USABILITY TESTING

Sherry Pagoto, PhD
University of Massachusetts Medical School

Kristin Schneider, PhD
Rosalind Franklin University
Usability is broadly defined as how easy it is to use a system, how efficiently a person can use the system once they learn it, and how pleasant it is to use.

Poor usability will lead to low usage, poor adherence, no matter how effective the app might be.
1. Paper prototypes
2. Heuristic evaluation
3. App development
4. Team alpha testing
5. Usability sessions (iterative)
6. Focus groups
7. Pre-pilot test / Small n studies
Create rough, even hand-sketched, drawings of an interface to use as prototypes, or models, of a design.

Then, observing a user undertaking a task using paper prototypes enables the testing of design ideas at an extremely low cost and before any coding has been done.
Paper prototyping of an iPhone application (image credit – Golem.de)
1. **Visibility of system status.** The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

2. **Match between system and the real world.** The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

3. **User control and freedom.** Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

4. **Consistency and standards.** Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform (e.g., iPhone) conventions.

5. **Error prevention.** Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

Jakob Nielsen, 10 usability heuristics for user interface design
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. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

7. **Flexibility and efficiency of use.** Accelerators -- 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.

8. **Aesthetic and minimalist design.** Dialogues should not contain information which is irrelevant or rarely needed. 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 codes), precisely indicate the problem, and constructively suggest a solution.

10. **Help and documentation.** Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.
App development

Programming phase in which the app is coded based on the preceding steps.
Team alpha testing

Each member of the team tests app, making notes on suggested changes.

You want to catch bugs, typos, wording issues, kinks and other snags before testing in target population.
Usability Testing: Steps

- Define test objectives
- Determine criteria for test participants
- Task selection: map onto test objectives
- Create & describe task scenarios
- Select measures & procedures for recording
- Prepare test materials & environment
- Prepare protocol for usability session
- Run participants; analyze data
Define Test Objectives

- Formative: improve the interface as part of iterative design process
  - What can be done to improve design?
  - Typically think-aloud protocol
    - Ask participants to use the app while continuously thinking out loud — that is, simply verbalizing their thoughts as they move through the user interface.

- Summative: assess overall quality of interface
  - Objectives may be more vague to get a sense for how user will interact with the interface with little guidance

- Create plan for achieving objectives
Selecting test tasks

- Formative: as representative as possible of app uses
- Example: physical activity app
  - “Assume you went for a 30 minute walk, go through the procedure for entering that activity into the app.”
  - “Assume that you are having difficulty in following an exercise routine. The app is designed to help you with this problem and suggest some potential solutions. We want you to use the app first with this problem in mind.”
Selecting test tasks

- Summative: user-defined tasks
- Example: physical activity app
  - “How would you use the app to help you increase your physical activity”
Pilot tests

- Prior to usability testing: pilot test procedure tasks
  - Ensure testing protocol maps onto test objectives
  - Ensure instructions are comprehensible
  - Clarify definitions of user error, task completion
  - Rough estimate of usability session length

- Not necessary to use population of interest
  - May only require 1 or 2 participants depending on complexity of interface and tasks
Aim: Find the most design flaws with fewest participants

User criteria

- Ideal to define based on target population
- Example: physical activity app developed for older adults (age ≥ 65 years)
  - recruit diverse group: males, females, age range 65+

Sample size: Nielsen & Landauer (1993)

- Usability problems found (i) = N (1-(1-λ)^i)
  - i=# of test users
  - N=Total # of usability problems
  - λ=Probability for finding any single problem with any test user

- N and λ will vary by project; estimate as you collect data
- Use similar projects as a guide for estimating N and λ
- Ultimately may be constrained by costs
Prior to Conducting Usability Testing

- Select experimenter who will lead participants through sessions
  - Require knowledge of the test tasks, objectives & interface
- Good to require programmer presence at session or nearby: can address system crashes and other large problems as they occur
- Have designer present to observe, but not direct the participant
  - Improve understanding of user issues
- Ensure space and equipment is sufficient for testing
  - Video recorder
  - Any special space or equipment
    - Physical activity app: may need space for them to move around or use physical activity equipment.
Usability Testing Sessions: Introduction

- Introduce purpose
  - Prime participants that problems will arise
  - Experimenter will provide little guidance
  - If think-aloud protocol, encourage them to talk through the tasks
- Questionnaires
  - Describe sample; familiarity with apps
- Basic training on use of interface may or may not be required for participant
  - If you intend to recruit naïve mobile app users, they might need basic education on how to navigate app
During Usability Testing

- Begin with easy tasks so that user experiences “success”
- Go through tasks one at a time
- Do not let on if a participant is slow or making mistakes
- Balance between allowing participants to fail and preventing frustration
  - May need to provide ‘hints’
- If think-aloud protocol: may need to repeatedly cue ‘thinking out loud’
- Debrief
  - Any last comments?
  - Post-testing questionnaires
Usability Possible Outcomes

- Time user takes to complete a task
- # of tasks completed within a given time limit
- Ratio of successful interactions and errors
- Time spent recovering from errors
- # of user errors
- # of immediately subsequent erroneous actions
- # of features that were utilized by the user
- # of features that were never used by the user
- # of system features the used remembers using during debriefing
- Proportion of user statements that were positive vs critical toward the system
- # of times the user expresses clear frustration
- # of times the user used an inefficient strategy (as opposed to efficient strategy)
- # of times user gets sidetracked while completing a task.
Survey: Technology Acceptance Model

**Barriers**
- Difficulties in finding and installing the app has a negative impact on my usage.
- Difficult configuration of the app has a negative impact on my usage.
- Poor performance of the app has a negative impact on my usage.

**Behavioral Control**
- I can use the app without help from others.
- I have the means and resources to use the app.
- I have the knowledge and skills to use the app.

Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989;13(3):319-40.
Perceived Enjoyment
_____I think it is fun to use the app.
_____The app brings enjoyment.
_____I use the app to kill time.

Perceived Usefulness
_____The app is useful in my weight management efforts.
_____The app improves my efficiency.
_____Using the app saves time.

Intention to Use
_____If it were possible, I would use the app over the next two months.
_____If it were possible, I would use the app over the next year.
Iterative testing is necessary

Imperative to test in target population

Consider that usability issues may vary across age, gender, technological experience, socioeconomic status, etc.
Behavioral strategies in weight loss apps are narrow.

Patients who struggle with self-monitoring will get little to nothing out of a self-monitoring app.

In behavioral counseling when people aren’t adherent to self-monitoring, we do problem solving. In fact, problem solving is a key strategy reflected in every behavioral counseling visit. This is what counselors do!

Can we get an app to do this?
Smart Coach has an avatar-facilitated, idiographic problem solving feature that processes information intelligently to help patients identify solutions to their weight loss problems.

“Tech support” for weight loss
1. Need a database of weight loss problems and accompanying solutions
2. Need an algorithm of questions that narrows the problem down so that solutions can be tailored (e.g., “I hate exercise” is too broad, “I’m too tired to exercise after work” is better).
3. Each problem has to have multiple tailored solutions, not too few, not too many.
4. Solutions need to be useful, interesting, and specific.
5. Solutions need to be expert generated, however peer generated solutions may also be helpful.
6. Prompts are needed to remind the user to employ the solution at the right time.
Recruited members of the target population (30 obese adults of both sexes)

Had counselors conduct problem solving sessions using 5 step problem solving process (Nezu, Nezu, D’Zurilla, 2012)
1. Describe the problem
   ✓ What is the problem?
   ✓ Why does it happen?
   ✓ When does it happen/not happen?
   ✓ Where does it happen/not happen?
   ✓ With whom does it happen/not happen?
   ✓ What have you tried?
   ✓ What has worked/not worked?
2. Brainstorm – counselor will make a list of several possible solutions. If done in a group, the entire group can participate in generating solutions. Patient is asked not to weigh in on any solution, but to just listen.
3. **Pick a solution to try** – patient is asked to select ONE solution he/she is willing to try over the next week.

4. **Schedule the attempt** – The specific days and times the attempt will be made.

5. **Follow-up** – How did it go? If the solution was not helpful, the patient is asked to select another to try from the list. Purpose is to have patient always attempting new things until a solution is found that works.
Algorithms for the “Describe Problem” series of questions were developed separately for dietary and exercise problems, and designed around the most common characteristics of problems.
What is your biggest problem with exercise?

Don’t have time
Get bored
Have an injury
I find it hard to get started
Too tired
Exercise Algorithm

- What exercise are you currently doing?
- When do you do this exercise (separately for weekdays and weekends)?
- Where do you do this exercise (separately for weekdays and weekends)?
- How long do you do this exercise?
- Where do you prefer to exercise?
- When do you prefer to exercise?
- How do you feel about exercising at this time and at this place?
- Which exercises have you tried and liked?
- Which exercise have you tried and not liked?
- Do you have a gym membership?
- Do you have a physical therapist?
- Have you talked to a provider about your injury?
- Do you have childcare responsibilities?
## Solutions

Tailored to responses to the questions. Examples:

<table>
<thead>
<tr>
<th>Solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Get a referral for a physical therapist for advice on how to proceed with your injury.</td>
</tr>
<tr>
<td>Consult your primary care provider for further advice, or get a referral to a physical therapist.</td>
</tr>
<tr>
<td>Download an app such as MyFitnessPal to track your exercise.</td>
</tr>
<tr>
<td>Try to exercise at home when you are watching TV.</td>
</tr>
<tr>
<td>Start with 5 minutes of any exercise</td>
</tr>
<tr>
<td>Find some free exercise videos on Youtube</td>
</tr>
<tr>
<td>Reward yourself with money or a gift when you achieve a goal for your exercise.</td>
</tr>
<tr>
<td>Add the &lt;exercise&gt; schedule for the weekend activity plan.</td>
</tr>
<tr>
<td>Set a timer to get up and walk for 5 mins after sitting for 2 hours or more</td>
</tr>
<tr>
<td>Plan your exercise with a schedule. Write it in your calendar at the beginning of the week.</td>
</tr>
<tr>
<td>Play active games with your children.</td>
</tr>
<tr>
<td>Try to find some low-cost options for personal training, maybe through email or an app</td>
</tr>
<tr>
<td>Incorporate active play with your child such as dancing, active videogames, or nature hike</td>
</tr>
<tr>
<td>Get a babysitter to watch your child while you exercise.</td>
</tr>
</tbody>
</table>
Usability testing

4 sessions of 3 participants from the target population

“Think aloud” protocol

Technology Acceptance Model

In between sessions, the app was updated based on feedback.
My problem not captured in the algorithm

Not enough solutions, solutions aren’t novel enough

Would like more info re: solutions (too brief)

Would like to save solutions tried, and relevant info

Interface (sliding bar not obvious, no back button on certain pages)
PRACTICE
Members of target population asked to use the app for 1 week, then convene in a focus group to discuss experiences.

**Longer period allows feedback on** 1) whether they followed through with solutions, 2) better sense of the quality/quantity of solutions, 3) scheduling of solutions, 4) better sense of whether they would actually use the app over a period of time, 5) usability when they don’t have someone standing over them
We will enroll 10 people from target population to participate in a 6-week intervention.

6 in person groups along with the app. They will also be asked to use a tracking app (My Fitness Pal)

Focus group to follow to get feedback on the app after extended period of use and when accompanied by behavioral intervention (the intended use)
Next steps...

Fully powered randomized trial...

Does the tool have utility when used with a behavioral program?

When used alone?

Might it be an adjunct to apps that have other functions?