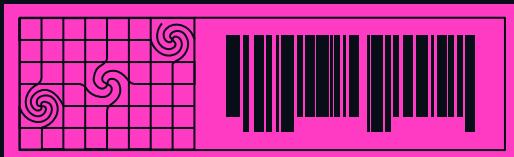




# AI-Capella

CS 147 2025au - Assignment 5





# OUR TEAM



Vaishnav G.



Bryant M.



Max R.



Marcus L.





# Value Proposition

## 1. Who's it for

- **A-capella singers** and vocal groups

## 3. What we deliver

- Live AI feedback as you sing
- Personalized coaching
- Daily confidence loop
- Tailored progress analytics.
- Practice anywhere



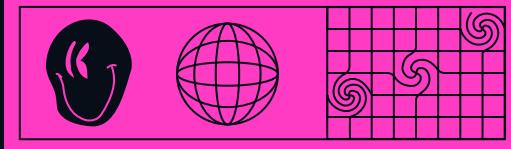
## 2. Why it matters

- Coaching is infrequent and harsh; Apps are static and not tailored to individual growth.
- Singers struggle with different memorization, harmonization and pitch-matching.

## 4. Why we're different.

- Live AI feedback
- Adaptive to user (not one-size) fit all
- Built specifically for A-capella singers

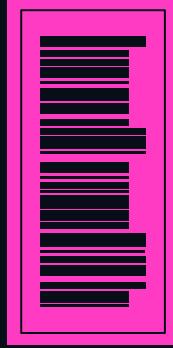
**AI-Capella - Live voice AI that builds confident a-capella singers**



# The Problem

A-Capella singers struggle to manage of memorization, harmonization, pitch-matching, and rhythm. To make matters worse, in-person voice coaching can too infrequent and too harsh to be conducive for improvement.





# The Solution

AI-Capella – your AI vocal coach.

- Generative AI for live feedback and encouragement
- Voice AI technology for adaptive real-time feedback and learnings
- 24/7 accessibility with short impromptu lessons tailored to your mistakes.





# OUTLINE OF TALK

<b>Sketching Explorations</b>	10-15 sketch overview; two fleshed realizations.
<b>Selected Interface &amp; Rationale</b>	key pros/cons; constraints/data; superiority case.
<b>Low-Fi Prototype Construction</b>	how built/operated; features & interactions.
<b>Low-Fi Prototype : 3 task flows</b>	simple/moderate/complex flows with captions/labels/transitions.
<b>Testing Methodologies</b>	participants(demo/recruit); environment/apparatus; procedure/roles; goals & metrics.
<b>Testing Results</b>	process trends; bottom-line KPIs; observations; goal attainment
<b>Discussion</b>	implications; design changes; what testing couldn't reveal
<b>Appendix</b>	full pros/cons; prototype link; test script/prep; critical incidents log; extra figures.





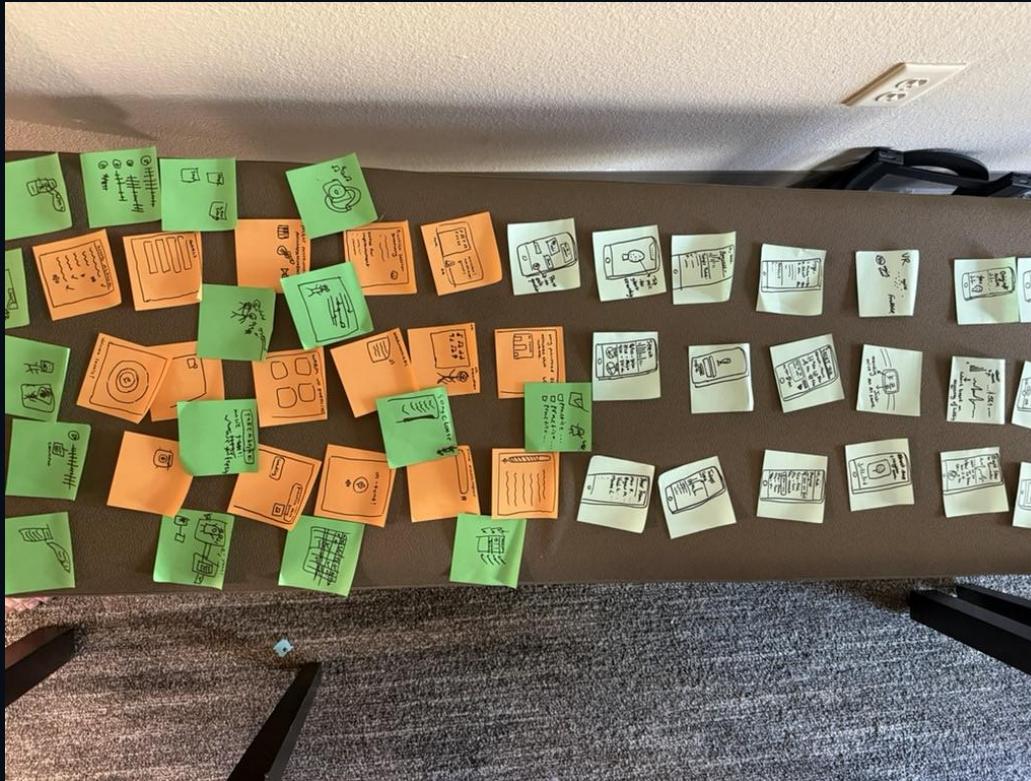
01

# SKETCHING EXPLORATIONS

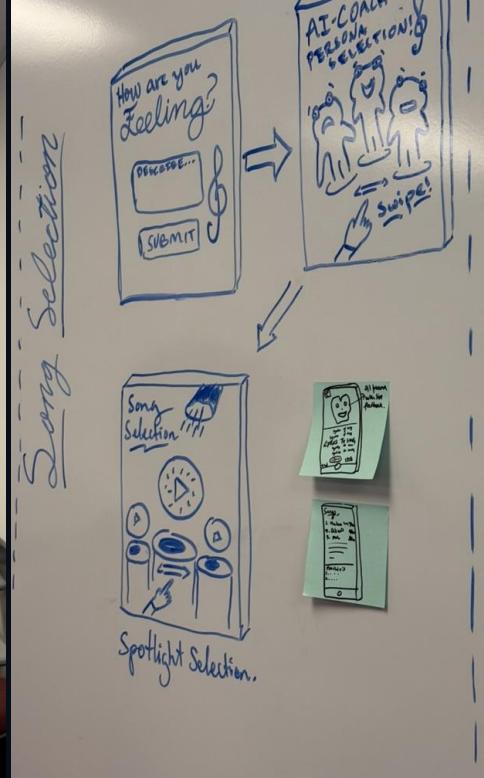
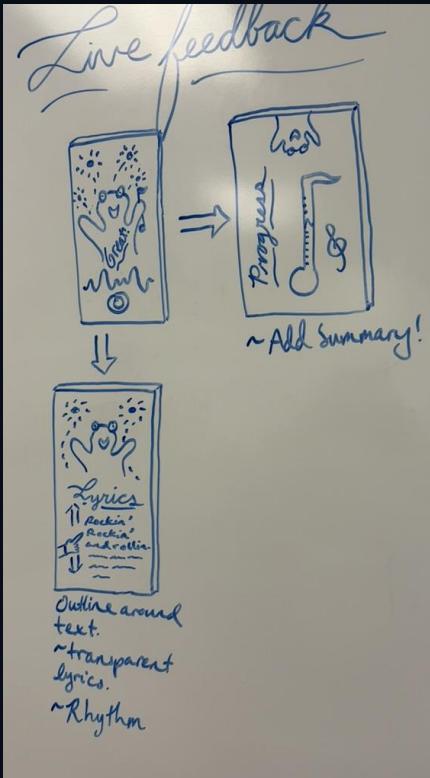




# Concept Sketches

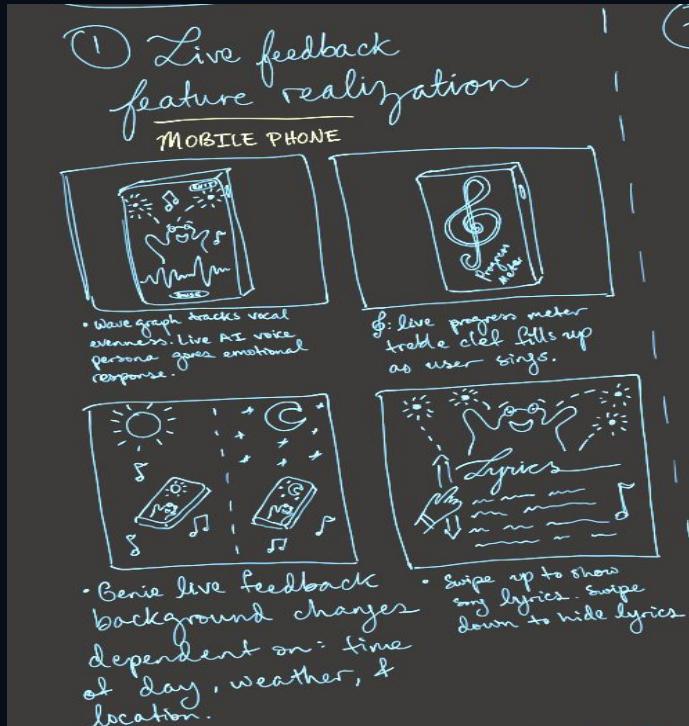


# Concept Sketches





# Realizations



Live  
Feedback  
Realization





# Realizations

## ② Haptic feedback for voice training

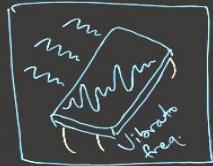
WATCH



~ Phone vibrates when you are singing on pitch!



• Watch gives off beat - signaling when to breath, how long to hold breath, and song tempo.



~ Phone emits different vibration frequencies dependent on vibration levels.



• Watch begins to vibrate if it senses you are straining your voice.

## ③ App/feedback personalization

MOBILE PHONE



• Adjust the personality of your AI-Genie based on your mood.



• AI-Genie adjusts instruction based on your needs.



• Tailored progress reports and trajectory





# Realization

## ④ VR/AR Realization



~ Genie follows user around as user practices.



~ Walk around and physically interact with songs during selection.



~ Ride the music coaster.  
• Singing with more quality and confidence makes the coaster go faster!



~ Hurdle over lyrics to memorize them actively!

## VR/AR Realization





# Realization

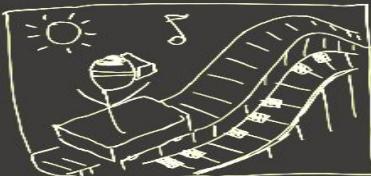
## ④ VR/AR Realization



~ Genie follows user around as user practices.



~ Walk around and physically interact with songs during selection.



~ Ride the music coaster.  
• Singing with more quality and confidence makes the coaster go faster!



~ Hurdle over lyrics to memorize them actively!

## VR/AR Realization





# Realization

## ④ VR/AR Realization



~ Genie follows user around as user practices.



~ Walk around and physically interact with songs during selection.



~ Ride the music coaster.  
• Singing with more quality and confidence makes the coaster go faster!



~ Hurdle over lyrics to memorize them actively!

## VR/AR Realization





02

Selected interface & rationale

## Selected Interface & Rationale



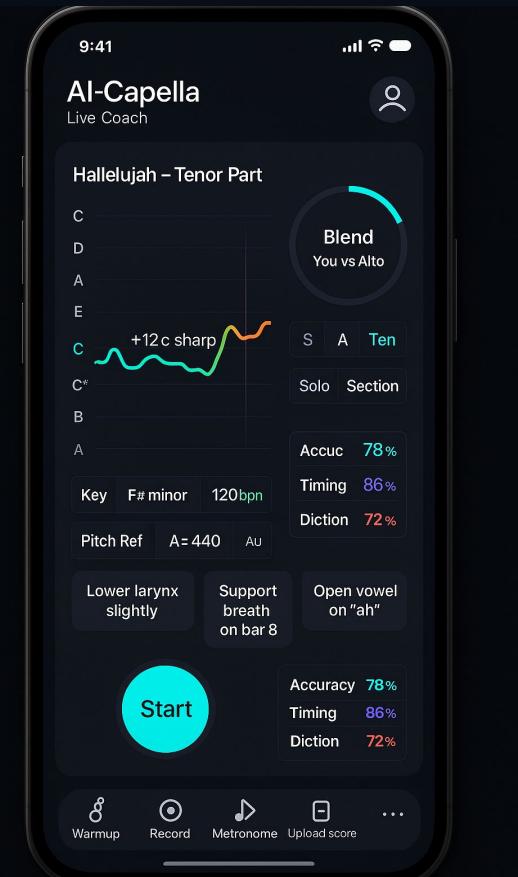
# OUR CHOSEN INTERFACE

## Our Choice :

- **Form factor:** Mobile phone app
- **Core loop:** “Press to practice → sing → instant guidance → quick drill → micro-summary”
- **Usage context:** private, quick sessions

## Why Mobile (vs. watch/AR/VR) :

- **Access & comfort:** Nearly everyone has a phone; singing alone with a phone feels natural
- **Frictionless setup:** Mic + screen + speaker in one; no pairing or calibration
- **Rich feedback canvas:** Enough space for lyrics, pitch line, and gentle guidance without clutter
- **Shareability:** Easy to export clips/reports to section leaders and group chats





# Pros & Cons

## Pros

- **Immediate, personalized guidance**
  - Remove the uncertainty of "am I improving".
  - No generalized lessons - custom to your voice and stage.
- **Private practice**
  - Lowers anxiety and improves confidence.
- **Engaging UI**
  - Progress reports, feedbacks, and gamification encourage consistency.
- **Self-paced learning**
  - The UI adapts to the lessons.

## Cons

- **Scope limits**
  - Tone quality/emotion detection is early → focus on pitch/rhythm first
- **Cognitive load**
  - Reading while singing can overwhelm → use minimal, glanceable cues
- **Human nuance gap**
  - Not a vocal coach replacement → positioned as "daily coach between lessons"





# CONSTRAINTS & OUR RESPONSES



## Small screen

- 3 elements max on practice screen
- “eyes-free” audio cues optional



## Phone speakers bleed into mic

- Headphones recommended flag
- echo-aware onset detector



## Mic noise/latency:

- Background noise cancelation on backend
- Keep guidance slightly buffered



## Reading while singing

- Big single action;
- 1-2 word cues;





# Findings & Superiority



## Findings That Shaped Decisions

- 1. Singers want **immediate feedback** that isn't mean; prefer **private practice** spots
- 2. **Confidence** is as important as accuracy—tone of feedback matters
- 3. Desire for **short, frequent sessions** over long lessons
- 4. Group needs: exportable **progress snippets** to align sections quickly



## Why This Design is Superior

- **Confidence-first:** Micro-wins and gamification reduce anxiety and increase practice
- **Right info and time:** Glanceable cues while singing. deeper analytics after takes
- **Team-friendly:** Shareable clips/metrics tighten rehearsal time for the whole group





03

# Low-Fi Prototype Construction



Paper Prototype with rectangular (phone shaped) paper cutouts; Paper overlays

### Form Factor



Facilitator (guides), Computer (swaps screens/overlays, times cues), Note-taker (quotes, hesitation).

### Roles

Warm-up → Simple → Moderate → Complex task; quick debrief after each.

### Session Flow





# PROTOTYPE INTERACTIONS AND FEATURES

## HOME

- SELECT SONG

User finds a song  
**Swipe action;**  
**Select action;**

- DRILLS

User picks a drill to practice more with  
**Select Action;**

- PROGRESS

View your progress  
**Select Action;**

## LIVE FEEDBACK

View live feedback as you sing  
**Select Action; View feedback**

## PRACTICE DRILLS

Watch progress update as you do drills  
**Select Action;**

## PROGRESS

Watch progress update as you do drills  
**Select Action;**





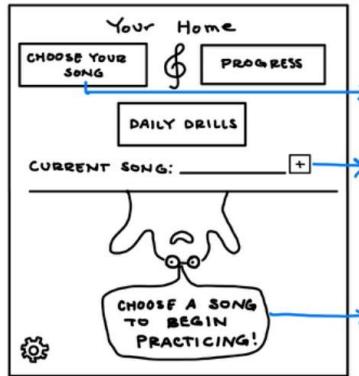
# 04 Low-Fi Prototype : 3 task flows

Blue arrows depict intended & quickest way to complete task. Green arrows indicate where users will be led to if they press other buttons. Task Flows with green arrows are in the appendix

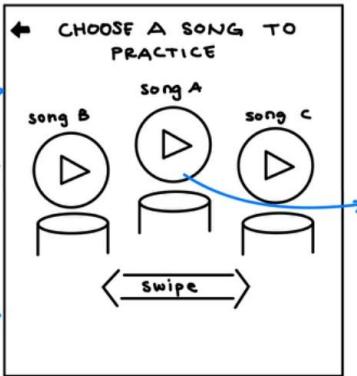


# Simple Task

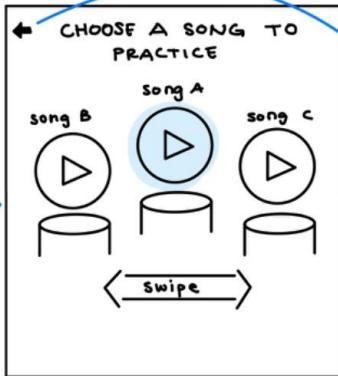
## SIMPLE TASK: SELECT A SONG & RECEIVE LIVE FEEDBACK



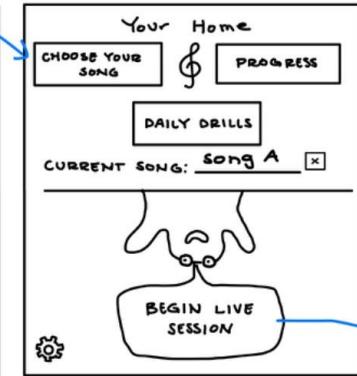
Home page has 3 main buttons (choose a song, progress, and daily drills). The coach avatar also serves as a button based on recommended task. There are multiple ways to choose a song.



The song library has different podiums you can swipe through to select a song. It might play a sample of the song.



Song is selected, user returns to homepage by pressing back button



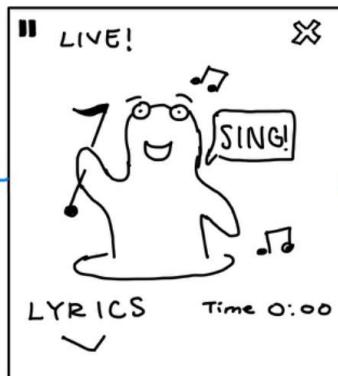
Coach button updates according to selected song. Activates live feedback session



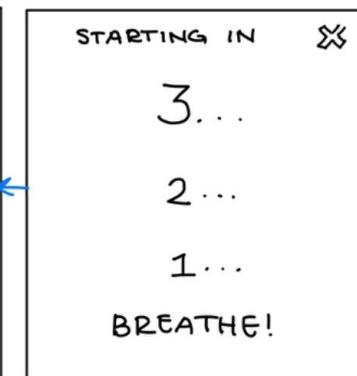
The simple task is complete when the track ends. User can press the finish button to return home, or can check progress after.



An example of the feedback a user might receive, user should address the problem to continue through session. This is done through vocal interaction.



User sings soon to track during live session. Has option to view lyrics and time elapsed. As well as pausing or exiting the session. No interaction with the screen is needed though to complete the session



Countdown screen allows user to prepare for session. No action is needed.

# Moderate Task

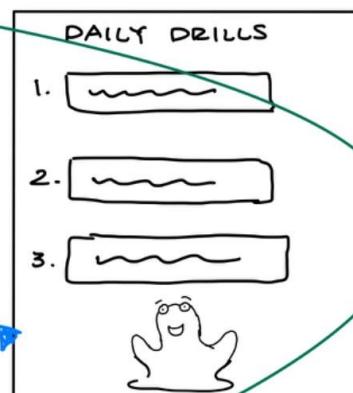
MODERATE TASK: REVIEW FEEDBACK & GENERATE DAILY DRILLS



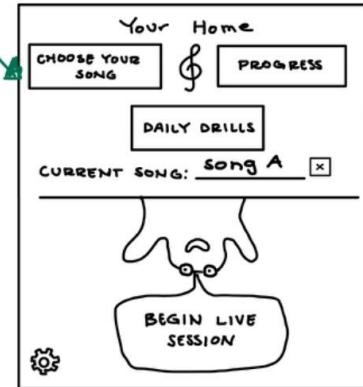
This task follows the simple task. Users can directly access feedback or finish the session



Session summary provides info on different aspects of singing: timing, pitch, for example. User can generate daily drills to address problems.



A list of generated daily drills

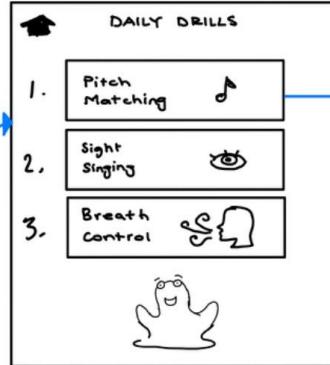


# Complex Task

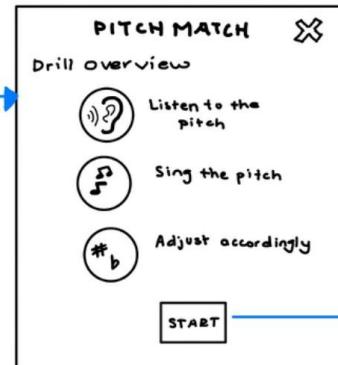
COMPLEX TASK: PRACTICE DAILY DRILLS & SEE PROGRESS UPDATE



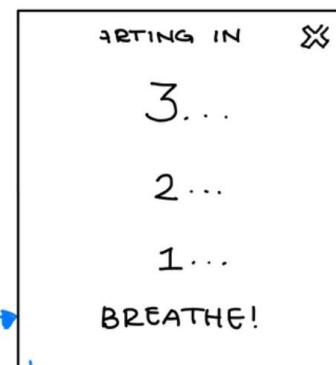
USER can view progress from home. But it will be empty without prior sessions



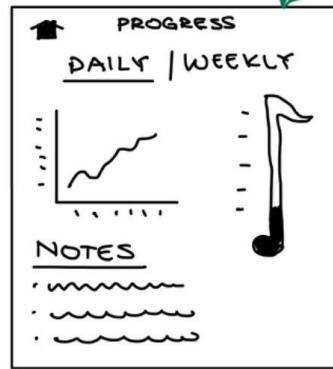
List of daily drills, generated earlier from previous live feedback session. This page is accessible from home page as well as when daily drills are generated.



Overview of drill, instructions.



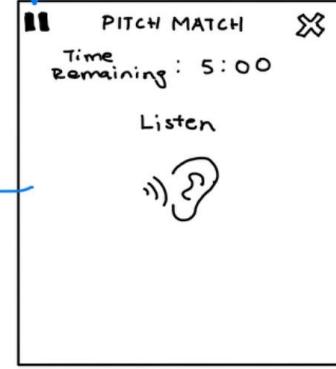
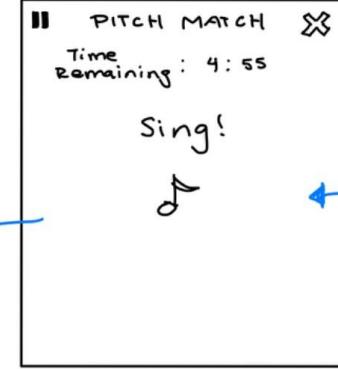
Countdown screen allows user to prepare for session. No action is needed.



The complex task is complete when the user sees their progress. Progress is measured from initial live session, then changes through daily drills and additional live sessions. Users can see daily and weekly updates



screen when drill is over.



Drill example



# 05 Testing Methodologies





# Participants

Participant	Background	A-Capella Experience	Design Experience	Recruited	Compensation
CL	CS Professional	Former	Some	Via friends	Verbal thanks
DB	CS-adjacent	Current	None	Via friends	None
PS	Design student	Casual Singer	Yes	Via friends	None
PJ	College student	Current	None	Via friends	None
RS	CS student	Current	Some	Via friends	None





# Environment & Apparatus

- Environment:
  - Tresidder: made it less stressful and easier for us to recruit and spread our screens out
- Apparatus:
  - We used a paper prototype with multiple cut and drawn screens with smaller overlay cut outs





# Procedure: Team Roles

**Facilitator:** Give Task instructions and provide guidance through interface if user gets stuck



**Computer:** Handles the prototype (i.e. flipping through screens, adding necessary overlays).



**Note-taker:** Keeps track of time, tracks number of mistakes a user makes.

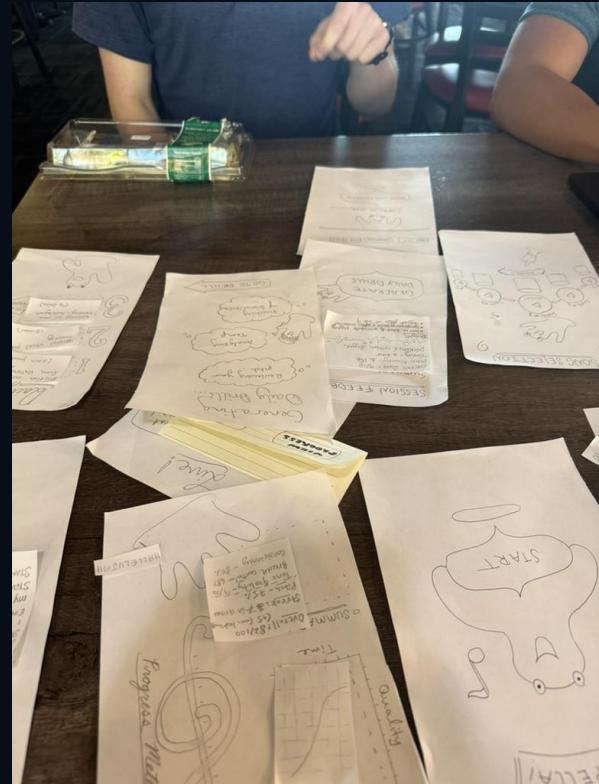
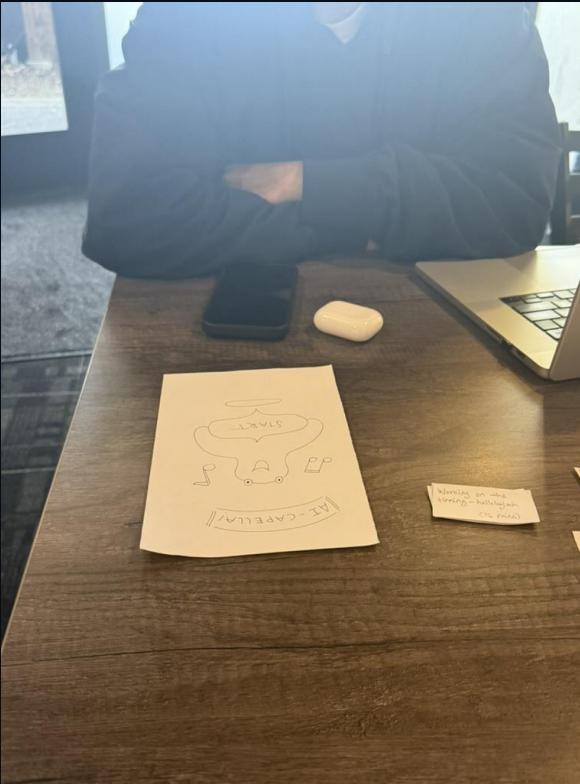


Note-taker

Computer



# Photos





# Usability Goals & Key Measurements

## Usability Goal 1:

Enhance user enjoyment through the live vocal feedback feature.

### Key Measurement:

- **Process Data:** Observable emotional cues (e.g., smiling, engaged behavior).
- **Self reported Data:** Pre/post-session ratings of enjoyment, frustration, and confidence, plus average lesson completion time.

### Justification:

The live feedback feature is being refined to keep users engaged without overwhelming them. Tracking emotional and behavioral responses will show whether it boosts enjoyment and motivation for continued use. This supports our goal of creating a positive, personalized practice environment that encourages consistent vocal training.





# Usability Goals & Key Measurements

## Usability Goal 2:

Improve learnability by enabling users to generate daily drills more efficiently across sessions.

### Key Measurement:

- **Process Data:** User confidence and navigation fluency (fewer pauses or misclicks).
- **Bottom-Line Data:** Reduction in time to generate drills between initial and later sessions, plus fewer unnecessary clicks or navigation errors.

### Justification:

Tracking speed and navigation confidence shows how quickly users internalize the app's workflow. As familiarity grows, faster drill generation reflects improved usability and intuitive design, ensuring the system supports skill growth without adding cognitive load.





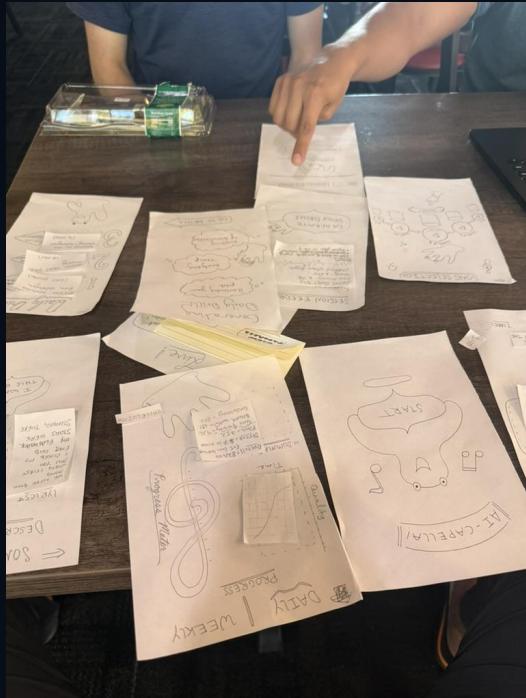
# 06 Testing Results





# Big Picture

- Core idea (live, encouraging feedback) resonated strongly
- Main friction: sequence confusion between song → *feedback* → *drills* → *progress*
  - Some confusion in getting to screens as well as simple steps
- Users asked “when does live feedback start?” and “what’s a daily drill?”
  - Major confusion about the daily drills
- Emotions trended **more excited & confident, less anxious & frustrated** as they progressed





# Bottom-Line

Metric	Target	Observed	Results
Task completion	100%	100%	Met
Avg. task time	CS-adjacent	~1:40	Met
Major errors	<2 per person	1-2 minor ones	Acceptable
Satisfaction (Excited / Happy / Confident)	$\geq 4/5$	4.5/4.25/3.75	Met, but not great
Frustration	$\leq 2/5$	1.25	Great





# Other Observations

**Flow Confusion:** Users skipped summary pages or jumped ahead

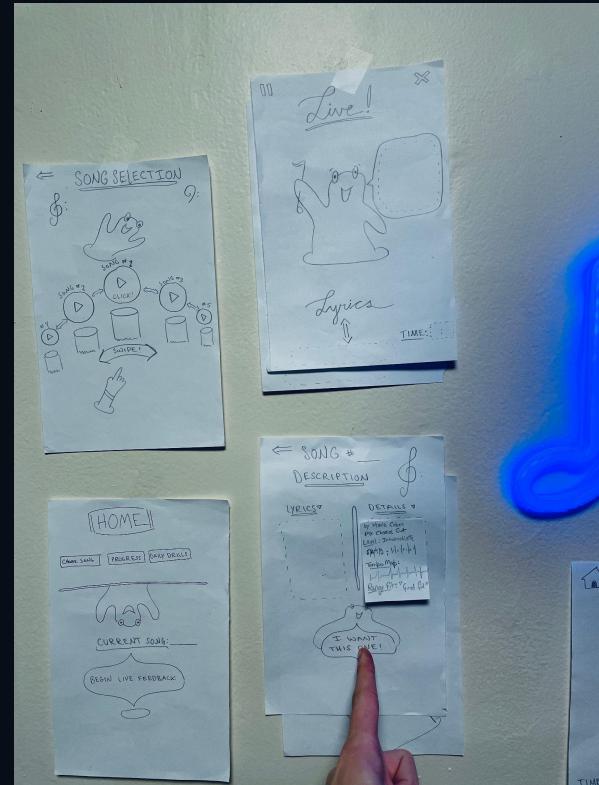
**Confusion:** "Daily drills," "pitch accuracy," "quality" unclear → need tooltips/tutorial

**Tone:** Universally described as *supportive* and *non-judgmental*

## Visual Feedback:

- Larger lyric text for readability
- Highlight selected song color
- Add confetti/progress animations for positive feedback

**Transparency:** Curiosity about how drills are generated → explain AI logic





# Usability Goals

Goal	Status	Next Step
Clear navigation flow	Partial	Improve clarity
Supportive Experience	Achieved	Keep/improve
Actionable Feedback	Partial	Clarify metrics for growth
Efficient task time	Achieved	Work on more clarity
Daily Drills vs Practice	Unsolved, but we're working on it	Define them better





# Key takeaways/goals

- **Clarify user flow**
- **Define our progress metrics**
- **Refine drill logic**
- **Enhance motivation visually**
- **Improve readability**
- **Keep positivity**



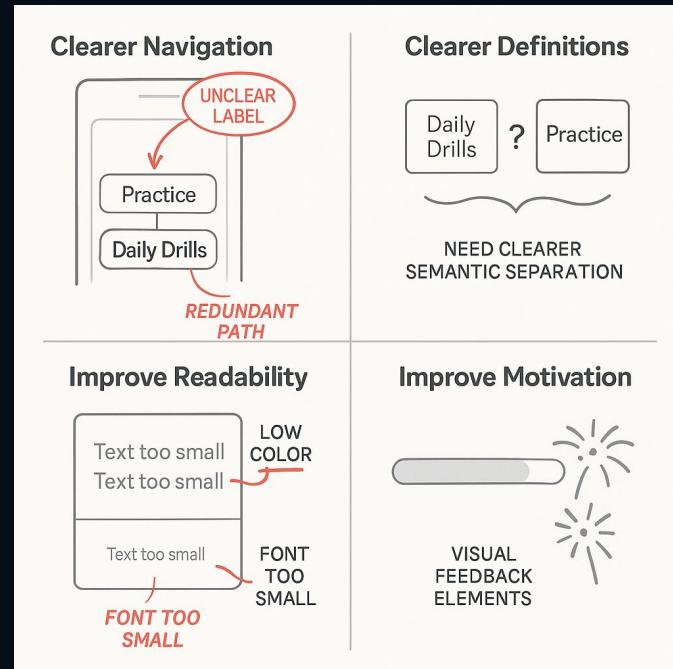


# 07 Discussion



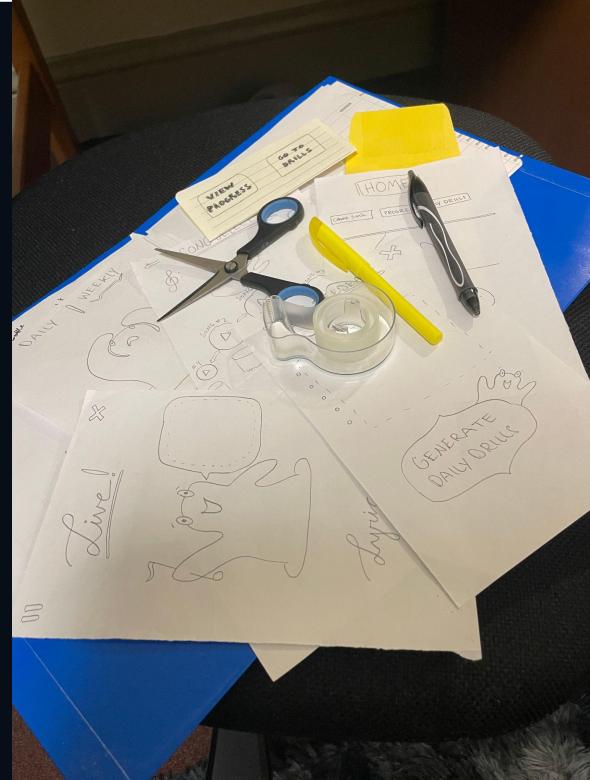
# Major Changes Needed

- Clearer Navigation:
  - Redo button and flows so it is more intuitive
- Clearer Definitions:
  - Need to finalize and redefine the difference between "Daily Drills" and "Practice"
- Improve readability:
  - Just need certain parts to be clearer and larger
- Improve Motivation:
  - Increase visual motivation metrics (i.e. progress bar, little programmed fireworks etc)



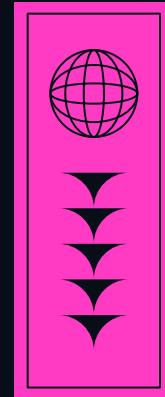
# Limits of Low-Fi Testing

- Audio and Live feedback:  
○ Paper prototype can't simulate music, audio, or feedback so users cannot really experience our project
- Observer Bias:  
○ There is a chance of possible observer bias as we tell them of their tasks and they don't get to fully explore on their own
- Limited Observation Window:  
○ Our product is meant to be practiced for multiple days but with the Low-fi prototype you can't really see long term use or problems with it





# Appendix





# Pros & Cons (live feedback concept)



- **Pros:**

- **Immediate personalized feedback:**
  - Knowing our interviewees, we knew they would want immediate feedback to help improve their skills without being mean, which is something that will help improve their abilities.
- **Convenient and can practice anywhere:**
  - Our interviewees wanted to be able to practice and receive lessons anywhere and preferably in private, which is something this app can ensure so people can feel comfortable practicing where they want to.
- **Engaging:**
  - The real-time feedback and fun screens we would have would help engage users and make the process more fun and exciting. We noticed that engagement and enjoyment must overcome the fear and anxiety of singing.
- **Supporting self-paced learning:**
  - We noticed with interviewees that they wanted the progression for learning to be tailored to them, and the app will be able to do that, as the feedback will lead to changing up lessons and managing them for growth depending on how they're

- **Cons:**

- **Limited scope of feedback with current technology:**
  - Most current AI voice detections can only really detect pitch and timing, but not tone quality, breath support, or emotional expression, which may frustrate users trying to improve overall.
- **Overstimulation with feedback:**
  - Users could be overwhelmed with receiving live feedback in text on the screen if it has too much other information displayed.
- **Removal of human interaction:**
  - There are parts of vocal coaching that technology cannot necessarily be aware of, like emotional phrasing, and users may reach a point where they abandon it after reaching that level.
- **Over Reliance on visual feedback:**
  - Having to focus on singing while also reading feedback may disrupt users and further increase their anxiety about practicing.





# Pros & Cons (haptic feedback)

- **Pros:**
  - **Real-time correction:**
    - Haptic feedback can provide immediate cues (like vibrations or pulses) when a singer goes off-pitch or uses incorrect breath support, allowing for faster learning than post-performance review.
  - **Multi-sensory learning:**
    - Engaging touch in addition to hearing can strengthen learning retention, especially for kinesthetic learners who benefit from “feeling” feedback.
  - **Physical feedback for focus:**
    - While singing, users might be too focused on lyrics or memorization to notice visual cues. Haptic feedback allows them to stay present in the music while still receiving guidance.
  - **Maintains emotional flow:**
    - Because feedback is silent and unobtrusive, it doesn't interrupt the expressive or emotional flow of singing the way visual or auditory corrections might. Error correction is not inherently or emotionally negative.
- **Cons:**
  - **Individual sensitivity differences:**
    - Signals might be too subtle for some users to distinguish between. It may also be difficult for users to memorize several haptic signals, so feedback may be limited to a few key things like intonation.
  - **Potential for emotional disconnect:**
    - While tactile, the feedback lacks the warmth and human intuition of a real instructor's presence or encouragement.
  - **Lack of contextual understanding:**
    - Haptic feedback can tell users that something is wrong (e.g., pitch off, poor breath support), but not why, without explanation, improvement might be surface-level.
  - **Possible desensitization:**
    - With prolonged use, users might begin to ignore or “tune out” the tactile cues, diminishing their effectiveness over time.





# Script

Hi, we're a team of Stanford CS students completing a survey-based project for an introductory HCI class, CS 147. We were wondering if you could spare 20 minutes of your time to test out our simple voice-AI paper prototype?

First please rate the extent to which you're feeling the following emotions (1-5): Excited: 1 2 3 4 5 , Anxious: 1 2 3 4 5 , Confident: 1 2 3 4 5 , Fulfilled: 1 2 3 4 5 , Frustrated: 1 2 3 4 5 , Happy: 1 2 3 4 5

Perfect! So we will be showing you various paper screens; treat them exactly as you would if they were screens from a mobile phone app. Feel free to try out any of the paper prototype buttons, drop downs, or swipe features; we will introduce new screens and overlays as you navigate through the app. Ultimately, we are going to have you try and complete 3 tasks. We acknowledge that this is a super weird format so please ask for help at any point if you get stuck! Let us know when you are ready to begin and we'll present the first task.

Ok, so your first task is to select a song and receive live voice-ai feedback while practicing that song. Press start to begin!

Great job! For the next task, please review your song-specific practice session feedback and generate personalized daily drills.

Now that you have your daily drills, your final task is to choose a daily drill to practice through a live feedback session and then check out your updated progress page!

Finally, now that you have completed our prototype test, we would really appreciate it if you could fill out our post-test emotion survey (the same one you filled out at the beginning of the session): Excited: 1 2 3 4 5 , Anxious: 1 2 3 4 5 , Confident: 1 2 3 4 5 , Fulfilled: 1 2 3 4 5 , Frustrated: 1 2 3 4 5 , Happy: 1 2 3 4 5 . Have a wonderful day!





# Usability Goals & Key Measurements

## Usability Goal 1:

Enhance user enjoyment through the live vocal feedback feature.

### Key Measurement:

Process Data: Observable emotional cues such as smiling and other behavioral indicators

Bottom-Line Data: Pre- and post-session ratings of enjoyment, frustration, and confidence, plus average time to complete a full lesson.

### Justification:

At this stage, the prototype's live feedback feature is being refined to ensure users feel engaged rather than overwhelmed. Measuring both emotional response and behavioral engagement will reveal whether the live feedback is enhancing the learning experience and motivating repeated use. This aligns with our focus on creating a positive, personalized practice environment that encourages consistent vocal training.

can you make this more concise without losing anything please?





# Usability Goals & Key Measurements

## Usability Goal 2:

Improve learnability by helping users generate daily drills more efficiently across Sessions.



**Key Measurement:** Reduction in total time required to complete the live feedback phase and generate new daily drills across repeated sessions.

- Process Data: User confidence and fluidity in navigation, indicated by fewer pauses or misclicks during task completion.
- Bottom-Line Data: Difference in average time spent generating new daily drills between the first and subsequent iterations, as well as the number of unnecessary button clicks or navigation errors.



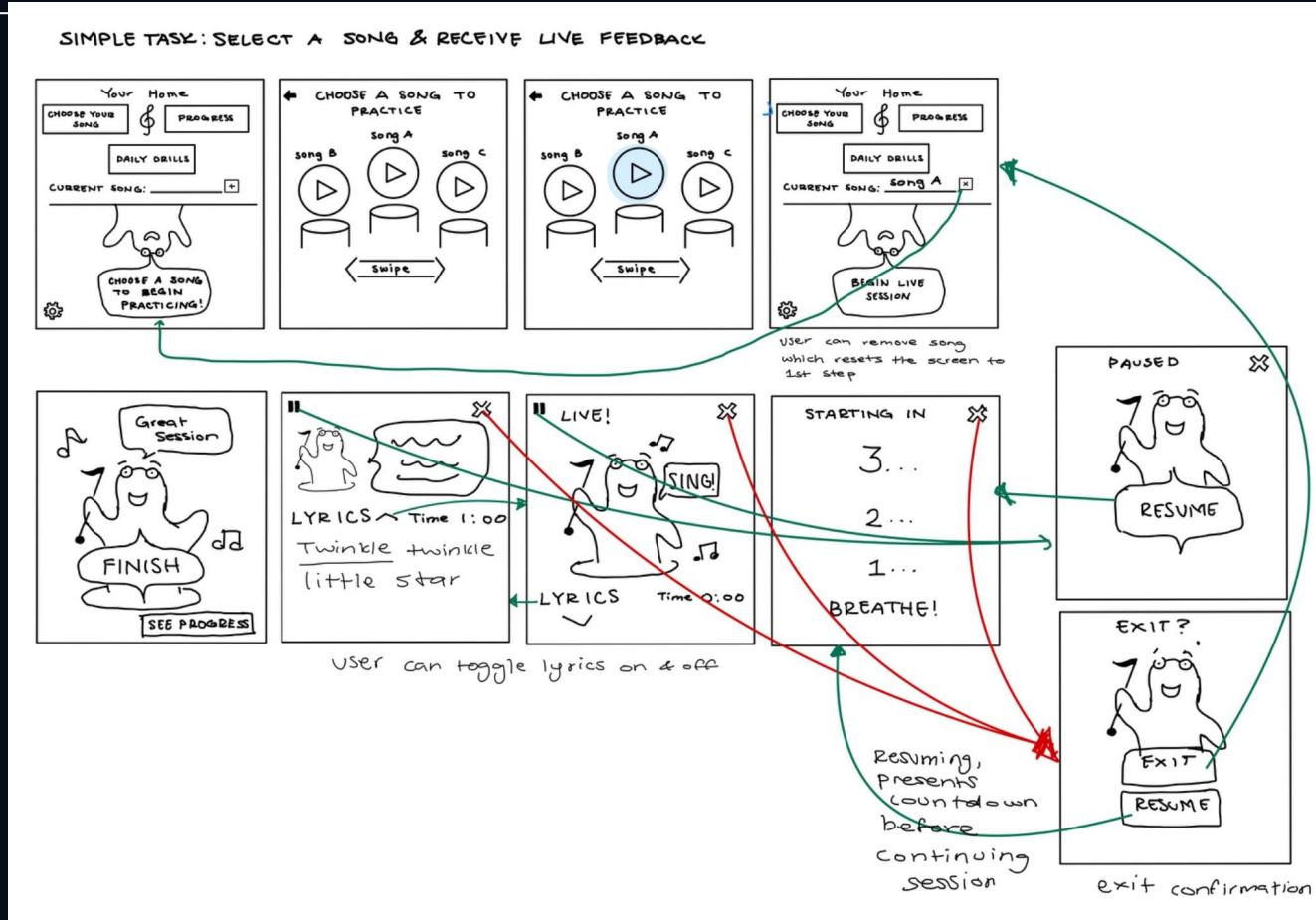
## Justification:

Since users must adapt to our mobile app workflow, tracking improvements in speed and navigation confidence provides insight into how quickly they are internalizing the app's structure. As users become more familiar with drill generation, efficiency gains will reflect improved usability and intuitive design. This metric ensures that the system supports skill progression without adding unnecessary cognitive load.

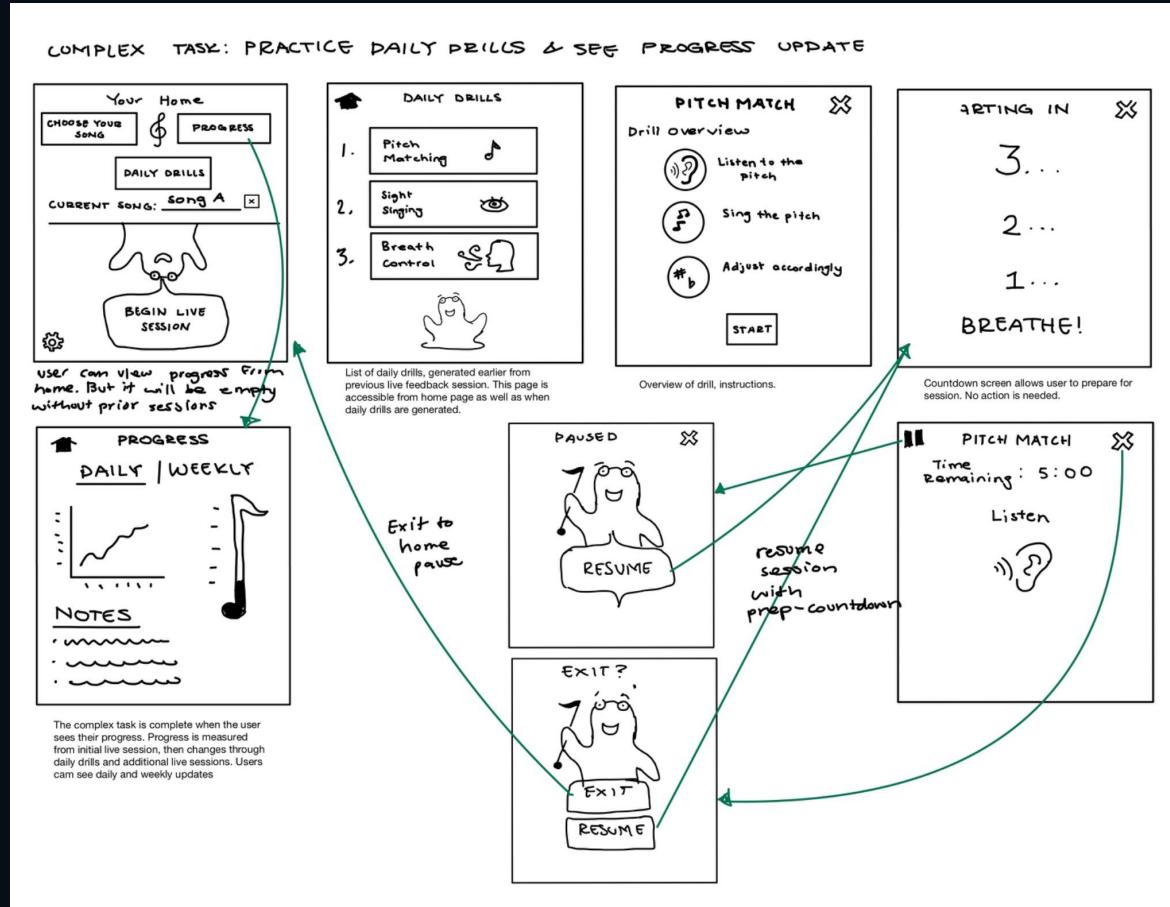




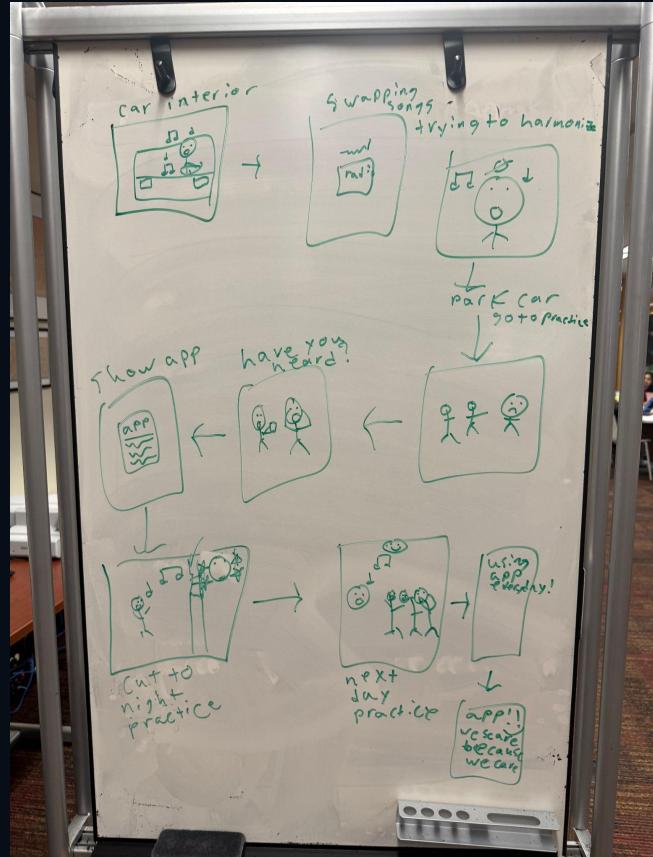
# Simple taskflow screens with all button interactions



# Complex taskflow screens with all button interactions



# First Draft Storyboard!



# Second Draft Storyboard! (Fullscreen)

*~ Drawn & Written by Max Rodriguez*

**1.** Main character driving to A-Capella practice.  
~ Outside - daytime for lighting

**2.** Plays song on Spotify  
~ Zoomed in

**3.** Struggles to sing along  
~ character is visually nervous.

**4.** Cut to walking to practice.  
~ Still outside for lighting.

**5.** Quick shot (close-up) to character walking in

**6.** Greets A-Capella coach.  
~ Introduces new character.  
~ Piano clearly seen in background to show this is ABOUT MUSIC

**7.** A-Capella coach instructs main character to sing the song "Ilona".  
~ Zoom focus on coach.  
~ They are seen.

**8.** Main character is nervous.  
~ Zoom in on emotion  
~ takes deep breath

**9.** Attempts to sing  
~ Main character trembles and voice cracks.

**10.** A-Capella coach not impressed and disappointed.  
~ Zoom into

**11.** A-Capella Coach introduces AI-Capella!  
~ Zoom in  
~ Glow effect

**12.** Main character starts the app.  
~ Wide shot.  
~ a little longer for suspense.

**TASK 1: Practice a Song w/ AI-Capella**

**13.** A-Capella Genie appears!  
~ they are dressed up  
~ Almost ecstatic  
~ Ready to help!

**14.** A-Capella Genie introduces himself to main character!  
~ very friendly interaction.

**15.** Main character sings the song again for the Genie.  
~ A little less anxious.

**16.** Genie listens and provides positive feedback and instructions.  
~ *"You're getting there, keep going with your solo voice!"*

**TASK 2: Receive & learn From the live Voice-AI Feedback**

**17.** Main character sings again with more confidence.  
~ they listened to the feedback!  
~ singing definitely improves

**18.** Main character sings again for A-Capella coach.  
~ A-Capella coach impressed!  
~ Main singing with more confidence.

**19.** Begin practice montage!  
~ Singing improving!  
~ More excitement  
~ Personalized feedback.

**20. FINAL SCENE**

~ Show main character and A-Capella coach singing together passionately!