Sonifying Gait Kinematics Using the Sound of Wading

A Study on Ecological Movement Representations

Kantan, Prithvi Ravi; Dahl, Sofia; Serafin, Stefania; Spaich, Erika G.

Published in:
The 28th International Conference on Auditory Display (ICAD 2023)

Publication date:
2023

Document Version
Publisher's PDF, also known as Version of record

SONIFYING GAIT KINEMATICS USING THE SOUND OF WADING: A STUDY ON ECOLOGICAL MOVEMENT REPRESENTATIONS

Prithvi Kantan
Sofia Dahl
Stefania Serafin

Aalborg University
A.C. Meyers Vænge 15
Copenhagen, Denmark
{prka,sof,sts}@create.aau.dk

Erika G. Spaich

Aalborg University
Selma Lagerløfs Vej 249
Gistrup, Denmark
espaich@hst.aau.dk

ABSTRACT

Interactive sonification of human movement can enhance motor learning by introducing an additional sensory feedback channel during rehabilitation. Successful past work on sonifying human walking has focused primarily on providing users with feedback on the stance (ground contact) phase of the gait cycle, whilst the equally important swing phase has been largely neglected. In this work, we developed and assessed a swing phase sonification algorithm that generates ecological feedback on limb swing in the form of wading sounds at two different depths. We evaluated the perceptual qualities of the algorithm output in a test with 16 respondents, as well as the user experience of 9 healthy participants in a real-life walking pilot test. Despite our simple approach of simulating wading acoustics using gyroscope-modulated liquid sounds, our results suggested that the simulation sounded perceptually natural, unintrusive during walking, and responded well to limb swing in real-time. Future work includes further improvements to the algorithm followed by rigorous user tests with both healthy and impaired individuals. We believe that this research can contribute to the development of meaningful auditory feedback schemes targeting limb swing.

1. INTRODUCTION

Walking (gait) is a crucial activity of daily living [1] and a highly complex combination of movements whose biomechanical characteristics are mediated by visual, auditory, proprioceptive, and tactile sensory feedback [2, 3]. Existing gait sonification systems have either sonified movement parameters such as ground reaction force and limb velocity [4, 5] or manipulated recorded / synthesized footstep sounds [6, 7, 8] to provide auditory feedback on the movement. Experiments have found that gait parameters, perceived agency, and self-perceptions of movement quality can be altered even by basic manipulations of real footstep sounds such as time delays [8] and spectral modification [7].

Feedback based on footstep sounds (as used in [4, 6, 9, 7, 10]) has been shown to be intuitive and meaningful. This makes sense as footstep sounds are a natural real-life consequence of walking, and thereby a sound morphology that people know how to engage with [11]. Ecological sounds of this nature have been shown to afford a greater degree of feedback comprehension and movement modification than abstract sounds due to superior matching between said sounds and the previous experiences of the learner [12, 13, 14]. However, these approaches only consider the portion of the gait cycle where the foot is in contact with the ground *(stance phase)* whilst disregarding the non-contact portion where the limb rotates and displaces through the air *(swing phase)*. The latter does not generate sound during the course of normal overground walking, so the notion of ecological feedback on swing kinematics is somewhat nebulous. In fact, to our knowledge, very few studies (e.g. [4]) have tackled the challenge of sonifying swing phase kinematic features [15] despite their importance to healthy gait; given the complex coordination required between several leg segments in order to successfully swing the limb forward [16], we believe that the potential of auditory feedback on swing dynamics warrants further investigation.

There is, however, one common real-life scenario where the swing phase *does* generate audible sound, and that is when walking through a shallow liquid (wading). When wading, the physical disturbance of the liquid mass by the swinging limb leads to bubble generation, resulting in a sound that reflects the properties of both the liquid mass and the physical excitation, i.e., the pushing force of the swinging limb [17]. Splashing sounds have previously been explored to signify the start and end of a movement, and shown to enhance motivation and relaxation [18]. As such, we propose that wading sounds may be an appropriate and ecological sound morphology to provide swing phase feedback. Swing kinematics and liquid sound generation are individually complicated phenomena, and simulating the natural acoustics of the physical interaction between the limbs and water is hence challenging. However, we posit that it is feasible to create a plausible simulation by sonifying limb angular velocity measured from *inertial measurement units (IMUs)*. The aim of this study was to develop a real-time gait sonification algorithm to simulate wading and evaluate its output in terms of perceptual characteristics and user experience.

2. IMPLEMENTATION

Given the complexity of precisely modelling the acoustics of wading, we made the following simplifying assumptions about the sound generation process:

1. The sound of wading is spectrally similar to the sound of flowing water.
2. Sound is only generated during the forward swing of the limb (captured by thigh rotation).
3. The intensity and timbre of the wading sound are directly correlated to the angular velocity trajectory of the limb.

As such, our system (shown in Fig. 1) was built to track the thigh's angular velocity in real-time and map it to the amplitude envelope and textural parameters of flowing water sounds. It comprised (A) two wireless IMU sensors to track thigh rotation, (B) custom PC software to process and map the IMU data in real-time (upgraded version of the technical framework described in [19]), and (C) a REAPER session with a custom-built signal chain for audio generation. The components communicated via *Open Sound Control (OSC)* messages at a fixed data sampling rate of 100 Hz, and the final audio output was 44.1 kHz / 24 bit stereo.

2.1. Sensing

For motion sensing and wireless transmission to the software, we used thigh-mounted (both legs, lateral placement) M5Stack Grey[1] microcontrollers equipped with MPU-9250 IMU units and enclosed in bespoke silicone casing. These devices transmitted motion data wirelessly to our custom JUCE software.

2.2. Real-time Data Processing

As shown in Fig. 1, the gyroscope signals from each thigh were first smoothed using 2nd order Butterworth low-pass filters with a 5 Hz cutoff frequency and Q-factor of 0.7. The angular velocity signal typically shows large positive excursions during the swing phase, whereas the stance phase is characterized by small negative excursions due to the slow backward rotation (see first graph in Fig. 3). The smoothed left and right signals were then processed through a real-time mapping architecture in the software. Specifically, the positive rotation velocity range was normalized to a 0-1 range, summed, and rescaled as shown in Fig. 1 to yield two control signals for sound generation. These signals were transmitted to REAPER using OSC for further processing.

2.3. Sound Mapping and Generation

We adopted a hybrid approach combining pre-recorded flowing water samples with the synthesized output of van den Doel's physical model [17]. The core of their algorithm calculates the impulse response $i(t)$ of a radially oscillating bubble as:

$$i(t) = asin(2\pi ft)e^{-dt}$$

where f is the resonance frequency, d is the damping factor, a is the amplitude and t is time. This simple equation, given the right parameters, can simulate the sound of a single bubble. Details on how to stochastically concatenate several virtual bubbles to create a stream of water are explained in [17]. Our aim was to leverage the textural realism of recorded sounds as well as the parametric control provided by physical models.

The *physical model component* was achieved using the Water[2] VST plugin (32 bit), which is a real-time implementation of the liquid synthesis algorithm described in [17]. As shown in Fig. 1, we mapped one control signal to the bubble generation rate parameter. The bounds were selected such that stationary standing (and double support periods during walking) resulted in no bubble generation, slow limb swing led to slow bubble generation, and fast limb swing led to rapid bubble generation. Audio examples of the physical model in isolation are provided in the supplementary materials[3].

The *recorded sample component* was achieved using recorded sound clips of flowing water from the FreeSound online library [4]. These clips were edited, crossfaded, looped, and summed with the physical model component. The second control signal, as shown in Fig. 1, was linearly mapped to a gain control applied to the summed physical model and recorded sample components. The control signal value ranges were tuned by an analysis-by-synthesis approach. Separate REAPER sessions were created for shallow and deep wading simulations, and sound examples at different speeds are provided in the supplementary materials. The system had an approximate round-trip latency of 100 ms.

3. EXPERIMENT 1: SOUND QUALITY ASSESSMENT

To assess whether the simulation would elicit the intended perceptions of depth, walking speed, naturalness and effort, we carried out a web-based perceptual evaluation of the shallow and deep wading sonifications. At the outset, we formulated the following hypotheses:

- **H1:** Audio clips corresponding to the deep wading simulation will elicit perceptions of greater water depth than those corresponding to the shallow simulation irrespective of the actual walking speed.
- **H2:** The perceived walking speed upon listening to the clips will directly mirror the actual walking speed irrespective of the simulated water depth.
- **H3:** The simulated depth will affect listeners' perceptions of perceived effort irrespective of walking speed.
- **H4:** The listeners' perceptions of naturalness will not vary between simulated depths and actual walking speeds.

3.1. Stimuli

We captured thigh IMU recordings from one healthy 30 y/o male walker (author PK) during overground back-and-forth walking on a straight 4 m path and sampled at 100 Hz. This was carried out at three walking speeds - slow (minimal speed), normal (brisk), and fast (threshold of running). Using our setup, we generated sonified sequences from the data corresponding to each of the simulated depths (shallow, deep). The sequences were adjusted by ear to have roughly equal loudness. We rendered a set of 7-sec excerpts (including turning pauses) as stereo 320 kbit MP3 files. Hence, the complete set of stimuli comprised six audio clips (2 simulated depths $\times$ 3 walking speeds) (link in footnote).

[1] https://shop.m5stack.com/products/grey-development-core

[2] https://www.kvraudio.com/product/water-by-xoxos

[3] https://doi.org/10.5281/zenodo.7952311

[4] www.freesound.org

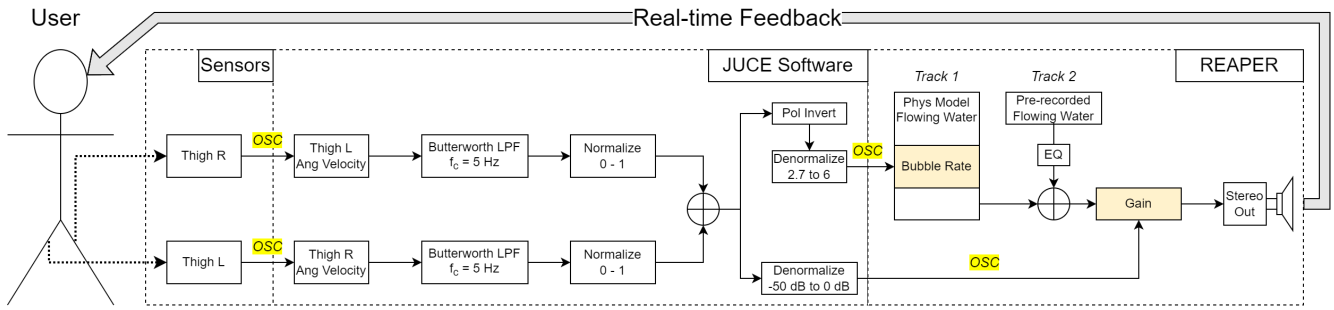

Figure 1: The overall block diagram of the feedback loop. The normalization was carried out such that 0 deg/sec - 100 deg/sec (empirically chosen range) of thigh rotation velocity in the positive (forward) direction was represented as a 0 - 1 range. LPF = Lowpass Filter, Pol Invert = Polarity Invert, OSC = Open Sound Control. The mapped audio processing parameters are shown in yellow.

3.2. Participants

A convenience sample of 16 participants (9 men, 7 women) aged 42 ± 15.15 years (ranging from 24-69) were invited to participate via mailing lists and social media. The assessment was conducted anonymously, and no sensitive information was collected from them. The social media invitation mentioned the purpose of the research and the planned use of their data.

3.3. Experimental Procedure and Outcomes

The evaluation was set up as a survey on Google Forms. Participants were initially briefed about the purpose of the research as well as the structure of the survey, but not about our hypotheses. After collecting basic information about participant age and gender, the survey was divided into four parts addressing the four outcomes of the study that took a total of 5-10 minutes to complete:

- Perceived Water Depth: Participants listened to each of the six audio clips (presented in a random order) and were asked to rate on an 11-point scale (0-10) their perceived depth of the water (0 = 0 m / 0 ft, 10 = 1 m / 3.28 ft).
- Perceived Walking Speed: They then listened to the same clips in a different random order and were asked to rate the perceived walking speed of the walker (0 = slowest possible, 10 = almost running).
- Perceived Effort: The procedure was repeated and participants were asked to rate the clips with regard to the physical activity of the walker in terms of perceived effort (0 = minimal effort, 10 = maximal effort).
- Perceived Naturalness: The participants were asked to assess the extent to which the sounds resembled those of walking through water in real life (0 = highly unrealistic, 10 = highly realistic).

3.4. Data Analysis

The age and gender data were aggregated, the rating data were exported from Google Forms in a comma-separated format, rearranged, and statistically analyzed in SPSS 27.0. In correspondence with the stated hypotheses, we aimed to analyze the effects of (a) Simulated Depth (deep, shallow), and (b) Actual Walking Speed (slow, medium, fast) on perceived *Depth, Walking Speed, Effort, and Naturalness* ratings. We first checked all data for normality (Shapiro-Wilks test) and homogeneity of variance (Levene's test) for each set of factors, and found the distributions to both exhibit significant deviations from normality and significantly non-homogeneous variance between factor levels. Therefore, we chose to adopt a non-parametric repeated measures analysis for each outcome. We first checked for main effects of Simulated Depth and Actual Walking Speed using Friedman tests. If significant effects were detected, planned pairwise comparisons specific to H1-3 were carried out using Wilcoxon signed-rank tests. A significance criterion $\alpha = 0.05$ was used for all statistical analyses. The reported *p*-values are those obtained post Holm-Bonferroni correction. The SPSS datasets are provided in the supplementary materials[5].

3.5. Results

A significant main effect of Simulated Depth and Actual Walking Speed on perceived depth ($\chi^2(7)$ =43.68, $p < 0.001$); walking speed ($\chi^2(7)$ =50.31, $p < 0.001$), and effort ($\chi^2(7)$ =17.33, $p =$ 0.004), but not for naturalness ($\chi^2(7)$ =7.19, $p = 0.207$) were shown. The results of the planned Wilcoxon signed-rank comparisons are depicted in Fig. 2.

There was a clear effect of *Simulated Depth* on perceived depth at all three walking speeds; in line with H1, the clips corresponding to the *deep* water simulation were rated as deeper than those of the *shallow* simulation for slow (Z = -3.192, $p = 0.009$), medium (Z = -3.246, $p = 0.009$), and fast (Z = -3.335, $p = 0.007$) walking. There were no significant differences in perceived depth ratings between the actual walking speeds within either of the simulated depths.

There was an effect of *Actual Walking Speed* on perceived walking speed; as stated in H2, the *medium* walking clips were rated as significantly faster than the *slow* clips for both deep (Z = -3.313, $p = 0.003$) and shallow (Z = -3.255, $p = 0.003$) simulated depths. The *fast* clips were rated as significantly faster than the *medium* clips for the deep water simulation (Z = -2.259, $p =$ 0.047) but not the shallow (Z = -1.349, $p = 0.177$).

There was an effect of *Simulated Depth* on perceived effort; as per H3, the clips corresponding to the *deep* water simulation were rated significantly higher than their *shallow* counterparts for each

[5] `https://doi.org/10.5281/zenodo.7952311`

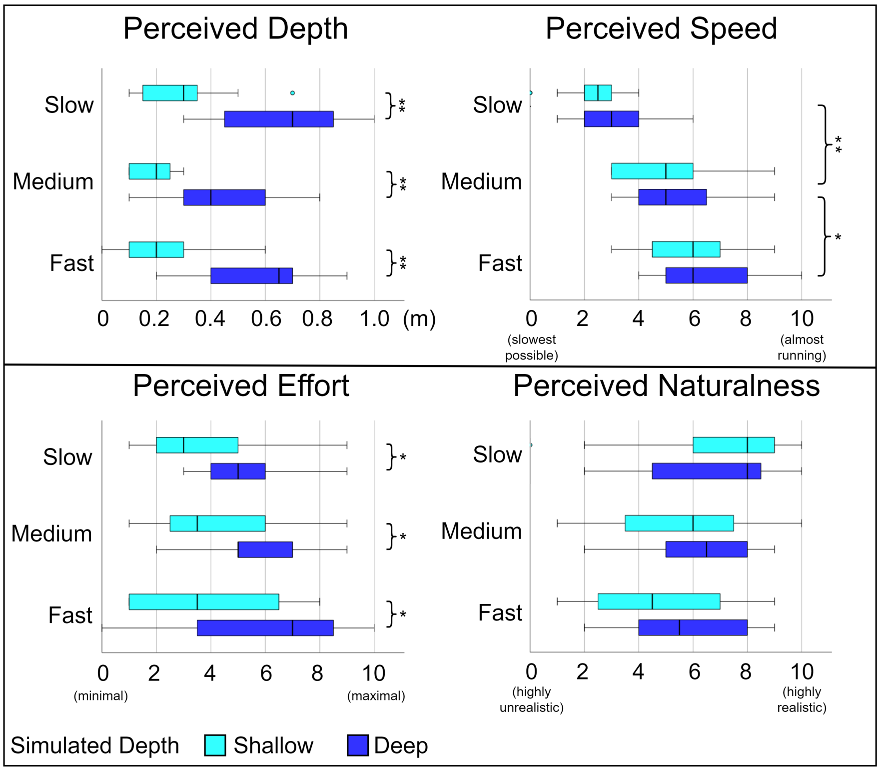

Figure 2: Boxplots visualizing the results of the survey for all factor levels and outcomes. Within each plot, the ratings are clustered vertically based on simulated depth (dark blue = deep, light blue = shallow). In each case, the boxes represent the interquartile range (IQR), and the notches within the boxes represent the median. The whiskers indicate variability outside the IQR. The small circles represent potential outliers ($>$ 1.5 IQR but $<=$ 3 IQR above (below) the upper (lower) quartile. The dots denote extreme values ($>$ 3 IQR above (below) the upper (lower) quartile). Significant differences are indicated by the asterisks between levels. * = p $<$ 0.05, ** = p $<$ 0.01

of the walking speeds - slow (Z = -2.214, $p = 0.026$), medium (Z = -2.381, $p = 0.034$), and fast (Z = -2.778, $p = 0.016$).

Although there was no significant main effect on perceived naturalness across the depth/speed combinations ($\chi^2(7)$ =7.19, $p = 0.207$) as per H4, it is worth noting that both simulated depths received high ratings of naturalness for slow walking (median rating of 8/10 in both cases). However, it is also evident in Fig. 2 that there was a lot of individual variability in the ratings, and the naturalness ratings tend to be somewhat worse for the faster walking speeds compared to the slowest.

3.6. Interim Discussion

The results indicate that the respondents were able to distinguish between the clips representing different walking speeds and simulated water depths. For each walking speed, they rated the perceived effort of walking to be greater for deep water, which is logical given the greater drag force that exists in real-life [20], although there was interestingly no difference between the walking speeds themselves, even though drag force is known to increase with speed [20]. In terms of perceived naturalness, we estimated that the fast walking clips received somewhat low ratings primarily due to the shape of their amplitude envelopes. Specifically, the slow walking sounds had a slow onset and decay analogous to the relatively gradual thigh acceleration and deceleration compared to fast walking, where the sounds sharply increased and decreased in amplitude. The latter, although in line with our second simplifying assumption, may have been perceived to be unnatural as real liquid sounds take some time to dissipate after the initial excitation has died away (due to the physics of bubble generation [17]). Fig. 3 (top pane) shows how the swing phase is characterized by rotations of the thigh as well as the shank. With no shank information used for sonification purposes, the wading sound was in its decay phase when the shank rotation reached its peak velocity (unnatural). To satisfy our overall aim of creating a plausible real-time interactive wading simulation, this would require immediate addressal before proceeding, the sound quality assessment validated H1-4 and provided positive early evidence in favor of our sonification methods, which we then modified and tested in an interactive setting.

4. EXPERIMENT 2: USER EXPERIENCE ASSESSMENT

The purpose of this test was to assess the experience of users interacting with the wading simulation system in real-time. We added a new mapping configuration by modifying the smoothing filter applied to the gyroscope signals to have its $f_c = 0.8Hz$ (empirically set for a natural-sounding decay) followed by a gain factor to compensate for the resulting amplitude decrease. The result of this is shown in the bottom pane of Fig. 3. As shown, the wad-

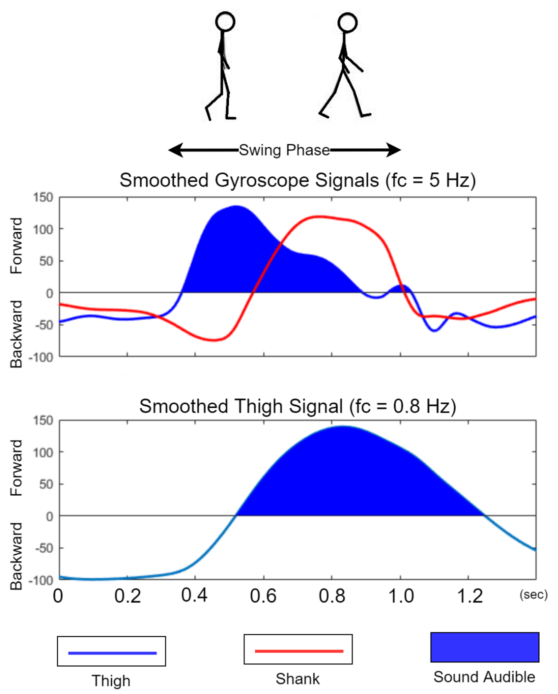

Figure 3: An illustration of the swing phase shown in context with smoothed thigh and shank angular velocity trajectories as well as the temporal extent and amplitude envelope of the resulting wading sound for two smoothing filter cutoff frequencies (f_c = 5 Hz and f_c = 0.8 Hz). Shank information was *not* used for sonification purposes and is plotted for context only.

ing sound is now more temporally aligned with the shank rotation due to the phase delay, and takes substantially longer to smoothly die out due to the attenuation of high frequency components in the angular velocity signal. We found that this led to a more natural wading sound. In addition to the configuration with 5 Hz cutoff $S_{5.0}$ and 0.8 Hz cutoff $S_{0.8}$, we created an asymmetric configuration *AS* which used different cutoff frequencies for the left and right thigh signals as summarized in Table 1. In all conditions, the sounds corresponding to the left and right limb were respectively panned 50% left and right in the stereo field. We proceeded to test user experience when walking with *AS*, $S_{5.0}$, and $S_{0.8}$ in a brief pilot study. Audio examples of each are provided in the supplementary materials.

4.1. Participants

A convenience sample of 9 participants (8 men, 1 woman) aged 29.66 ± 4.24 y/o from the student and staff population of Aalborg University, Copenhagen volunteered to participate. None of them reported having any form of documented hearing loss. All of them were briefed beforehand about the purpose and length of the experiment, and informed that they could withdraw at any time. Participants were told that there would be three identically structured experimental conditions, and the walking task details were explained to them verbally by the experimenter. Each of them provided informed consent prior to participation, and the experiment

Condition	Acronym	Fc - LEFT (Hz)	Fc - RIGHT (Hz)
Symmetric Minor Smoothing	$S_{5.0}$	5	5
Symmetric Major Smoothing	$S_{0.8}$	0.8	0.8
Asymmetric	*AS*	5	0.8

Table 1: A summary of the three smoothing configurations with their acronyms and additional filter cutoff frequencies (Fc) for the left and right thigh angular velocity signals.

was carried out according to the Helsinki declaration. Following this, they were asked to fill out basic information about themselves (age, gender, and documented hearing loss).

4.2. Setup

The test was carried out in a quiet laboratory at Aalborg University, Copenhagen. The laboratory has an approximately 6 m × 6 m vacant space with black synthetic flooring and surrounded on three sides by black curtains. An inner square was marked on the floor using tape to demarcate an approximately 1 m wide walking path around its outer perimeter. The start / end point for each walking lap was also marked. A pair of budget AirSonic wired headphones was suspended loosely from the center of the ceiling. We used headphones to ensure that participants perceived the stereo field consistently irrespective of head orientation or their physical location in the space.

Motion data streaming, mapping, and sound generation (simulated depth: *deep*) were carried out on a Dell Inspiron 15 7000 Windows laptop computer. The computer and experimenter were set up adjacent to the path. For low-latency sound output, an M-Audio M-Track Solo[6] audio interface was used.

4.3. Procedure

The IMU sensors were calibrated for static bias compensation and mounted laterally on the participants' thighs. During each of the three conditions (order counterbalanced across participants), they were asked to walk a total of *six* laps of the square path with the sensors and headphones on while their movements were recorded. They were told that during the first three laps *(baseline)*, no sound would be audible whereas there would be some walking-generated feedback during the last three *(feedback)*. No information was given to them about the specifics of the sound mapping. During each lap, they started walking from the start point (alternating clockwise and anticlockwise paths around the space) until they reached the start point again. They then paused for ∼2 sec, turned around, and started in the opposite direction. After three laps, the water sound feedback was provided at a standardized volume (approximately 80 dB SPL) and the participant walked three more laps. Once complete, the participant filled out a Google Forms survey about various aspects of user experience - perceived agency, naturalness, effort, speed, comfort, change in movement (identical for $S_{5.0}$, $S_{0.8}$ and slightly different for *AS* as not all items were deemed relevant to this condition) which had a series of 7-point rating scale items:

[6] `https://www.m-audio.com/m-track-solo`

- ($S_{5.0}$, $S_{0.8}$) Whether the sound perfectly followed/responded to their movements (1 = strongly disagree, 7 = strongly agree).
- ($S_{5.0}$, $S_{0.8}$, *AS*) Whether they noticed a change in their walking after the sounds started (1 = no change, 7 = extremely drastic change).
- ($S_{5.0}$, $S_{0.8}$) How natural they felt the sounds were (1 = highly unnatural, 7 = highly natural).
- ($S_{5.0}$, $S_{0.8}$) Whether the sounds changed their perception of the muscular effort they applied in order to walk (1 = much less effort, 4 = no change, 7 = much more effort).
- ($S_{5.0}$, $S_{0.8}$) How disturbing the sounds were while walking (1 = not at all disturbing, 7 = highly disturbing).
- ($S_{5.0}$, $S_{0.8}$, *AS*) How comfortable they felt when walking with the sounds (1 = highly uncomfortable, 7 = highly comfortable).
- ($S_{5.0}$, $S_{0.8}$, *AS*) Whether there was a change in their walking speed after the sounds started (1 = major decrease, 4 = no change, 7 = major increase).

Some participants also provided spontaneous verbal comments, which were noted. In all, the experiment took $\sim$25 min per participant.

4.4. Data Analysis

Due to the small sample size ($n = 9$), we used non-parametric tests for all statistical analyses. The survey responses were exported to IBM SPSS 27.0 for analysis. Each rating scale was compared between conditions using (a) Wilcoxon Signed-Rank (WSR) tests where only $S_{5.0}$ and $S_{0.8}$ were assessed, and (b) Friedman's ANOVA conditionally followed by post-hoc WSR tests if all conditions were applicable. Additionally, Spearman correlation coefficients were computed on the long form Likert data for all responses between each pair of items. A significance criterion $\alpha = 0.05$ was used for all statistical analysis. The SPSS datasets are provided in the supplementary materials.

4.5. Results

The survey results are depicted in Fig. 4. In terms of perceived naturalness, the WSR test revealed no significant difference (Z = -1.781, p = 0.075) between $S_{0.8}$ and $S_{5.0}$, although $S_{0.8}$ had a greater sample median. There was no significant difference between disturbance level ratings either (Z = -1.725, p = 0.084), though in this case, $S_{5.0}$ had a greater sample median. In terms of comfort level ratings, the Friedman test did not reveal a significant main effect across conditions ($\chi^2(2)$ =4.067, p = 0.131), but $S_{0.8}$ received the highest sample median rating (6) and *AS* received the lowest (3). We found significant negative correlations between disturbance and comfort ratings (ρ = -0.688, p = 0.002) and between disturbance and naturalness ratings (ρ = -0.504, p = 0.033).

For both $S_{0.8}$ and $S_{5.0}$, the participants tended to agree with relative uniformity that the sound followed their movement (Z = -0.171, p = 0.865) with sample median ratings of 6 and 5 respectively. They also expressed that the sound feedback in general led to changes in their walking compared to the baseline laps (mid-bottom of Fig. 4), although their estimated magnitude of walking change did not significantly vary between conditions ($\chi^2(2)$ =0.074, p = 0.964). However, we did observe that $S_{0.8}$ seemed to elicit somewhat lower walking change ratings than the other conditions. There was no significant difference in terms of speed changes across conditions ($\chi^2(2)$ = 4.727, p = 0.094). Between $S_{0.8}$ and $S_{5.0}$, there were no significant differences in terms of changes in muscular effort (Z = -1.0, p = 0.317), where the median rating was 4 (no change) in both cases. There were significant positive correlations between rated walking change and speed change (ρ = 0.543, p = 0.003) and between rated walking change and whether the sound was perceived to follow the movement (ρ = 0.528, p = 0.024).

Several participants commented verbally that they tried to adapt their walking to attain a more comfortable sound. Some stated that they felt a greater awareness of their walking and sense of rhythm during the $S_{5.0}$ condition, whereas they felt a sense of imbalance in the *AS* condition that they needed to compensate for.

5. GENERAL DISCUSSION

In this study, we developed a real-time gait sonification algorithm to simulate wading sounds and evaluated it in terms of sound quality and interactive user experience. Although the user experience assessment did not yield significant differences between the data filtering conditions, we believe that this was due to the small sample size and conservative nonparametric analysis. Whilst the observed tendencies in the results should be interpreted with caution, they yield a set of interesting insights about the sonification design problem at hand. The participants' overall agreement that $S_{0.8}$ and $S_{5.0}$ sounds followed their movements (i.e. agency) supports the notion that the wading simulation was movement-congruent in terms of spatial (L-R panning) and temporal contiguity [21]. This coupled with the high ratings of naturalness in both experiments suggests that the movement and sound were also perceived as semantically congruent. This is further supported by how effectively the participants in Expt. 1 were able to distinguish between the actual walking speeds and depths based on sound alone.

Based on the results of Menzer et al. [8], one may have expected $S_{0.8}$ to elicit lower ratings of agency than $S_{5.0}$ due to the delayed onset of sound relative to the start of the swing phase, but no such tendency was evident. This could be because Menzer et al. experimented with delayed footstep sounds (impulsive events by nature) where temporal discrepancies are more easily perceptible than with gradual onset sounds such as ours. The negative correlations seen between disturbance and both naturalness and comfort ratings, whilst not directly indicative of causal relations, may point to the importance of sonic naturalness when designing an application for long-term use. One interpretation is that unnatural sounds are perceived as less semantically congruent with the movement, which in turn causes users to be less comfortable and find them more disturbing than natural sounds [6, 22]. This would also explain why the explicitly incongruent *AS* configuration elicited the lowest median comfort ratings.

Across conditions, the participants reported only minimal change in perceived walking speed and muscular effort during the feedback laps versus the baseline. Even though wading entails greater effort due to fluid drag forces [20], the participants carrying out the task also received contradictory information from other modalities (vision, touch, proprioception). Based on known mechanisms of intersensory bias [23], the auditory information was most likely suppressed in their overall perception of effort. This is, however, at odds with similar studies that have demonstrated changes in body and movement perception due to altered auditory

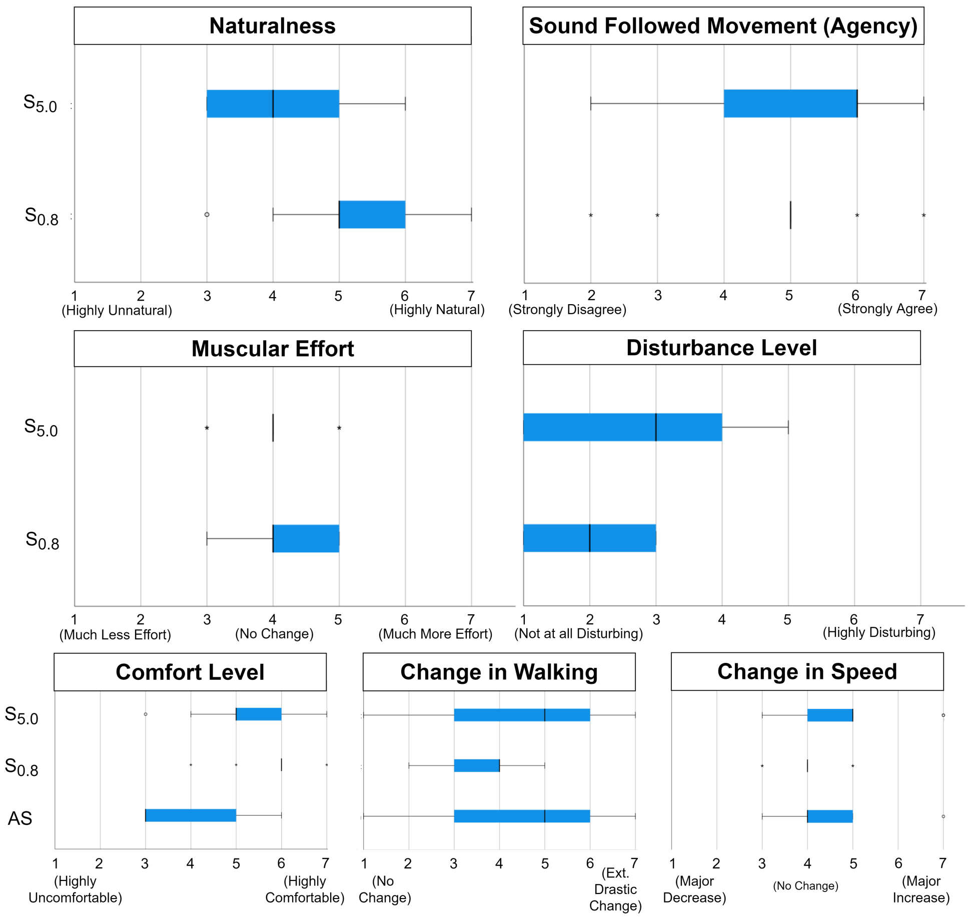

Figure 4: Boxplots visualizing the rating results. In each case, the boxes represent the interquartile range (IQR), and the notches within the boxes represent the median. The whiskers indicate variability outside the IQR. The small circles represent potential outliers (> 1.5 IQR but <= 3 IQR above (below) the upper (lower) quartile. The dots denote extreme values (> 3 IQR above (below) the upper (lower) quartile).

feedback [7, 18]. A follow-up study with a larger sample will be necessary to confirm this finding. Participants did, however, report some amount of perceived change in their walking over the baseline in each of the feedback conditions. Future studies will include a kinematic analysis to ascertain the precise nature of this change. In any case, the positive correlation between agency and perceived walking change ratings indicates that a strong feeling of agency is necessary in order to elicit motor change, a notion well aligned with theories in existing literature [11, 14].

On the whole, our results indicate that the wading simulation is plausible and can be improved with adjustments to the envelope characteristics of the sound. Adding a shank sensor coupled with envelope follower processing can help capture the swing phase duration and dynamics more accurately in the resulting sound. There also exist more advanced physical models that can more realistically simulate the acoustic interaction between an excitation and water. Key limitations of the study include the small and skewed convenience sample, lack of shank consideration in the algorithm, and lack of investigation of the kinematic effects of the feedback. Given the small sample size, it may have been more informative to record the detailed subjective impressions of the participants. Our use of a 7-point Likert scale in Experiment 2 (as opposed to the 11-point scale in Experiment 1) may have resulted in a loss of resolution in the responses of the participants, partially contributing to the observed lack of difference between conditions.

6. CONCLUSION

Despite the complexity of wading acoustics in real-life, the results of our sound quality assessment and user experience assessment indicate that we were able to generate a plausible simulation using gyroscope-modulated liquid sounds to convey feedback about thigh dynamics during walking in an ecological, meaningful, and minimally disturbing manner. Future work will include making the simulation more realistic by accounting for knee and shank kinematics, as well as more extensive real-life tests to assess the effect of the feedback (and its parameters) on gait biomechanics. Although the feedback parameters and use-cases will take more investigation to clearly define, we believe that this research can contribute to the development of meaningful auditory feedback schemes targeting limb swing.

7. ACKNOWLEDGMENTS

We would like to thank the experiment participants for their time. We would also like to thank Cumhur Erkut and Daniel Overholt for their kind co-operation with regard to laboratory use. NordForsk's Nordic University Hub, Nordic Sound and Music Computing Network NordicSMC, project number 86892. The wading sonification idea was initially conceived of by authors PK and EGS. Algorithm design, testing, and analysis were primarily carried out by author PK. Authors SD, SS, and EGS supervised the project and assisted with writing.

8. REFERENCES

[1] S. Studenski, S. Perera, K. Patel, C. Rosano, K. Faulkner, M. Inzitari, J. Brach, J. Chandler, P. Cawthon, E. B. Connor, *et al.*, "Gait speed and survival in older adults," *Jama*, vol. 305, no. 1, pp. 50–58, 2011.

[2] M. Alfuth and D. Rosenbaum, "Effects of changes in plantar sensory feedback on human gait characteristics: a systematic review," *Footwear Science*, vol. 4, no. 1, pp. 1–22, 2012.

[3] T. Lam and K. G. Pearson, "The role of proprioceptive feedback in the regulation and adaptation of locomotor activity," *Sensorimotor Control of Movement and Posture*, pp. 343–355, 2002.

[4] M. W. Rodger, W. R. Young, and C. M. Craig, "Synthesis of walking sounds for alleviating gait disturbances in parkinson's disease," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 3, pp. 543–548, 2013.

[5] J. Reh, T.-H. Hwang, G. Schmitz, and A. O. Effenberg, "Dual mode gait sonification for rehabilitation after unilateral hip arthroplasty," *Brain Sciences*, vol. 9, no. 3, p. 66, 2019.

[6] L. Turchet, S. Serafin, and P. Cesari, "Walking pace affected by interactive sounds simulating stepping on different terrains," *ACM Transactions on Applied Perception (TAP)*, vol. 10, no. 4, pp. 1–14, 2013.

[7] A. Tajadura-Jiménez, M. Basia, O. Deroy, M. Fairhurst, N. Marquardt, and N. Bianchi-Berthouze, "As light as your footsteps: altering walking sounds to change perceived body weight, emotional state and gait," in *Proceedings of the 33rd annual ACM conference on human factors in computing systems*, 2015, pp. 2943–2952.

[8] F. Menzer, A. Brooks, P. Halje, C. Faller, M. Vetterli, and O. Blanke, "Feeling in control of your footsteps: conscious gait monitoring and the auditory consequences of footsteps," *Cognitive neuroscience*, vol. 1, no. 3, pp. 184–192, 2010.

[9] R. Bresin, A. Dewitt, S. Papetti, M. Civolani, and F. Fontana, "Expressive Sonification of Footstep sounds," *Proceedings of the Interaction Sonification Workshop (ISon) 2010*, 05 2010.

[10] J. Maculewicz, C. Erkut, and S. Serafin, "How can soundscapes affect the preferred walking pace?" *Applied Acoustics*, vol. 114, pp. 230–239, 2016.

[11] J. F. Dyer, P. Stapleton, and M. Rodger, "Mapping sonification for perception and action in motor skill learning," *Frontiers in neuroscience*, vol. 11, p. 463, 2017.

[12] J. P. Thomas and M. Shiffrar, "I can see you better if i can hear you coming: Action-consistent sounds facilitate the visual detection of human gait," *Journal of vision*, vol. 10, no. 12, pp. 14–14, 2010.

[13] M. Murgia, I. Santoro, G. Tamburini, V. Prpic, F. Sors, A. Galmonte, and T. Agostini, "Ecological sounds affect breath duration more than artificial sounds," *Psychological research*, vol. 80, no. 1, pp. 76–81, 2016.

[14] A. Pizzera and T. Hohmann, "Acoustic information during motor control and action perception: a review," *The Open Psychology Journal*, vol. 8, no. 1, 2015.

[15] D. Linnhoff, S. Alizadeh, N. Schaffert, and K. Mattes, "Use of Acoustic Feedback to Change Gait Patterns: Implementation and Transfer to Motor Learning Theory—-A Scoping Review," *Journal of Motor Learning and Development*, vol. 1, no. aop, pp. 1–21, 2020.

[16] P. M. Mills and R. S. Barrett, "Swing phase mechanics of healthy young and elderly men," *Human movement science*, vol. 20, no. 4-5, pp. 427–446, 2001.

[17] K. v. d. Doel, "Physically based models for liquid sounds," *ACM Trans. Appl. Percept.*, vol. 2, no. 4, p. 534–546, oct 2005. [Online]. Available: https://doi.org/10.1145/1101530.1101554

[18] J. Ley-Flores, F. Bevilacqua, N. Bianchi-Berthouze, and A. Taiadura-Jimenez, "Altering body perception and emotion in physically inactive people through movement sonification," in *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 2019, pp. 1–7.

[19] P. R. Kantan, E. G. Spaich, and S. Dahl, "A Metaphor-Based Technical Framework for Musical Sonification in Movement Rehabilitation," in *The 26th International Conference on Auditory Display (ICAD 2021)*, 2021.

[20] A. Ivaniski-Mello, M. Zimmermann Casal, R. R. Costa, C. L. Alberton, F. G. Martinez, and L. A. Peyré-Tartaruga, "Quantifying physiological and biomechanical responses of shallow water walking: a systematic review and meta-analysis," *Research in Sports Medicine*, pp. 1–15, 2022.

[21] B. E. Stein and T. R. Stanford, "Multisensory integration: current issues from the perspective of the single neuron," *Nature reviews neuroscience*, vol. 9, no. 4, pp. 255–266, 2008.

[22] A. Tajadura-Jiménez, N. Bianchi-Berthouze, E. Furfaro, and F. Bevilacqua, "Sonification of surface tapping changes behavior, surface perception, and emotion," *IEEE MultiMedia*, vol. 22, no. 1, pp. 48–57, 2015.

[23] R. B. Welch and D. H. Warren, "Immediate perceptual response to intersensory discrepancy." *Psychological bulletin*, vol. 88, no. 3, p. 638, 1980.