Introduction to HCI
Fall 2021

Evaluation
Usability Testing

Mahmood Jasim
UMass Amherst
mjasim@cs.umass.edu
https://people.cs.umass.edu/~mjasim/

Learning Goals

- Understand the role of usability testing in HCI
- Be able to define usability testing (nelson's definition vs others)
- Understand how usability testing is different from other evaluation methods
- Explain when usability studies are typically conducted and why
  - Give examples of locations, tasks, metrics, evaluation methods that might be involved
- Explain how to plan and conduct a usability study

What is the role of usability in HCI?

- Usability: a primary focus of HCI
  - Evaluate system usability
    - How easy it is for the user to get the system to do what s/he needs it to do
  - Design for usability
  - Establish/apply metrics and standards for usability

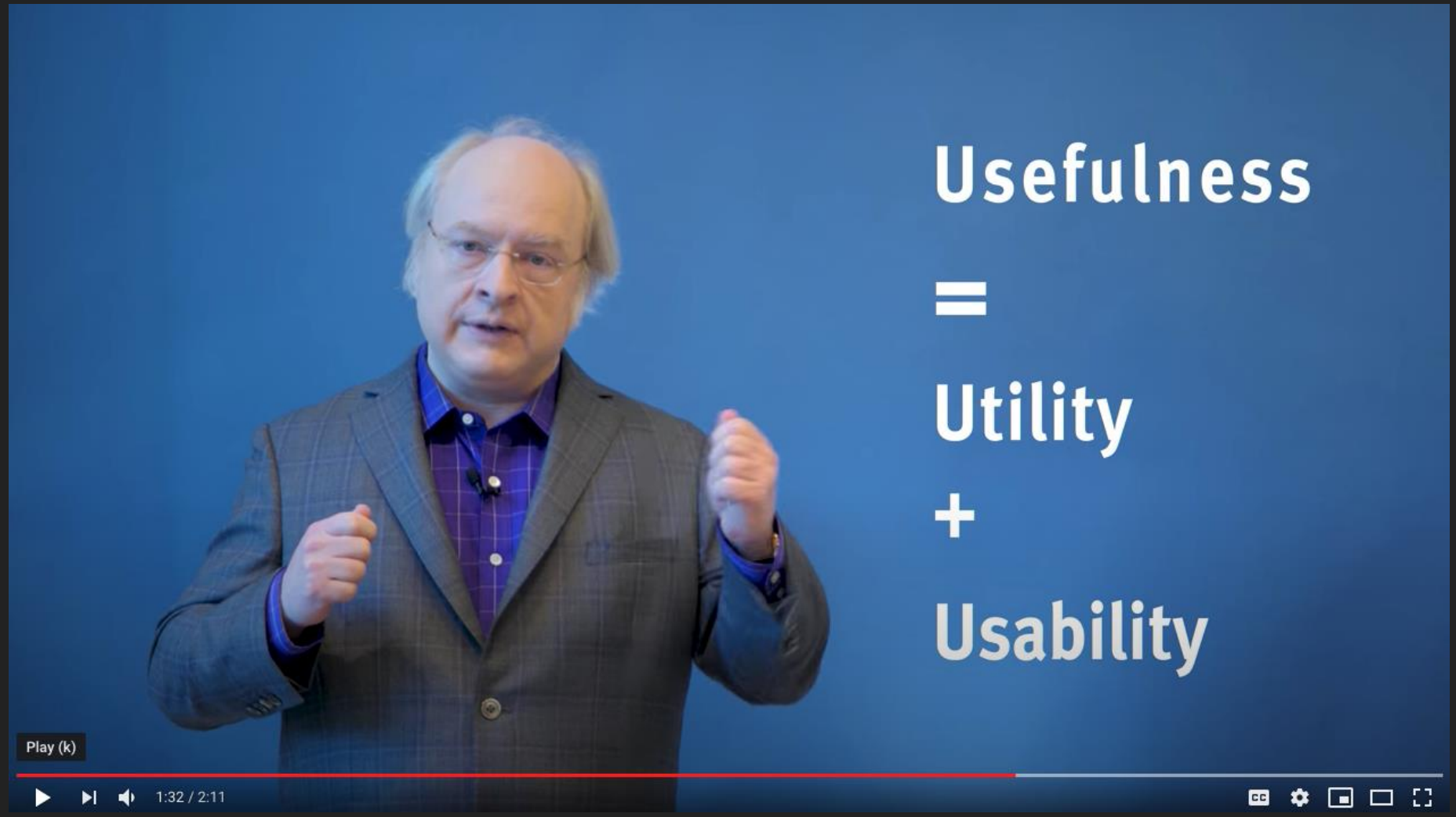
What is the role of usability in HCI?

- HCI starts with understanding the problems that users are having

Then designing a system that solves these problems

- requirements, task examples specify what it should do
- decide on conceptual/interface design for how system will do it

Usability studies: see if we succeeded



https://www.youtube.com/watch?v=VwgZtqTQzg8

Usability (Nielsen's definition)

- Learnability
  - easy to learn so a user can rapidly start to use it
- Efficiency
  - once the user has learned the system, a high degree of productivity is possible (better known as performance)
- Memorability
  - the user should be able to return to the system and not have to learn again
- Errors
  - users should make few errors and recover easily
- Satisfaction
  - the system should be pleasant to use

- Usability study/test evaluates an interactive system/prototype with respect to all/some of these elements, always involving real users

Elements of a usability test

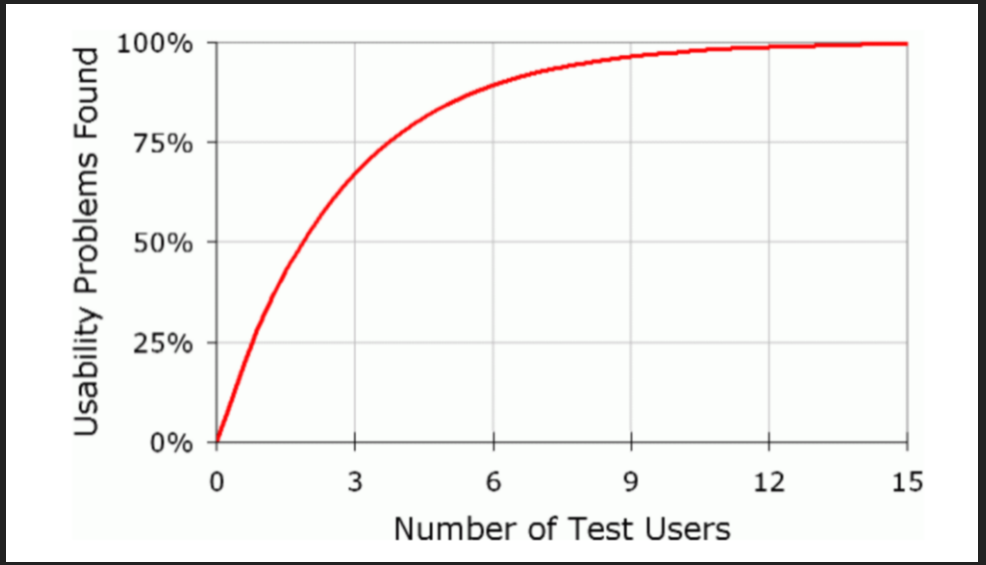
- Interactive system / prototype
- Evaluation goals
- Tasks
- Measures/metrics
- Data collection/recording methods
- Participants

When designing a usability test:

- Choice of methods: triangulate
  - Typically: one instrument counts something, while another interprets what was counted
- Choice of metrics: driven by your requirements & eval goals
  - As well as basic usability principles
- How many users: should be representative of your user groups
  - e.g. if you want to support both expert and novice users, should have good numbers of both!
  - Within a demographic, < 4-5 is dubious; often >10-12 is of marginal value
- Sometimes constraints dictate low numbers
  - If you have to generalize, consider who your test users are, and how representative they are

How many users?

- Tom Landauer and Jakob Nielsen
- The number of usability problems found in a usability test with *n* users:
- $N(1-(1-L)^n)$, where N = #problems, L = proportion of problems found while testing with a single user, n = #users



https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/

https://www.nngroup.com/videos/usability-testing-5-users-information-foraging/

Tasks

- Generally: user researcher specifies the task
  - At quite low level; e.g. The subtask that will take you from one screen to the next.
  - Or, at entire task level: see if someone can figure it out, start to finish, and watch /count / measure the challenges s/he has

Methods: examples of common ones

- Observational techniques:
  - silent
  - think aloud
  - constructive interaction

- Query techniques:
  - Interview
  - survey
  - questionnaire

Metrics Examples of common ones

- Time:
  - To complete a task (entire, or a portion)
  - Learn a task
  - Resume a task after interruption
  - Find something on a screen
  - Attain specified degree of proficiency

- Errors:
  - Number per task or unit of time
    - Different types: e.g., Navigation, selection, interpretation
  - Number of users making the error
  - Alternately: number of successes

Metrics: Examples of common ones

Events of interest:

- page views or clicks
- access of particular tools
- timeouts
- questions asked or help tools consulted
- # users willing to recommend

Subjective factors:

- task level satisfaction
- perception of aesthetics
- perceived ease of use
- perceived preference
- (all can be measured on a Likert or semantic rating scale)

Alternatives to usability testing

- Usability testing requires users, relatively refined prototypes, and usually focusses on measuring something.

- "Discount" methods can also target prototypes at various stages and be done without users
    - Heuristic evaluation
    - Cognitive walkthrough

- Because you might not have access to users . . .
    - Can do it first (before a usability study)
    - Possible to apply these methods yourself while iterating on a design (before it's totally finished)

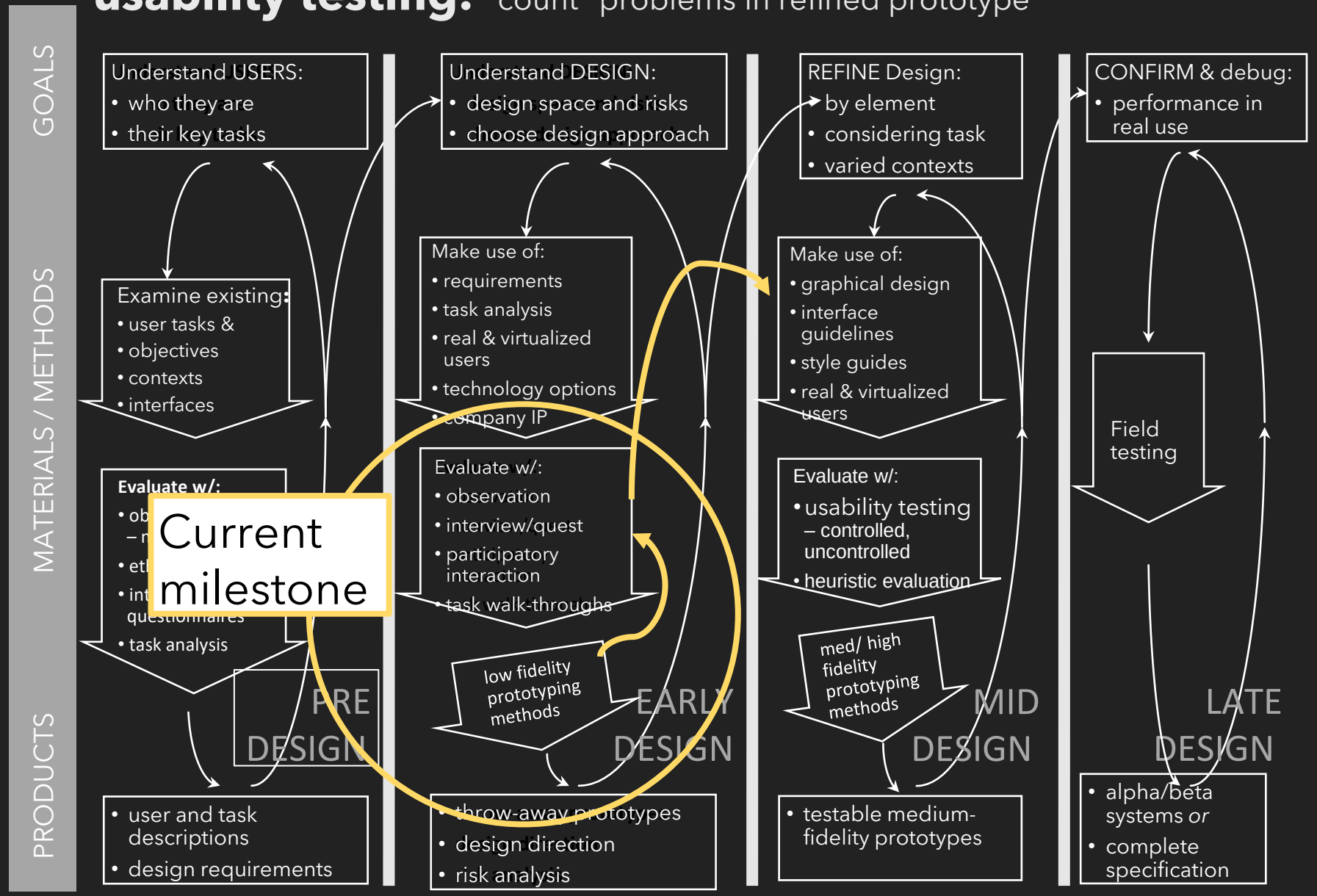
Biggest differences with alternatives

- Usability testing requires:
- A refined interface.
  - This could be a medium fidelity prototype.
  - Or it could be the bad old interface, which you plan to revise or replace
    i.e., Might be “evaluate for understanding the problem”
- Measured outcomes.
- Users (participants).

Note on terminology

- Not entirely standardized…
- User Study – very general. Any study that involves actual or prospective users. Can be anytime -- from before a system is built (Empathize / Pre-Design) right to a controlled experiment.
- Usability Study – more specific. Requires a system for which task performance can be measured (usually Mid / Late Design, but can be Pre-Designing for a system being re-designed)
- Controlled Experiment – a specific type of usability study with hypotheses and statistical testing, often comparing alternate designs Informal / Small User Study – often used before a usability study, not ready to measure things yet, interested in higher-level feedback. (Early design).

usability testing: "count" problems in refined prototype



Usability testing in your project

- Evaluation goals?
  - You will likely want to draw from your requirements and task examples; may need to prioritize;
  - Test how well your system supports what you intended it to
  - Metrics, evaluation methods, etc. Should follow

- Hi fidelity prototype scope?
  - Prototype should be a working system
  - It should do enough to test if your design will meet your goals (and be achievable in the time available)

In-class activity

- Work in groups
  - Design user tasks for your own project
  - Identify the following
    - Evaluation goals
    - Prototype scope
    - Evaluation metrics

- https://tinyurl.com/ykws6a54

Optional Reading

- Usability 101
  - https://www.nngroup.com/articles/usability-101-introduction-to-usability/

- How many test users in a usability study?
  - https://www.nngroup.com/articles/how-many-test-users/