

CTT534 – Thiết Kế Giao Diện  
HK II 2015 – 2016

## Đánh giá thiết kế giao diện

Bài giảng có tham khảo các nguồn tài liệu sau  
- MIT CS Course 6.813/6.831  
- USC's CSCI 588

## Nội dung

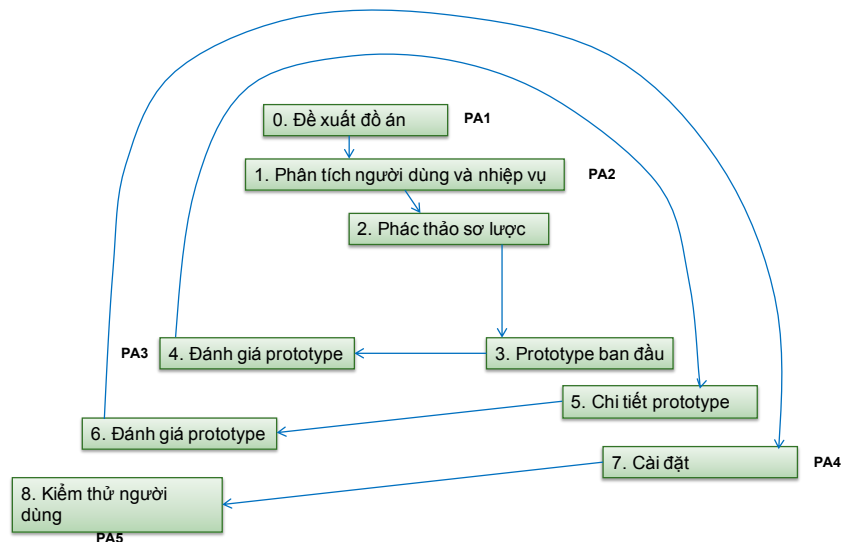
- Tổng quan về đánh giá thiết kế giao diện
- Các phương pháp đánh giá
  - Chuyên gia
    - Đánh giá heuristic
  - Người dùng
    - Field studies
    - Formative evaluation
    - Controlled experiment

5/24/16



2

## Quy trình áp dụng cho đồ án



5/24/16



3

## Đánh giá – Why? What? Where? When?

- Thiết kế tương tác và đánh giá là quá trình liên tục, xem xét các khía cạnh sau:
  - Why: kiểm tra xem người dùng có thể dùng sản phẩm ko, họ có thích sản phẩm ko
  - What: thiết kế ý tưởng ban đầu, làm prototype và sau đó là chi tiết, hoàn thành prototype
  - Where: trong môi trường tự nhiên hoặc có thiết lập
  - When: xuyên suốt quá trình thiết kế; sản phẩm hoàn thành có thể được đánh giá để thu thập thông tin mô tả về sản phẩm đó

5/24/16



4

## Một vòng thiết kế đảm bảo tính khả dụng

- Thiết kế ý tưởng (conceptual design)
  - Phát triển và đánh giá các ý tưởng thiết kế ban đầu cùng với người dùng
- Thiết kế prototype ban đầu (early prototype design)
  - Qui trình phát triển liên tục cùng với người dùng
    - Cách ứng xử của hệ thống có khớp với yêu cầu nghiệp vụ của người dùng ko?
    - Có vấn đề gì với thiết kế hiện tại hay ko?
- Thiết kế prototype hoàn chỉnh
  - Kiểm thử chấp nhận (acceptance testing) giúp xác định hệ thống có thỏa mãn các yêu cầu hiệu năng của người dùng mong đợi ko

5/24/16



5

## Mục đích của đánh giá

- Đánh giá hiệu quả của tương tác với người dùng
  - Tính khả dụng và sự thỏa mãn của người dùng
- Đánh giá phạm vi của chức năng hệ thống
- Xác định các vấn đề cụ thể trong bản thiết kế hiện tại
  - Xác định các khía cạnh thiết kế gây ra nhầm lẫn hoặc kết quả ko mong đợi

5/24/16



6

## Các phương pháp đánh giá

- Đánh giá heuristic
  - Do chuyên gia thực hiện
- Do người dùng thực hiện việc kiểm tra
  - Field studies
  - Formative evaluation
  - Controlled experiment

5/24/16



7

## Nội dung

- Tổng quan về đánh giá thiết kế giao diện
- Các phương pháp đánh giá
  - Chuyên gia
    - Đánh giá heuristic
  - Người dùng
    - Field studies
    - Formative evaluation
    - Controlled experiment



5/24/16



8

## Các hướng dẫn/heuristics về tính khả dụng

- Nhiều heuristics dựa vào
  - 10 heuristics do Nielsen đề nghị
  - Các luật của Norman trong Design of Everyday Things
  - 8 luật vàng của Shneiderman
- Heuristics
  - Giúp người thiết kế chọn lựa giữa các giải pháp thiết kế
  - Giúp người đánh giá tìm ra các vấn đề trong thiết kế, đánh giá các giao diện (do đó gọi là đánh giá dựa vào các heuristics)



## Các nguyên tắc thiết kế UI đã học

- Tính dễ học(learnability)/dễ nhớ (memorability) – Viết tắt là L
- Tính trực quan (visibility) – V
- Tính đơn giản (simplicity) – S
- User control – UC
- Xử lý lỗi (error handling) – ER
- Tính hiệu quả (efficiency) – E
- Thiết kế đồ họa (graphic design) – GD



## Các heuristics của Nielsen

- Khớp với đời thực (L)
- Nhất quán và tuân theo chuẩn (L)
- Trợ giúp và tài liệu hóa (L)
- User control và tự do (UC)
- Thể hiện trực quan trạng thái hệ thống (V)
- Linh động và hiệu quả (EF)
- Ngăn ngừa lỗi (ER)
- Nhận dạng tốt hơn hồi tưởng (ER)
- Báo cáo lỗi, chẩn đoán lỗi, phục hồi lỗi (ER)
- Thiết kế có thẩm mỹ và tối thiểu hóa (GD)



## Các luật của Norman

- Dấu hiệu tương tác (affordances) (L)
- Ánh xạ tự nhiên (natural mapping) (L)
- Tính trực quan (V)
- Phản hồi (S)



## 8 luật vàng của Shneiderman

1. Tính nhất quán (L)
2. Shortcut (EF)
3. Phản hồi (V)
4. Dialog closure (V)
5. Xử lý lỗi đơn giản (ER)
6. Hành động có thể quay lui (UC)
7. Đặt người dùng trong sự kiểm soát (UC)
8. Giảm thời gian nạp bộ nhớ ngắn hạn (ER)



## Đánh giá heuristics

- Phát minh và đánh giá bởi Jakob Nielsen
  - Chi phí thấp hơn và hiệu quả hơn các phương pháp khác
- Là một kỹ thuật thanh tra (inspection)
- Thường được thực hiện bởi các chuyên gia
- Các bước cơ bản
  - Duyệt lại (review)/thanh tra (inspect) giao diện
  - So sánh giao diện với các heuristics đang có
  - Ghi nhận lại bằng văn bản những lỗi về tính khả dụng
  - Lí giải và biện hộ các lỗi dựa vào các heuristics



## Hướng dẫn đánh giá heuristics-1

- Biện hộ vấn đề với 1 hoặc nhiều heuristics
  - “Dialog A có button ‘Hủy’ trong khi nút cùng chức năng như vậy trong Dialog B lại có tên là ‘Đóng’” (Tính nhất quán)
  - Đừng bao giờ bảo “Tôi ko thích layout này” (“I don’t like the layout”)
- Liệt kê các vấn đề
- Duyệt qua giao diện nhiều lần
- Đừng giới hạn mình vào 10 heuristics của Nielsen
  - Dĩ nhiên, ko phải tất cả heuristics có thể áp dụng cho 1 giao diện người dùng



## Hướng dẫn đánh giá heuristics-2

- Phải có nhiều người đánh giá
  - Nhiều người sẽ tìm ra nhiều vấn đề khác nhau hơn là chỉ có 1 người
  - Cùng 1 vấn đề có thể xuất hiện nhiều lần (duplicated)
  - Nielsen đề nghị cần có 3-5 người đánh giá
- Kết hợp đánh giá dựa vào heuristics và các phương pháp đánh giá khác
  - Mỗi phương pháp sẽ tìm ra các lỗi khác nhau
  - Đánh giá dựa vào heuristics có chi phí thấp hơn



## Đánh giá website HCMUS



## Đánh giá prototype

- Đánh giá dựa vào heuristic có thể áp dụng vào
  - Bản phác thảo
  - Prototype trên giấy và prototype trên máy
  - Code minh họa (demo code)
- Một số heuristics ko thể áp dụng được ở đây
  - Phản hồi
  - Vấn đề thiếu thành phần ("Missing-element")
  - Trợ giúp và tài liệu hóa

5/24/16



18

## Quá trình đánh giá chính qui

- Huấn luyện
  - Tổ chức meeting để giới thiệu về ứng dụng, các giao diện của ứng dụng
  - Giải thích về các loại người dùng (user population), lĩnh vực (domain), các kịch bản (scenario)
- Đánh giá
  - Những người đánh giá sẽ làm việc độc lập để ghi nhận các vấn đề và viết report
- Đánh giá mức độ nghiêm trọng (severity rating)
  - Người đánh giá đưa ra độ nghiêm trọng cho các vấn đề tìm thấy
- Mô tả ngắn gọn
  - Đội đánh giá và đội thiết kế thảo luận trên kết quả và đưa ra các giải pháp đề nghị

5/24/16



19

## Mức độ nghiêm trọng

- Có thể dùng các thang đo khác nhau, chẳng hạn
  - Cosmetic/trivial: tầm thường, ko quan trọng, độ ưu tiên thấp
  - Minor: lỗi nhỏ, ko nghiêm trọng, độ ưu tiên thấp
  - Major: lỗi lớn, khá nghiêm trọng, độ ưu tiên cao
  - Critical: lỗi cực kì nghiêm trọng, độ ưu tiên cao nhất
- Khi đánh giá mức độ nghiêm trọng, hãy xem xét các nhân tố sau
  - Tần suất: tần suất xuất hiện của lỗi?
  - Ảnh hưởng/tác động: có khó giải quyết ko?
  - Sự dai dẳng: lỗi lặp lại nhiều lần ko?

5/24/16



20

## Viết đánh giá heuristics-1

- Người đánh giá đưa ra mô tả về vấn đề tìm thấy
  - Để giảm thời gian cho cả đội thiết kế và đội đánh giá
- Bao gồm những nhận định tích cực lẫn sự phê bình
  - Nhấn mạnh vào sự phê bình và các đề nghị
- Không ko nhất thiết là quá gay gắt
- Phải cụ thể
  - Ko cụ thể: “text khó đọc”
  - Cụ thể: “text có độ tương phản kém”



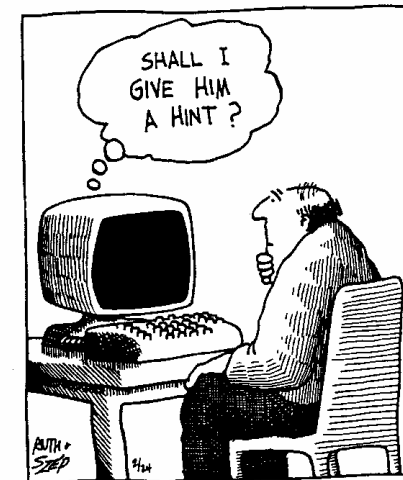
## Viết đánh giá heuristics-2

- Báo cáo đánh giá bao gồm các phần sau
  - Tiêu đề của lỗi
  - Mô tả lỗi
  - Heuristic
  - Mức độ nghiêm trọng
  - Đề nghị và các hình chụp (nếu có)



## Cognitive walkthrough

- Là kĩ thuật thanh tra tập trung vào tính dễ học
- Đầu vào
  - Prototype
  - Các tác vụ
  - Chuỗi hành động
  - Phân tích người dùng
- Người đánh giá đặt ra các câu hỏi theo luồng sau của hành động
  - Người dùng có tìm thấy chức năng trên giao diện?
  - Người dùng có biết phải làm gì ko?
  - ...



## Nội dung

- Tổng quan về đánh giá thiết kế giao diện
- Các phương pháp đánh giá
  - Chuyên gia
    - Đánh giá heuristic
  - Người dùng
    - Field studies
    - Formative evaluation
    - Controlled experiment



## Field studies

- Môi trường sử dụng thật sự
  - Quan sát các hiệu ứng xảy ra trong môi trường thực tế
  - Không thể nói giao diện là tốt cho đến khi đưa cho người dùng sử dụng và họ thấy thỏa mãn với nó
  - Khó để dự đoán những gì người dùng thật sự muốn
    - Vấn đề được xác định là những thứ phát sinh khi người dùng sử dụng ứng dụng
  - Tính khả dụng có thể được đo lường thông qua các kiểm tra bằng cách tính toán hiệu năng sử dụng từ phía người dùng



## Field studies (tt)

- Vấn đề
  - Tốn thời gian, chi phí thực hiện
  - Có thể không tổng quát hóa được



## Quy trình thực hiện

1. Chuẩn bị đề xuất kiểm tra
2. Chọn thành viên tham gia vào quy trình
3. Chọn các mục cần kiểm tra
4. Thực hiện kiểm tra
5. Lượng giá các kết quả kiểm tra



## 1. Chuẩn bị đề xuất kiểm tra

- Đề xuất kiểm tra (test proposal)
  - Mục đích kiểm tra (test objective)
  - Các thành viên tham gia kiểm tra (test participants)
  - Môi trường và nguyên liệu kiểm tra (test environment & materials)
  - Phương pháp luận kiểm tra (test methodology) sẽ áp dụng
    - Mô tả chi tiết về hệ thống/tác vụ/thuộc tính sẽ được kiểm tra
  - Lượng giá kiểm tra (test measures)
    - Dữ liệu kiểm tra cần được thu thập và phân tích



## 2. Chọn thành viên tham gia

- Đại diện từ phía người dùng
  - Người dùng phải tham dự vào quá trình đánh giá sản phẩm
  - Các tác vụ liên quan, kiến thức chuyên biệt về nghiệp vụ
- Chọn các đối tượng có liên quan hoặc gần liên quan
  - Hệ thống dành cho bác sĩ
    - Chọn sinh viên y khoa
  - Hệ thống dành cho kỹ sư
    - Chọn sinh viên kỹ sư



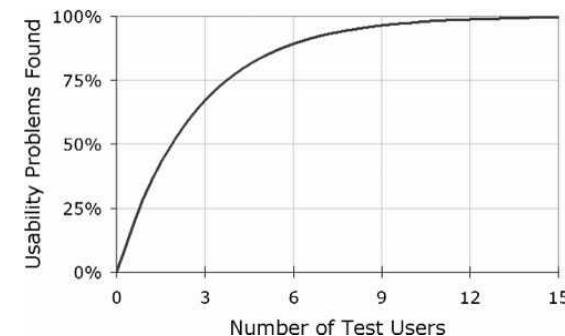
## 2. Chọn thành viên tham gia (tt)

- Chọn bao nhiêu người là đủ?
  - Quá nhiều người
    - Hao tổn chi phí – quan sát quá nhiều người dùng làm tốn thời gian và chi phí
  - Quá ít người
    - Biến động cao
    - Vấn đề về sự chủ quan của người dùng – Mỗi cá nhân có những vấn đề khác biệt
- Giải pháp cục bộ
  - Chọn số người phù hợp
    - Vấn đề lớn thường được phát hiện bởi 1 nhóm người
    - Các vấn đề nhỏ và cần sự đo lường chính xác thì cần nhiều người dùng tham gia



## 2. Chọn thành viên tham gia (tt)

- Bao nhiêu người là đủ?
  - J. Nielsen chỉ ra rằng các kiểm tra thực hiện với 5 người dùng sẽ phát hiện ra ít nhất 80% vấn đề về tính khả dụng
  - 3 test với 5 người thì tốt hơn 1 test với 15 người





## Xem xét khía cạnh đạo đức Ethical considerations

- Người dùng là cũng là con người!
- Testing là trải nghiệm mệt mỏi
  - Áp lực thực hiện → lỗi chắc chắn xảy ra
  - Cảm giác ko đầy đủ, thỏa đáng
- Bạn, người tổ chức việc test, cho nhiệm vụ làm giảm bớt sự căng thẳng đó:
  - Cần có sự đồng thuận thực hiện từ tester
  - Giúp tester cảm thấy thoải mái, tránh áp lực
  - Để họ biết, họ có thể ngừng việc test nếu muốn
  - Nhấn mạnh rằng chúng ta đang test hệ thống, ko phải đang test tester
  - Thu thập dữ liệu một cách ẩn danh nếu có thể

■ Luật thường yêu cầu sự đồng thuận từ phía tester

33

## Xem xét khía cạnh đạo đức (tt)

- Mục tiêu của việc test, là để xác định các vấn đề trong tính khả dụng, ko phải để kiểm tra năng lực của người dùng

5/24/16



34

## 3. Chọn các mục cần kiểm tra

- Định nghĩa mục tiêu chính xác, rõ ràng của mỗi test
  - Việc có mục tiêu và hướng dẫn test rõ ràng làm cho việc đo lường trên kết quả tester cũng dễ dàng hơn
- Mục tiêu test cần phản ánh các tác vụ thực tế
- Tránh “uốn cong” tác vụ theo hướng sao cho nó phù hợp nhất với thiết kế của bạn
- Đừng chọn tác vụ quá phân mảnh
- Một lần nữa, mục tiêu là đánh giá phần mềm, ko phải đánh giá người dùng!

5/24/16



35

## 4. Tiến hành test

- Các phương pháp đánh giá
  - Thanh tra (inspection)
  - Quan sát trực tiếp (direct observation)
  - Các kĩ thuật truy vấn (query techniques)
    - Phỏng vấn/bảng câu hỏi/bảng khảo sát

5/24/16



36

## Thanh tra (inspection)

- Áp dụng trên hệ thống hoặc trên prototype
  - does the system “feel right”?
- Lợi ích
  - Phát hiện ra các vấn đề lớn, ngay từ các phiên bản đầu
  - Hỗ trợ kỹ thuật walkthrough mà lấy tác vụ làm trung tâm
- Vấn đề
  - Ko đáng tin cậy vì hoàn toàn chủ quan

5/24/16



37

## Quan sát trực tiếp (direct observation)

- Người đánh giá quan sát và ghi nhận cách người dùng tương tác với hệ thống
- Trong môi trường thực nghiệm tính khả dụng
  - Người dùng hoàn thành tập tác vụ định trước
- Trong môi trường thực tế
  - Người dùng đi qua các nhiệm vụ cơ bản
- Lợi ích
  - Cực tốt khi muốn xác định vấn đề rõ ràng (gross) về giao diện/thiết kế excellent
  - Giá trị (validity) phụ thuộc vào cách tình huống được kiểm soát thế nào



5/24/16



38

## Kỹ thuật truy vấn (query techniques)

- Người đánh giá hỏi ý kiến của người dùng về hệ thống
- Phương pháp thông dụng nhất
  - Phỏng vấn
  - Bảng câu hỏi/Bảng khảo sát



"We also have some part-time positions available for people who only want to work 60 or 80 hours a week."

5/24/16



39

## Kỹ thuật truy vấn: Interview

- Người đánh giá lần lượt hỏi người dùng
  - Tập câu hỏi chuẩn bị trước
    - Có cấu trúc, ko cấu trúc hoặc kết hợp
  - Ko chính qui, chủ quan, chi phí thấp
- Tốt để tìm ra các vấn đề riêng biệt đối với từng người dùng
- Linh động vì người đánh giá có thể hỏi sâu vào các vấn đề quan tâm được phát sinh

5/24/16



40

## Interview (tt)

- Chuẩn bị tập các câu hỏi trọng yếu
  - Vài câu hỏi giúp khởi đầu thuận lợi buổi interview
- Có thể dựa trên kết quả quan sát người dùng
  - Phỏng vấn hậu quan sát (post-observation interview)
- Từ phản hồi của người dùng, ta sẽ có những câu hỏi tiếp theo đó
  - Tránh để người dùng dẫn dắt buổi test



## Interview (tt)

- Post-observation Interview
  - Có người dùng thực hiện test có quan sát
  - Tạo bản ghi video của test
  - Phỏng vấn người dùng với bản ghi video đó
    - Người dùng cho nhận xét về các sự kiện đã xảy ra trong quá trình sử dụng hệ thống
      - Chẳng hạn, tôi đã dùng button này bởi vì...
    - Người dùng thường đưa ra đề nghị cụ thể
      - Chẳng hạn, nên làm button này nhỏ lại...



## Interview (tt)

- Ưu điểm
  - Có thể thay đổi để phù hợp ngữ cảnh
  - Phát hiện vấn đề một cách đầy đủ hơn
  - Làm lộ ra góc nhìn/quan điểm của người dùng và tìm ra các vấn đề ko tiên đoán trước được
- Nhược điểm
  - Cực kì chủ quan
  - Tốn thời gian



## Kỹ thuật truy vấn: Survey

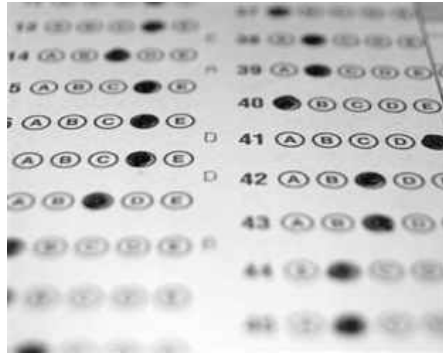
- Chuẩn bị tập cố định các câu hỏi cho người dùng
  - Cần phải thiết kế cẩn thận
- Tốn chi phí chuẩn bị, nhưng chi phí quản trị thấp administration
  - Sử dụng các phương pháp ít tốn chi phí như phone, mail, email, online form
  - Ko cần sự xuất hiện của người đánh giá
- Dữ liệu có thể định lượng được



## Survey (tt)

### ■ Các dạng câu hỏi khảo sát

- Open-ended
- Rating scalar
- Multi-choice
- Ranked



## Survey (tt)

### ■ Câu hỏi đóng-mở (open-ended questions)

- Xác định quan điểm tự nhiên
  - Bạn có thể đề nghị giúp vài cải tiến cho hệ thống này?
- Đùng giới hạn câu trả lời và kì vọng của người đánh giá
- Tốt cho thông tin chủ quan chung
  - Nhưng khó phân tích chặt chẽ

## Survey (tt)

### ■ Rating scalar questions

- Yêu cầu người dùng đánh giá 1 phát biểu cụ thể dựa vào thang đo số  
“Hệ thống dễ dàng phục hồi sau sự cố”  
Ko đồng ý 1 2 3 4 5 Đồng ý
- Thang đo thường tương ứng với mức độ đồng ý hay ko đồng ý với phát biểu
  - Thường chọn thang đo 5-7 mức
- Dễ phân tích nhưng cần thiết kế tỉ mỉ

## Survey (tt)

### ■ Câu hỏi nhiều lựa chọn (multi-choice questions)

- Người trả lời sẽ đưa ra các chọn lựa từ tập xác định  
“Có cách rõ ràng nào để quay về điểm bắt đầu hay menu chính ko?”  
A. Có, rất rõ ràng  
B. Có thể có, nhưng khó xác định  
C. Không, tôi ko thấy

## Survey (tt)

### ■ Ranked questions

- Người trả lời sắp xếp thứ tự các item trong danh sách
- Hữu dụng để khai thác sở thích người dùng
- Thuộc dạng chọn lựa bắt buộc (forced choice)

Hãy sắp xếp mức độ hữu dụng của các phương pháp nhập liệu text sau trên thiết bị di động

(1.hữu dụng nhất, 2. hữu dụng tiếp theo, ..., 0 ko hữu dụng)

\_\_ keypad

\_\_ bàn phím QWERTY

\_\_ bàn phím ảo trên màn hình có thể touch



## Survey (tt)

### ■ Ưu điểm

- Nhanh và áp dụng được cho nhóm đông người dùng
- Ko cần sự có mặt của người đánh giá
  - Có thể phân tích và quản trị một cách tự động
- Có thể phân tích một cách chặt chẽ
  - Kết quả định lượng được

### ■ Nhược điểm

- Chuẩn bị tốn kém
- Chất lượng giới hạn của phản hồi

### ■ Phụ thuộc vào việc ta có chuẩn bị câu hỏi đúng ko



## Nội dung

### ■ Tổng quan về đánh giá thiết kế giao diện

### ■ Các phương pháp đánh giá

- Chuyên gia
  - Đánh giá heuristic
- Người dùng
  - Field studies
  - Formative evaluation
  - Controlled experiment



## Formative evaluation

### ■ Các bước cơ bản: tương tự field studies

### ■ Khác biệt

- Người dùng được gán vào các tác vụ cụ thể chứ ko chọn lựa ngẫu nhiên
- Người đánh giá sẽ
  - Chọn thành viên tham gia
  - Gán tác vụ cho mỗi thành viên
  - Xem quá trình thực hiện tác vụ của các thành viên

### ■ Áp dụng kĩ thuật “think aloud”



## Think aloud

- Mục tiêu: ko chỉ biết được người dùng đang làm gì, mà còn biết được người dùng đang nghĩ gì
- Đầu tiên, yêu cầu người dùng nói lên suy nghĩ của mình khi đang thực hiện tác vụ. Chẳng hạn:
  - Họ đang cố làm gì
  - Sau họ chọn hành động này
  - Cách họ hiểu hệ thống đang làm gì
- Yêu cầu người dùng tiếp tục nói trong khi thực hiện
  - – “tell me what you are thinking”
- Ghi nhận lại suy nghĩ người dùng
  - Note, audio, video

5/24/16



53

## Think aloud (tt)

- Đây là cách đánh giá được sử dụng rộng rãi trong giới công nghiệp
- Ưu điểm
  - Đơn giản – yêu cầu sự giám định ít (little expertise)
  - Thấu hiểu suy nghĩ người dùng
  - Cho thấy hệ thống được sử dụng thế nào
- Nhược điểm
  - Chủ quan
  - Chọn lọc
  - Việc vừa làm vừa mô tả có thể làm giảm hiệu năng thực hiện tác vụ (task performance)

5/24/16



54

## Ví dụ

- <http://www.youtube.com/watch?v=Aa-svs5mQD8&feature=related>

5/24/16



55

## Nội dung

- Tổng quan về đánh giá thiết kế giao diện
- Các phương pháp đánh giá
  - Chuyên gia
    - Đánh giá heuristic
  - Người dùng
    - Field studies
    - Formative evaluation
    - Controlled experiment



5/24/16



56

## Controlled experiment

- Phương pháp khoa học truyền thống được ưa thích
- Kết quả thuyết phục rõ ràng về vấn đề cụ thể
- Benchmark tests
- Giúp so sánh hệ thống, tinh chỉnh các chi tiết



## Controlled experiment (tt)

- Giải quyết các vấn đề thực tế, sử dụng phương pháp khoa học và framework lí thuyết
- Theory-driven, hypothesis testing
  - Nghiên cứu các mối liên hệ bằng cách sử dụng các biến độc lập (**independent variables**)
  - Quan sát hiệu ứng trên 1 hoặc nhiều biến phụ thuộc (**dependent variables**)
    - Thay đổi các biến độc lập → quan sát và đo lường các biến phụ thuộc



## Controlled experiment (tt)

- Qui trình thực hiện
  - Chuẩn bị đề xuất
  - Chuẩn bị môi trường
  - Chọn thành viên tham gia
  - Xác định các dạng, các thuộc tính cần dùng
  - Xác định chiến lược đánh giá



## 1. Chuẩn bị đề xuất

- Test proposal bao gồm
  - Mục đích kiểm tra (test objective)
  - Các thành viên tham gia kiểm tra (test participants)
  - Môi trường và nguyên liệu kiểm tra (test environment & materials)
  - Phương pháp luận kiểm tra (test methodology) sẽ áp dụng
    - Mô tả chi tiết về hệ thống/tác vụ/thuộc tính sẽ được kiểm tra
  - Lượng giá kiểm tra (test measures)
    - Dữ liệu kiểm tra cần được thu thập và phân tích



## 2. Chuẩn bị môi trường

- Môi trường kiểm tra tính khả dụng
- Trang thiết bị
  - Máy tính
  - Camera
  - Truy cập internet



- Video:

<http://www.youtube.com/watch?v=XtXYchttqQ4&feature=related>

## 3. Chọn thành viên tham gia

- Thành viên tham gia phải được chọn lựa để đại diện cho cộng đồng người dùng mong đợi, xem xét các khía cạnh sau:
  - background, kinh nghiệm với tác vụ, trình độ học vấn...
- Người tham gia tình nguyện và được thông báo rõ ràng

## 4. Xác định các dạng thức và thuộc tính

- Các dạng khác nhau
  - Paper mockups
  - Special usability testing
  - Remote usability testing
  - ...
- Quay video quá trình thực hiện tác vụ của người tham gia

## 5. Chiến lược đánh giá

- Phương pháp khoa học áp dụng vào HCI bao gồm các bước chính sau:
  - A. state lucid and testable hypothesis
    - Đưa ra giả thuyết rõ ràng, có thể test được
  - B. identify independent variables that are to be manipulated
    - Xác định các biến độc lập
  - C. choose dependent variables that will be measured
    - Xác định các biến phụ thuộc để đo lường
  - D. select subjects and assign subjects to groups
    - Chọn chủ thể và gán chủ thể vào nhóm



## 5. Chiến lược đánh giá (tt)

- Phương pháp khoa học áp dụng vào HCI bao gồm các bước chính sau:
  - E. control for biasing factors
    - Điều kiện các nhân tố thiên lệch
  - F. apply statistical methods to data analysis
    - Sử dụng các phương pháp thống kê để phân tích dữ liệu
  - G. interpret experimental results
    - Diễn đạt kết quả thực nghiệm

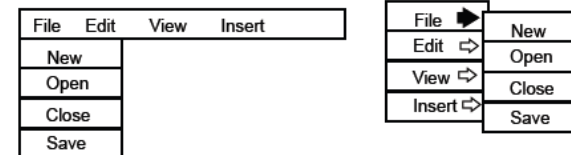


## 5. Chiến lược đánh giá (tt)

### ■ Giả thuyết (hypothesis)

- Ví dụ, test menu

“Ko có sự khác biệt về hiệu năng người dùng (xét trên khía cạnh thời gian và lỗi) khi chọn 1 item đơn từ 1 pop-up hoặc chọn 1 pull-down menu có 4 item, bất luận chủ thể trước đó đã có kinh nghiệm sử dụng chuột hay sử dụng các kiểu menu khác”



## 5. Chiến lược đánh giá (tt)

### ■ Các biến độc lập (independent variables)

- Những thứ ko phụ thuộc vào hành vi của người dùng
- Chẳng hạn, trong ví dụ test menu ở trên:

“Ko có sự khác biệt về hiệu năng người dùng (xét trên khía cạnh thời gian và lỗi) khi chọn 1 item đơn từ 1 pop-up hoặc chọn 1 pull-down menu có 4 item, bất luận chủ thể trước đó đã có kinh nghiệm sử dụng chuột hay sử dụng các kiểu menu khác”

#### **Biến độc lập:**

- Loại menu: pop-up và full-down
- Loại chủ thể: chuyên gia và người mới sử dụng



## 5. Chiến lược đánh giá (tt)

### ■ Biến phụ thuộc

- Là các biến phụ thuộc vào hành vi (behavior) hay phản ứng (reaction) của người dùng đối với các biến độc lập
- Những thứ cụ thể mà bạn thiết lập để đo lường/quan sát
- Chẳng hạn, trong ví dụ test menu ở trên:

#### **Biến phụ thuộc:**

- Tốc độ: thời gian chọn 1 item
- Tỉ lệ lỗi: số lỗi gây ra
- Dễ học: thời gian để sử dụng menu thành thực



## 5. Chiến lược đánh giá (tt)

### ■ Chủ thể (subjects)

- Những người thực hiện tác vụ
- Hãy cẩn thận trong việc chọn lựa chủ thể và gán chủ thể vào các nhóm



## 5. Chiến lược đánh giá (tt)

### ■ Phân tích dữ liệu thống kê (statistical data analysis)

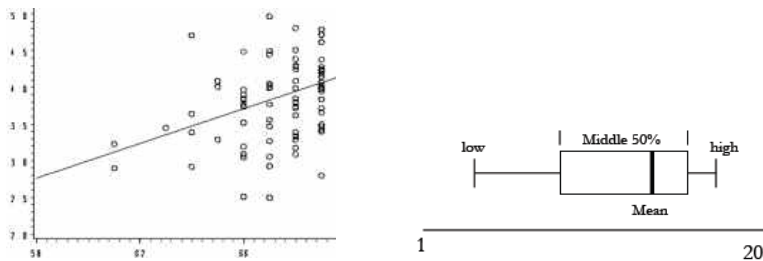
- Áp dụng các phương pháp thống kê
- Phân tích đơn giản
  - Xác định mean và median
  - So sánh giá trị mean và median giữa các phương pháp với nhau
  - Sử dụng các phương pháp kiểm tra giả thuyết (chẳng hạn, student t-test, ANOVA)



## 5. Chiến lược đánh giá (tt)

### ■ Phân tích dữ liệu thống kê

- Trình bày dưới dạng biểu đồ, bảng biểu hoặc đồ thị



## 5. Chiến lược đánh giá (tt)

### ■ Thể hiện kết quả thực hiện (interpreting experimental results)

- Bạn nghĩ rằng kết quả đưa ra ý nghĩa gì
- Hàm ý của chúng đối với thiết kế của bạn
- Làm sao để tổng quát hoá kết quả
- Các giới hạn và phê bình



## Tổng kết

- Tổng quan qui trình đánh giá giao diện
  - Thực hiện xuyên suốt trong quá trình thiết kế giao diện để đảm bảo tính khả dụng
- Qui trình field studies
  - Chuẩn bị đề xuất kiểm tra
  - Chọn thành viên tham gia
  - Chọn thuộc tính sẽ kiểm tra
  - Thực hiện kiểm tra. Kỹ thuật:
    - Thanh tra (inspection)
    - Quan sát trực tiếp (direct observation)
    - Kỹ thuật truy vấn (Query techniques)
  - Đo đạc các kết quả kiểm tra

## Tổng kết (tt)

- Formative evaluation
  - Kỹ thuật “think aloud”
- Controlled experiment
  - Các bước chính trong controlled experiments
    - Đề ra giả thuyết
    - Xác định các biến độc lập, các biến phụ thuộc
    - Chọn lựa chủ thể
    - Tiến hành phân tích thống kê
    - Trình bày kết quả