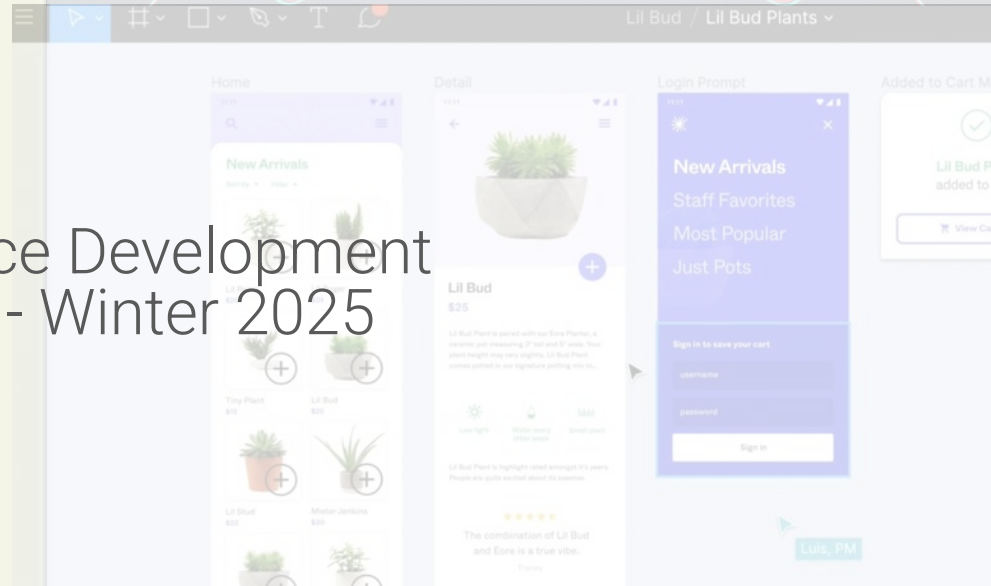


The DESIGN
of EVERYDAY
THINGS

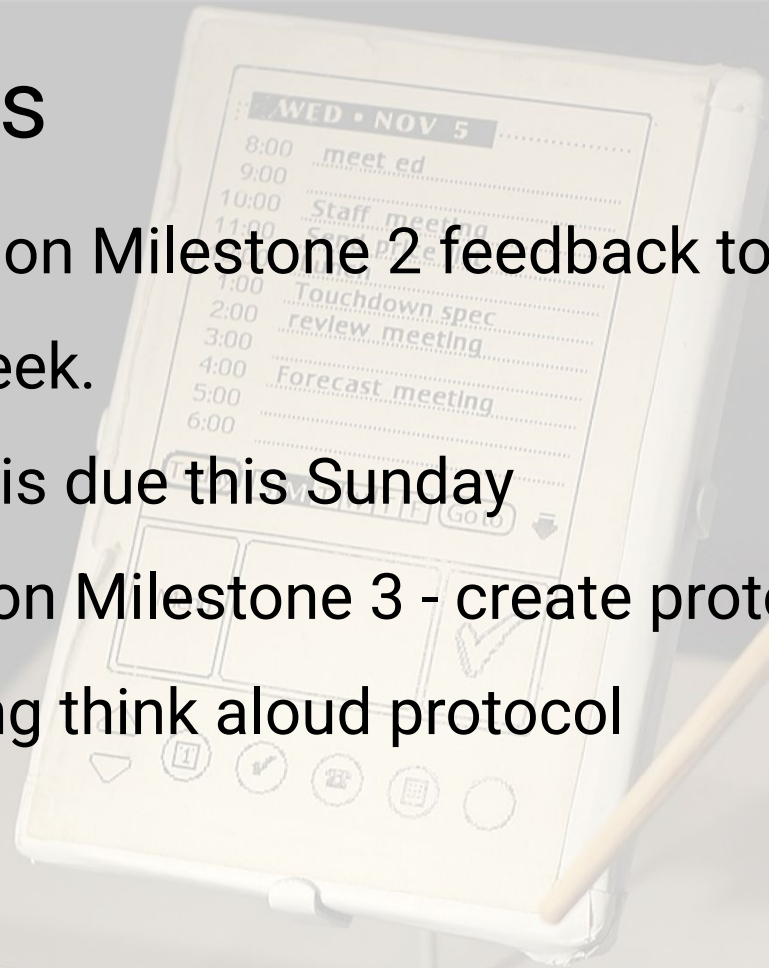
Think Aloud Protocol

User Interface Development
EECS 493 - Winter 2025



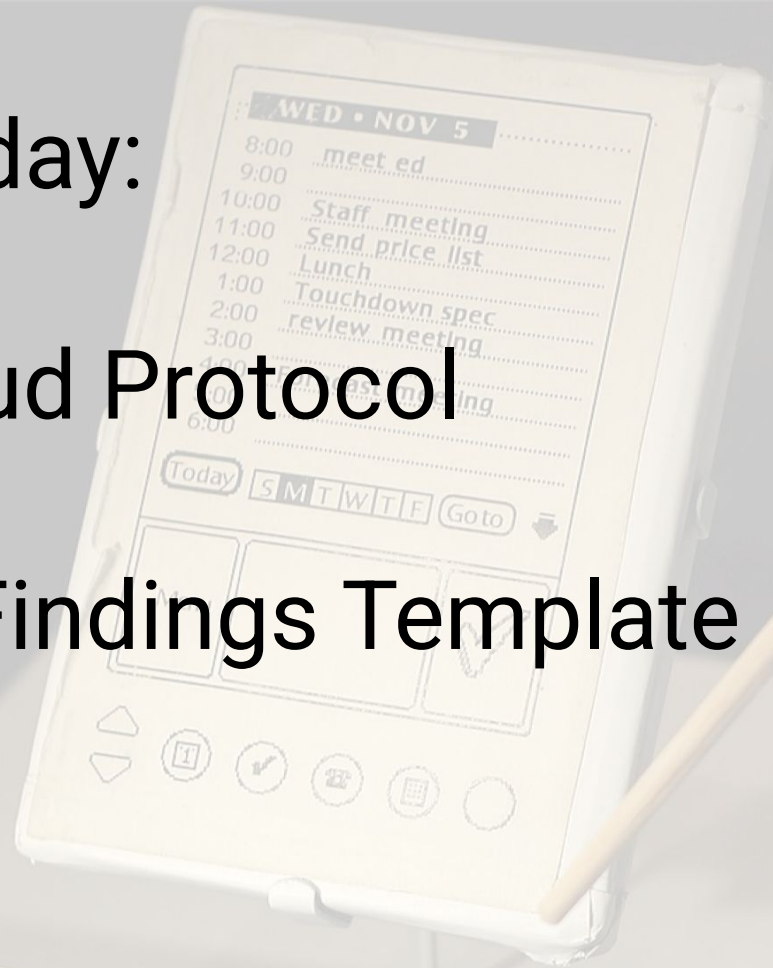
Current items

1. We're working on Milestone 2 feedback to be sent out by EOD this week.
2. Assignment 4 is due this Sunday
3. Start working on Milestone 3 - create prototype and run user tests using think aloud protocol



Goals for today:

- Think-Aloud Protocol
- Usability Findings Template

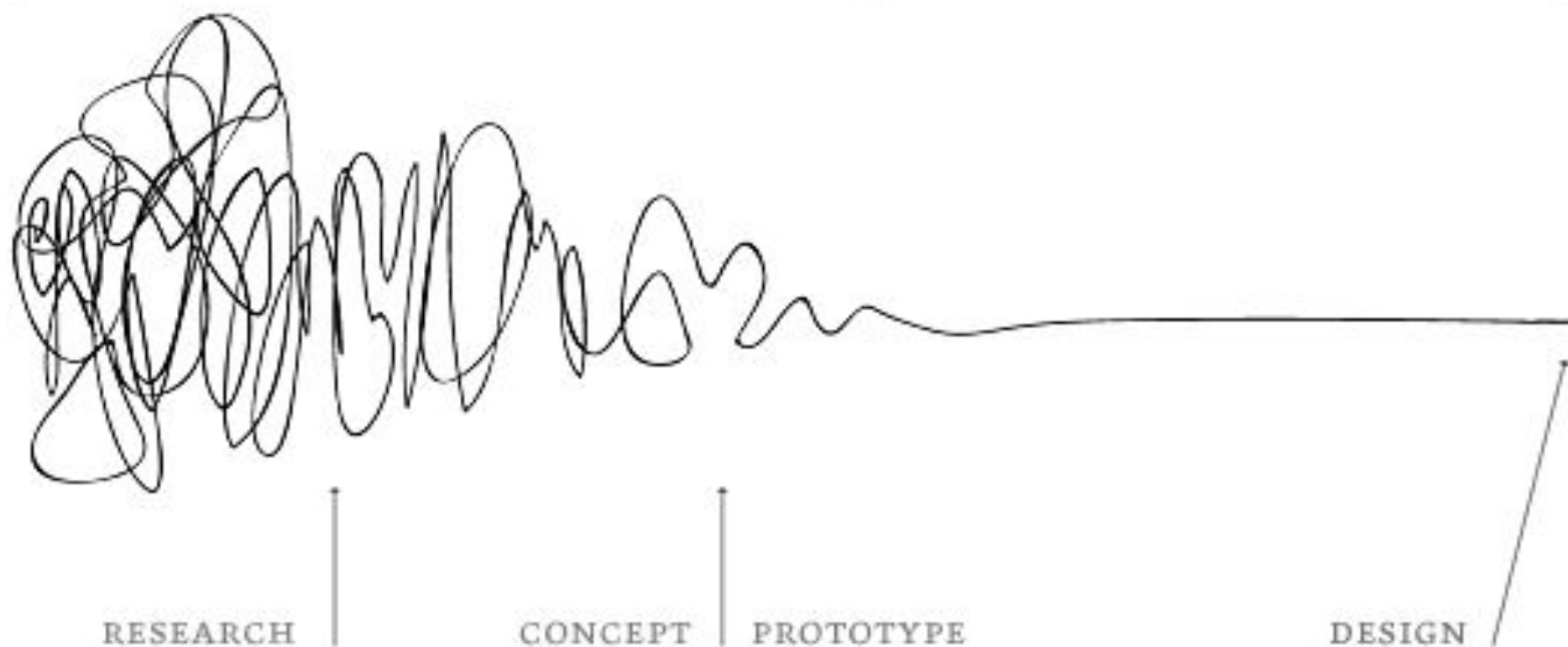


Milestone 3 -> 1st round of prototyping & evaluation

- From Milestone 2, your team have synthesized findings and converged on a design solution
- Design & implement a high-fidelity prototype using Figma
 - The amount of details should be at least 4x those of Assignment 4
 - You can use and build on external templates, but need to document clearly which come from you.
- Run 2 think-aloud sessions using the Figma prototype.
- For each think-aloud session, have the user do the key tasks.
- Summarize usability findings in UFT templates, need to have 6-8 UFTs
- Briefly summarize what you will do in the next milestone to improve your app.

UNCERTAINTY / PATTERNS / INSIGHTS

CLARITY / FOCUS



Heuristic Evaluation vs. User Testing

- Heuristic Evaluation is much faster than user testing
 - 1-2 hours each evaluator vs. days- weeks
- Heuristic Evaluation does not require interpreting users' actions
- User testing is more accurate and realistic (by definition)
 - Takes into account actual users and tasks
 - Heuristic Evaluation may miss problems
- Good to alternate between Heuristic Evaluation and user testing
 - Find different problems
 - Don't waste participants

Think-Aloud Protocol

- Empirical technique – most commonly used for assessing the usability of a prototype of an interface.
- Rooted on psychology research, based on think-aloud protocol analysis & critical incident analysis.
- **Ask the user think-aloud while performing tasks on a system.**
- You watch & learn how the user thinks about the task, what is their mental model, how do they approach the task and does the system provide affordance to accomplish it?



An example of a user thinking out loud
while navigating a website

<https://www.youtube.com/watch?v=g34tOmyKaMM>

Example Think-Aloud

Use Google maps to figure out how you would go from
Beyster to Trader Joe's by bus

Get into pairs

- One person acts as the interviewer
 - Use Google maps to figure out how you would go from Beyster to Trader Joe's by bus
 - Listens, don't need to ask additional questions.
- One person as interviewee- doing the think-aloud
 - Open Google maps, think-aloud, display your thought process while completing the task.

Why Think-Alouds?

- It helps uncover the users' thought processes while they're interacting with your software.
- It helps you identify the mismatches between your system design with the user's mental models.
- It helps you identify usability issues, e.g., some features are not discoverable, users are not clear what they are supposed to do on each page, users may accidentally make errors.

Use TAP Iteratively at Different Stages

Most often used to **assess the Usability** of higher fidelity prototype but can be used at all stages of iterative design:

- Paper (low-fidelity) prototypes
- Wireframes
- High-fi prototypes
- Pre-launch
- Post-launch
- Live
- All of the above



Typical Procedure

1. Design the study/ write the script
2. Recruit subject
3. Conduct the think-aloud
4. Analyze the results
5. Summarize and consolidate into report

Typical Procedure

1. **Design the study/ write the script**
2. Recruit subject
3. Conduct the think-aloud
4. Analyze the results
5. Summarize and consolidate into report

Design the study- Select tasks

- Should reflect what real tasks will be like
- Tasks shouldn't be too long or complex; Tasks shouldn't require background that the test user won't have
- Avoid choosing tasks in the direction of what your design best supports
 - You want to identify problems.
- Concrete task
 - Use Google maps to figure out how you would go from Beyster to Trader Joe's by bus
 - Use Google maps to find routes <- too general, bad task

Example think-aloud task

Student Answers

[View by Student](#) [View by Rubric](#) F [Mode Switch](#) [Previous](#) 1 / 67 [Next](#)

Student 1

Summarization of the article from Article Review

A

Segmented student example

The article, The American Exception, by Somini Sengupta focuses on comparing climate footprints of citizens of different countries, specifically addressing that wealthier Americans contribute the most to climate change. The lifestyle choices and habitats of Americans cause greater emissions per capita, resulting in an overwhelming generation of pollution.

Give an example that incorrectly identifies shortage as a market failure

B

Query for new examples with textual prompts

Submit

Cancel

(GPT) The article, The American Exception, by Somini Sengupta focuses on comparing climate footprints of citizens of different countries, specifically addressing that wealthier Americans contribute the most to climate change. The lifestyle choices and habitats of Americans cause greater emissions per capita, resulting in an overwhelming generation of pollution. This leads to a shortage of clean air, which is a clear example of market failure.

The housing market has seen a rapid increase in price, and it likely will soon see a corrective decrease in price. There is a **shortage of supply** in the housing market, even with the less demand as a result of increasing prices.

Market Failure Identified from Article Review

C

Generated example and retrieved student example through query

In this case, the market failure is climate footprint and pollution. This is a negative that the decisions made by American citizens negatively impact the environment and have a far greater impact than citizens in many other countries around the world, who all ultimately face the consequences.

Create Questions

Question 1 / 1 [+ Question](#) [Delete](#)

Question Stem:

Which of the following is a good example of a Market Failure?

Simplify

Formalize

Correct Answer:

The article, The American Exception, by Somini Sengupta focuses on comparing climate footprints of citizens of different countries, specifically addressing that wealthier Americans contribute the most to climate change. The lifestyle choices and habitats of Americans cause greater emissions per capita, resulting in an overwhelming generation of pollution.

Simplify

Formalize

Contextualize with the Question Stem

+ Feedback

Incorrect Option 1:

Simplify

Formalize

Contextualize with the Question Stem

E

Additional AI supports

Get Option Suggestion

+ Feedback

Incorrect Option 2:

Question List Name:

Type in name here

Save Question List

- Task 1: design a multiple-choice question using this interface
- Task 2: design a multiple-choice question using this interface with the question stem “which of the following is a good example of a market failure” -> more specific, yields better data

Typical Procedure

1. Design the study/ write the script
- 2. Recruit subject**
3. Conduct the think-aloud
4. Analyze the results
5. Summarize and consolidate into report

Typical Procedure

1. Design the study/ write the script
2. Recruit subject
- 3. Conduct the think-aloud**
4. Analyze the results
5. Summarize and consolidate into report

Conduct the think-aloud - Introduction

- Describe the purpose of the evaluation
 - “I’m testing the product; I’m not testing you”
- Tell participants they can quit at any time
- Get participants’ consent to record the session, **need to record screen.**

Conduct the think-aloud - Introduction (continued)

- Briefly introduce your app/product
- Explain how to think-aloud
 - <https://www.youtube.com/watch?v=g34tOmyKaMM>
 - They don't need to explain, just need to verbalize their thoughts.
- Explain that you will not provide help (unless hopelessly stuck)
- Describe the task
 - Give written instructions, one task at a time

Tips for doing think-alouds

- Prompt the user “Please keep talking”
- If there are questions, leave until the end, don’t interrupt the user
- Don’t ask the user to explain during the think-aloud, that’ll introduce extra cognitive load.

Things to look for during a think-aloud

- What happened?
- Users' strategies (i.e., undo, reset, etc.)
- Users' mental model
- Intentions and goals leading to incident
 - "Oh i was looking for xxx here"
- Challenges that users have encountered, errors
 - "It doesn't let me ..."
- User's expectations
- User's emotional state, statements of distress/confusion
 - "This is hard to find.."

Tips for doing think-alouds

- Don't interrupt with the user's flow
- Leave questions until the end
- If the user can't complete a task, wait before intervening, let them explore a bit
- Don't ask questions
 - "Can you explain why you did that?" -> NO
- The only prompt: "Please keep talking"

Debriefing (post think-aloud)

- After the user completes all your tasks
- **Ask questions**
 - **Ask for explanation of problems encountered**
 - Often users don't remember, so best to remind them, demonstrate or show video segments.
 - **Ask for more information about comments made during test or behaviors**
 - E.g., "I noticed that you used xx sometimes, and chose to use a different feature other times, can you explain why you made the choice?"
 - **Ask for comments on specific features**
 - E.g., can you share what you like or dislike about the feature xxx?

Which of the following is NOT correct about think-aloud protocols? *

- ☐ I will pick a concrete task that reflects what real users will do in this app.
- ☐ If I observed the user made an unexpected action during the think-aloud, I will ask the user to further elaborate why they made the decision
- ☐ When the user hasn't been talking for 2 minutes, I'll prompt the user by saying "please keep talking"
- ☐ I will look for user-made errors, confusions, unexpected actions when analyzing the think-aloud recording.

Typical Procedure

1. Design the study/ write the script
2. Recruit subject
3. Conduct the think-aloud
- 4. Analyze the results**
- 5. Summarize and consolidate into report**

View Behavior and Document them in UFTs - Usability Finding Template

Watch the think-aloud screen recording

1. As you view behavior & hear evidence that aspect of the interface causes a problem or enables a positive experience, start new UFT
2. Give it a unique UFT ID and if problem or good feature
3. Give it a short name
4. In the end, connect all UFTs and find relationships, if any.

	Critical	Severe	Major	Minor	Positive
Priority	1	2	3	4	5

#1 Concise but descriptive name

Page / feature where the problem or positive experience occurs

Screen Capture

Caption: Short sentence to summarize the usability finding - positive or problematic.

Findings

EVIDENCE: Present your findings and describe the evidence for that finding in detail. Address the following aspects when relevant to the finding:

- o **Frequency:** common or rare? How many users are likely to experience the problem? Why?
- o **Impact:** Easy or difficult for the user to overcome? Will the user be able to achieve goals given this issue? Why?
- o **Persistence:** Is it a one-time problem that people will learn how to overcome, or will they be repeatedly bothered by it? Why?

Recommendations

- **SOLUTION:** Suggest possible solutions to the problem or recommendations on how to move forward. You can list multiple suggestions.
- **TRADE OFF:** You must include trade-offs to be credible. If you can't think of any bad trade-off, note so. If this is a good aspect, trade-offs are still appropriate.

UFT - Findings (severity of the problem, 3 factors)

- **Frequency: how common this is going to be among users.**
 - E.g., if it's a login issue, everyone is going to experience it, high frequency; if it's an issue that'll only occur with a small group of user (e.g., a short-cut issue), less frequent
- **Impact: easy or difficulty for users to overcome? Will users be able to achieve goals with this issue?**
 - E.g., if submit button on Canvas is disabled, high-impact; for course registration, you have to remember the course number (not ideal), but users can overcome it, low-impact
- **Persistence: is this a one-time problem that people will know how to recover from it?**
 - E.g., the first time you login to Canvas, you don't know where to find the quiz, but after the first time you know it, low-persistence; If DUO (2-factor authentication service doesn't allow you to remember a device, you'll encounter this issue repeatedly, high-persistence;



Highlight & Comment

Essay

Table

Preview

▼ Expand All

Sort by: Rubric Order ▼

Assignment (531 words)

To Governor of California,

< After the rise in popularity of almond milk, we have seen a rise in alternative milk substitutes like oak and soy milk. > Alternative milk substitutes have started taking market share from almond milk, so the increase in popularity of alternative milk substitutes affects the markets for alternative milk substitutes, almond milk and water. Almond milk and alternative milks are substitutes in consumption, so when the price of one substitutes increases, the demand for the other increases. Before the government gets involved with the market, the demand for almond milk would increase because of consumer preferences, the demand for alternative milk would decrease because it's a substitute in consumption to almond milk and the demand for water would increase because water is an input for almond milk.

As the newly elected governor of California, you should support the almond farmers and their families and consider passing a law to institute a binding price floor in the almond milk market to support farmers. A price floor sets a minimum limit on the price in the almond milk market, but at this price, there will be a surplus of almond milk. This will result in you purchasing the resulting surplus of almond milk.

A price floor would impact the almond milk market because if the price is set above the equilibrium, prices would increase and the quantity demanded would decrease. Quantity supplied would also increase because prices increase. If the quantity supplied outweighs the quantity demanded, then we will have a surplus of almond milk and the governor will have to purchase the surplus. Because of the higher prices for almond milk, we will see a decrease in quantity demanded because consumers will pivot to cheaper alternative milks. The price floor will help almond farmers because it gives them a minimum price to work with, but price floor could lead to inefficiencies like deadweight loss from inefficiently low quantities or wasted resources, among others.

A price floor would impact alternative milk markets because the demand will increase. As the price of almond milk increases due to the price floor, consumers will want cheaper alternative milks. Since the demand increases for alternative milk markets, we can raise the prices, which is known as cross-price elasticity. With alternative milk markets becoming more popular, farmers who produce oats for oat milk or soybeans for soy milk will see an

Feedback☒ Know the difference bewteen [demand] and [quantity demanded].

Necessary Basic Economic Cx ▼ ▼

☒ Know the difference between [supply] and [quantity supplied].

Necessary Basic Economic Cx ▼ ▼

☒ Should not mention that the market equilibrium is always efficient

Necessary Basic Economic Cx ▼ ▼

☒ Know Supply and Demand represents producers (or sellers) and consumers (or buyers), respectively.

Necessary Basic Economic Cx ▼ ▼

☒ Show understanding of the Law of Demand, that is, as the price of the good increases, the quantity demanded of the good or service decreases.

Necessary Basic Economic Cx ▼ ▲

Historic Feedback

Not available

GPT Feedback

Add to feedback

You've effectively applied the Law of Demand to the almond milk market. This understanding is crucial for analyzing consumer behavior.

◀ 1/2 ▶

You've mentioned the water needs of almond farmers, but how does this compare to oat or soy farmers? Consider exploring this comparison.

menting box)

346/368 Graded

Essay

Table

Preview

Expand All

Sort by: Rubric Order

Expand All

Sort by: Highlight Order

Feedback

- ☒ Know the difference between [demand] and [quantity demanded].
Necessary Basic Economic Co
- ☒ Know the difference between [supply] and [quantity supplied].
Necessary Basic Economic Co
- ☒ Should not mention that the market equilibrium is always efficient.
Necessary Basic Economic Co
- ☒ Know Supply and Demand represents producers (or sellers) and consumers (or buyers), respectively.
Necessary Basic Economic Co
- ☒ Show understanding of the Law of Demand, that is, as the price of the good increases, the quantity demanded of the good or service decreases.
Necessary Basic Economic Co

Historic Feedback

Not available

GPT Feedback

Add to feedback

You've effectively applied the Law of Demand to the almond milk market. This understanding is crucial for analyzing consumer behavior.

< 1/2 >

You've mentioned the water needs of almond farmers, but how does this compare to oat or soy farmers? Consider exploring this comparison.

	Critical	Severe	Major	Minor	Positive
Priority	1	2	3	4	5

#1 Concise but descriptive name

Page / feature where the problem or positive experience occurs

Screen Capture

Caption: Short sentence to summarize the usability finding - positive or problematic.

Findings

EVIDENCE: Present your findings and describe the evidence for that finding in detail. Address the following aspects when relevant to the finding:

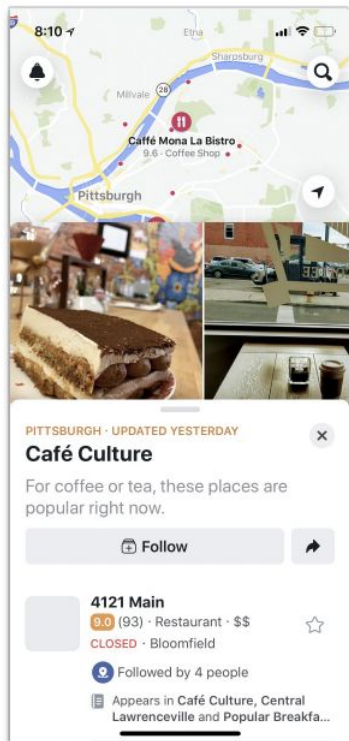
- o **Frequency:** common or rare? How many users are likely to experience the problem? Why?
- o **Impact:** Easy or difficult for the user to overcome? Will the user be able to achieve goals given this issue? Why?
- o **Persistence:** Is it a one-time problem that people will learn how to overcome, or will they be repeatedly bothered by it? Why?

Recommendations

- **SOLUTION:** Suggest possible solutions to the problem or recommendations on how to move forward. You can list multiple suggestions.
- **TRADE OFF:** You must include trade-offs to be credible. If you can't think of any bad trade-off, note so. If this is a good aspect, trade-offs are still appropriate.

#8 Lack of information on location on event page

Individual Event Page



During 0:57 - 1:02, the user was confused on what to do next after selecting an event that interested her

Findings

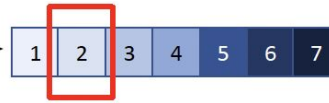
- The user was confused and unsure about the event because she didn't know where the event was at and for what purpose
- From the recommended list, she selected "Cafe Culture" as it sounded interesting, but immediately left after 5 seconds as she was unsure of where the event was
 - Frequency:** Users are likely to experience this issue as a lot of the event pages do not have a clear information hierarchy for each of the hosted events such as addresses, contact, etc.
 - Impact:** It may be difficult for users to overcome this issue because the only actions that a user can take from here is either to follow the event or share the event to other friends, which isn't solving their needs on figuring out what the address of the event is
 - Persistence:** Users may be repeatedly bothered by it as it doesn't display the address of the event, instead it uses Facebook's integrated map which doesn't clearly show the users on what next steps to take to get to that event. Another user in [Slide 7] also struggled with this issue.

Recommendations

- Solution:** Include the address of the publicly hosted event and the point of contact for that event in case of help
- Trade-offs:** Linking another Facebook user's profile as a point of contact may cause privacy concerns

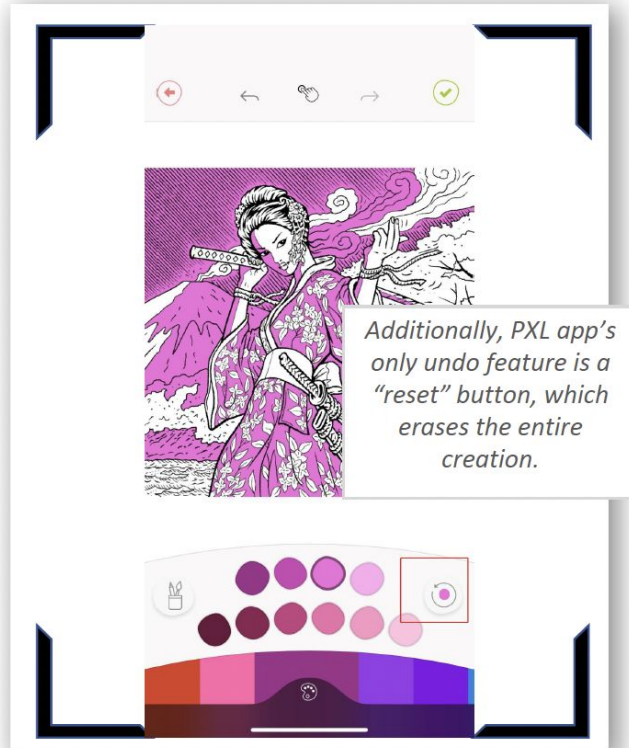
Based on UAR by Brad A. Myers & Bonnie John

Severity Rating →



#6 Undo icon's function is not obvious to user

Colorfy drawing toolbar once user start drawing



Findings

- **Evidence:** Users #4 & #5 couldn't find the undo button to erase their misplaced color when they wanted to eliminate the color of an area.
- **Frequency:** Both users had problems figuring out how to fix their mistakes because the erase icon on the right is not clearly read as an undo button.
- **Impact:** It is easy to overcome because once users know what the button does, they will be able to use it.
- **Persistence:** It is a one-time problem that users will learn to overcome. It will only appear the first time the user tries to undo one of their colors.

Recommendations

- **Solution:** Change the undo icon to a more commonly used icon, such as a left arrow or an eraser.
- **Trade-off:** Users who haven't used similar software, like photo editing apps, may not know what an eraser icon does.

UFT - Template

<https://docs.google.com/presentation/d/1TvXOGqCPUHwCEiV58HA64S4mpcZnTT1VORml9S6oR-Y/edit?usp=sharing>

Watch this think-aloud

Context:

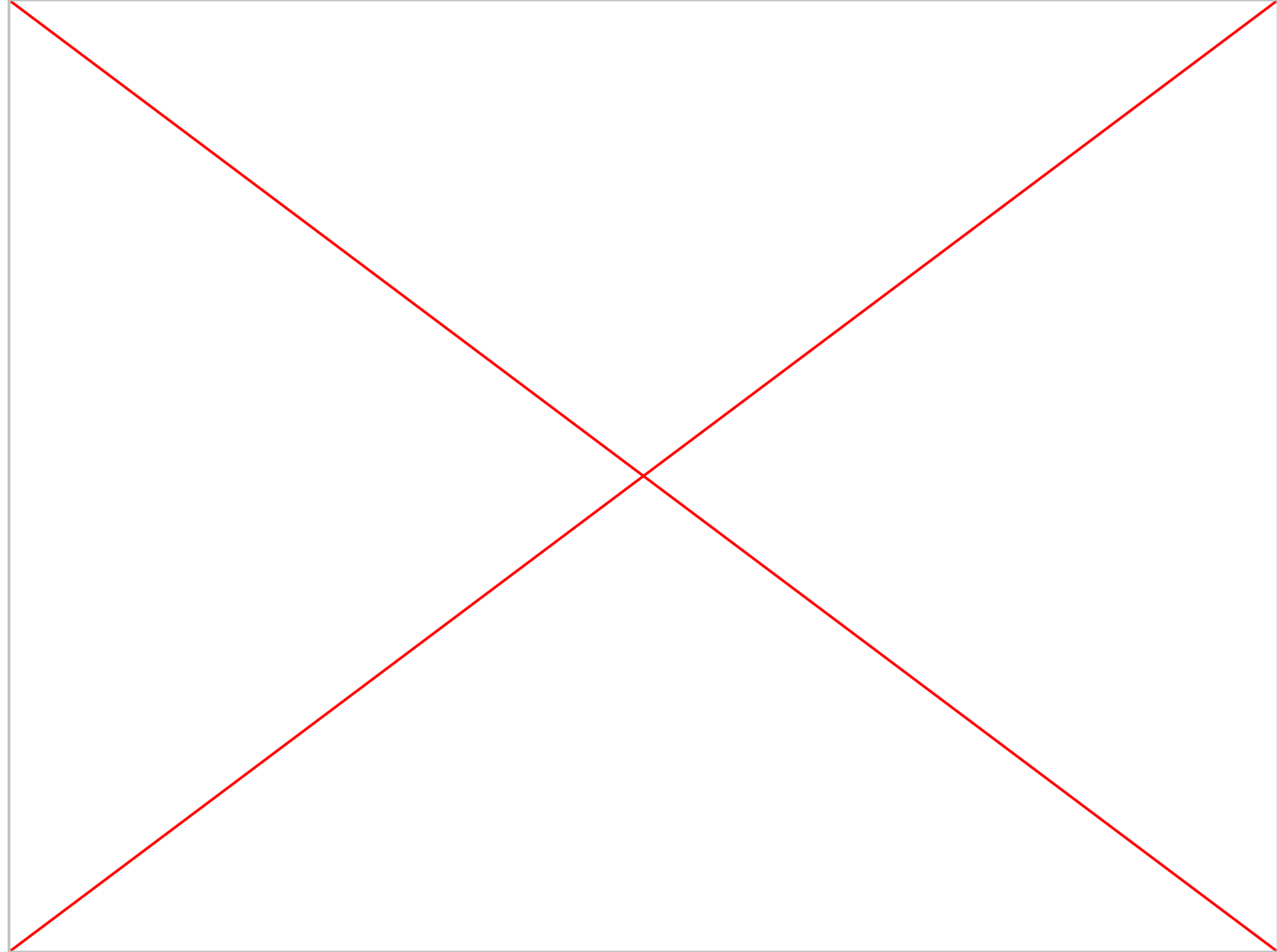
The designer is trying to improve an app helping people navigate public arts in Pittsburgh.

The designer does a think-aloud with the app.

There are three tasks

- 1) Task 1: looking at this piece of art (Liberty Avenue Musicians), what do you want to learn about it? [context: they're actually standing in front of the art]
- 2) Task 2: Find information about the art piece Magnolias in the app
- 3) Task 3: Look for other pieces of art you might visit in the app

Take notes while watching! [Video link](#)



Question 1

What did you notice the experimenter do in the think-aloud that you think were good practices?

Question 2

Get into pairs and think of two usability problems of the app.

	Critical	Severe	Major	Minor	Positive
Priority	1	2	3	4	5

#1 Concise but descriptive name

Page / feature where the problem or positive experience occurs

Screen Capture

Caption: Short sentence to summarize the usability finding - positive or problematic.

Findings

EVIDENCE: Present your findings and describe the evidence for that finding in detail. Address the following aspects when relevant to the finding:

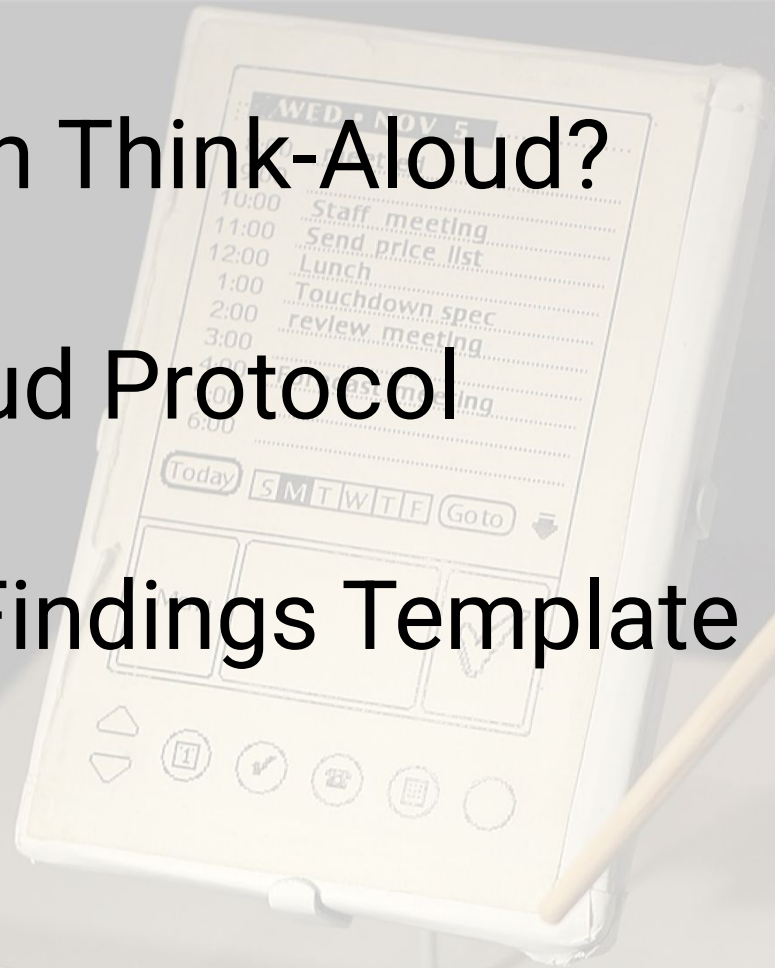
- o **Frequency:** common or rare? How many users are likely to experience the problem? Why?
- o **Impact:** Easy or difficult for the user to overcome? Will the user be able to achieve goals given this issue? Why?
- o **Persistence:** Is it a one-time problem that people will learn how to overcome, or will they be repeatedly bothered by it? Why?

Recommendations

- **SOLUTION:** Suggest possible solutions to the problem or recommendations on how to move forward. You can list multiple suggestions.
- **TRADE OFF:** You must include trade-offs to be credible. If you can't think of any bad trade-off, note so. If this is a good aspect, trade-offs are still appropriate.

Questions on Think-Aloud?

- Think-Aloud Protocol
- Usability Findings Template



PalmPilot wooden model

Milestone 3 -> 1st round of prototyping & evaluation

- From Milestone 2, your team have synthesized findings and converged on a design solution
- Design & implement a high-fidelity prototype using Figma
 - The amount of details should be at least 3x those of Assignment 4
 - You can use and build on external templates, but need to document clearly which come from you.
- Run 2 think-aloud sessions using the Figma prototype.
- For each think-aloud session, pick 4 tasks.
- Summarize usability findings in UFT templates, need to have 6-8 UFTs
- Briefly summarize what you will do in the next milestone to improve your app.

- The prototype should have a significant amount of work or details. Compared to Assignment 4, your prototype should have at least 4x the amount of work or details. This includes but not limited to the number of frames, features, functions, supported interactions, and more. If you have more than 4 members in your team, it should be proportionally more.
 - We will primarily evaluate this from the perspective of “tasks”/“critical user journeys”, i.e., a *sequence* of user interactions your prototype supports for users to accomplish one of their primary goals. For example, in Assignment 4, you designed a task of favoriting, accepting and completing quests (i.e., engage with and complete a quest). This task contains a series of *subtasks*, including browsing quests, viewing quest details, favoriting a quest, accepting a quest, completing a quest, receiving rewards, etc, where the user could perform one or multiple interactions to complete the subtasks.
 - For this milestone, your prototype should support **2 key tasks**, with **at least 8 subtasks** per task. **Subtasks need to be part of the task rather than just a combination of different features.**
 - Both tasks need to support the same goal (e.g., “improve resume”) and the goal should not be too general (e.g., “find a job”, since there will be many directions you can go under this goal). If your task has less than 8 subtasks, it may indicate the task is too specific (e.g., “favorite the quest”, “accept the quest”).
 - The task should be key goals users need to accomplish using your prototype, such as 1) sign up for a career fair event, find it on your calendar, and cancel it, or 2) carry out a consultation session with an AI agent to do a mock interview and provide feedback (with many conversation turns to concretely achieve something). In the context of Gmail, a task would be to compose and send an email, or create a folder and sort emails into it, where each of them would include a series of steps and interactions.
 - Unless highly relevant and really needed, we want to discourage you from prototyping or design tasks for obvious things that already exist, such as account management related tasks.
 - You are welcome to send us a link of your prototype as you go, or come to office hours, and we'll be able to give you suggestions on whether they feel sufficient or not.

The task should be key goals users need to accomplish using your prototype, such as 1) sign up for a career fair event, find it on your calendar, and cancel it, or 2) carry out a consultation session with an AI agent to do a mock interview and provide feedback (with many conversation turns to concretely achieve something). In the context of Gmail, a task would be to compose and send an email, or create a folder and sort emails into it, where each of them would include a series of steps and interactions.

Unless highly relevant and really needed, we want to discourage you from prototyping or design tasks for obvious things that already exist, such as account management related tasks.

You are welcome to send us a link of your prototype as you go, or come to office hours, and we'll be able to give you suggestions on whether they feel sufficient or not.

Methods we have learnt (need-finding, ideation, prototyping and testing)

Methods we have learnt (need-finding, ideation, prototyping and testing)

- Observations
- Semi-structured interviews
- Storyboards and speed dating
- Low-fidelity prototyping
- High-fidelity prototyping
- Heuristic evaluation
- Think-aloud protocol

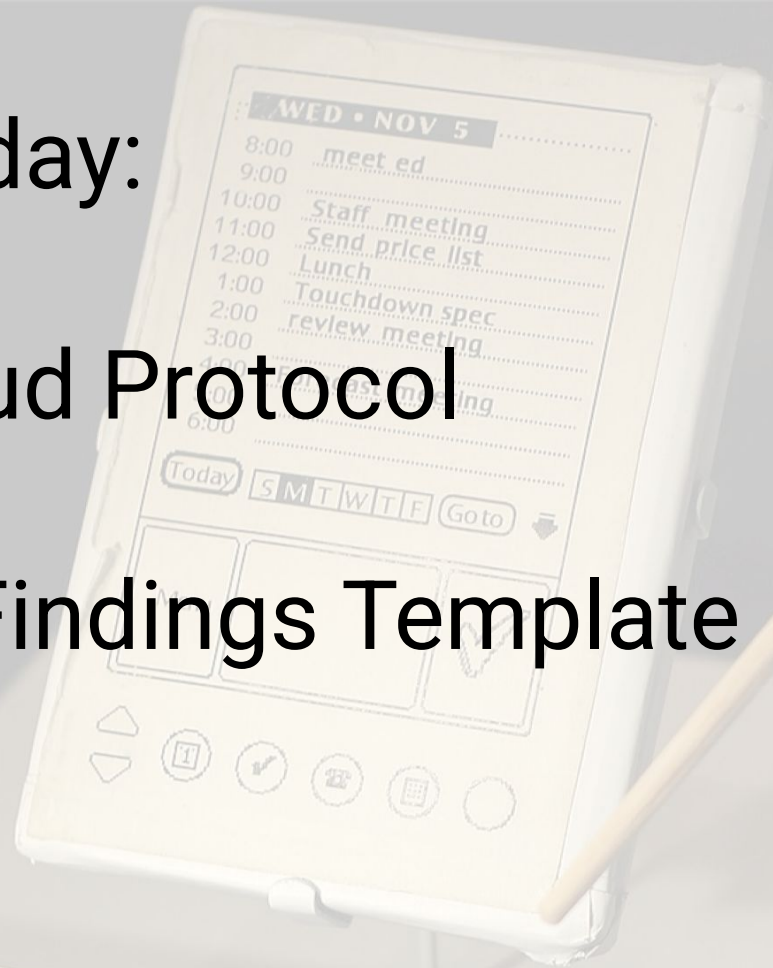
Other Methods?

Types of Usability Evaluation Methods

- **Involving users or not**
 - Heuristic evaluation (non-user)
 - Think-aloud (w/ user)
- **Quantitative vs. Qualitative**
 - Think-aloud, interview (qualitative)
 - Surveys, experiments with quantitative outcomes, log data analysis (quantitative)
- **Earlier vs Later stage in evaluation**
 - Heuristic evaluation (earliest)
 - Think-alouds (earlier)
 - A/B testing, field deployment studies (later)

Goals for today:

- Think-Aloud Protocol
- Usability Findings Template



PalmPilot wooden model