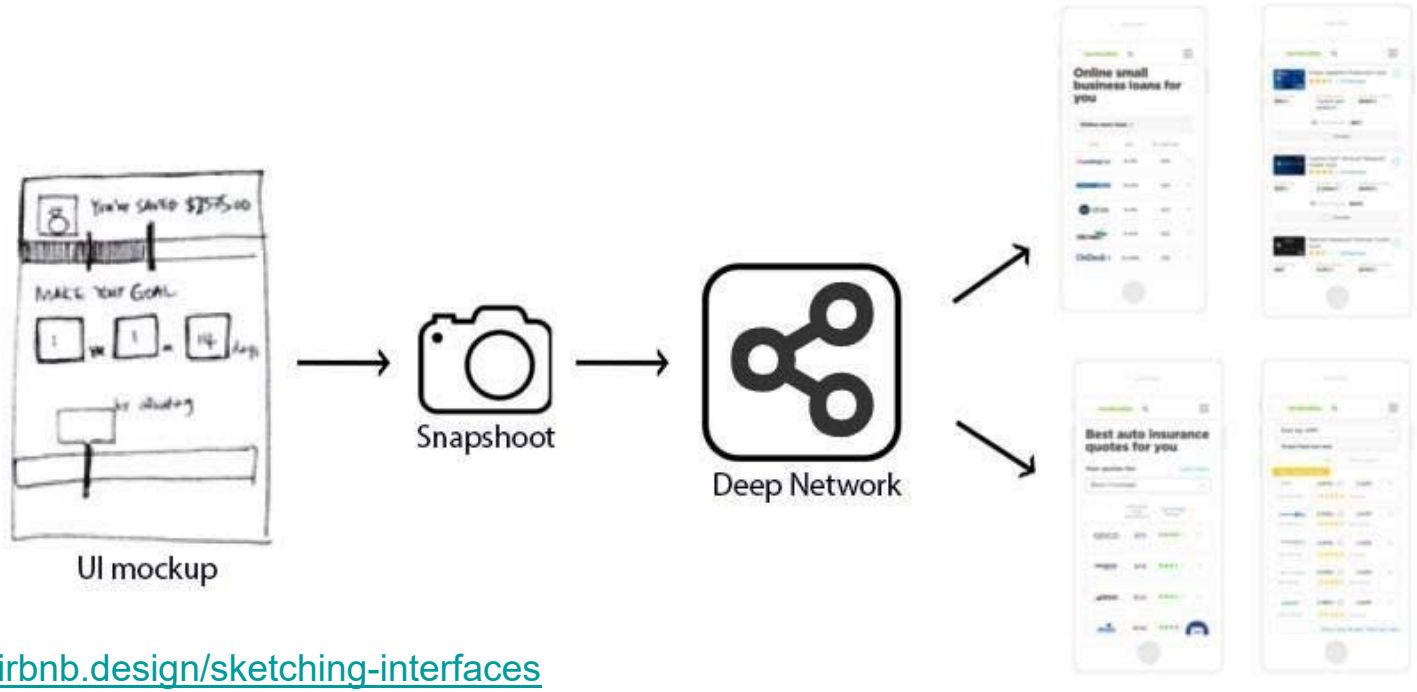
- This is different from our topic in week 4: Intelligent User interface or AI-enabled user interface!
- It means how we may use AI to improve the way we create UI/UX
- Do you have any idea?



Example #1

Q: Can AI help us to reduce routine works in UI design?

This preliminary work demonstrates the idea of directly generating a user interface with the associated code from an input low-fi prototype design (as an image).



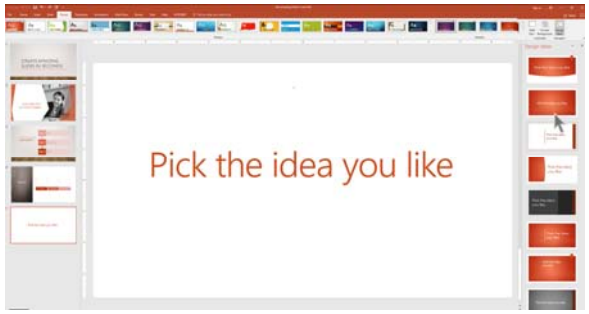
Source:
https://airbnb.design/sketching-interfaces
https://xd.adobe.com/ideas/process/wireframing/wireframing-automation-ai-ux-design/

Example #2

Q: Can AI's generative power help to produce UI contents?

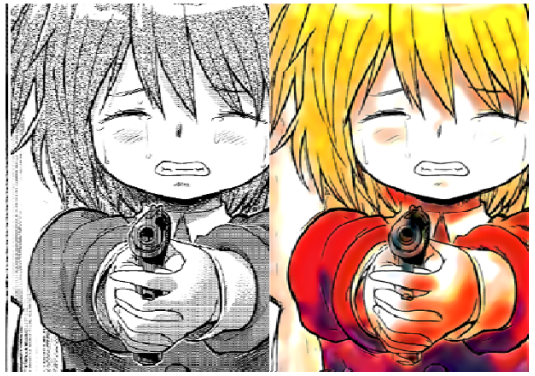
Examples:

- PowerPoint designer (link)
  - We provide contents
  - Software tools (or AI) produce layout and make suggestions!



- AI generates contents
  - E.g., https://thispersondoesnotexist.com/
  - Adobe uses this technology here: https://uifaces.co/plugin-adobe-xd
- AI modifies contents:
  - Photo -> Cartoon -> Manga

https://medium.com/syncedreview/reproducing-japanese-anime-styles-with-cartoongan-ai-cf30d583736e



https://www.cse.cuhk.edu.hk/~ttwong/papers/screenstyle/screenstyle.html

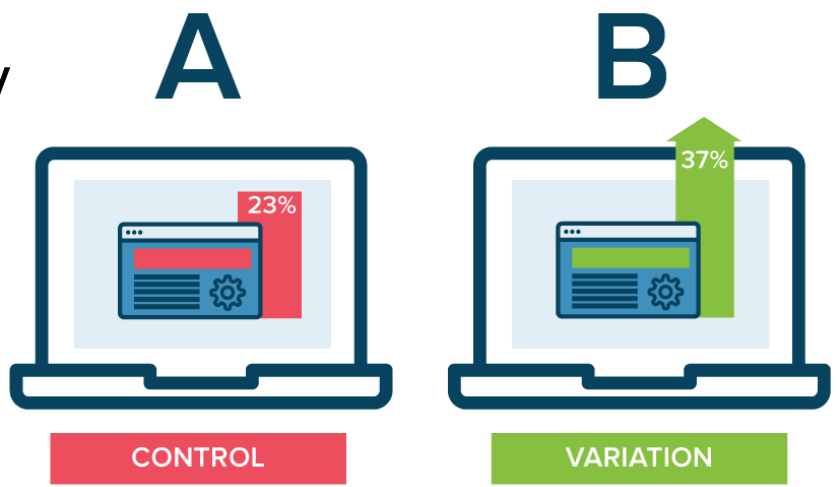
Example #3

Q: Can AI help us to continuously evaluate & refine websites?

Preliminary works are not yet AI but mainly **crowdsourcing**:

- Show two UI variants to massive users (A/B test)
- Find out the user behavior
- Refine the website accordingly

Then, later…
Can we combine crowdsourcing with generative power of AI?



Source:
https://uxdesign.cc/a-b-testing-to-drive-product-adoption-8126ca8be27e
http://go.refuel4.com/analyze-my-ads-1
https://netflixtechblog.com/its-all-a-bout-testing-the-netflix-experimentation-platform-4e1ca458c15
https://www.youtube.com/watch?v=VvTYuQPINec (Netflix use A/B test + AI to improve their production)

Example #4

Q: Can AI help us to automatically check a user interface?

Remember?

- Guideline review
- Consistency check
- … and beyond

In recent years, many software tools and companies have come into play, e.g.,

- https://applitools.com/applitools-ai-and-deep-learning/ (https://www.youtube.com/watch?v=eIPFaD1WWaY)
- https://stackshare.io/applitools/alternatives

Final note…

1. Things happen very fast
2. Creativity and idea are the key
3. This is a fairly unexplored area in the past

Further reading:
https://blog.adobe.com/en/publish/2019/10/01/the-state-of-ai-in-ui-ux-design.html#gs.shxvtw

Summary

- Evaluation of UI is necessary
- Should be part of a project budget
- Several different approaches outlined
  - Expert Review (6):
    - 6 methods
  - Usability Testing and Labs (4):
    - various tests
    - physical env. & setup
    - participants
    - how to record
  - Survey Instruments
  - Acceptance Tests (5)
    - based on 5 measurable human factors
    - not to detect flaws but to verify the requirements
  - Evaluation during Active Use (6)
    - 6 strategies
- Aspects that AI can help UI/UX design & evaluation?

Basic Terminologies you've learnt

- Expert review

Heuristic evaluation (Discount Evaluation), Guidelines review, Consistency inspection, Cognitive walkthrough, Metaphors of human thinking, and Formal usability inspection

- Nielsen's 10 usability heuristics
- Usability lab.
- Informed consent
- Logging software
- Think-aloud (or talk-aloud)
- Eye tracking and Heat map
- Usability tests (different variants): A/B test, mockup test, etc.
- Survey instruments and Likert scale
- Acceptance tests: objective and measurable goals
- Continuous Evaluation, Focus group, Public Test Server