The Effect of Rhythm in Mid-air Gestures on the User Experience in Virtual Reality

Vincent Reynaert, Florent Berthaut, Yosra Rekik, Laurent Grisoni

The Effect of Rhythm in Mid-air Gestures on the User Experience in Virtual Reality

Vincent Reynaert[1], Florent Berthaut[1], Yosra Rekik[2], and Laurent Grisoni[1]

[1] Univ. Lille, UMR 9189 CRIStAL, F-59000 Lille, France, vincent.reynaert, florent.berthaut, laurent.grisoni@univ-lille.fr

[2] Univ. Polytechnique Hauts-de-France, LAMIH, CNRS, UMR 8201, F-59313, Valenciennes, France, yosra.rekik@uphf.fr

Abstract. In this work, we examine the effect of mid-air gesture rhythm on user experience in Virtual Reality. In particular, we investigate gesture regularity, speed and highlighting with a sound guide. We measure the effect of these components on the perceived fatigue, presence, difficulty, success and helpfulness. Our findings indicate that an irregular and slow rhythm leads to a lower arm fatigue. We also find that such an irregular rhythm could increase the user perceived difficulty of the task and the absence of a sound guide could decrease the sense of presence.

Keywords: Virtual Reality, Mid-air gestures, Muscular Fatigue, Presence, Rhythm, Sound Guide, User Experience

1 Introduction

VR headsets have become increasingly ubiquitous, with VR applications that support long period VR experience like immersive cinematographic content, video games or creative applications. A major feature of VR applications is their ability to enable users to directly manipulate information with their hands through mid-air gestures. However, interacting through mid-air gestures during a long period of time can affect the perceived arm fatigue due to the "gorilla-arm effect" [2,10] which can eventually lead to physical injury and consequently can deteriorate the user experience. Thus, designing gestures for a VR application can prove a challenging task. The current practice of mid-air gestures design and VR applications has outlined several guidelines to assist practitioners in this regard [6,10,20]. Researchers have also reported users preferences in arm position [2,4,8,10] and provided specific insight on how alternative gesture sets should be designed [10] to reduce fatigue. Different arm fatigue measures have been also proposed to characterize the gorilla-arm effect including subjective ones (like Borg [1], NASA-TLX [9] or Likert ratings) and quantitative ones like the "Consumed Endurance" metric [10].

Aside from understanding how to measure such fatigue, it is also important to investigate approaches to diminish it. Remarkably, perceived fatigue has been shown to be affected by music, especially when synchronised with the user's movements. For example, Szmedra et al. [18] demonstrated that listening to music while running decreases perceived muscular fatigue. Williams et al. [21] showed that generative music with a rhythm synchronised to the cadence of runners can improve their performance but also decrease their perceived effort. However, and despite that rhythm and audio have been proved useful to reduce fatigue in real scenarios, no study have examined the effect of gesture rhythm on user experience during mid-air gestures in VR, in particular of components of rhythm such as regularity, speed and highlighting using sound guides. Previous work on rhythm in gestural and 3D interfaces have focused on uses other than fatigue reduction.

In most interactive systems, users have full control over application automata execution, through the time that interaction events are generated. In VR and for some specific use-cases, it is occasionally possible that some time constrains may exist, *e.g.*, with timers for gaze-based selection or (standard or serious) games. For interaction situations in which it is reasonable to have the application influence the rate at which user is interacting, and in order to explore approaches to control perceived interaction fatigue, gestural rhythm is one of the promising exploration roads, as advocated by Costello [12]. For example, Mueller et al. [14] proposed to "Help players identify rhythm in their movements", and to acknowledge the rhythm in a sequence of gestures. Rhythm has been also used as a way to input commands through sequences of tapping gestures [7] or with micro-gestures performed at a given tempo [5] or for target selection by following trajectories at a given speed with mid-air gestures [3,19]. In 3D User Interfaces, sound guides were primarily investigated to help visually impaired users find targets in a 3D environment [13,15].

In this context, among other fundamental questions, one can for instance ask the following: Does imposing a rhythm in the production of the gesture have an effect on the fatigue? What causes articulation difficulty and what triggers perceived fatigue? How does gesture speed affect the user experience? Is there a link between the regularity of the speed and the perceived fatigue? Do sound guides helps users in increasing the perceived presence?

We argue that rhythm in mid-air gestures is an important factor for improving the user experience in VR applications, including perceived fatigue and presence, that has been little explored so far and, consequently, is little understood. In this paper, we conducted an experiment to examine the effect of rhythm (in terms of speed, regularity and highlighting using a sound guide) in mid-air gestures on the user experience in VR and we provide the community with some insight on this phenomenon. We used a pointing task where the participants were asked to follow a target moving to a predefined rhythm in the presence or not of a sound guide.

2 Experiment

We conducted an experiment to measure the effect of rhythm during mid-air pointing task in VR on user experience in terms of user perceived fatigue, presence, difficulty, success and helpfulness.

Label	**Question**
Post. Fatigue	How tiring did you find this task? (Borg CR10)
Presence	Average result of the Adaptation/Immersion presence questionnaire [16].
Difficulty	How difficult did you find this task?
Success	How successful did you feel in performing the task?
Liveliness	How lively did you find the rhythm?
Helpfulness	How much did the sound help you in following the target?

Table 1: Questions employed to elicit user experience. For the fatigue we used the Borg CR10 [1] and for the remainder questions, we used a 7-points Likert-scale question (strongly disagree to strongly agree).

Participants. 15 participants (12 males, 3 females) volunteered to take part in our experiment. They were aged between 18 and 45 years ($mean = 28.4, s.d = 6.8$). Thirteen participants were right handed and two were left handed. Two participants considered themselves as experts in understanding the rhythm, seven as proficient, three as intermediate and three as novices. Five participants defined themselves as VR experts, three participants used it frequently, six others occasionally and the last one had never tested it.

Due to the COVID-19 pandemic, part of this experiment was done remotely with seven participants who owned a VR headset. They were recruited via mailing-lists and forums. These participants downloaded the experiment software which was a Godot application build for SteamVR and Oculus Quest. Then, they performed the task while videoconferencing with one of the authors. The remainder eight participants carried out the experiment in our laboratory and used the SteamVR version. We follow all necessary sanitary precautions, in particular cleaning the equipment before and after use. The headsets used by the participants were: 2 × Oculus Quest 2, 1 × Oculus Quest 1, 1 × Oculus Rift S, 2 × HTC Vive Pro and 8 × Valve Index in the lab.

Design. The experiment used a 2 × 2 × 2 × 5 within-subjects design for the factors: *regularity*, *speed*, *sound* and *time period*. The *regularity* of the rhythm of the target's displacement covers two conditions: *regular* (there is no change in speed) and *irregular* (speed is randomly reduced by 0%, 5%, 10%, 25%, 50%, but never increased in order to avoid higher speeds). We chose a random variation to prevent any learning of changes in speed that participants might develop. The *speed* of the target to follow covers two conditions: *slow* (1s = 4 beats = 1 trajectory, *speed*=1 m/s) and *fast* (700ms = 4 beats = 1 trajectory, *speed*=1.4 m/s). The *sound* played at the same rhythm as the target displacement covers two conditions: *playing* (audio feedback is used to highlight the rhythm, consisting in four sounds played per trajectory, one at each beat) and *muted* (no audio feedback). The *time period* corresponds to a re-sampling of the task in five successive periods of time: 30-60, 60-90, 90-120, 120-150, 150-180 seconds. We voluntarily removed the first period of time 0-30 from our analysis to reduce the bias in speed and accuracy when starting to follow the target at the beginning of the task.

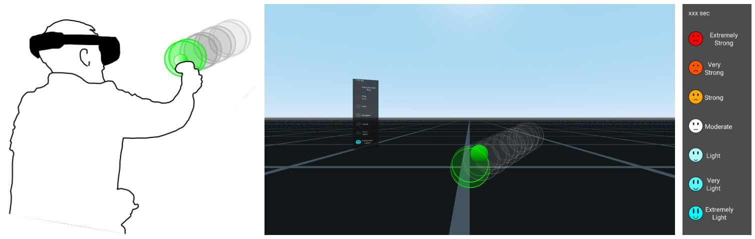

Fig. 1: Experimental setup. Left: Seated participant with VR headset. Center: what they saw, i.e. the moving Borg CR10 scale on the left and the spherical cursor and cylindrical target with its trajectory. Right: Borg Scale Panel

Task. We use a pointing task in which participants had to keep a spherical cursor inside a cylindrical target moving on the XY plane in front of them. The trajectory of the target was shown as a trail behind the target, moving towards the user. The target the participant had to follow was a cylinder with a $radius = 25cm$ and a $height = 10cm$. The controller was represented by a sphere of $radius = 5cm$. In addition, the target was constrained in a circle whose center was located at ($x = 0cm$, $y = 100cm$, $z = -46cm$) and with $radius = 50cm$. The result is depicted in Figure 1 and in a short video (download or watch).

Procedure. At the beginning of the experiment, participants signed an informed consent agreement. Participants were then instructed to seat on a chair without armrest and to make sure to have a clear area in front of them in order to avoid risk of collisions when interacting with the application. All participants then put their VR headset on. Distant participants were instructed to calibrate as accurately as possible the ground and the center of their VR headset so that the interaction area was the same for all. When participants opened the virtual environment, the task was explained to them both orally and with a text on a virtual panel. Participants then filled a brief demographic questionnaire.

The experiment started with a training phase composed of one training block before moving to the experimental phase composed of eight blocks. In both phases, participants were instructed to follow the target as accurately as possible with their dominant hand. In parallel they had to use their non-dominant hand to report their perceived arm fatigue on the Borg scale by using the joystick on the controller. This scale was displayed on a panel that appeared every 15 seconds on the left side of the interaction area and slowly moved out of the field of view of participants while fading out, in order to remind participants to provide ratings regularly.

After each trial participants had to complete a questionnaire composed of the Adaptation/Immersion part of the presence questionnaire [16] with additional questions on their final level of fatigue and on their perception of the task. Table 1 provides the list of corresponding questions. At the end of the questionnaire, participants were instructed to rest as much as possible in order to reduce the effect of accumulated fatigue. Participants then reported their initial level of fatigue for the following trial. During tasks, we also recorded the hand and target positions, in order to retrieve the distance to the target.

The training phase was designed so that participants familiarize themselves with the task without creating a learning or order effect. Therefore, all participants performed a first trial in a condition which was not part of the ones we wanted to compare. More specifically, in this training phase, we alternated each 8 beats (2 trajectories) between *sound playing* and *sound muted* conditions, we set a medium *speed* (850ms = 4 beats = 1 trajectory) and the *regularity* was *regular*. This allowed participants to understand the relation between their movements and the imposed rhythm and the sound, and to get used to reporting their level of fatigue on the Borg scale.

In the experiment phase, participants then successively performed eight blocks (2 *regularity*× 2 *speed*× 2 *sound*). The order of the 8 blocks was counterbalanced across participants through a balanced latin square, in order to avoid an order effect on fatigue or engagement. For each block, participants completed the six periods of time for a total of 180s. After each block, participants took a break. The experiment took around one hour.

3 Results

In this section, we present the results for the performances and the questionnaire. A Shapiro-Wilk normality test [17] showed that the data was not normal for all the measures of performance ($.39 < W < .93$ and $p < .0001$), consequently, we used the ARTool [11,22] test to perform an ANOVA on non-parametric data, followed by post-hoc ART-Contrast for statistically significant main effects or interactions. In addition, as the answers to the questionnaire correspond to ordinal data, the same ARTool method was applied. In the following, we report significant results.

3.1 User Performances

We measured the following depending variables: the *perceived arm fatigue* (evolution of the fatigue on the Borg CR10 scale during the task), the *distance to target* (distance between the hand and the target) and the *accuracy* (ratio of time spent in the target). Figure 2 shows bar-plots of all conditions for results with statistically significant differences.

During Task Fatigue. There were significant main effects of *regularity* ($F_{1,546} = 13.21$, $p < .0001$) and *time period* ($F_{1,546} = 181.34$, $p < .0001$) on *fatigue*. Post-hoc tests showed that *regular* tasks were rated significantly more fatiguing than *irregular* ones ($p = .0003$). We also found that the perceived fatigue increased significantly across increasing period of time ($p < .001$).

Distance to Target. There was a significant main effect of *speed* ($F_{1,546} = 94.52$, $p < .0001$) on *the distance of the hand to the target*. Post-hoc tests revealed that *fast speed* increased significantly the distance between the hand and the target than when using *slow speed* ($p < .0001$).

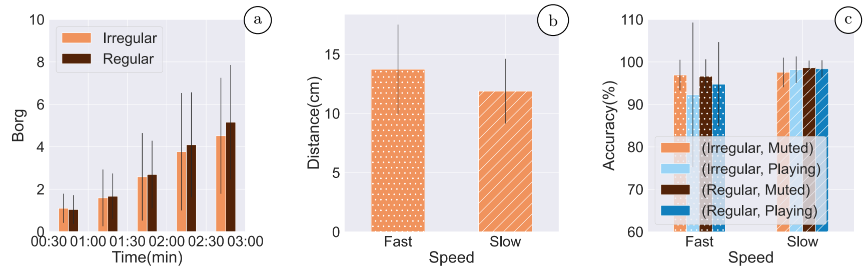

Fig. 2: Results for each of the dependent measures: (a) Average Level of Arm Fatigue by *regularity* and *time period*, (b) Average Distance between Hand and Target by *speed*and (c) Average Accuracy by *regularity*, *sound* and *speed*.

Accuracy. There were significant main effects of *regularity* ($F_{1,546} = 7.49$, $p = .0064$) and *speed* ($F_{1,546} = 45.91$, $p < .0001$) on *accuracy* with significant *regularity* × *speed* ($F_{1,546} = 4.66$, $p = .0313$), *speed* × *sound* ($F_{1,546} = 4.29$, $p = .0388$), *regularity* × *speed* × *sound* ($F_{1,546} = 8.15$, $p = .0045$) interactions. Post-hoc tests revealed that when using either *regular regularity* in *sound muted* or *irregular regularity* in *sound playing*, *fast speed* was significantly less accurate than *slow speed* ($p < .05$). However, the accuracy remained high in all cases (min=92.26%, max=98.62%).

3.2 Questionnaire

Figure 3 shows bar-plots of all conditions for questions with statistically significant differences.

Post-Task Fatigue. There were significant main effects of *speed* ($F(1, 98) = 12.91, p = .0005$) and *sound* ($F(1, 98) = 5.96, p = .0165$) on *fatigue*. Post-hoc tests revealed that both the *fast speed* and the *sound playing* significantly implied more perceived fatigue than respectively the *slow speed* and the *sound muted* ($p < .05$).

Presence. There was significant main effect of *sound* ($F(1, 98) = 7.30$, $p = .0081$) on *presence*. Post-hoc tests revealed that the *sound playing* implied a significantly greater sense of presence than the *sound muted* ($p < .05$).

Difficulty. There were significant main effects of *regularity* ($F(1, 98) = 11.93$, $p = .0008$) and *speed* ($F(1, 98) = 6.05$, $p = .0157$) on difficulty. Post-hoc test revealed that the *irregular regularity* conditions were rated significantly more difficult than the *regular* ones ($p < .001$). In addition, the *fast speed* was rated, significantly more difficult than the *slow speed* ($p < .05$).

Feeling of Success. There was a significant main effect of *speed* ($F(1, 98) = 4.96$, $p = .0282$) on *feeling of success*. Post-hoc tests showed the *slow* speed increased significantly the feeling of the success than *fast* speed ($p < .05$).

Liveliness. There was a significant main effect of *sound* ($F(1, 98) = 30.72$, $p < .0001$) on *liveliness*. Post-hoc revealed that the *sound playing* helped participants feel the rhythm and made it more lively than the *sound muted*($p < .0001$).

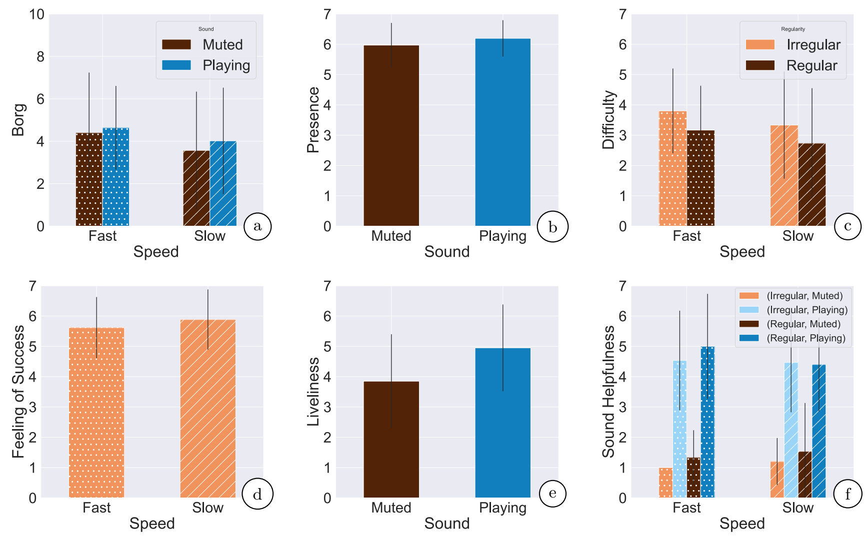

Fig. 3: Results for each of the dependent measures: (a) Average Level of Arm Fatigue at the End of Tasks by *sound* and *speed*, (b) Average Sense of Presence by *sound*, (c) Average Perceived Difficulty by *regularity* and *speed*, (d) Average Feeling of Success by *speed*, (e) Average Liveliness by *sound* and (f) Average Feeling of Helpfulness of the Sound by *regularity*, *speed* and *sound*.

Sound Helpfulness. There was a significant main effect of *sound* ($F(1, 98) = 221.76$, $p < .0001$) on *sound helpfulness* with significant *regularity* $\times$ *sound* ($F(1, 98)= 7.09$, $p = .0091$) and *sound* $\times$ *speed* ($F(1, 98) = 4.48$, $p = .0369$) interactions. Post-hoc tests revealed that for both *regular* and *irregular regularity*, *sound playing* was perceived as helping target following when compared with *sound muted* ($p < .0001$). We also found that for both *fast* and *slow speed*, *sound playing* was perceived as helping target following when compared with *sound muted* ($p < .0001$).

3.3 Discussion of Experiment Results

Effect of the *sound*. In terms of perceived fatigue, our findings indicate that the *sound* does not have an impact during the task. However and contrarily to related work [18,21], after tasks, we found an increase in perceived fatigue when the *sound* was played. These findings could be explained by a greater sense of presence and liveliness in *sound playing* . However, only one participant explicitly confirmed this. Five other participants found that they were more concentrated on their arm fatigue during *sound muted* and consequently they felt more fatigued. In addition, even if the *sound playing* is perceived helpful to follow the target, it did not seem to impact the perception of success or difficulty of tasks. These results seem to contradict previous research in which music helped increase the endurance [21]. However, we used basic sounds which only served to highlight the target displacement and could therefore reinforce the perception of effort, compared to more complex music which might instead "mask" the effort.

Effect of the *speed*. In terms of perceived fatigue, as for the *sound*, participants felt an impact of the *speed* only after the task but unsurprisingly a *fast speed* implied more fatigue than a *slow* one. However, during interviews at least five participants felt that the *slow speed* was more fatiguing that the *fast* because it was less "fun". The *fast speed* also had a bad impact on the distance between the hand and target and on the accuracy (although it remained high). In the same way, *fast speed* increased the perception of difficulty and decreased the feeling of success. In other words, the *fast speed* had a negative impact on users performances and on their perceived arm fatigue.

Effect of the *regularity*. In terms of perceived fatigue, contrarily to speed and sound, we found that the *regularity* has an impact during the task but not after, with the *irregular* condition being less fatiguing than the *regular* one. This difference in terms of fatigue could be explained by the slower average speed induced by the irregularity. However, the *irregular* condition also seemed to increase the perceived difficulty, which would not be the case if the difference in *regularity* was only a difference in speed. In addition, three participants suggested that the in the *irregular* condition was funnier and more engaging and therefore that they were less focused on the fatigue.

4 Conclusion

Through an experiment, we explored how the rhythm of mid-air gestures impacted the user experience in Virtual Reality. We found that, if the aim is to reduce fatigue, the use of *irregular* and *slow* gestures without a *sound* highlighting the gesture speed may be an interesting choice. However, our results also suggest that the absence of *sound* reduces the sense of presence and liveliness. Our experiment suffers from some limitations which could be lifted in future studies. A first bias may come from the lack of diversity in participants' gender and from the remote and therefore little controlled conditions of the experiment. Our results should be refined with more participants and by looking at the impact of gender, level of VR expertise and dominant hands. As pointed out earlier, we observed differences with related work on the effect of the sound on perceived fatigue. Future work should investigate the role of sound/music complexity and the effect of mapping the sound with either the wanted trajectory or the user's gesture. While our results suggest a large effect of the *speed* and *regularity* on the user experience, variations on these parameters should be investigated, such as slower and faster speeds and patterns in irregular rhythms. We believe that our results are applicable beyond VR, in mid-air interactions where the temporality of gestures is controlled or when a long-term use is frequent, such as games or creative applications. In particular, rehabilitation exercises require series of repetitive movements whose rhythm could be adjusted to control fatigue.

References

1. Borg, G.: Borg's perceived exertion and pain scales. Borg's perceived exertion and pain scales, Human Kinetics, Champaign, IL, US (1998), pages: viii, 104
2. Boring, S., Jurmu, M., Butz, A.: Scroll, tilt or move it: using mobile phones to continuously control pointers on large public displays. In: Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7. pp. 161–168 (2009)
3. Carter, M., Velloso, E., Downs, J., Sellen, A., O'Hara, K., Vetere, F.: Pathsync: Multi-user gestural interaction with touchless rhythmic path mimicry. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. pp. 3415–3427 (2016)
4. Cockburn, A., Quinn, P., Gutwin, C., Ramos, G., Looser, J.: Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback. International Journal of Human-Computer Studies **69**(6), 401–414 (2011)
5. Freeman, E., Griffiths, G., Brewster, S.A.: Rhythmic micro-gestures: Discreet interaction on-the-go. In: Proceedings of the 19th ACM International Conference on Multimodal Interaction. pp. 115–119 (2017)
6. Fuller, J.R., Lomond, K.V., Fung, J., Côté, J.N.: Posture-movement changes following repetitive motion-induced shoulder muscle fatigue. Journal of Electromyography and Kinesiology **19**(6), 1043–1052 (2009)
7. Ghomi, E., Faure, G., Huot, S., Chapuis, O., Beaudouin-Lafon, M.: Using rhythmic patterns as an input method. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1253–1262 (2012)
8. Harrison, C., Ramamurthy, S., Hudson, S.E.: On-body interaction: armed and dangerous. In: Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction. pp. 69–76 (2012)
9. Hart, S.G., Staveland, L.E.: Development of nasa-tlx (task load index): Results of empirical and theoretical research. In: Advances in psychology, vol. 52, pp. 139–183. Elsevier (1988)
10. Hincapié-Ramos, J.D., Guo, X., Moghadasian, P., Irani, P.: Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In: Proceedings of the 32nd annual ACM conference on Human factors in computing systems. pp. 1063–1072. ACM (2014)
11. Kay, M., Wobbrock, J.O.: Package 'artool' (2020)
12. Mary Costello, B.: Paying attention to rhythm in hci: Some thoughts on methods. In: 32nd Australian Conference on Human-Computer Interaction. pp. 471–480 (2020)
13. Mereu, S.W., Kazman, R.: Audio enhanced 3d interfaces for visually impaired users. ACM SIGCAPH Computers and the Physically Handicapped pp. 10–15 (1997)
14. Mueller, F., Isbister, K.: Movement-based game guidelines. In: Proceedings of the sigchi conference on human factors in computing systems. pp. 2191–2200 (2014)
15. Sánchez, J., Sáenz, M.: 3d sound interactive environments for problem solving. In: Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility. pp. 173–179 (2005)
16. Schneider, D.K.: Presence Questionnaire (PQ) - EduTech Wiki (Jul 2019), http://edutechwiki.unige.ch/en/Presence/Presence_Questionnaire_(PQ)
17. Shapiro, S.S., Wilk, M.B.: An analysis of variance test for normality (complete samples). Biometrika **52**(3/4), 591–611 (1965)

18. Szmedra, L., Bacharach, D.: Effect of music on perceived exertion, plasma lactate, norepinephrine and cardiovascular hemodynamics during treadmill running. International journal of sports medicine **19**(1), 32–37 (1998)
19. Velloso, E., Carter, M., Newn, J., Esteves, A., Clarke, C., Gellersen, H.: Motion correlation: Selecting objects by matching their movement. ACM Transactions on Computer-Human Interaction (TOCHI) **24**(3), 1–35 (2017)
20. Wang, H.: How will different control/display ratios influence muscle fatigue and game experience in virtual reality games? (2016)
21. Williams, D.A., Fazenda, B., Williamson, V.J., Fazekas, G.: Biophysiologically synchronous computer generated music improves performance and reduces perceived effort in trail runners. In: Michon, R., Schroeder, F. (eds.) Proceedings of the International Conference on New Interfaces for Musical Expression. pp. 531–536. Birmingham City University, Birmingham, UK (Jul 2020), https://www.nime.org/proceedings/2020/nime2020_paper102.pdf
22. Wobbrock, J., Findlater, L., Gergle, D., Higgins, J.: The aligned rank transform for nonparametric factorial analyses using only anova procedures. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (2011)