Multisensory Trajectory Control at One Interaction Point, with Rhythm

Alessio Bellino
Pontificia Universidad Católica de Chile
Santiago, Chile
abellino@uc.cl

Davide Rocchesso
Università degli Studi di Milano
Milan, Italy
davide.rocchesso@unimi.it

Rosario Mulè
Università degli Studi di Palermo
Palermo, Italy

Luca D'Arrigo Reitano
Università degli Studi di Palermo
Palermo, Italy

ABSTRACT

We investigate navigation in two dimensions with velocity control, using a single-button interface. Users adjust the controlled-object speed through rhythmic tapping and its direction by pressing and tilting, releasing the button to finalize the rotation. Feedback to control action and object motion is provided by integration of multiple sensory modalities. Tactile pulses are delivered at 30-degree intervals during rotation, emulating the detents of a rotary encoder. Simultaneously, a sonic glissando accompanies rotation, thus rendering upward or downward rotation. Both visual and auditory cues are used to provide absolute positional feedback. Discrete notes are rhythmically played, whose pitch indicates vertical position, while stereo audio panning follows horizontal position. The rhythmic pace aligns with on-screen object speed, as dictated by tapping time intervals. Two studies were designed around a target-following task, under different sensory conditions. Study 1 has shown that target tracking can be effectively performed with multisensory rhythmic interaction. This is true also when the controlled object is intermittently hidden to view, although it was not possible to measure the advantage provided by auditory-tactile feedback. In study 2, no significant performance differences were observed between auditory and tactile conditions in situations of intermittent visual feedback, indicating that, if the two kinds of non-visual feedback are effective, they are essentially equivalent.

CCS CONCEPTS

• **Human-centered computing** → **Interaction techniques**; **Auditory feedback**; • **Hardware** → **Tactile and hand-based interface**.

KEYWORDS

Multisensory interaction, Rhythmic interaction, Two-dimensional navigation

AM '24, September 18–20, 2024, Milan, Italy

ACM ISBN 979-8-4007-0968-5/24/09
https://doi.org/10.1145/3678299.3678340

ACM Reference Format:
Alessio Bellino, Davide Rocchesso, Rosario Mulè, and Luca D'Arrigo Reitano. 2024. Multisensory Trajectory Control at One Interaction Point, with Rhythm. In *Audio Mostly 2024 - Explorations in Sonic Cultures (AM '24), September 18–20, 2024, Milan, Italy.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3678299.3678340

1 INTRODUCTION

Human and animal locomotion often produce audible rhythms that communicate a great deal of information about the moving creature (e.g, a horse), the kind of gesture (e.g., galloping), the expressive intentions [7], and the motion speed [10]. In everyday activities, we often produce and perceive movement in space through rhythm. Therefore, the introduction of rhythm in interfaces, for controlling and monitoring moving objects, is ecologically sound. Despite this observation, however, the exploitation of rhythms for the control of virtual or physical moving objects remains marginal in the interaction design literature, even in the niche of rhythmicity for interaction.

The navigation of two-dimensional spaces using rhythmic patterns generated by bimanual tapping, effectively governing the orthogonal components of velocity, has been proposed and investigated [13]. In our current work, we propose a technique that relies on rhythm as part of a streamlined interaction mechanism. Notably, we now use a single interaction point (i.e., a button) instead of the previously explored two-button setup, and rhythmic tempo is used to control and non-visually display the movement speed. While the two-button velocity control is based on two concurrent rhythmic streams, and direction of motion can be deduced from comparison of the two politemporal rhythms, in the one-button realization there is only one rhythmic stream, so that only the absolute value of speed can be rendered through rhythm alone. Positional and directional cues must be deduced from other features of multisensory feedback, such as sound pitch and audio spatialization. Overall, these studies aim at finding effective ways to control trajectories in spacetime, with a minimal number of interaction points and exploiting redundancy in multisensory feedback. The control actions should be discrete and sparse, only to impart deviations from inertial Galilean motion, but it should be possible to continuously monitor motion in space, relying on one or more of the available senses. Having a single interaction point and sparse control actions may be beneficial in all those contexts where minimality of the interface is a value, for reasons of limited motor abilities of the user or for the smallness of the interactive surface, as in wearables.

The one-button system has been first demonstrated in a conference setting [2], but in this contribution we present a user study investigating the relative effectiveness and appreciation of different motion-feedback modalities.

In section 2 we briefly describe how rhythm has been previously used in human-computer interaction. In section 3 we present the one-button design for velocity control and explain how and why we arrived at such realization. In section 4 and 5 we present two studies using the proposed one-button velocity-control prototype. The primary objectives of this research were to assess participants' ability to track a moving target on a two-dimensional plane using a single-button interface, and to investigate how integrating visual, auditory, and tactile feedback affects user performance. The studies aimed to determine how different sensory feedback mechanisms could enhance the precision and responsiveness of interactions within a controlled setting. The first study was aimed at showing that object tracking is feasible with a one-button interface, as well as to show how users are capable of exploiting non-visual feedback to compensate for temporarily missing visual feedback. The second study compares auditory feedback to vibro-tactile feedback. Both the reported studies were developed around a target-following task where the target moved along a path reminiscent of a regatta. Participants were tasked with tracking the target as closely as possible, with the process illustrated in Figure 1. A video demonstration of the task and the path visualization is available online[1], and a static example is shown in Figure 2. All participants signed an informed consent form before engaging in the studies, and the experimental protocols received approval from the ethical committee at the University of Palermo

2 MULTISENSORY RHYTHMS IN HCI

The rhythmic interaction with devices or technology-augmented objects has been studied in a wide range of contexts and scales. The amodality of rhythms [6] has produced studies and solutions for one or more of the senses of touch, hearing, vision, and proprioception.

In sonic interaction design, systems and interfaces that support rhythmicity and afford the development of virtuosity have been proposed [4]. The role of rhythm in multisensory continuous interaction has been investigated with design exercises [14]. Film editing and rhythmic interaction design techniques have been proposed for use in cinematic virtual reality [3]. Rhythmic tutoring has been proposed for interaction by handclapping [9]. Additionally, rhythmic patterns, incorporating short and long taps and breaks, have been proposed as an input method to replace single commands. Evaluation has been carried out on recall efficacy, revealing similarities to keyboard shortcuts [8]. Rhythmic microgestures have been introduced as a means of non-visual interaction in mobility scenarios [5]. In the tactile domain, rhythmic tactons have been implemented in the context of the car driving environment [11].

Rhythm and motion have been extensively investigated for human walking, especially for the purpose of recreating and manipulating the experience of virtual locomotion, as well as to augment walking experiences [15]. Horse gait patterns have been used to augment human locomotion by biking, so that one can get the bike to walk, trot or gallop [10].

The exploration of control selection through visual rhythmic patterns and motion synchronization has been investigated across a diverse range of users [1]. Motion synchronization, largely regulated through cycles, pulses, and rhythms, has been considered as crucial to design interfaces for interpersonal interaction [12].

3 INTERACTION DESIGN RATIONALE

In designing an interface for trajectory control we took a minimalist standpoint, starting with the constraint of having a single interaction point, that is the single button. We also wanted to exploit the redundancy of multisensory displays, and augmented the button and the visual display with tactile and auditory feedback. In contrast to TickTacking [13], where users could regulate the speed of an object by adjusting the tapping rate using two buttons, our approach exploits tapping on a single button, with the additional capability of holding down the button to make the controlled object rotate. The proposed interaction is discrete, as changes in direction and speed are imparted sparingly, only at the times when the inertial Galilean motion of the controlled object needs to be adjusted. However, within a single direction-change action, a continuous and sustained rotation control is applied. The multisensory display and auditory iteration of the imparted rhythmic cell enable continuous monitoring of the object motion.

In the early iterations of the design, the rotation was clockwise and speed was kept constant while the button was held down. The action of keeping the button pressed is rhythmically consistent, as it is comparable to that of a drummer striking a cymbal (similar to tapping the button) and then manually damping the vibration (similar to holding the button) to muffle the sound. Going through design refinement, we aimed at improving control precision and immediacy, and introduced the tilting motion of the finger on the button to control the rotation direction and speed. A right tilt elicits clockwise rotation, and a left tilt elicits counterclockwise rotation. The amount of tilt is proportional to the rotation speed. Multisensory feedback on rotation, besides rotating the visual icon of the controlled object, was introduced with touch and sound. A tactile pulse, at intervals of 30 degrees of rotation, simulates the detents of a rotary encoder. A raising glissando accompanies upward rotation, and a lowering glissando accompanies downward rotation.

A simple kinematic auditory display took the form of sequences of piano-like tones, to convey two aspects of the controlled-object motion: (1) the position in the two-dimensional space and (2) the speed. In representing position, we use pitch and spatialization. The vertical limits of the 2D plane correspond to a range of two octaves, but the pitch is not limited to discrete semitones, as it is made to correspond to the object position measured in pixels. Lower pitches indicate a lower position, and higher pitches indicate a higher position. The simplest possible sound spatialization was used to represent the horizontal position of the object, i.e., stereo amplitude panning. The rightmost position produces tones exclusively on the right channel, and the leftmost position produces tones exclusively on the left channel. Speed is conveyed by the repetition rate (*tempo*) of the tones, a higher tempo corresponding to a faster motion.

[1] https://youtu.be/uLVLdQcD7l8

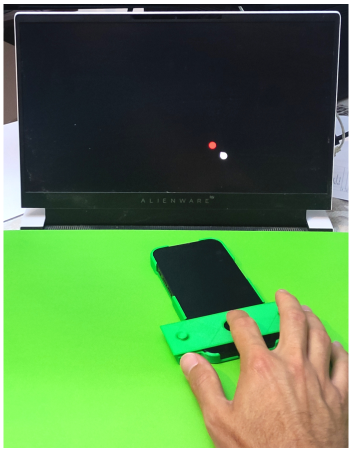

Figure 1: The experimental setup, as used in object following. The user controls the gray nail-like object with the goal of staying as close as possible to the red dot. The iPhone simulates the button using its Taptic Engine, delivering vibrotactile pulses.

When complementing visual information, sound and vibration may not directly enhance precision, but nevertheless provide additional information. They may increase awareness of the position/rotation of the controlled element, that is particularly beneficial when the element cannot be seen, either because the user is visually impaired or because their visual attention is directed somewhere else.

The experimental apparatus uses an Apple iPhone to simulate the button as illustrated in Figure 1. The iPhone, natively equipped with its Taptic Engine, provides effective tactile feedback. The smartphone is covered by a layer of plastic, featuring the designated hole, which indicates the point where users should interact with their finger. Tap, hold, and tilt actions are detected by the touchscreen. In particular, the finger inclination is detected by analyzing the slight variation in its relative position.

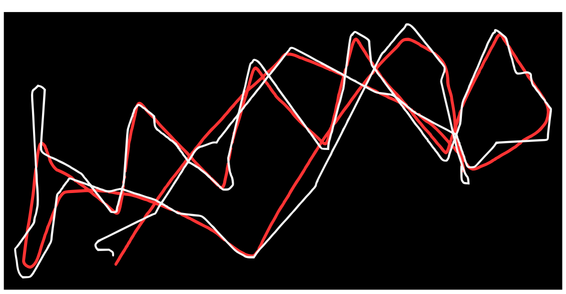

Figure 2: Visualization of path following data.

4 STUDY 1

This study aims to assess the ability of users to control a trajectory using one-button rhythmic tapping and holding, and to explore how the combination of auditory and tactile feedback improves user precision in scenarios where the cursor is intermittently absent (visible for 5 seconds, then not visible for 5 seconds). Visual, auditory, and tactile feedback, when present, work as described in section 3.

4.1 Participants

We recruited 12 participants (8 males and 4 females) from the student population at the Department of Mathematics and Computer Science, University of Palermo. All participants were native Italian speakers. Five reported no visual impairments, while the remaining seven had conditions such as myopia, astigmatism, or hyperopia. One participant reported a 20% hearing loss in the left ear. The median age was 23.5 years with an interquartile range of 3.25 years. Two participants reported playing a musical instrument.

4.2 Experimental Setting

We utilized two distinct interfaces: a computer-based Audio-Visual Interface and an iPhone-based Controller Interface. Each was developed using different technologies tailored to the specific functionalities required for the experiment.

4.2.1 Audio-Visual Interface. The computer interface was developed using standard HTML5, CSS, and JavaScript technologies. It operated on an Asus Zenbook UX530UX equipped with a 15.6" NanoEdge anti-glare, LED-backlit IPS display, which rendered visuals at a resolution of 1920x1080 pixels and a refresh rate of 60 fps. The visual setup featured a simplistic design with a black background and two primary elements: a white nail-shaped controlled object that indicated direction, and a red circular target. This interface was powered by a Windows 10 OS with a Google Chrome browser, and audio feedback was delivered through JBL Tune 510BT headphones set at a 50% volume level for all participants.

4.2.2 iPhone Controller Interface. The controller interface was implemented on an iPhone 8 Plus using the JavaScript-based Capacitor framework[2], which facilitated the integration of native functionalities not accessible via standard HTML5 APIs, such as the Apple Taptic Engine[3]. The interface displayed a black screen and a designated 20mm x 20mm square area for button-based interactions. The primary function of the iPhone interface was to manage the control of the object displayed on the computer interface. Tactile feedback was a critical feature, achieved through the iPhone's Taptic Engine, which generated vibrations corresponding to each touch input and every 30° rotation of the controlled object.

4.3 Experimental Conditions

Three conditions were tested to explore the impact of various feedback combinations on participant performance:

(1) **Visible-Auditory-Tactile:** In this condition, the controlled object was always visible, accompanied by continuous auditory feedback and tactile vibrations, providing participants with a full sensory experience.
(2) **Intermittent-Auditory-Tactile:** Here, the controlled object alternately appeared and disappeared every 5 seconds.

[2] https://capacitorjs.com/
[3] https://capacitorjs.com/docs/apis/haptics

Despite its intermittent visibility, continuous sound and haptic feedback were provided to assist participants in tracking the object.

(3) **Intermittent-NoAuditory-NoTactile:** This condition featured intermittent visibility without auditory or tactile feedback. Participants faced the greatest challenge as they had to rely solely on visual cues during the moments the object was visible.

These conditions were deliberately structured to assess how different feedback modalities, particularly the absence or presence of visual, auditory, and tactile cues, affect the ability to monitor and manipulate a dynamically moving target in the experimental interface.

4.4 Hypotheses

The study evaluated the following hypotheses:

- **HP1:** Users can effectively control the movement of an object in two-dimensional space using rhythmic button interactions.
- **HP2:** There is a significant difference in user performance between the Intermittent-Auditory-Tactile condition and the Intermittent-NoAuditory-NoTactile condition, with the former yielding better results than the latter.

4.5 Procedure

Participants in the study went through a structured sequence of phases designed to familiarize them with the experimental interface and evaluate their performance. Initially, they received a brief introduction that included detailed instructions and demonstrations about the functionalities of the interface and controller, lasting approximately five minutes. This was followed by a training phase where participants had five minutes to interact with the controls without the distraction of the moving target, allowing them to freely explore and become comfortable with the system. Finally, the testing phase involved exposing each participant to the three distinct conditions. Each condition lasted about five minutes and thirty seconds, and the order of exposure was counterbalanced using a Balanced Latin Square to minimize any carryover effects from previous conditions. Additionally, we mirrored the x and/or y axes across conditions to prevent users from remembering the trajectory across the different conditions.

4.6 Data Analysis

The performance of participants was assessed by measuring the Euclidean distance between the controlled object and the target at each timestamp, and calculating the average distance across all timestamps.

4.7 Results

4.7.1 Quantitative analysis. The analysis revealed that the average distance from the target in tracking performance varied significantly across conditions (see figure 3). Taken the average distances of all participants, and summarizing across participants, the Visible-Auditory-Tactile condition recorded a mean distance of 51.78 pixels (median 46.53), considerably less than the Intermittent-Auditory-Tactile and Intermittent-NoAuditory-NoTactile conditions, which had mean distances of 83.52 pixels (median 81.76) and 89.08 pixels (median 89.55), respectively. An ANOVA test was conducted and the null hypothesis—that distance is independent of feedback modality—was rejected ($F(2,22) = 22.32$, $p < 0.001$). Subsequent post-hoc comparisons using Bonferroni correction showed significant differences between the Visible-Auditory-Tactile condition and both the Intermittent-Auditory-Tactile condition ($p < 0.01$) and the Intermittent-NoAuditory-NoTactile condition ($p < 0.001$). However, no significant differences were found between the Intermittent-Auditory-Tactile and Intermittent-NoAuditory-NoTactile conditions ($p > 0.05$).

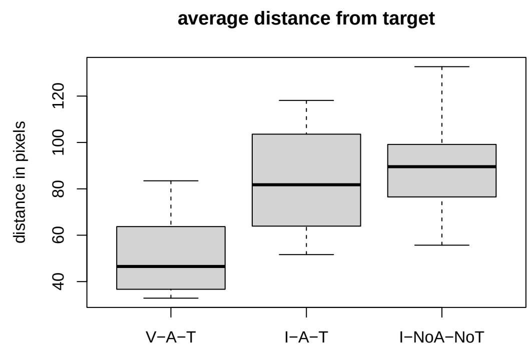

Figure 3: Study 1: Experimental average distance from target in the three conditions: Visible-Auditory-Tactile; Intermittent-Auditory-Tactile; Intermittent-NoAuditory-NoTactile

Upon completing the experiments, participants were asked to fill out NASA-TLX questionnaires to evaluate mental, physical, and temporal demands, as well as their performance, effort, and overall frustration. The data from these questionnaires were analyzed using non-parametric Friedman tests, which revealed significant differences across the three feedback conditions—Visible-Auditory-Tactile, Intermittent-Auditory-Tactile, and Intermittent-NoAuditory-NoTactile.

Significant differences in mental demand were detected, with median loads of -3, 0.75, and 3.25 for the three conditions (Friedman chi-squared = 17.636, df = 2, $p < 0.01$). Subsequent post-hoc analysis with Bonferroni correction showed that these differences were particularly pronounced between the Visible-Auditory-Tactile and Intermittent-NoAuditory-NoTactile conditions ($p < 0.01$), and between the Intermittent-Auditory-Tactile and Intermittent-NoAuditory-NoTactile conditions ($p < 0.01$).

In terms of performance load, significant differences were also noted, with median loads of -5, -3.5, -0.25 for the three conditions (Friedman chi-squared = 6.4, df = 2, $p < 0.05$), in particular between the Visible-Auditory-Tactile and Intermittent-NoAuditory-NoTactile conditions ($p < 0.05$), with participants rating their performance more satisfactorily in the former.

Lastly, the required effort showed significant variability, with median loads of -1, 1, 3.25 for the three conditions (Friedman chi-squared = 7, df = 2, $p < 0.05$), with significant difference between the Visible-Auditory-Tactile and Intermittent-NoAuditory-NoTactile conditions ($p < 0.05$). Participants reported exerting more effort in the Intermittent-NoAuditory-NoTactile condition compared to the Visible-Auditory-Tactile condition.

The questionnaires indicate that additional effort is required when the image of the controlled object visually disappears, as if a change of strategy were necessary to deal with the situation of partial visual deprivation.

4.7.2 Qualitative analysis. The qualitative feedback from the participants revealed a range of opinions that are noteworthy. Concerning the role of audio, it was reported that the function of sound in the experiment was not clearly defined. Three participants explicitly stated that the sound did not significantly enhance their performance, whereas five others found it beneficial. Interestingly, two participants mentioned that although they found the sound annoying, it still facilitated the tracking process in the absence of visual feedback. Additionally, three out of twelve participants reported that the sound was bothersome.

Opinions on the effectiveness of vibration feedback were mixed. Two participants recognized its utility, but six had negative impressions, particularly criticizing the rotation mechanism as non-intuitive and suggesting improvements. Moreover, another participant mentioned that due to cold hands, they could not fully perceive the vibrations.

Regarding tactics adopted during the experiment, two participants described a specific strategy used when the object to be controlled became invisible. They rotated on the spot, waiting for the object to reappear before resuming the tracking process.

5 STUDY 2

The purpose of Study 2 is to determine whether there are performance differences between tactile and auditory feedback among users. The experimental setup is largely similar to that of Study 1, with modifications below.

5.1 Participants

A new cohort of 12 participants (5 males and 7 females) was selected from the student body at the University of Palermo. Three reported no visual impairments, with the remaining nine indicating either myopia or astigmatism. No auditory impairments were reported. The median age was 23.5 years with an interquartile range of 2.5 years. Three participants reported playing a musical instrument.

5.2 Experimental Conditions

Two conditions were tested in Study 2 to assess the differential impact of auditory and tactile feedback on participant performance with an intermittently visible target:

(1) **Intermittent-Auditory-NoTactile:** In this condition, participants relied solely on auditory feedback to track the intermittently visible object. No tactile feedback was provided, testing the effectiveness of sound cues alone.

(2) **Intermittent-NoAuditory-Tactile:** This condition removed auditory feedback, allowing participants to use tactile vibrations to track the intermittently visible object. This setup assessed the utility of tactile cues in the absence of auditory information.

5.3 Hypothesis

We speculate that auditory feedback may prove to be more effective than tactile feedback alone. This assumption is based on the application of the feedback types within the experiment: auditory feedback is continuously available throughout the experiment, potentially providing more consistent cues for tracking the intermittently visible target. In contrast, tactile feedback is only utilized during the rotation of the element, which may limit its effectiveness in assisting participants with trajectory tracking.

5.4 Procedure

The procedure for Study 2 closely resembles that of Study 1, with a simplification in the experimental design from three to two conditions. To minimize carryover effects, the sequence in which the conditions were presented was alternated among participants.

5.5 Results

5.5.1 Quantitative analysis. No significant differences were observed between the Intermittent-Auditory-NoTactile and Intermittent-NoAuditory-Tactile conditions in Experiment 2 (paired t-test, $t = 0.5281$, $p = n.s.$, see figure 4). Participants demonstrated the ability to track a target with an overall mean distance of 122.45 pixels (median 114.9) in the Intermittent-Auditory-NoTactile condition and 116.38 pixels (median 105.81) in the Intermittent-NoAuditory-Tactile condition, suggesting similar performance across these feedback modalities.

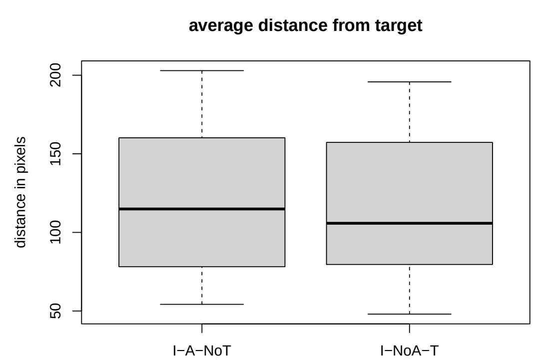

Figure 4: Study 2: Experimental average distance from target in the two conditions: Intermittent-Auditory-NoTactile; Intermittent-NoAuditory-Tactile

Regarding the Raw-NASA-TLX questionnaire, a non-parametric Wilcoxon signed-rank test was conducted, which revealed a single significant difference concerning mental demand, with a median load of -2 for Intermittent-Auditory-NoTactile, and a median load of 0 for Intermittent-NoAuditory-Tactile (V = 4.5, p = 0.03614). As a result, participants generally perceived a slightly lower mental demand in the Intermittent-Auditory-NoTactile condition compared to the Intermittent-NoAuditory-Tactile condition, which was considered neutral.

5.5.2 Qualitative analysis. Participants generally expressed a high level of engagement and interest. However, one participant reported experiencing mild eye strain by the end of the test, likely due to prolonged focus on the moving target. Another participant noted difficulty in accurately perceiving vibrations, attributing this to having cold hands. Regarding auditory feedback, while five participants found the sound useful, one considered it annoying. Another participant thought the sound was beneficial but not essential. Conversely, two participants described the sound as confusing and bothersome. As for tactile feedback, it was well received by six participants, particularly praised for its effectiveness during rotation. Nonetheless, one participant deemed it helpful but not crucial, and another gave a negative assessment of the tactile feedback.

6 DISCUSSION AND CONCLUSIONS

Designing an interface for trajectory control at one interaction point proved to be a challenging exercise, but exploiting multisensory feedback and rhythmicity in interaction we could show that a task as difficult as following a moving object is indeed feasible, with just a discrete and sparse set of actions on a single button. However, the findings from our studies are partially inconclusive. In Study 1, Hypothesis 1 was indeed confirmed: participants effectively controlled the movement of an object in two-dimensional space through rhythmic button interactions, maintaining a mean distance of 51 pixels from the target under full visual, auditory, and tactile feedback. This indicates better precision compared to previous studies, such as TickTacking, where the best condition—full visual and auditory feedback—resulted in a mean distance of 105 pixels [13]. However, Hypothesis 2 was not confirmed. The only significant difference observed was in mental demand—participants experienced higher mental demand under the Intermittent-NoAuditory-NoTactile condition compared to the Intermittent-Auditory-Tactile condition. This solitary distinction might be attributed to the limited number of participants or the balanced interval of visibility and invisibility (5 seconds each) in the intermittent feedback. These settings may not adequately challenge participants to depend extensively on auditory cues.

In Study 2, no significant differences in performance were observed, suggesting that user performance may not depend on the specific type of feedback provided. Similar to Study 1, this finding could also be attributed to the balanced proportion of hide and show times.

Given the outcomes, we plan to conduct further experiments using a modified duration ratio, such as 9 seconds of hiding followed by 1 second of visibility. This change is intended to accentuate the potential effects of auditory feedback and could yield clearer insights into how different feedback modalities influence user performance. Additionally, expanding the number of participants could enhance the robustness of these findings and allow us to detect more nuanced effects of auditory and tactile feedback.

ACKNOWLEDGMENTS

The second author is supported by MUR-PRIN - grant 2022RB939W.

REFERENCES

[1] Alessio Bellino. 2022. Rhythmic-Synchronization-Based Interaction: Effect of Interfering Auditory Stimuli, Age and Gender on Users' Performances. *Applied Sciences* 12, 6 (2022). https://doi.org/10.3390/app12063053

[2] Alessio Bellino and Davide Rocchesso. 2024. Controlling Trajectories with OneButto and Rhythm. In *17th International Conference on Advanced Visual Interfaces*. Genova, Italy.

[3] Cumhur Erkut. 2017. Rhythmic interaction in VR: interplay between sound design and editing. In *IEEE 3rd VR Workshop on Sonic Interactions for Virtual Environments (SIVE)*. 1–4. https://doi.org/10.1109/SIVE.2017.7901611

[4] Cumhur Erkut, Antti Jylhä, and Davide Rocchesso. 2013. Heigh Ho: Rhythmicity in Sonic Interaction. In *Sonic Interaction Design*, Karmen Franinović and Stefania Serafin (Eds.). MIT press, 341–350.

[5] Euan Freeman, Gareth Griffiths, and Stephen A. Brewster. 2017. Rhythmic Micro-Gestures: Discreet Interaction on-the-Go. In *Proceedings of the 19th ACM International Conference on Multimodal Interaction* (Glasgow, UK). 115–119. https://doi.org/10.1145/3136755.3136815

[6] Euan Freeman, Graham Wilson, Dong-Bach Vo, Alex Ng, Ioannis Politis, and Stephen Brewster. 2017. *Multimodal feedback in HCI: haptics, non-speech audio, and their applications*. Association for Computing Machinery and Morgan & Claypool, 277–317. https://doi.org/10.1145/3015783.3015792

[7] Alfred Gabrielsson. 2009. The relationship between musical structure and perceived expression. *The Oxford handbook of music psychology* (2009), 141–150.

[8] Emilien Ghomi, Guillaume Faure, Stéphane Huot, Olivier Chapuis, and Michel Beaudouin-Lafon. 2012. Using Rhythmic Patterns as an Input Method. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA). 1253–1262. https://doi.org/10.1145/2207676.2208579

[9] Antti Jylhä, Inger Ekman, Cumhur Erkut, and Koray Tahiroğlu. 2011. Design and Evaluation of Human-Computer Rhythmic Interaction in a Tutoring System. *Computer Music Journal* 35, 2 (2011), 36–48.

[10] Hanna Landin, Sus Lundgren, and Johannes Prison. 2002. The Iron Horse: A Sound Ride. In *Proceedings of the Second Nordic Conference on Human-Computer Interaction* (Aarhus, Denmark). 303–306. https://doi.org/10.1145/572020.572075

[11] Toni Pakkanen, Roope Raisamo, and Veikko Surakka. 2014. Audio-Haptic Car Navigation Interface with Rhythmic Tactons. In *Haptics: Neuroscience, Devices, Modeling, and Applications*, Malika Auvray and Christian Duriez (Eds.). Springer Berlin Heidelberg, 208–215.

[12] Michal Rinott and Noam Tractinsky. 2022. Designing for interpersonal motor synchronization. *Human–Computer Interaction* 37, 1 (2022), 69–116. https://doi.org/10.1080/07370024.2021.1912608

[13] Davide Rocchesso, Alessio Bellino, and Antonino Perez. 2023. TickTacking–Drawing trajectories with two buttons and rhythm. In *Proceedings of the Sound and Music Computing Conference*. Stockholm, Sweden.

[14] Davide Rocchesso, Pietro Polotti, and Stefano Delle Monache. 2009. Designing continuous sonic interaction. *International Journal of Design* 3, 3 (2009).

[15] Y. Visell, F. Fontana, B.L. Giordano, R. Nordahl, S. Serafin, and R. Bresin. 2009. Sound design and perception in walking interactions. *International Journal of Human-Computer Studies* 67, 11 (2009), 947–959. https://doi.org/10.1016/j.ijhcs.2009.07.007