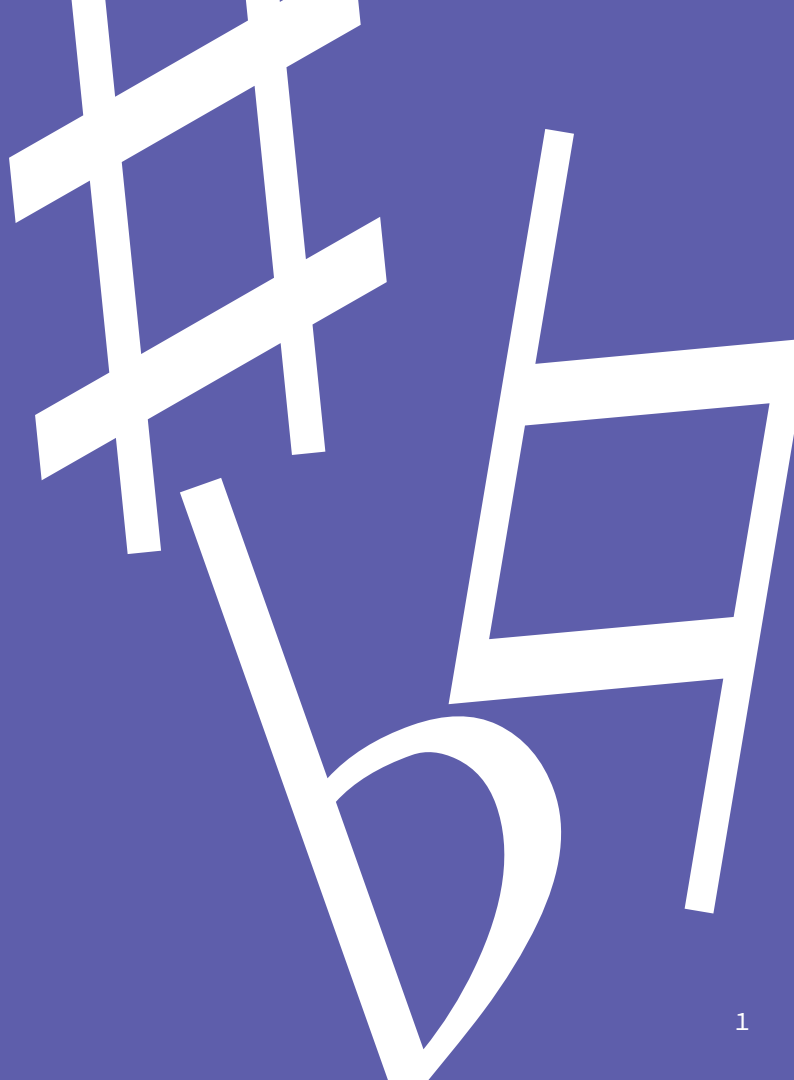


Non-Visual Flute Lessons - How Possible is it?

Dani Leinwander

Georgia Institute of Technology - College of Design - Music Technology - CCML (Computational & Cognitive Musicology Lab)

Advisor

Claire Arthur

Computational & Cognitive Musicology Lab (CCML)

Codesigner

Alonso Zubillaga

Special thanks to:

Abigale Stangl

MultiSense Design Research Studio

Content

1

Recap / Phase 1

2

Phase 2 & Phase 3

3

TakeAways & Timeline

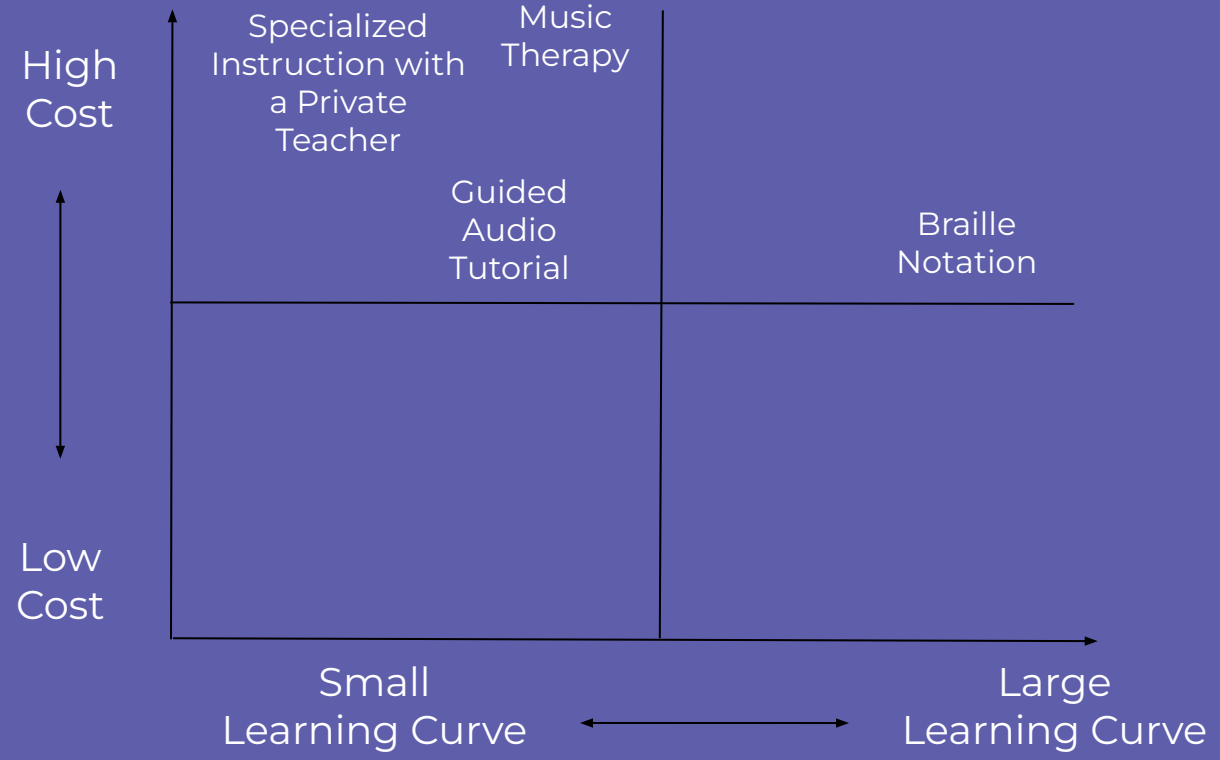
Project Goals

Alonso's Knowledge of his lived experiences and currently available tools for non-visual individuals

Dani's Music Education knowledge and Access to music technology knowledge

Existing Technologies

for non-visual instrument learning

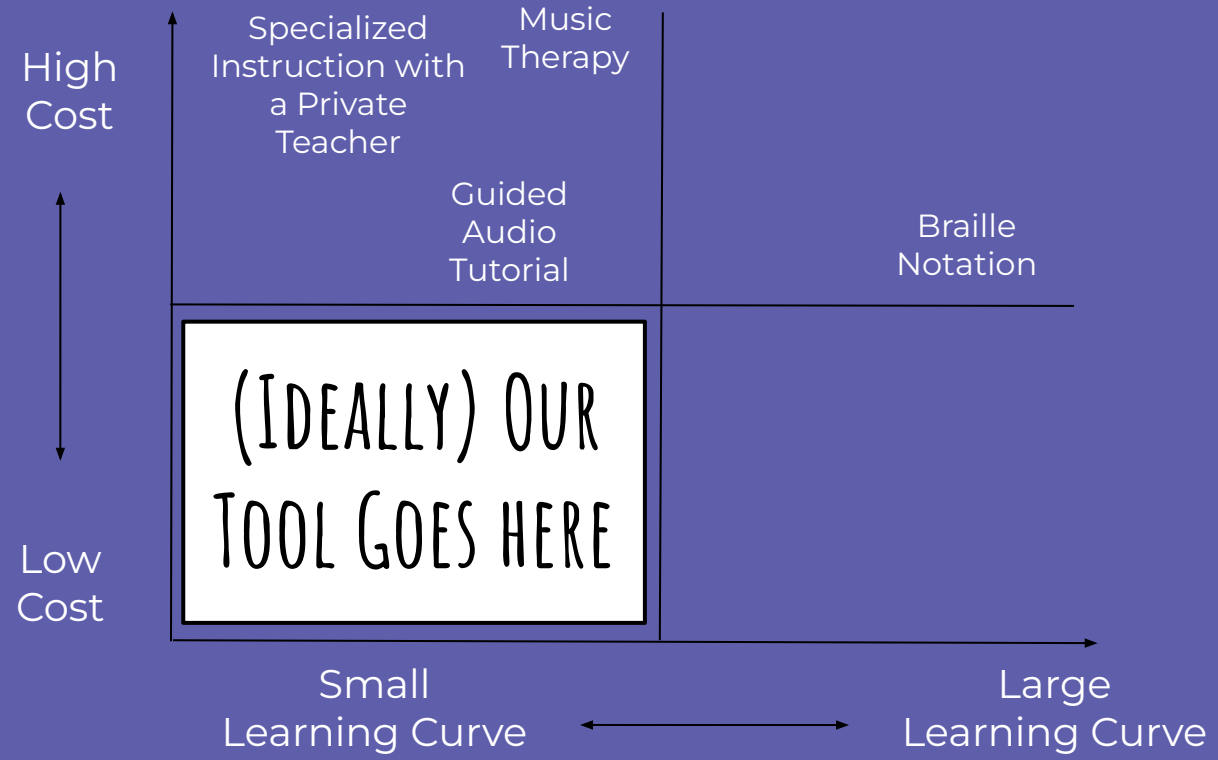


"We Avoid PDFs": Improving Notation Access for Blind and Visually Impaired Musicians (2023 William Payne & Amy Hurst)

"Why are there so many steps?": Improving Access to Blind and Low Vision Music Learning through Personal Adaptations and Future Design Ideas (2023 Leon Lu et al)

Existing Technologies

for non-visual instrument learning



Identified Mismatches

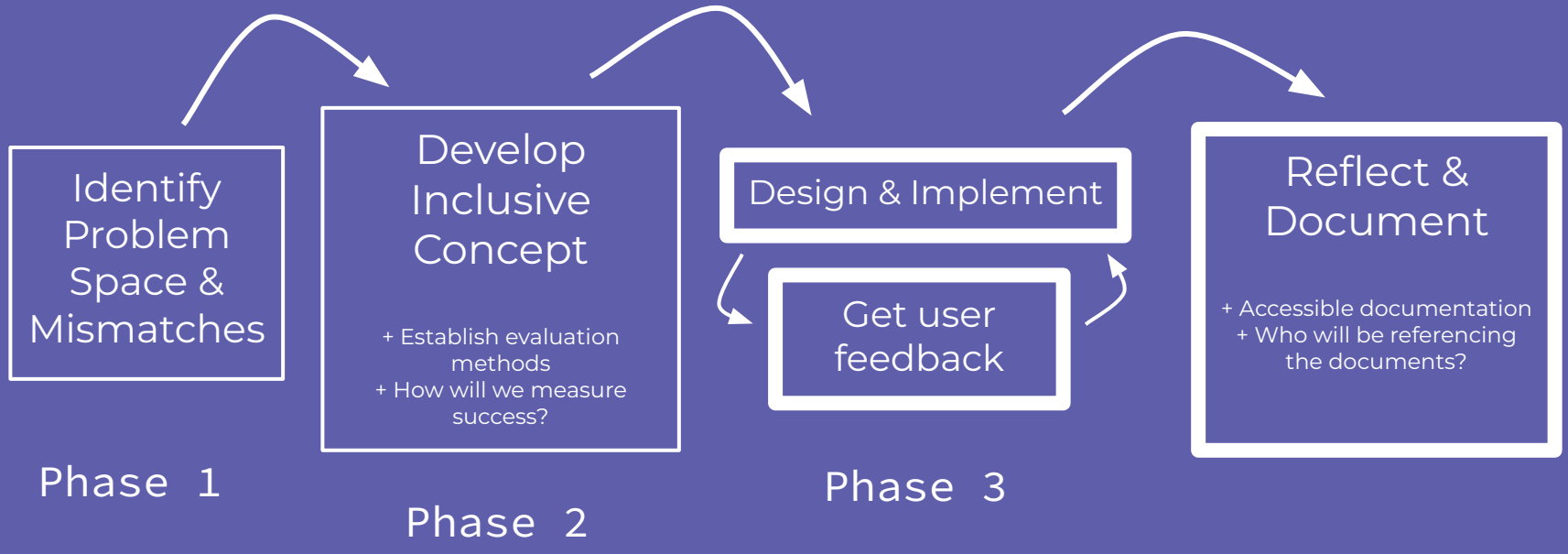
1
Sheet music
(assumes sight)

2
Fingering Charts
(assumes sight)

3
Visual Cues
(assumes sight)

Our Codesign Framework

Through semi-structured lessons and interviews:



Phase 1 - Design Approach

Identify Problem Space & Mismatches

Phase 1

In which ways are instructional methods for flute adapted or reconceptualized for non-visual instruction?

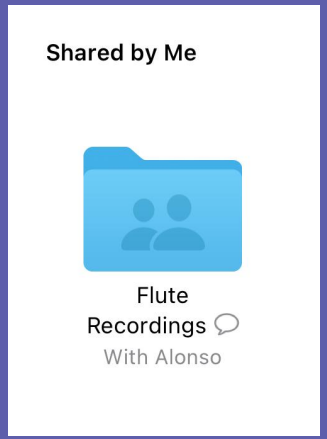
- Conducted 7 sessions: 1 initial meeting, 5 lessons, and 1 interview (with the purpose of re-collecting lost data)

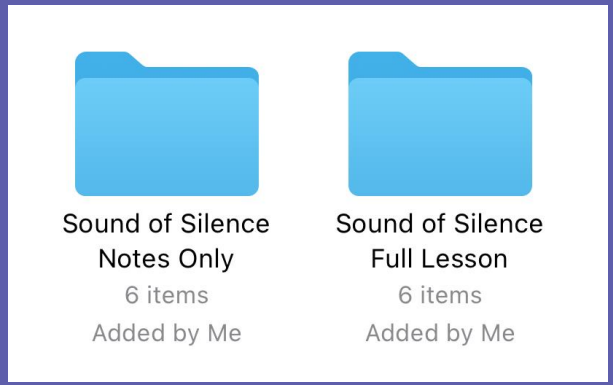
Phase 1 - Technology Approach

Hand made audio files

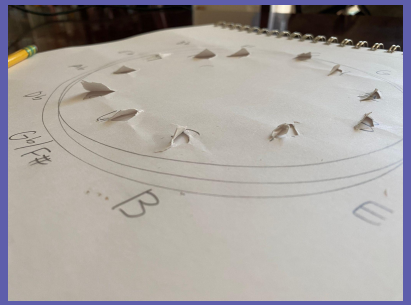
- Shared with Alonso through a shared folder on iOS Files app (ended up being a very accessible app for this purpose)
  - Small exception of the voice over playing over the audio files at the beginning

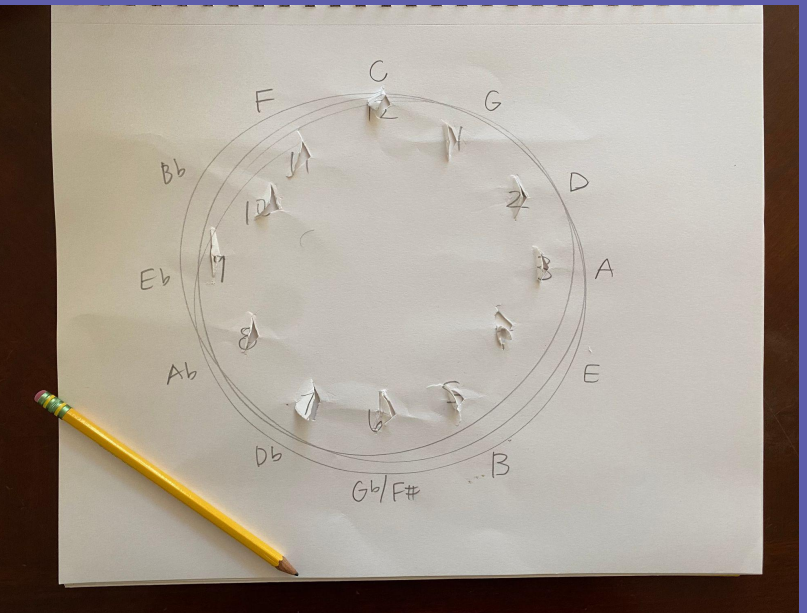
Using the Message app on iOS to send musical information as plain text (such as the sequence of pitches in a song)





Phase 1 - Instructional Approach





What adaptations can be made to the instructional design of private lessons for non-visual learners?

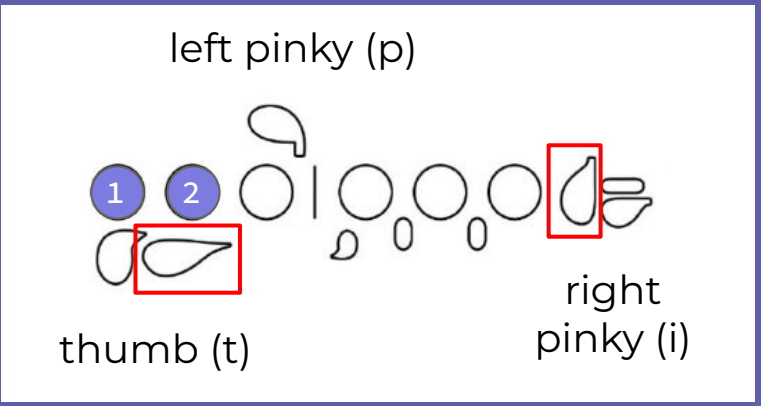
Tactile guidance

- Adapting the instrument
- Adapting instructional tools

[Bo Pep Thumb Rest]

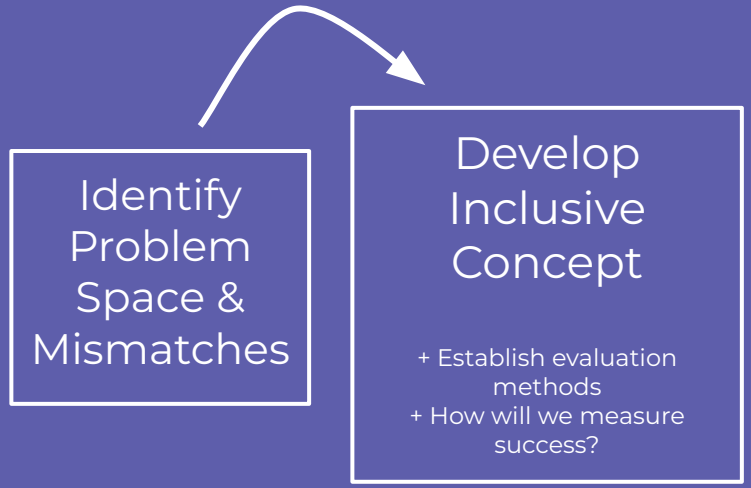
Phase 1 Fingering Codes

								n = none
audio file name = unique name						n = none		p = pinky on d sharp key
						t = thumb	n = none	s = pinky on c sharp key
					0 = none	h = thumb b key	p = pinky on g# key	c = pinky on both cs and next key
					(n = none?)			b = pinky on both cs and c and b key
unique ID	unique name	pitch (sharps)	alt pitch (flats)	octave #	homekeys (0,1,2,3,4,5,6)	thumb (n,t,h)	left pinky (n,p)	right pinky (n,i,cs,c,b)
1	B3	B	n	3	123456	t	n	b
2	C4	C	n	4	123456	t	n	c
3	Csharp4	Csharp	Dflat	4	123456	t	n	s
4	D4	D	n	4	123456	t	n	n
5	Dsharp4	Dsharp	Eflat	4	123456	t	n	i
6	E4	E	n	4	12345	t	n	i
7	F4	F	n	4	1234	t	n	i
8	Fsharp4	Fsharp	Gflat	4	1236	t	n	i
9	G4	G	n	4	123	t	n	i
10	Gsharp4	Gsharp	Aflat	4	123	t	p	i
11	A4	A	n	4	12	t	n	i
12	Asharp4	Asharp	Bflat	4	14	t	n	i
13	B4	B	n	4	1	t	n	i
14	C5	C	n	5	1	n	n	i
15	Csharp5	Csharp	Dflat	5	0	n	n	i
16	D5	D	n	5	23456	t	n	n
17	Dsharp5	Dsharp	Eflat	5	123456	t	n	i
18	E5	E	n	5	12345	t	n	i
19	F5	F	n	5	1234	t	n	i
20	Fsharp5	Fsharp	Gflat	5	1236	t	n	i
21	G5	G	n	5	123	t	n	i
22	Gsharp5	Gsharp	Aflat	5	123	t	p	i
23	A5	A	n	5	12	t	n	i
24	Asharp5	Asharp	Bflat	5	14	t	n	i
25	B5	B	n	5	1	t	n	i
26	C6	C	n	6	1	n	n	i



11	A4	A	n	4	12	t	n	i

Phase 2 - Design Approach

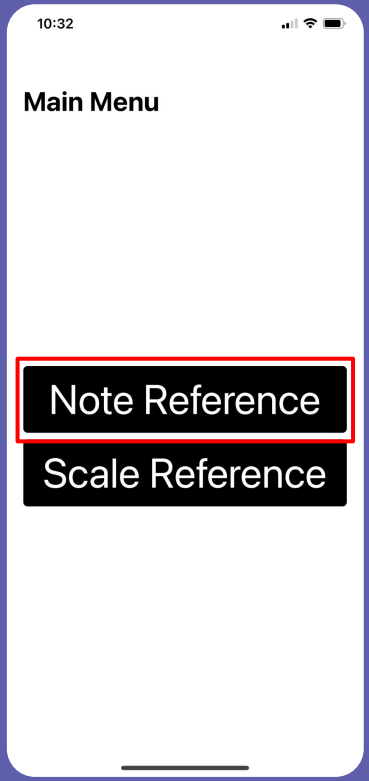


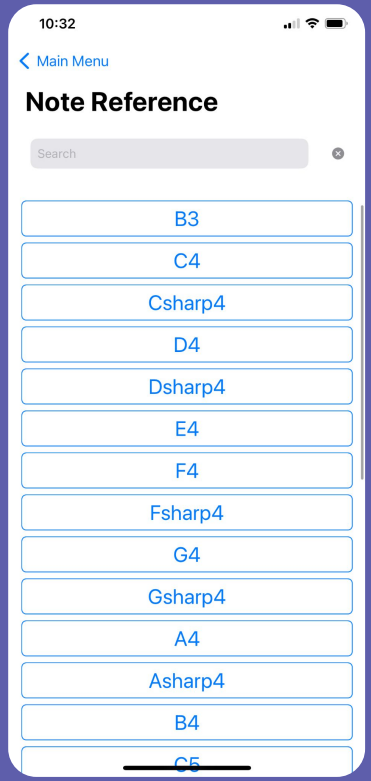
Phase 2

What non-visual heuristics does the learner expect during instruction, and what solutions should be prioritized based on their expectations as opposed to accepted practices in private instrument instructional design?

- Purpose: to use what we have learned about the process of private flute lessons from phase 1 to develop an inclusive concept
- Conducted 3 co-design reflections: 2 interviews and 1 session conducted by third author

Phase 2 - Technology Approach





Non-visual aural learning app

How to get sheet music and fingering chart information into usable form so we can present it back non-visually?

(remember, we are looking for solutions that use iOS and are feasible for me (Dani) to implement)

- 1st option: Computer vision
- 2nd option: Pitch recognition
- 3rd option: Segmentation & Presentation in multiple sound files
  - MIDI → python backend for processing into segments → move segmented audio files by hand to iOS app
  - Trying to automate what I was doing with the hand made audio files

Phase 3 - Technology Approach

MIDI Processing Approach

Initially processing of MIDI data done with Python libraries and plain text, but in the future I am considering re-writing using HumDrum and kernel notation

- music21 - process MIDI data
- fluidsynth - convert MIDI to .wav

Segmentation Approach

- difflib - segment similarity

https://github.com/dleinwan/blvmusic

```
class Codebook:
```

Can generate audio files of reference information given an input .csv file

```
class Segmentor:
```

For segmenting sequences and analyzing their similarity

Alonso's Ideal Aural Learning App

Accessed using iOS Voice Over on iPhone **(portability)**

- Portable, always has his iPhone anyway

Easily understood by teachers, other students, their family and friends **(interdependence/human intervention in automated tasks)**

Note & Scale References **(variety of content)**

+ Aural clips stitched together from premade clips:
  + note names
  + fingerings
  + sounding pitches

User has control over dynamically generating content based on individual needs **(flexibility & customization)**

Integration with other applications like Spotify for ease of song and lyric learning **(variety of content)**

- Would like similar controls over navigating audio files as he currently is able to do with Spotify

Would like gamification for learning music theory concepts and for drilling **(engagement)**

- To what extent can we automate in the app the instructional methods that we use during lessons?

WCAG 2.2 = Web content accessibility guidelines
(also includes mobile design guidelines)

WCAG standards do not sufficiently cover how to effectively deliver musical material to users non-visually

§ **Success Criterion 1.4.7 Low or No Background Audio**

(Level AAA)

Understanding Low or No Background Audio
How to Meet Low or No Background Audio

For prerecorded audio-only content that (1) contains primarily speech in the foreground, (2) is not an audio CAPTCHA or audio logo, and (3) is not vocalization intended to be primarily musical expression such as singing or rapping, at least one of the following is true:

No Background
The audio does not contain background sounds.

Turn Off
The background sounds can be turned off.

20 dB
The background sounds are at least 20 decibels lower than the foreground speech content, with the exception of occasional sounds that last for only one or two seconds.

NOTE

Per the definition of "decibel," background sound that meets this requirement will be approximately four times quieter than the foreground speech content.

§ **Success Criterion 1.2.1 Audio-only and Video-only (Prerecorded)**

(Level A)

Understanding Audio-only and Video-only (Prerecorded)
How to Meet Audio-only and Video-only (Prerecorded)

For prerecorded audio-only and prerecorded video-only media, the following are true, except when the audio or video is a media alternative for text and is clearly labeled as such:

Prerecorded Audio-only
An alternative for time-based media is provided that presents equivalent information for prerecorded audio-only content.

§ **Success Criterion 1.4.2 Audio Control**

(Level A)

Understanding Audio Control
How to Meet Audio Control

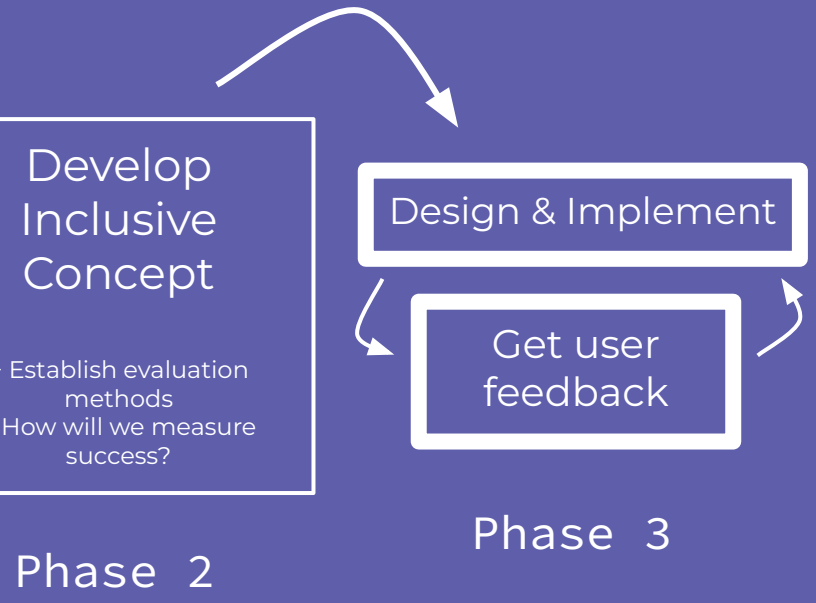
If any audio on a Web page plays automatically for more than 3 seconds, either a mechanism is available to pause or stop the audio, or a mechanism is available to control audio volume independently from the overall system volume level.

https://www.w3.org/TR/WCAG22/

If the (WCAG) standards do not already exist for the functions we are trying to implement, then we need to research *why* and **more importantly**:

How *we could create additional standards to help designers deliver audio content non-visually?*

Phase 3 - Design Approach



When using the app for the first time during orientation, during instruction, and during independent use, which usability issues arise and what features impact the instructional interactions in lessons and during independent use?

- So far: have conducted 2 lessons
- Planned: 1 (or 2?) more lesson(s)

Phase 3 - Pitch Segmentation

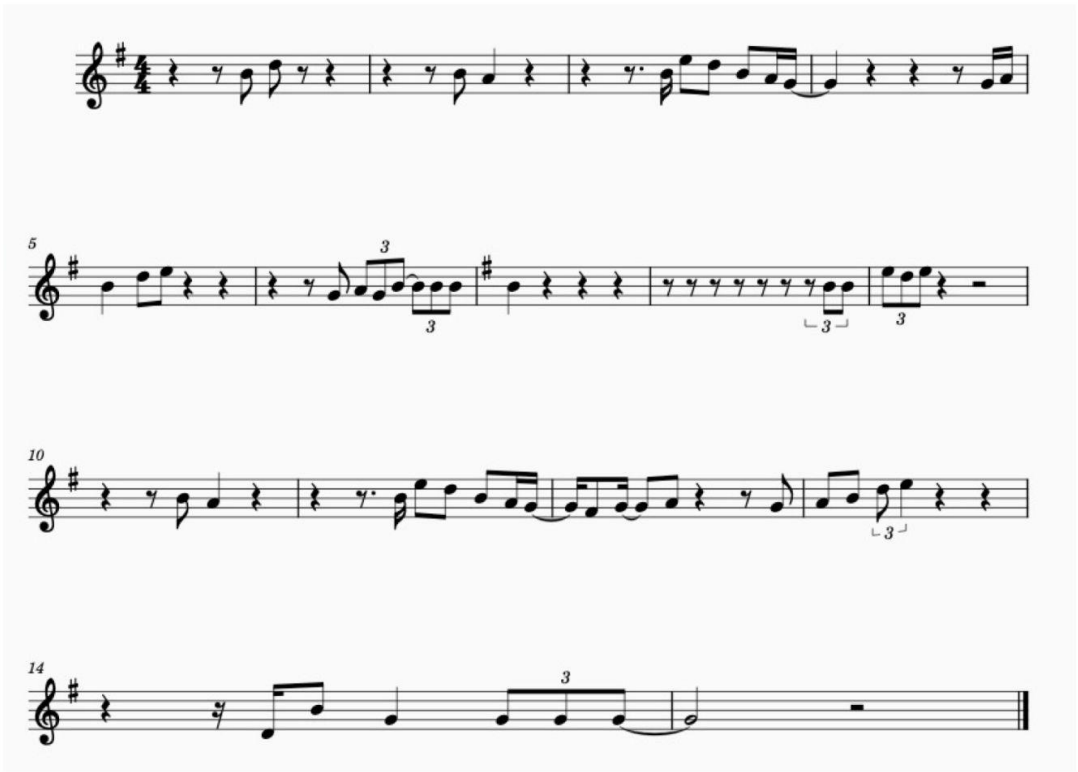
Segment pitch information sequentially (using difflib.SequenceMatcher)

Group pitches into certain # of notes, test how similar the segments are to each other

Can possibly find most useful groupings

Future: need to figure out how to test similarity of uneven groupings

NOTES	OFFSET 0	1	2	3
	0.00	0.00	0.00	0.00
	0.00	0.00	0.00	0.00
2	0.02	0.06	0.00	0.00
3	0.02	0.00	0.04	0.00
4	0.04	0.00	0.02	0.00
5	0.00	0.04	0.02	0.02
6	0.00	0.00	0.00	0.00



Take-Away Summary

1. WCAG can be improved by adding new standards for audio delivery
2. We need more non-visual development tools!
3. Aural learning tools can provide a platform for non-visual mental practice
4. Human-centered case studies can inspire technology design innovations.
    a. These innovations may be simple but have a large effect!

Check out:

WCAG Standards
www.w3.org

The World Wide Web Consortium (W3C) develops standards and guidelines to help everyone build a web based on the principles of accessibility, internationalization, privacy and security.

ASSETS - Accessibility Conference
assets24.sigaccess.org

The 26th International ACM SIGACCESS Conference on Computers and Accessibility.

The ASSETS conference is the premier forum for presenting research on the design, evaluation, use, and education related to computing for people with disabilities and older adults. For those in Europe and Oceania, ASSETS is rated as Core A — a designation for the top academic conferences that are "highly respected in a discipline area" (Core A; Top 16%).

Timeline

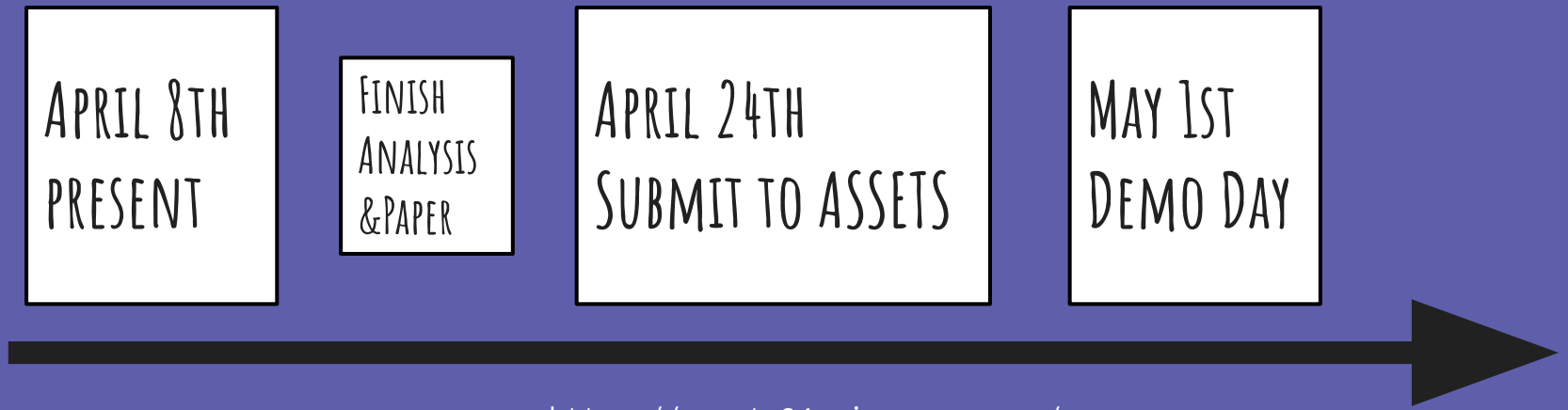
	November	December	January	February	March	April
App & Note Reference	Basic Generative Framework for note and scale learning (almost there)	Design & implement user input of generative parameters	Complete Initial Design Recommendation	Finish baseline implementation (iOS, SwiftUI)	Adjusting input parameters, getting feedback, rinse and repeat	End interviews and reflections Document Findings
Melody Segment-ation	Finalize data structure for performance criterion symbolic data representation + How do we represent this data for our algorithm to process?	Gather and/or curate MIDI (or MusicXML?) data for input	Devise baseline rule based system for giving difficulty score for each transition + How will we use the performance difficulty criterion to give a difficulty score? + How to use that score to inform segmentation?	Complete design recommendation Implement Algorithm	Adjusting input parameters, getting feedback, rinse and repeat	

Design — Implement — Document

Timeline

Identifying and Addressing Mismatches in Private Instrument Lessons for Blind Adults:

A Codesign case study



https://assets24.sigaccess.org/

Citations

Bennett, C. L., Peil, B., & Rosner, D. K. (2019). Biographical Prototypes: Reimagining Recognition and Disability in Design. *Proceedings of the 2019 on Designing Interactive Systems Conference*, 35–47. https://doi.org/10.1145/3322276.3322376

Frid, E. (2019). Accessible Digital Musical Instruments—A Review of Musical Interfaces in Inclusive Music Practice. *Multimodal Technologies and Interaction*, *3*(3), Article 3. https://doi.org/10.3390/mti3030057

Fyans, A. C., Marquez-Borbon, A., Stapleton, P., & Gurevich, M. (2012). *Ecological Considerations For Participatory Design Of Dmis*. https://doi.org/10.5281/ZENODO.1178257

Holmes, K. (2020). *Mismatch: How Inclusion Shapes Design*. MIT Press.

Kulik, J. A., Kulik, C.-L. C., & Cohen, P. A. (1979). Research on audio-tutorial instruction: A meta-analysis of comparative studies. *Research in Higher Education*, *11*(4), 321–341. https://doi.org/10.1007/BF00975623

Lu, L., Cochrane, K. A., Kang, J., & Girouard, A. (2023). "Why are there so many steps?": Improving Access to Blind and Low Vision Music Learning through Personal Adaptations and Future Design Ideas. *ACM Transactions on Accessible Computing*, *16*(3), 22:1-22:20. https://doi.org/10.1145/3615663

Mankoff, J., Hayes, G. R., & Kasnitz, D. (2010). Disability studies as a source of critical inquiry for the field of assistive technology. *Proceedings of the 12th International ACM SIGACCESS Conference on Computers and Accessibility*, 3–10. https://doi.org/10.1145/1878803.1878807

Payne, W. C., Xu, A. Y., Ahmed, F., Ye, L., & Hurst, A. (2020). How Blind and Visually Impaired Composers, Producers, and Songwriters Leverage and Adapt Music Technology. *Proceedings of the 22nd International ACM SIGACCESS Conference on Computers and Accessibility*, 1–12. https://doi.org/10.1145/3373625.3417002

Payne, W., & Hurst, A. (2023). "We Avoid PDFs": Improving Notation Access for Blind and Visually Impaired Musicians. In I. Sserwanga, A. Goulding, H. Moulaison-Sandy, J. T. Du, A. L. Soares, V. Hessami, & R. D. Frank (Eds.), *Information for a Better World: Normality, Virtuality, Physicality, Inclusivity* (pp. 581–597). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-28032-0_44

Quintero, C. (2022). A review: Accessible technology through participatory design. *Disability and Rehabilitation: Assistive Technology*, *17*(4), 369–375. https://doi.org/10.1080/17483107.2020.1785564

Saha, A., McHugh, T. B., & Piper, A. M. (2023). Tutoria11y: Enhancing Accessible Interactive Tutorial Creation by Blind Audio Producers. *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–14. https://doi.org/10.1145/3544548.3580698

Sharif, A., McCall, A. L., & Bolante, K. R. (2022). Should I Say "Disabled People" or "People with Disabilities"? Language Preferences of Disabled People Between Identity- and Person-First Language. *The 24th International ACM SIGACCESS Conference on Computers and Accessibility*, 1–18. https://doi.org/10.1145/3517428.3544813

Storytelling to Sensemaking: A Systematic Framework for Designing Auditory Description Display for Interactives | Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. (n.d.). Retrieved October 24, 2023, from https://dl.acm.org/doi/abs/10.1145/3313831.3376460

Tigwell, G. W. (2021). Nuanced Perspectives Toward Disability Simulations from Digital Designers, Blind, Low Vision, and Color Blind People. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–15. https://doi.org/10.1145/3411764.3445620

Thank you for hearing me today