

## Milestone IV: Evaluation and Recommendations

NORAH RIDLEY, ANA-PIETJE DU PLESSIS, MAIZE MUGOT, DECLAN HILLS, and DEREK KUNKEL

### ACM Reference Format:

Norah Ridley, Ana-Pietje Du Plessis, Maize Mugot, Declan Hills, and Derek Kunkel. 2023. Milestone IV: Evaluation and Recommendations. 1, 1 (April 2023), 12 pages. <https://doi.org/10.1145/nnnnnnn.nnnnnnn>

---

Authors' address: Norah Ridley, [norah.ridley@usask.ca](mailto:norah.ridley@usask.ca); Ana-Pietje Du Plessis, [ap.duplessis@usask.ca](mailto:ap.duplessis@usask.ca); Maize Mugot, [mariaisabelle.mugot@usask.ca](mailto:mariaisabelle.mugot@usask.ca); Declan Hills, [declan.hills@usask.ca](mailto:declan.hills@usask.ca); Derek Kunkel, [ddk960@mail.usask.ca](mailto:ddk960@mail.usask.ca).

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2023 Association for Computing Machinery.

Manuscript submitted to ACM

Manuscript submitted to ACM

## 1 PROBLEM AND MOTIVATION

For many music creators, a problem they would often encounter is involving many applications into developing music, causing a disjointed workflow. For example, for a creator with an iPhone, they could be using the Notes app to write lyrics, the Voice Memo app to record specific instruments, the camera to record instrument setups, and so on. Additionally, these applications do not offer an usable interface to gather these files into one place, much less sharing with other people and marking versions of a track. Solving this problem would mean a more streamlined process that makes creating music more focused on fun collaboration, rather than spending more time that should be needed organising disparate files into one place.

## 2 RELATED LITERATURE AND BACKGROUND

In a survey of musicians in Belgium and The Netherlands during the pandemic, Onderdijk et. al. [4] found that the majority of these artists had adopted alternative methods to live practicing and recording with other members. These methods were primarily based in recording individual parts and circulating them through online platforms. This research complements Settles and Dow's [5] findings, which suggested that online collaborations are successful in music when artists have similar interests, ample communication, and matching passions for music, which often leads to the formation of teams. In the case of Belgians and Netherlandsers, this successful online collaboration was the result of teams, or bands, already being formed. This move towards online, collaborative musical workspaces that transcend physical space limitations necessitates the creation of portable applications that are accessible, usable, and intuitive.

Novice users are often a target audience for creative support tools since they help users with limited experience navigate the difficult aspects of the creative process. Davis et. al. [2] developed a creativity support tool with automation that supports novice users during their experimentation with visual narratives. In addition to being able to tell the computer the amount of work they'd like to perform themselves, users can also dynamically offload to the computer work they would prefer not to do or lack the skill-set for. This support tool allows domain novices generate creative content without being restricted by their level of technical skill. Huang et. al. [3] introduced ChordRipple, a creativity support tool that uses an adapted neural network from natural language processing to makes chord recommendations. The aim of this tool is to assist novice composers in developing their compositions beyond common chord progressions. Amateur songwriters often struggle to creative distinctive and style-specific accompaniments, which can ultimately lead to frustration and project abandonment. In response to this problem, Chuan and Chew [1] proposed a hybrid system to generate style-specific accompaniments to melodies.

While our app aims to mitigate user frustration by providing a unified tool for the many aspects of songwriting, we focus on providing support to users of all experiences. Compared to other work, our app provides general support for users rather than providing support for a specific task.

## 3 DESCRIPTION OF YOUR SYSTEM

The redesign of our prototype leaned more vertically, specifically into the project functionality and collection of different files all in one place. This implementation aligns with our motivation to create a more streamlined application for music creators. We implemented a more in depth version of a project tree, with the ability to import a text file from another source. The decision to develop functionality into importing a text file was because of the tasks we outlined in Milestone 2, which meant we already had discussed a more detailed flow and feel of the task. More generally, we added a more final look to the overall tree screen, with four buttons that allowed adding text, recording, image and

sheet music files. Then, we added a “Riff Bank” component to the application, which is a space to collect and import files from other applications. We implemented the Riff Bank with the idea of supporting different file types in mind to eliminate the task of switching between applications. That way, it is much easier to organize versions of a song if everything is in one place.

#### 4 EVALUATION WITH USERS REPORT

We followed a multi-pronged approach for user evaluation; combining multiple questionnaires to gather quantitative data regarding user experience with a follow-up interview for the collection of qualitative information.

The questionnaires given to users gathered data on their ability to complete tasks, their satisfaction levels, their aesthetic design preferences, and their demographic information. These followed an “A-B” pattern, wherein a user was given one-of-two interfaces, completed several tasks with it, and answered a questionnaire on their experience using that interface. The user was then given a second, slightly different interface, completed the same task, and answered an additional questionnaire. These were used to find a favoured implementation. Both questionnaires left a blank section for users to input qualitative insights.

The follow-up interviews contained a series of open and close-ended questions, providing the user with an opportunity to provide dynamic feedback and verbalise their experiences using the app beyond the two tasks given. These interviews were recorded and transcribed to provide additional feedback. The interview questions were weighted towards open-ended, subjective inquiries to gather user insights.

Goals:

- To discover which of two music-project creation interface options our testers preferred.
- To quantitatively measure our systems usability, design, and user-satisfaction level. This helped us focus our project on user preferences as opposed to forcing them to adapt to our design, as well as providing proof-of-concept (or disproof-of-concept) for our project.
- To understand what improvements may be made to our system through qualitative, user-specific feedback.

##### 4.1 Participant Pool

This evaluation was undertaken by 6 musicians ranging from ages 25 to 37 in Saskatoon. We had 2 additional testers booked, but were unable to meet up due to logistical constraints. These individuals were selected to represent a variety of musical genres, levels of performance experience, and identities.

##### 4.2 Rationale

The combination of observational methods and questionnaires used in this evaluation were used to capture a wide range of responses individual users may have to our prototype, quantifiable or otherwise. The combination of these methods was somewhat time-consuming for broad testing but worked well within our constrained group of 6 users.

We made a concentrated effort to gather qualitative data during the follow-up interview portion of the evaluation, and asked more open-ended questions to users during this time. The choice to focus on user-subjectivity at this point came at the expense of analytics-friendly data that could have been gathered from closed-questions. This lack of objective information was offset by the questionnaire, which provided ample qualitative information for objective system assessment.

Our user group consisted of 6 participants with the majority of users identifying as “Band member” (Figure 1).

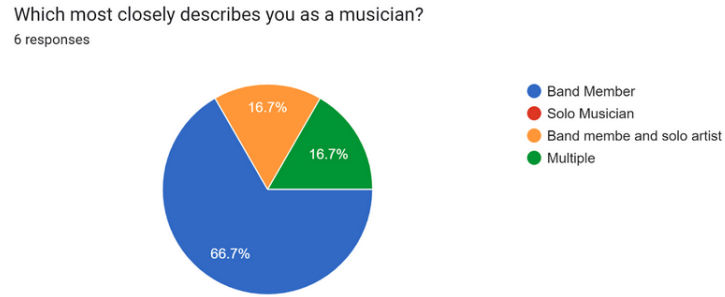


Fig. 1. Participant musician identification

All of our users said that they use their phone to record music, which suggests that there is a need for our app (Figure 2).

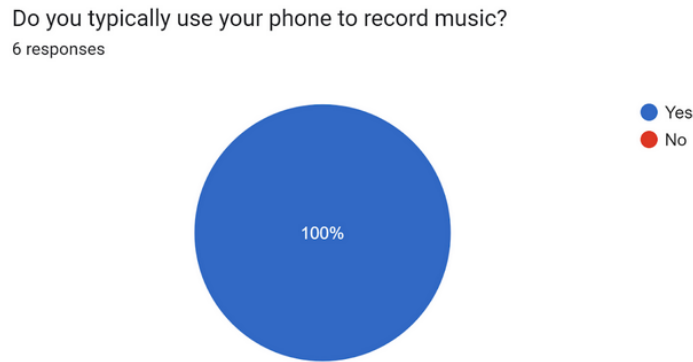


Fig. 2. Participants' use of phone for music recording

Half of the users completed the user evaluation starting with version A and then version B of the app while the other half started with version B then version A.

Table 3 suggests that users had a much stronger negative response to the tree extending upwards.

We asked the users to rate the app's ease of use ranging from "very difficult", "difficult", "neutral", "easy" and "very easy" for version A (Table 1) and version B (Table 2), respectively. Generally, users found the app easy to use. Satisfaction levels were equal for A and B, each receiving three "Very satisfied"s and three "Satisfied"s. This also indicates a high level of satisfaction among users.

All but one user preferred the first version they tested which indicates that recency bias was not a contributing factor in determining preference (Table 4. These results are also in line with preferences extracted during each phase.

Figure 5 indicates that most users found the app easy or very easy to navigate with one user finding it difficult.

All of the users expressed a desire to use the app after completing their evaluation (Figure 6. All of the users indicated an interest in using the app frequently (Figure 7

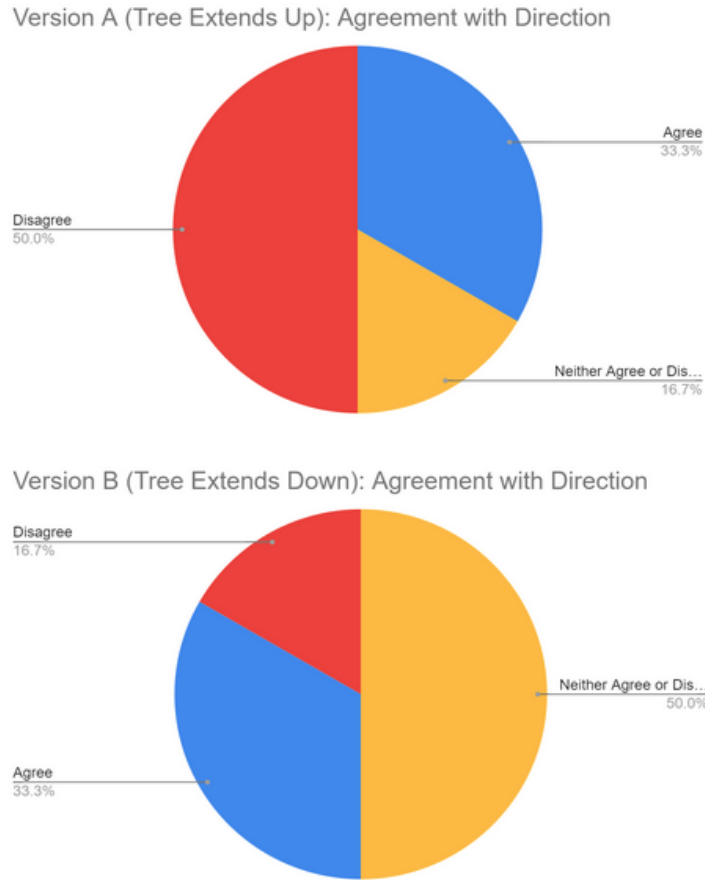


Fig. 3. Participants' agreement with the tree direction

Table 1. Indication of ease of use (version A)

| Question  | Very difficult | Difficult | Neutral | Easy | Very easy |
|---|----------------|-----------|---------|------|-----------|
| How easy is it to find a project?   | 0              | 0         | 0       | 4    | 2         |
| How easy is it to start a new project?  | 0              | 0         | 1       | 4    | 1         |
| How easy is it to create a new element?                                       | 0              | 0         | 1       | 3    | 2         |
| How easy is it to move an existing element from outside the app to a project? | 0              | 1         | 2       | 2    | 1         |

Table 2. Indication of ease of use (version B)

| Question  | Very difficult | Difficult | Neutral | Easy | Very easy |
|---|----------------|-----------|---------|------|-----------|
| How easy is it to find a project?   | 0              | 0         | 1       | 2    | 3         |
| How easy is it to start a new project?  | 0              | 0         | 2       | 3    | 1         |
| How easy is it to create a new element?                                       | 0              | 0         | 2       | 1    | 3         |
| How easy is it to move an existing element from outside the app to a project? | 0              | 1         | 3       | 0    | 2         |

Figure 8 users did indicate that there was room for the tool to grow in terms of additional features even though they would use the app.

Which version did you prefer?

6 responses

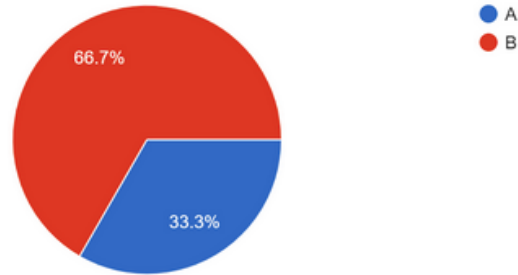


Fig. 4. Participants' app version preferences

How easy is it to navigate through the app?

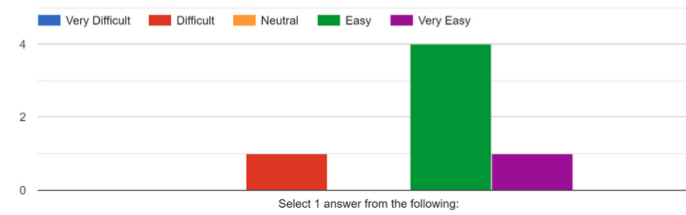


Fig. 5. Ease of navigation through the app

Would you use this app?

6 responses

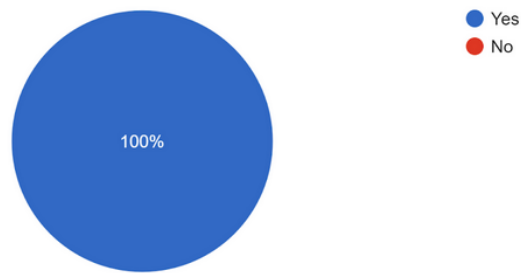


Fig. 6. Participants' desire to use the app

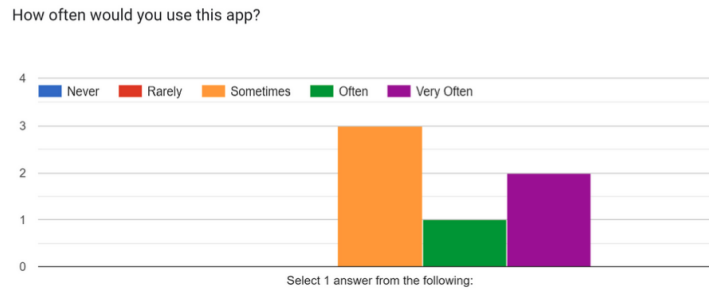


Fig. 7. Participants' projected app use frequency

Do you think Whole Node has all the features you need from a tool to assist you in tracking changes in your music over time?

6 responses

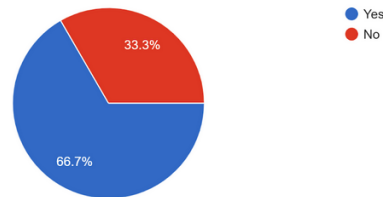


Fig. 8. Participants' opinion on the app's features

## 5 FINAL RECOMMENDATIONS

### 5.1 Formulate Final Project Conclusions and Design Recommendations

#### General outcome: (b) minor adjustments needed, overall approach validated

User testing confirmed that there is a need for a tool that allows musicians to store/organise files in a centralised location. Our evaluation also highlighted the following points:

- User testing indicated that the design is:
  - Easy to use (Easiness score / "I quite liked it, it makes it really easy to look at each change somebody's made")
  - Includes useful features
  - Something participants would use
  - A solution to the users' problem of coordinating media files
- Though users indicated a preference for project trees starting at the top and moving down, there was one user that strongly preferred the opposite (based on comments), this indicates that we may want to explore giving users the option for either. However, since a user may never realise they have a preference for that structure if never presented with this option, we may wish to do further testing in which users only see one version of the interface.

- 100% of participants would use the app “Sometimes” to “Very Often”, but 33.3% of users indicated that the app does not contain all the features they would like.
- The following answers to “What is your favourite feature of the app?” demonstrate that users appreciate the collaborative nature of the tool and the flexibility of use in creating and organising creative elements.
  - “Importing text from notes”
  - “Ability to organize progression of ideas collaboratively”
  - “Riff bank, having somewhere to collect random ideas and being able to easily pull them into projects is fantastic”
- The following answers to “What is your least favourite feature of the app?” demonstrate that the aesthetics and perceived affordances need refining.
  - “The start new project button isn’t labeled”
  - “A lil crashy tbh”
  - “Buttons dont look like buttons”
- The following answers to “What is your least favourite feature of the app?” demonstrate that we need to improve the user’s ability to navigate in the Riff Bank.
  - “Riff bank not being organized”
  - “No zoom in or out on riff bank”

There were limitations to our evaluation. No solo musicians were tested, as all users were also in a group. Therefore we cannot determine if this tool would be useful for solo artists without further testing. Furthermore, we were only able to organise six user evaluations, and these participants were mostly male. This is not a sufficient number of participants to generalize preferences to all musicians.

Overall, user feedback suggests that the quality of the design is high for a medium fidelity prototype and generated useful feedback for further design iterations and testing.

While our overall approach was successful, there are some general concepts and more detailed steps that should be taken in order to increase the quality of the application. For the most part, we need to focus on feedback and documentation. Our users had a somewhat difficult time because our interactive elements, such as the red button in the landing page, did not indicate well that they were tappable or clickable. In addition, some of our design elements afforded a different function than what actually was given. For example, the button for attaching sheet music was a purple button with a yellow eighth note, which was unrecognizable to our testers. Proper documentation, feedback and addition of appropriate idioms would help users more easily understand and adapt to the application. Finally, while not directly relevant to our main motivation, we could explore design elements such as which color scheme is beneficial for people with and without colour-vision deficiency to increase accessibility, and applying aesthetics from established apps like the Voice Memo app for easier adaptation.

For more specific improvements and additions, there were several elements that we identified from our analysis as well as suggestions from users. Since the results from our A/B testing concluded that there was no meaningful difference between the direction of the project tree, but that everyone had preferences, there could be an option to let the user decide the direction in the settings page. We also want to flesh out functionality for connecting the Riff Bank and the projects together; more specifically we want to add files in the Riff Bank to projects. In terms of user feedback, a way to immediately reverse decisions such as deleting or adding files would be beneficial for user control. Riff Bank navigation could be improved by adding zoom in/zoom out functionality, or changing the structure of the



Riff Bank to be more rigid, like a list. We could expand the types of files that can be imported into the application; testers identified snippets of an existing song on services such as Spotify, and digital audio workstation (DAW) files. Random name generation should be added with project creation to help people who could struggle with creating a meaningful project name. Plagiarism checkers could be added as a way for songwriters to help avoid copyright. Finally, there should be a way to directly edit sheet music files into the application for convenience.

In relation to our user evaluation plan, there are two improvements that we could take to enhance our approach to user testing. The first improvement would be expanding our user sample size. Our testing group was limited in that we only could gather six people, and they were all a part of a band. Expanding our research to include songwriters who are not part of one singular group could help generalize the results that we collected from this round of user evaluation. The second improvement would be to make the structure of the demographic questions more closed, such as implementing the Likert scale. Some of the questions in that section ended up receiving results that were difficult to quantify, so enforcing a scale would be beneficial for gathering analyzable data.

## 5.2 Reflect on Your Design Process

The next iteration of our design would include changes influenced by user feedback. There was a strong leaning within the group towards trees that extend upward. This conflicted with user feedback that the preferred direction for the trees was downward. However, this was exactly what we hoped to determine by conducting A/B Testing. Additionally, we designed the Riff Bank as an unorganised space to help express the freedom to create elements without constraints - something akin to a desk space where you can organise items according to your liking to support your creativity. Feedback from the users demonstrates that they would like this space to be more constrained and more consistent with organisation in other parts of the system.

Our choice of evaluation methods helped us to answer the questions we set out to explore:

- Would this tool be useful? Yes
- Does a tree structure work for tracking changes in musical projects? Yes
- Which direction should the tree extend? Down

Additionally, Figma was very useful as a mobile tool that we could use to conduct evaluations in the wild. We would have preferred to gather data from a larger test group, however there was limited time in which to schedule user evaluations. If using an expanded participant pool, we would aim to better reflect the age and gender distributions of creatives in the music industry, and we would have liked to include solo musicians to determine if the tool is perceived as useful by people who won't be collaborating with others.

Sketching was a very valuable part of the process as it enforced early walkthroughs that highlighted features to include that we hadn't previously considered. Getting user feedback similarly provided insights into use cases we wouldn't have considered.

In future interface designs, we would aim to condense the low-fidelity/sketching portion of the process to leave more time for user evaluations. Prototyping within a team can be very rapid, however adding external parties adds friction to the process and requires more flexibility in scheduling.

## REFERENCES

- [1] Ching-Hua Chuan, Elaine Chew, et al. 2007. A hybrid system for automatic generation of style-specific accompaniment. In *Proceedings of the 4th international joint workshop on computational creativity*. Goldsmiths, University of London London, 57–64.

- [2] Nicholas Davis, Holger Winnemöller, Mira Dontcheva, and Ellen Yi-Luen Do. 2013. Toward a Cognitive Theory of Creativity Support. In *Proceedings of the 9th ACM Conference on Creativity and Cognition* (Sydney, Australia) (*CampC '13*). Association for Computing Machinery, New York, NY, USA, 13–22. <https://doi.org/10.1145/2466627.2466655>
- [3] Cheng-Zhi Anna Huang, David Duvenaud, and Krzysztof Z. Gajos. 2016. ChordRipple: Recommending Chords to Help Novice Composers Go Beyond the Ordinary. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (Sonoma, California, USA) (*IUI '16*). Association for Computing Machinery, New York, NY, USA, 241–250. <https://doi.org/10.1145/2856767.2856792>
- [4] Kelsey E. Onderdijk, Freya Acar, and Edith Van Dyck. 2021. Impact of Lockdown Measures on Joint Music Making: Playing Online and Physically Together. *Frontiers in Psychology* 12 (2021). <https://doi.org/10.3389/fpsyg.2021.642713>
- [5] Burr Settles and Steven Dow. 2013. Let's Get Together: The Formation and Success of Online Creative Collaborations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) (*CHI '13*). Association for Computing Machinery, New York, NY, USA, 2009–2018. <https://doi.org/10.1145/2470654.2466266>

## APPENDICES

### A Questionnaire

**(Do this section twice - once for each version)**

Immediate Recall Questions:

- (1) Which version did you do? [A or B]
- (2) When adding new elements (audio, images, etc.) to a project, how much did you enjoy the design? [Likert scale]
- (3) When adding new elements (audio, images, etc.) to a project, I agree with the direction (up or down) of adding a new element: [Likert scale]
- (4) How easy is it to find a project? / It is easy to find a project. [Likert scale]
- (5) How easy is it to start a new project? [Likert scale]
- (6) How easy is it to create a new element? [Likert scale]
- (7) How easy is it to move an existing element from outside the app to a project? [Likert scale]
- (8) How much did you enjoy your experience? [Likert scale]
- (9) **(Ask below after both versions have been completed)** Preference: Which version did you prefer? [First or second]

More General Questions

- (1) What is your favourite feature of the app?
- (2) What is your least favourite feature of the app?
- (3) Do you think Whole Node has all the features you need from a tool to assist you in tracking changes in your music over time?
- (4) Are there other features you would like to see?
- (5) Did you enjoy the colour scheme?
- (6) How easy is it to navigate through the app? [Likert scale]
- (7) Would you use this app?
- (8) How often would you use this app? [Likert scale]
- (9) Do you have any other comments about the app?

Demographic Questions

- (1) Which most closely describes you as a musician? [Band member, solo musician, other]
- (2) How many hours per week do you spend writing songs or developing music on your own?
- (3) How many hours per week do you spend working with other musicians?

- (4) Do you typically use your phone to record music?
- (5) How many hours per week do you use your phone to develop music?
- (6) How many hours per week do you use a phone or a computer for work or leisure?
- (7) What is your gender? [Man, woman, non binary, prefer not to say, not listed above]
- (8) If your gender was not listed, please enter it below:
- (9) What is your age?

## B Raw data

Timestamp,1. Which version did you do?,"When adding new elements (audio, images, etc.) to a project, how much did you enjoy the design? [Select 1 answer from the following:],"When adding new elements (audio, images, etc.) to a project, I agree with the direction (up or down) of adding a new element: [Select 1 answer from the following:],"When adding new elements (audio, images, etc.) to a project, how much did you enjoy the design? [Select 1 answer from the following:],"How easy is it to find a project? [Select 1 answer from the following:],"How easy is it to start a new project? [Select 1 answer from the following:],"How easy is it to create a new element? [Select 1 answer from the following:],"How easy is it to move an existing element from outside the app to a project? [Select 1 answer from the following:],"How much did you enjoy your experience? [Select 1 answer from the following:],"Which version did you do?";2. When adding new elements (audio, images, etc.) to a project, how much did you enjoy the design? [Select 1 answer from the following:],"When adding new elements (audio, images, etc.) to a project, I agree with the direction (up or down) of adding a new element: [Select 1 answer from the following:],"When adding new elements (audio, images, etc.) to a project, how much did you enjoy the design? [Select 1 answer from the following:],"How easy is it to find a project? [Select 1 answer from the following:],"How easy is it to start a new project? [Select 1 answer from the following:],"How easy is it to create a new element? [Select 1 answer from the following:],"How easy is it to move an existing element from outside the app to a project? [Select 1 answer from the following:],"How much did you enjoy your experience? [Select 1 answer from the following:],"Which version did you prefer?,"What is your favourite feature of the app?,"What is your least favourite feature of the app?,"Do you think Whole Node has all the features you need from a tool to assist you in tracking changes in your music over time?,"Are there other features you would like to see?,"Did you enjoy the colour scheme?,"How easy is it to navigate through the app? [Select 1 answer from the following:],"Would you use this app?,"How often would you use this app? [Select 1 answer from the following:],"Do you have any other comments about the app?,"Which most closely describes you as a musician?,"How many hours per week do you spend writing songs or developing music on your own?,"How many hours per week do you spend working with other musicians?,"Do you typically use your phone to record music?,"How many hours per week do you use your phone to develop music?,"How many hours per week do you use a phone or a computer for work or leisure?,"What is your gender?,"If your gender was not listed, please enter it below:","What is your age?","", 4/4/2023 18:52:50,A,Satisfied,Agree,Very Satisfied,Very Easy,Easy,Very Easy,Neutral,Very satisfied,B,Satisfied,Agree,Satisfied,Very Easy,Easy,Very Easy,Very Easy,Very Easy,Very satisfied,A,The simplicity of starting a new project with the red button,The start new project button isn't labeled, Yes,Refreshing button in case you deleted something you didn't want to,Yes,Easy,Yes,Very Often, You should have a random band name generator for new bands having hard time finding a name, Band member and solo artist,6/9hrs,2 to 6 hrs depending if we have shows in the future,Yes,3hrs, Way to much,Male,,36,,", 4/4/2023 22:32:13,A,Satisfied,Agree,Satisfied,Easy,Easy,Easy,Easy, Very satisfied,B,Neutral,Neither Agree or Disagree,Neutral,Neutral,Neutral,Neutral,Neutral,Very satisfied,A, When it went up,When it went down,Yes,Plagiarism checker,Yes,Very Easy,Yes,Often,Up. Not down. Also help me not plagiarize. Thank

you.,Band Member,2,3,Yes,10 mind total,Unquantifiable. Very large amount.,Male,,25,,,,, 4/4/2023 23:24:52,A,Very satisfied,Neither Agree or Disagree, Satisfied,Very Easy,Very Easy,Very Easy,Very Easy,Very satisfied,B,Very satisfied,Neither Agree or Disagree,Very Satisfied,Very Easy,Very Easy,Very Easy,Very Easy,Very Easy,Very satisfied,B,"Riff bank, having somewhere to collect random ideas and being able to easily pull them into projects is fantastic",No zoom in or out on riff bank,Yes,Blank sheet music where you could basically write out parts,Yes,Easy,Yes,Very Often,,Band Member,0-40,0-6,Yes,0-6,Constantly,Male,,27,,,,, 4/4/2023 22:17:01,B,Satisfied,Neither Agree or Disagree,Neutral,Easy,Easy,Easy,Neutral,Satisfied,A,Satisfied,D,Satisfied,B,Importing text from notes,Riff bank not being organized,Yes,Ability to reference snippets of songs in Spotify / Apple Music etc.,No,Easy,Yes,Sometimes,,Band Member,6,2,Yes,02-Jan,30 +,Male,,27,,,,, 4/4/2023 23:00:35,B,Satisfied,Agree,Satisfied,Very Easy,Easy,Very Easy,Neutral,Satisfied,A,Neutral,Disagree,Satisfied,Easy,Easy,Easy,Easy,Satisfied,B,Ability to organize progression of ideas collaboratively,A lil crashy tbh,No,Tracking changes DAW files,No,Easy,Yes,Sometimes,,Band Member,10 hrs,3-4 hrs,Yes,Very,30 hrs,Prefer not to say,,26,,,,, 4/5/2023 0:01:13,B,Satisfied,Disagree,Satisfied,Easy,Neutral,Neutral,Difficult,Satisfied,A,Easy,Easy,Neutral,Difficult,Satisfied,B,Pretty colors,Buttons dont look like buttons,No,Yes,Yes,Difficult,Yes,Sometimes,Buttons don't look clickable. Shadows or indentations? Record button is in the way and always on screen. Why? Want to add more than one thing to node. Whole note should go something on tree page? Plus sign doesn't look like see more details.,Multiple,Lots,3,Yes,Not very many,Many,Prefer not to say,,27,,,,,