Prototype Evaluation

Professor Tapan Parikh
School of Information, UC Berkeley

FORMATIVE vs SUMMATIVE EVALUATION

Formative evaluation - Discover product fit and usability issues as part of an iterative design process. Goal is to learn as much as possible.

Summative evaluation - Assess the performance of a prototype, or compare alternatives. Goal is a reliable, statistically valid comparison.

TESTING PROTOTYPES

What do you want to learn?

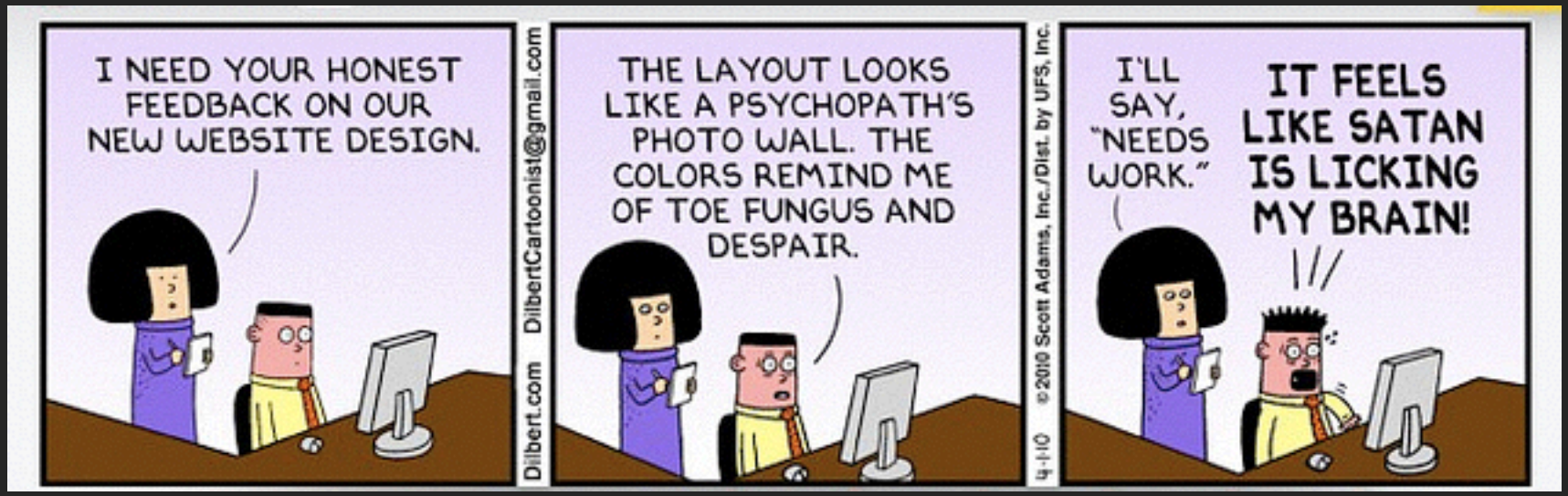
Design representative tasks / scenarios

Identify potential users / experts

Observe the users interacting with the prototype

Analyze the resulting data

Adapted from Jake Wobbrock



http://www.flickr.com/photos/aimeetwilliamson/4481888431/in/photostream/

SELECTING USERS

"Should be as representative as possible of the intended users"

If testing with a small number of users, avoid outlier groups

If testing with a larger number of users, aim for coverage of all "personas"

Include novices, probably experts too

It helps if users are already familiar with hardware (if not part of your design)

Adapted from Jake Wobbrock

THINKING ALOUD

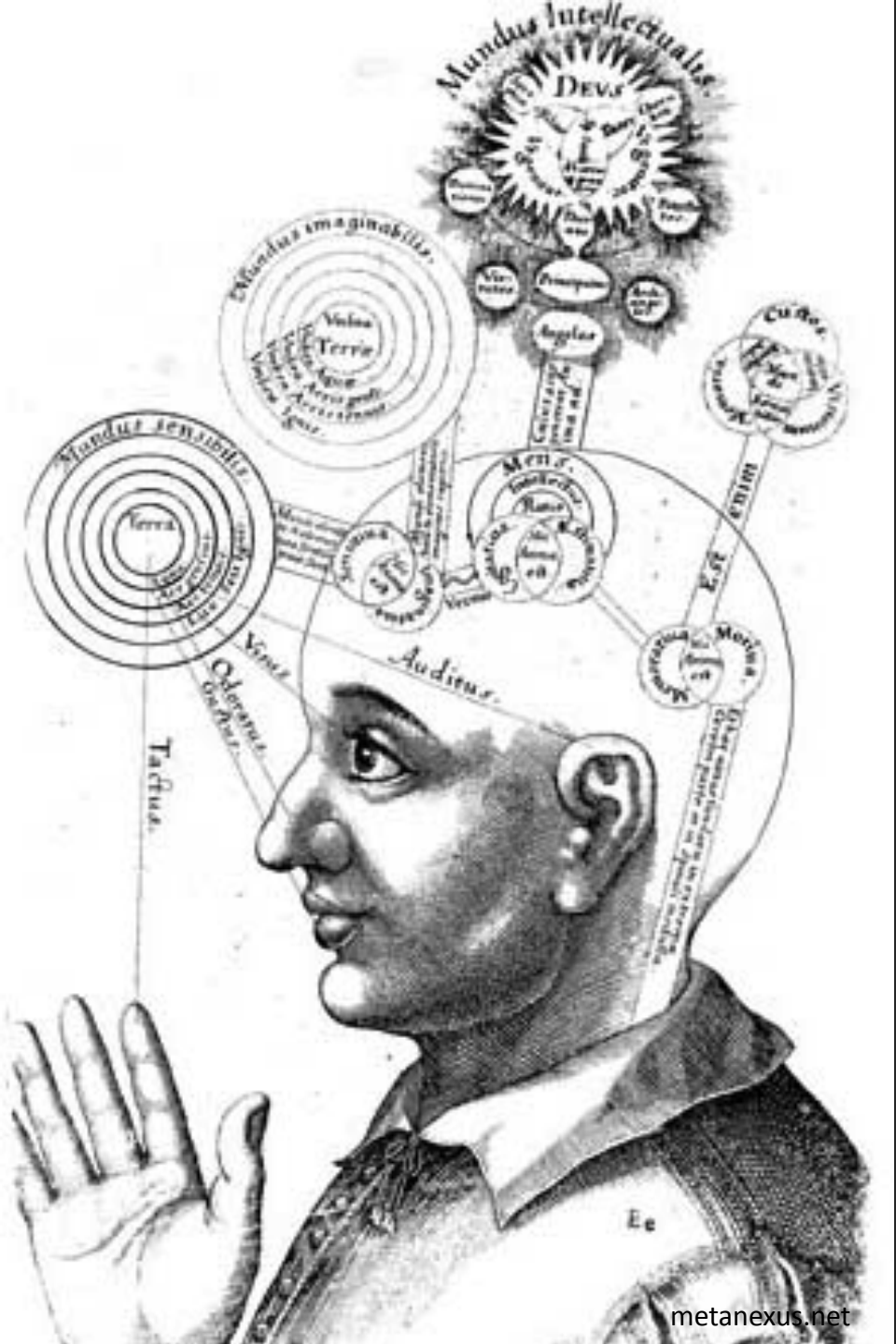
"Having a test subject use the system while continuously thinking aloud"

Useful for formative evaluation

Understand how users view the system by externalizing their thought process

Generates a lot of qualitative data from relatively small number of users

Focus on what the user is concretely doing and saying




0:17 / 4:19
CC

EXAMPLE TASK

Motivating Scenario: "You are using a mobile phone for accessing and editing contact information."

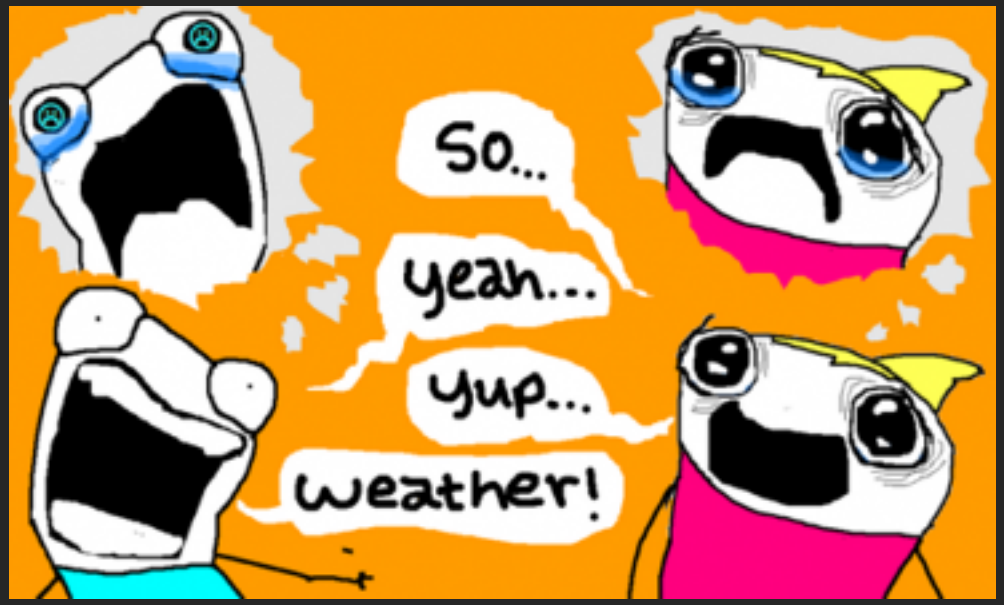
Tasks:

1. Find the contacts list in the phone.
2. View the contact information for John Smith.
3. Change John Smith's number with area code "510".
4. ...

Adapted from Jake Wobbrock

GETTING USERS TO OPEN UP

Thinking aloud can be unnatural and awkward



Requires prompting by the experimenter to ensure that the user continues to externalize their thought process

May slow them down but thats fine

EXAMPLE PROMPTS

"Please keep talking."

"Tell me what you are thinking."

"Tell me what you are trying to do."

"Are you looking for something? What?"

"What did you expect to happen just now?"

"What do you mean by that?

Adapted from Jake Wobbrock

POINTS TO REMEMBER

Do not make value judgments

User: "This is really confusing here."
Tester: "Yeah, you're right. It is." (BAD)
Tester: "Okay, I'll make a note of that." (GOOD)

Video or audio record (with user's permission), and/or take good notes

Screen captures / Eye tracking

When the user is thinking hard, don't disturb them with a prompt - wait!

Adapted from Jake Wobbrock

THINK ALOUD VARIANTS

Co-Discovery: Two users work together
- Can spur more conversation
- Needs 2x more users

Retrospective: Think aloud after the fact, reviewing a video recording
- Doesn't disturb the user during the task
- User may forget some thoughts, reactions

Coaching: Expert coach guides user, answering questions
- Identify training, help and documentation needs

Adapted from Jake Wobbrock

HOW TO TREAT USERS

Train them if you will assume some basic skills (ex. using a mouse)

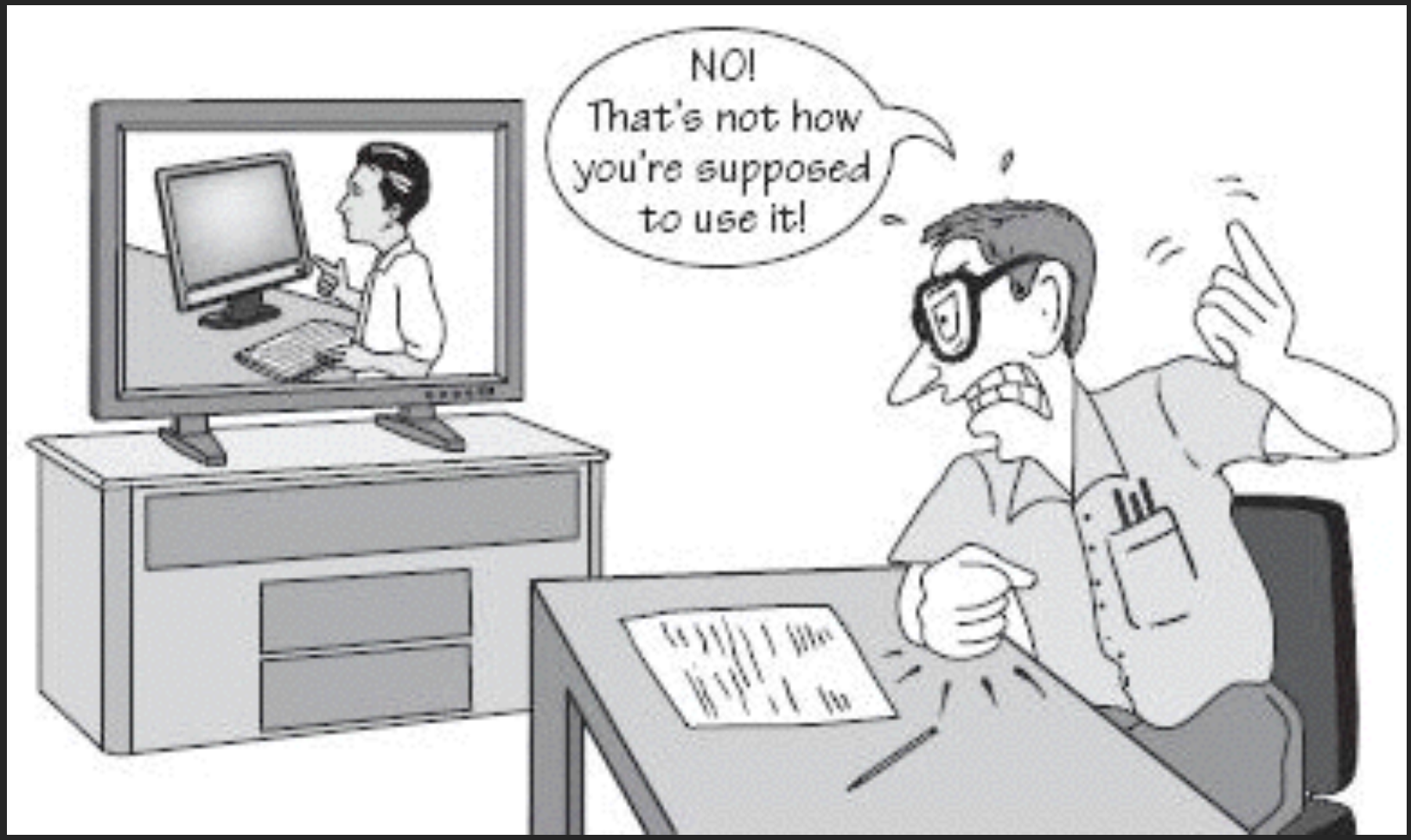
Do not blame or laugh at the user

Make it clear that the system is being tested, not the user

Make the first task easy

Inform users that they can quit anytime

After the test, thank the user



http://2.bp.blogspot.com/_A1dCJRsvsgQ/SwWVn3x0_8I/AAAAAAAAAXU/cTHK8BJGkA0/s1600/Usability_testing.JPG

HELPING USERS

Decide in advance how much help you will provide (depending on how complete your prototype is, and whether or not you plan to measure performance)

For the most part, you should allow users to figure things out

If they get stuck, give a few hints to get them going again

Terminate the test if the user is unhappy and not able to do anything

User can always voluntarily end the test

DESIGNERS AS EVALUATORS

Usually system designers are not the best evaluators

Potential for helping users too much, or explaining away usability problems

Evaluator should be trained in the evaluation method, and also be an expert in the system being tested

Can be a team of a designer and an evaluator, who handles user relations

For next time

Executive Summary and Analysis of Alternatives / Status Quo due

Project Pitches in class