# Evaluating an Auditory Display System for Ballet Training

Jocelyn Kavanagh
Georgia Institute of Technology
North Ave NW
Atlanta, Georgia 30332
Jkavanagh3@gatech.edu

## ABSTRACT
This study investigated the clarity and appropriateness of a dance training sonification system. 3 intermediate dancers participated in a training phase, a 2-trial evaluation phase, and an interview phase. Response time, accurate interpretations of the sonification, and NASA TLX results were recorded. The primary research question was *Is the sonification clear, easy to understand, and appropriate enough to provide use in dance training?* The study showed that the sonification system may be useful, but that it needs some modifications and further study. After two trials, the participants were able to interpret about 79.5% of the information presented in the sonification, and 2 of the 3 participants thought this would be a useful tool in the studio.

## CCS Concepts
CCS → Human-centered computing → Human computer interaction (HCI) → Interaction devices → **Sound-based input / output**

## Keywords
Ballet; sonification; body awareness; training; auditory display

## 1. INTRODUCTION
This study is an evaluation of a sonification system for dance training. The primary research question is **(RQ1) Is the sonification clear, easy to understand, and appropriate enough to provide use in dance training?** This question is answered by the culmination of the subsequent trials and interviews outlined in the study. I also proposed more pointed research question about the clarity of the sonification. **(RQ2) Will the dancer be able to distinguish between the different muscle group sonifications?** Related to this question **(RQ3) Will the sonification be too difficult to interpret in real-time?** And, I proposed a more qualitative question, **(RQ4) What are the dancer's aesthetic opinions about the sonification, and do they expect it to be helpful?**

## 2. METHODS
### 2.1 Participants
I recruited 3 intermediate dancers to participate in the study. All participants were young women that have studied dance for more than 5 years and are actively studying dance at a post-secondary institution.

### 2.2 Setting
The study was conducted in person following COVID-19 protocol as suggested by the CDC.

### 2.3 System
#### *2.3.1 Equipment*
The sonification system is hosted on Processing. Each study took place on a MacBook Pro using Sennheiser over-ear headphones. All participants were recorded using the recording app on a Google Pixel 2.

#### *2.3.2 Sonification Description*
The sonification system is meant to support dancers while training ballet fundamentals. As such, the system sonifies the muscular data of 4 muscle groups, weight distribution in the XYZ plane, and body alignment. Muscular contraction data is mapped to a group of sine and sawtooth waves, where each muscle has a discrete pitch and the contraction level is represented by the amount of sawtooth wave present in the mix (the more contracted, the louder the sawtooth wave). Weight distribution is mapped to 3 manipulations of those sine/sawtooth waves. The x-direction is represented by left-right spatial panning, the y-direction is represented by the presence and frequency of an amplitude modulator, and the z-direction is represented by the overall loudness of the mix. Body alignment (or posture) is mapped to the presence and volume of an additional auditory icon (as alignment gets worse, the tone sounds louder). Finally, the system also includes text-to-speech alerts for potentially at-risk muscle groups due to their tension levels.

Outside of the mappings, the system simulates two modes for the intended user, a full mode and a minimal intervention mode. The full mode includes all of the mappings detailed above and is intended for use while focusing on fundamental movements without external music. Minimal intervention mode takes away the sounds of the muscular contraction sine/sawtooth waves and leaves only the alerting sounds. This mode is intended for use outside of focused fundamental training and could be used with external music.

In a fully implemented state, the system should be used continuously. The dancer will rehearse fundamental skills while hearing the movement sonified in real-time. The dancer will reference the sound of their individual "perfect" form for a movement as given by an instructor, and they will attempt to match their movement sounds to that ideal.

#### *2.3.3 Sonification Controls*
The study showcased all available aspects of the sonification simulator, including Wizard-of-Oz controls and JSON loaded events.

The Wizard-of-Oz technique was used in the study to help the participants learn the mappings of the system. A setting in the simulator offers the experimenter control of the various mappings both in and out of context of the sonification as a whole.

2 sets of 5 JSON loaded events were used for the evaluation portion of the study. Two fundamental ballet skills, plies (like a squat) and sautés (like a vertical jump), were sonified through the 2 JSON event sets. For both plies and sautés, a "perfect" sonification was given as well as 4 variations. To simulate the comparison and thought process needed to adjust between proper and improper form, the 4 variations added discrepancies to each "perfect" JSON event. The discrepancies were added both

randomly and with common dance issues in mind, and, across the sets, the number of discrepancies ranged from 1-10.

The JSON events were triggered by the experimenter via labeled buttons, and each event was paired with a matching video of the fundamental movement. The JSON events were played both in and out of minimal intervention mode to highlight both intended aspects of the system, as well as to aid in answering the research questions.

### *2.3.4 Procedure*

The recruited participants heard a brief description and introduction of the sonification system. They were then briefed on the purpose of the study and potential risks, and they filled out consent documentation.

The study represents a within-subjects design, where each participant completed two sonification trials, a NASA TLX survey, and a post-task interview.

After gaining consent, the study began with a training phase to teach the participants how to interpret the auditory displays. The experimenter explicitly explained and demonstrated the system mappings via the simulator's Wizard-of-Oz controls.

Following the training phase, the participants began the first trial to evaluate the system. To begin, the experimenter triggered the "perfect" plie event 3 times in succession for the participants reference. Then, the experimenter triggered each subsequent variation of the plie. To test the participants' understanding of the sonifications and their ability to recognize improper form, they were asked to think aloud and describe how the subsequent events varied from the "perfect" plie. After completing the plie events, the same procedure was followed for the sauté events.

The next trial during the evaluation phase tested the participants ability to recognize the differences in the muscle contraction sonification specifically. Beginning again with plies, the experimenter went back and forth triggering the "perfect" plie and the improper events. The participants were asked to describe how the subsequent events varied from the plie, specifically in relation to the contraction level changes. After completing the plie events, the same procedure was followed for the sauté events.

The final trial of the evaluation phase tested the minimal intervention mode and was more qualitative. All events with tension and/or alignment alerts were played, and the participants were asked to describe what they heard. They were also asked their overall impressions of this mode of the system.

The interview phase was the final phase of the study. Participants were asked to complete the NASA Task Load Index to help assess the difficulty of interpreting the sonification. The participants were then asked the following interview questions (see Table 1). The participants were then asked to share any final comments and were debriefed.

**Table 1. Interview Questions**

| **Interview Questions** |
|---|
| *Please discuss the difficulty of the task. Were any aspects too hard to understand or frustrating?* |
| *Did you find the system to be overwhelming?* |
| *Would you see this providing use in the studio? Would it be distracting?* |
| *Did you enjoy the sounds of the sonification? Were any annoying or intrusive?* |
| *What were the most helpful parts of the system? Least helpful? What were the clearest aspects, and what were the most confusing?* |

### *2.3.5 Measures*

During the training phase, the elapsed time to complete was recorded to assess how difficult the system was to learn and understand. The pace of the training phase was set by the participants, and the experimenter followed a script for consistency across participants.

During the evaluation phase, the participants' answers were recorded and their percentage correct was calculated. For example, if there were 2 discrepancies and the participant reported 1, their score for that event would be .5. The participants' answer response time was also recorded to assess how easy it was to interpret the sonification in real-time.

During the interview phase, the NASA TLX scores were recorded, and the interview portion of the phase was audio recorded for future review.

## 3. RESULTS

In this section, I will present my findings in relation to my proposed research questions.

**RQ1: Is the sonification clear, easy to understand, and appropriate enough to provide use in dance training?** And

**RQ2: Will the dancer be able to distinguish between the different muscle group sonifications?**

Below are the raw results and means from Trial 1 (Table 2) and Trial 2 (Table 3) of the Evaluation phase. These results show the accuracy with which the participants understood the sonification. The negative numbers shown in parentheses show the number of incorrect statements made about the sonification.

**Table 2. Evaluation Trial 1. For each variation test from "perfect" form, the number of discrepancies (B), the number of discrepancies found by the participant (F), and the percentage found (%) were recorded.**

| | | P 1 | | P 2 | | P 3 | | Mean |
|---|---|---|---|---|---|---|---|---|
| | B | F | % | F | % | F | % | |
| Plie Test 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Plie Test 2 | 3 | 3 | 1 | 1 | .333 | 2 | .667 | .667 |
| Plie Test 3 | 5 | 3 | .6 | 3 | .6 | 3 | .6 | .6 |
| Plie Test 4 | 9 | 5 | .556 | 3 | .333 | 4 | .444 | .444 |
| Sauté Test 1 | 5 | 3 (-1) | .6 | 3 | .6 | 4 | .8 | .667 |
| Sauté Test | 1 | 1 | 1 | 1 (-1) | 1 | 1 (-2) | 1 | 1 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 2 | | | | | | | | |
| Sauté Test 3 | 2 | 2 | 1 | 0 | 0 | 1 | .5 | .5 |
| Sauté Test 4 | 10 | 7 | .7 | 5 | .5 | 4 | .4 | .533 |
| Mean | | | .807 | | .546 | | .676 | |

**Table 3. Evaluation Trial 2. For each variation test from "perfect" form, the number of discrepancies (B), the number of discrepancies found by the participant (F), and the percentage found (%) were recorded.**

| | | P 1 | | P 2 | | P 3 | | Mean |
|---|---|---|---|---|---|---|---|---|
| | B | F | % | F | % | F | % | |
| Plie Test 1 | 1 | 1 | 1 | 1 (-1) | 1 | 1 (-2) | 1 | 1 |
| Plie Test 2 | 3 | 3 | 1 | 2 | .667 | 3 (-1) | .1 | .889 |
| Plie Test 3 | 5 | 4 | .8 | 3 | .6 | 3 | .6 | .667 |
| Plie Test 4 | 9 | 8 | .889 | 3 | .333 | 4 | .444 | .556 |
| Sauté Test 1 | 5 | 3 | .6 | 3 | .6 | 4 | .8 | .667 |
| Sauté Test 2 | 1 | 1 | 1 | 1 (-1) | 1 | 1 | 1 | 1 |
| Sauté Test 3 | 2 | 2 | 1 | 2 | 1 | 2 (-1) | 1 | 1 |
| Sauté Test 4 | 10 | 7 | .7 | 7 | .7 | 4 | .4 | .6 |
| Mean | | | .851 | | .756 | | .774 | |

Below are the training times (Table 4) recorded for each participant.

**Table 4. Elapsed time during training phase in minutes.**

| | Participant 1 | Participant 2 | Participant 3 | Mean |
|---|---|---|---|---|
| Time | 5 | 6.2 | 6.733 | 5.978 |

**RQ3: Will the sonification be too difficult to interpret in real-time?**

Below are the raw results of the answer response time for each of the participants for both Trial 1 (see Table 5).

**Table 5. Participant Response time for Trial 1 in seconds.**

| | Participant 1 | Participant 2 | Participant 3 | Mean |
|---|---|---|---|---|
| Plie Test 1 | .5 | 5 | 16 | 7.167 |
| Plie Test 2 | 2 | 7 | 1 | 3.333 |
| Plie Test 3 | 1 | 3 | .5 | 1.5 |
| Plie Test 4 | .5 | 5 | .5 | 3 |
| Sauté Test 1 | 3 | 5 | .5 | 2.833 |
| Sauté Test 2 | 2 | 4 | .5 | 2.167 |
| Sauté Test 3 | 1 | 3 | .5 | 1.5 |
| Sauté Test 4 | 3 | 8 | 1.5 | 4.167 |
| Mean | 1.625 | 5 | 2.625 | |

Below are the NASA TLX scores from each participant (see Table 6).

**Table 6. NASA TLX Ratings where R is the rating per category, W is the weight of the category, and A is the adjusted rating.**

| | Participant 1 | | | Participant 2 | | | Participant 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | R | W | A | R | W | A | R | W | A |
| Mental | 65 | 4 | 260 | 70 | 5 | 350 | 60 | 1 | 60 |
| Temporal | 80 | 4 | 320 | 30 | 3 | 90 | 30 | 5 | 150 |
| Performance | 30 | 1 | 30 | 50 | 4 | 200 | 40 | 4 | 160 |
| Effort | 70 | 3 | 210 | 70 | 1 | 70 | 50 | 2 | 100 |
| Frustration | 35 | 3 | 105 | 30 | 2 | 60 | 70 | 3 | 210 |
| Rating | | | 61.67 | | | 51.33 | | | 45.33 |

**RQ4. What are the dancer's aesthetic opinions about the sonification, and do they expect it to be helpful?**

Below are highlights and excerpts from the interview phase of the evaluation.

*The callouts* (in reference to the text-to-speech and auditory icon) *are helpful, but the tones* (in reference to the sine/sawtooth waves) *are hard to single down.*

*Lots of information presented would be hard to adapt to while rehearsing.* (Spoken in reference to Plie and Sauté 4)

*The jumping shouldn't be silenced. You know when you are in the air, so it would be more helpful to hear your legs to prepare for landing.*

*If you aren't really musical, it can be hard to hear the different muscle contractions. And, it could take a while to internalize.*

Multiple times during the evaluation the participants expressed physically and verbally a struggle with recognizing the different pitches and contraction levels of the muscle groups. When an

unexpected muscle group was alerted through TTS (like core), the participants initially could not understand and were confused. The participants had no negative comments about the aesthetic quality of the sonification. 2 of the 3 participants said they could see using this system while training.

# 4. DISCUSSION

In this section, I will discuss my findings in relation to the proposed research questions.

**RQ1: Is the sonification clear, easy to understand, and appropriate enough to provide use in dance training?**

According to the results, it appears that the system was moderately successful. In reference to Table 2, the participants were able to interpret about half of the information presented in the sonifications during the Trial 1. Then, in Trial 2 (see Table 3), after a reminder of the "perfect" form, the participants were able to interpret more information from the system, over 70%. When breaking down the results, in both Trial 1 and Trial 2, the participants easily identified the alignment auditory icon, the TTS alerts, and the weight distribution information. By far, it was hardest for the participants to understand the change in contraction levels indicated by the sine/sawtooth mixes. These observations in scores are also backed by the comments made by the participants. The participants enjoyed hearing the alert sounds, but they were frustrated by the muscular contraction tones.

The participants took an average of 5.978 minutes to learn how to use the system. This could show how easy it is to understand and learn how to use, however, more exploration would be necessary. There was no baseline of control for a reference length in the study, the participants may have felt rushed to complete the training phase, and the participants were not widely successful in Trial 1 so they may not have understood the sonifications to the fullest level.

**RQ2: Will the dancer be able to distinguish between the different muscle group sonifications?**

As observed in the participants accuracies from Trials 1 and 2 and from their comments, the muscular contraction sonification through the sine/sawtooth waves was hard to understand. After a reminder of the "perfect" form in Trial 2, some participants were able to hear the difference in the contraction levels, but it was still the most missed category. Additionally, some of those correct answers may have been from context or luck. This may be true because any incorrect guess in Trials 1 and 2 was usually relating to the muscle contraction.

As some participants suggested, it may be too difficult to interpret the muscle contraction information simultaneous to the other alerts and manipulations. Perhaps the simulator should have isolated the muscle group contraction for a cleaner study to compare whether or not participants could detect a difference.

**RQ3: Will the sonification be too difficult to interpret in real-time?**

The response time data from the trials (see Table 5) suggests that the system may be too complex or hard to understand in real-time. With a range of 0.5-7s, the time needed for processing may be too great for use in real-time. However, this data may not be the best representation of the participants interpretation in real-time. The participants may have recognized features in real-time but had to recall them when giving their answer, or the participants may have needed time to think about some aspects even though they immediately interpreted others.

A better way to study the real-time aspect would be by having dancers actually use the system or by slowing the event data. Some participants mentioned that they thought they could adapt to and learn from the system if they were physically controlling it and if they could move slowly. By slowing down the event data, the participants could state the changes they recognized in real-time followed by a summary of all the features they interpreted after the sonification completed.

**RQ4. What are the dancer's aesthetic opinions about the sonification, and do they expect it to be helpful?**

The participants had no negative remarks during the interview phase about the sounds of the system. During Trials 1 and 2, when the discrepancy count was higher, some participants mentioned feeling slightly overwhelmed by the amount of information in such little time. Additionally, the participants initially had a difficult time understanding the TTS alerts. One participant mentioned that the sonification of the sauté should not have used an envelope to control the Z direction of a jump. Finally, the participants shared that the alerting sounds were most helpful, but they would like more information from them. It was hard for the participant to tell whether or not the alert was due to too much or too little tension, which they found confusing and misleading. However, overall, I think the participants were neutral toward the sonification sound design.

2 out of the 3 participants thought the system would be helpful for use in the studio. All thought that if they had more time to play with the system and physically work with it, they would have better understanding of the system's usefulness. The NASA TLX ratings show that the system was not extremely load intensive, suggesting that the system may be at least helpful in the background of the studio. However, with no baseline rating for each participant and so few participants, the ratings are not significantly meaningful.

# 5. CONCLUSION

The study showed that the sonification system may be useful, but that it needs some modifications and further study. After two trials, the participants were able to interpret about 79.5% of the discrepancies from "perfect" form, and 2 of the participants thought this would be a useful tool in the studio. The participants were able to recognize and enjoyed the auditory icon for alignment, the panning and amplitude modulation for weight distribution, and the TTS for tension flags. The changes in muscle contraction levels were the most difficult for the participants to interpret, so moving forward this aspect of the system should be amended.

In a future study, the current sonification of the muscles should be evaluated in isolation and compared to an improved sonification. Additionally, a different text-to-speech engine should be used and the envelope for Z direction should be removed. There should be an investigation to improve the detail and dimensionality of the alerting sounds, and a future study should incorporate the participants physically using the system.