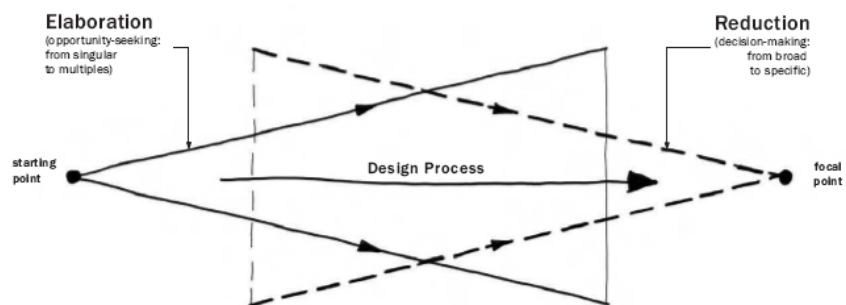# COMS W4170
UI Design and Evaluation

Steven Feiner
Department of Computer Science
Columbia University
New York, NY 10027

September 25, 2018
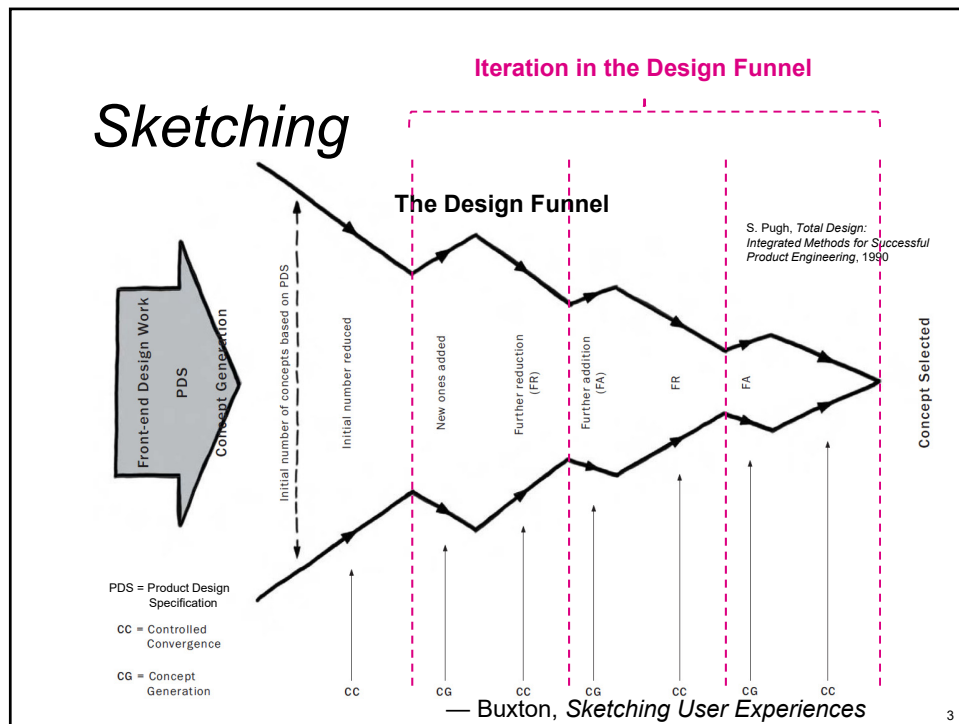
1

Sketching

P. Laseau, *Graphic Thinking for Architects and Designers*, 1980

— Buxton, *Sketching User Experiences*

2

Sketching

The Design Funnel

S. Pugh, *Total Design: Integrated Methods for Successful Product Engineering*, 1990

Front-end Design Work

PDS

Concept Generation

Initial number of concepts based on PDS

Initial number reduced

New ones added

Further reduction (FR)

Further addition (FA)

FR

FA

Concept Selected

PDS = Product Design Specification

CC = Controlled Convergence

CG = Concept Generation

CC CG CC CG CC CG CC

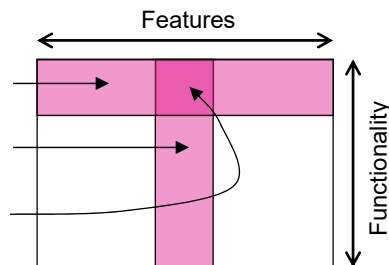— Buxton, *Sketching User Experiences*

3

Discount Usability Engineering
J. Nielsen

- Cost-cutting approach to evaluating usability "on the cheap"
    - Use lo-fi approaches
    - Sacrifice statistical significance (e.g., fewer participants)
- Ingredients
    - *Scenarios* reduce complexity
        - Horizontal prototype: Full UI with all features / reduced functionality
        - Vertical prototype: Partial UI with partial features / full functionality
        - Scenario: Partial features / partial functionality (whether on computer or on paper)

Features

Functionality

4

Discount Usability Engineering
J. Nielsen

- Ingredients (cont.)
  - Simplified "Thinking aloud" protocol
    - Participant "thinks aloud" while using system
    - Experimenter takes notes instead of recording electronically for later analysis
  - Heuristic evaluation

Heuristic Evaluation
http://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/

- *Evaluators* (3–5) individually and systematically inspect the UI, comparing it with a set of general (and, optionally, domain-specific) evaluation heuristics
  - May need to document a specific task to evaluate and steps for performing it if evaluators are not familiar with the domain
  - Go through UI at least twice
  - Note each problem individually
- *Observer* may take notes
  - Eliminates note-taking burden for evaluator
  - Observer is similar to an experimenter, but
    - Records, rather than interprets (i.e., evaluator does the evaluation)
    - Provides help (especially if evaluator is not familiar with the domain)

Heuristic Evaluation

http://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/

- Evaluators and observers may communicate afterwards, rate severity of problems found
  - Frequency of occurrence, user impact, persistence, market impact
    - https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/
- Evaluation heuristics
  - http://www.nngroup.com/articles/ten-usability-heuristics/
  - http://asktog.com/atc/principles-of-interaction-design/

Obtaining User Input

- During design ←→ After design
- In-person ←→ Remote
- Conscious user involvement ←→ Automated data gathering

Obtaining User Input: Interviews

- One-to-one (interviewer-to-interviewee)
  - Structured interview
    - Rigorously standardized questions and order
    - Easier to compare across participants
  - Semi-structured interview
    - Some questions planned in advance, but interviewer can develop new questions on the fly
    - Allows emphasis on interesting topics, exploration of unanticipated directions
  - Unstructured interview
    - Free-form

9

Obtaining User Input: Focus Groups

- One-to-many, several-to-many (moderator(s)-to-respondents)

http://www.pra.ca/en/focus-group-facility

  - + Diversity of opinions, participants can feed off each other
  - – Individuals can dominate or be intimidated
  - "We don't do focus groups. They just ensure that you don't offend anyone, and produce bland inoffensive products."
    — Jony Ive, Apple CDO

http://www.macworld.com/article/1141509/jonathan_ive_london.html

10

Obtaining User Input: Surveys

- Document user demographics
  - Age, gender, experience, personality
- Capture subjective reactions to system
  - Likert scales (developed by psychologist Rensis Likert ["Lick-urt"])

Label text matters!	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree

Label text matters!
- Extremes
- Biases

Label placement matters!

Poor   Excellent

1 2 3 4 5 6 7

- Free-form comments

Typically use an odd number of choices to allow neutrality, but sometimes use an even number to cause *forced choice*.

Note potential problems of scale inversion and inconsistency with multiple questions

Treat as *ordinal* (not *interval*) data—can't assume users consider values equidistant, but only that $n+1 > n$. Note that middle is *usually* (but not always!) neutral.

Obtaining User Input: Surveys

- Example: QUIS (Questionnaire for User Interaction Satisfaction) [See S Table 5.1]

Overall reactions to the system:

terrible   wonderful
    1 2 3 4 5 6 7 8 9   NA

frustrating   satisfying
    1 2 3 4 5 6 7 8 9   NA

dull   stimulating
    1 2 3 4 5 6 7 8 9   NA

difficult   easy
    1 2 3 4 5 6 7 8 9   NA

    ...

Ethnographic Observation

- Observation of users, based on methods used by *ethnographers* in field studies of cultures
  - *Understand* users, tasks, tools, interactions
- Holistic approach
  - Observe in natural habitat (home or work)
  - May ask questions (interview), participate in activities
  - Acquire subjective/objective data
    - qualitative anecdotes ↔ quantitative reports
- Differences with classical ethnography
  - Shorter immersion periods (e.g., hours/days vs. weeks/months)
  - Culture is often closer to that of the observer
  - Emphasis on interface design, rather than cultural understanding

13

User-Centered Design (UCD)

- General name for design processes that place users' needs at the forefront, from initial conception on…
- *Consider* users and *involve* users

14