

Self-Generated Auditory Feedback as a Cue to Support Rhythmic Motor Stability

Stas S. Krupenia^{*}, Pablo F. Hoffman^{*†}, Hagar Zalmanov[‡], Veronica .K.E. Krupenia[‡],
Daniel Gopher[‡]

(^{*}) *Thales Research and Technology, the Netherlands*

(^{*†}) *Aalborg University, Denmark*

([‡]) *Israel Institute of Technology, Israel*

*E-mail: stas.krupenia@d-cis.nl, pfh@es.aau.dk, hagar.zalmanov@gmail.com,
veronica@krupenia.com, dgopher@ie.technionl.ac.il*

Abstract

A goal of the SKILLS project is to develop Virtual Reality (VR)-based training simulators for different application domains, one of which is juggling. Within this context the value of multimodal VR environments for skill acquisition is investigated. In this study, we investigated whether it was necessary to render the sounds of virtual balls hitting virtual hands within the juggling training simulator. First, we recorded sounds at the jugglers' ears and found the sound of ball hitting hands to be audible. Second, we asked 24 jugglers to juggle under normal conditions (Audible) or while listening to pink noise intended to mask the juggling sounds (Inaudible). We found that although the jugglers themselves reported no difference in their juggling across these two conditions, external juggling experts rated rhythmic stability worse in the Inaudible condition than in the Audible condition. This result suggests that auditory information should be rendered in the VR juggling training simulator.

1. Introduction

Training in virtual domains can have many benefits compared to real world training [1, 2, 3]. In general, a richer sensory VR environment supports better learning; however, care must be taken to ensure a good match between the real and VR items [4]. The SKILLS VR-based juggling training simulator is intended to provide a non-juggler with the skills needed to juggle the three ball cascade (a basic juggling pattern; 3BC). Of interest to SKILLS is the value of multimodal information for assisting skill acquisition. In this study we examined the role of sounds generated when balls hit the hand as a cue to support motor stability.

For juggling, there are four sensory modalities that can provide feedback about performance; visual (location of balls, arms, and hands), kinesthetic/proprioceptive (the relationships between limbs/appendages), haptic (the feel of the ball entering/leaving the hands), and auditory (the sound of the ball hitting the hands). It is clearly necessary for juggling to use the visual modality and indeed for SKILLS this was used to represent the trainee's hands and three balls within VR [4]. In the current study we examined whether it is necessary or not to render the sound of balls hitting hands in the virtual environment.

An important distinction made in relation to juggling skill is between spatial and temporal accuracy. To examine the spatial and temporal constraints associated with juggling the 3BC, [5] had three intermediate jugglers (could juggle no more than three balls) and three expert jugglers (could juggle up to five balls) juggle a 3BC under spatially or temporally constrained conditions. Their goal was to examine the spatiotemporal characteristics of the pattern and to test the assumption that jugglers have an internally represented spatial clock. Van Santvoord and Beek examined if spatial variability increased with temporal constraints and vice versa for temporal variability [5]. Contrary to expectations, spatiotemporal variance associated with temporal constraints was similar to the variance associated with spatial constraints. The authors stated that despite this, the idea of a spatial clock remained valid because participants appeared to minimize the variability of flight time by throwing as consistently as possible. In other words, it appeared that juggling to an externally specified height more closely approximated the task of juggling than juggling to an external beat. Finally, the authors stated that the results of their study suggest that in learning to juggle, it may be more beneficial to provide spatial

feedback than temporal feedback however; the goal should be a combination.

The distinction between spatial and temporal factors is consistent with the informal view in the juggling community of two different juggling styles; technical and artistic. In technical juggling, the goal is to juggle many balls whereas in artistic juggling, the goal is to juggle ‘beautiful’ patterns. It can be argued that spatial skills are more critical for technical juggling whereas temporal skills are more critical for artistic juggling.

Previous research has shown that the role of intrinsic information, such as vision and proprioception, is essential for learning [6]. Many studies have examined general patterns of temporal and spatial skill acquisition for juggling, but few studies have examined training methods to support skill acquisition [5]. One study that partly examined juggling training methods was [7] who examined juggling skill acquisition with or without a metronome to provide rhythmic cues. The authors found no effect of auditory cue on skill acquisition; a finding consistent with [5] who noted the comparative greater importance of learning correct trajectories than correct rhythm (indicated by the metronome).

Clearly spatial and temporal components of juggling are intrinsically linked but literature suggests a dominance of spatial versus temporal components. Consistent with this, spatial feedback may be more important for skill acquisition than temporal feedback because jugglers appear to form a spatial map of the pattern more so than a temporal map. However, given the intrinsic link between the two components, and that auditory cues can deliver both temporal and spatial information, it was predicted that the absence of auditory feedback would result in poorer juggling.

To develop the SKILLS juggling training simulator we needed to know whether or not the sound of the virtual balls hitting the virtual hands must be rendered in VR. To address this question, we asked 24 jugglers to juggle two different juggling patterns under normal (audible) or inaudible (wearing headphones) juggling conditions. Before examining this, however, it was first necessary to determine whether jugglers could indeed hear the sound of the ball hitting their hand.

2. Pilot

To examine if jugglers can hear the sounds of balls hitting their hands, binaural audio was recorded while a juggler completed a 3BC. Microphones were placed at the entrance of a juggler’s ear canals. A spectrogram of these recordings is presented in Figure 1. The ordinate indicates frequency and the abscissa indicates time with top and bottom panels representing left and right ears respectively. Arrows

on the spectrogram indicate ball-hitting-hand events. The grey scale indicates acoustic energy with darker areas representing higher energy. The sound produced when a ball is caught by the hand is a soft impact sound—energy is briefly distributed along a wide frequency range (see arrows in Figure 1). These recordings show that under typical sound levels, the sound of the ball hitting the hands is likely audible.

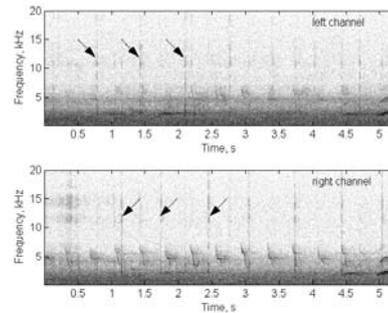


Figure 1. 5-second spectrogram of binaural recordings during three-ball cascade juggling from left (Top) and right (Bottom) ears.

This was supported by informal reports from the juggler indicating that he could hear the sounds. On the basis of these data, we predicted that when audio feedback is removed, there will be a decrease in performance on spatial and temporal accuracy.

3. Experiment 1

Participants juggled two patterns under different auditory conditions. Each participant completed four two-minute juggling blocks. All participants were video-recorded and performance was rated using self-report and external rating measures. The aim was to examine the value of auditory feedback on successful juggling to determine whether or not to include auditory rendering in the SKILLS training platform.

3.1 Methods

Participants. Twenty four males attending the 16th Israeli Juggling Convention (Gan Hashlosha, Israel) voluntarily participated in this experiment. The average age was 21.2 years ($SD=4.47$) and the average number of years juggling was 6 ($SD=2.94$, here, $N=11$). The decision to include only men was made because males are more represented within this community; there were only four females present in random sample of 45 people attending the 15th Israeli Juggling Convention

Juggling Task. Participants were asked to juggle two different juggling patterns; the 3BC and the Newest Trick (NT) that the participant had learnt.

The NT was included to avoid a ceiling level of performance. Almost all participants completed a different trick for the Newest Trick.

Auditory Conditions. Participants juggled the two different juggling patterns (3BC, NT) in two auditory conditions; Inaudible (I) or Audible (A). For the Inaudible conditions, participants wore a pair of Panasonic RP HNJ50 headphones and listened to pink noise presented diotically (48 kHz). It was not empirically tested whether or not the pink noise (combined with the headphones) wholly masked the juggling sounds, however, given the exploratory nature of this study, it was decided not to pursue more technical methods for controlling the inner ear sound pressure level and auditory masking. Instead, it was decided to rely on participant feedback regarding the degree of auditory masking. Participants could adjust the level of the pink noise so that it was both comfortable and masked external sounds.

Experiment Design. This 2 x 2 (Juggling Task x Auditory Condition) experiment was a within-subjects design, thus each person completed all four conditions; 3BC-A, 3BC-I, NT-A, and NT-I. The presentation of conditions was fully counterbalanced. Videos were made of each participant juggling each condition (96 video in total). At the conclusion of the experiment, a forty second sample taken from the middle of the two minute video recording was extracted.

Self Report Ratings. At the end of the experiment participants rated their juggling performance for each trial from 1 (extremely badly) to 100 (extremely well).

Expert Ratings. The 96 video clips were given to two juggling experts (MO [Tusen Konster och en boll] and AD [The University of Western Australia Juggling Club]) that were unfamiliar with the experiment. The two experts were asked to rate the juggling in each video on three scales; Technical Accuracy, Rhythmic Stability, and Overall Performance. Technical Accuracy was defined as “was the ball released at the correct position, was the ball caught in the correct position, did the arms move in the best possible pattern for the intended pattern, did the juggler ensure that their movements were as efficient as possible... Compare the juggling of the juggler to how a juggling simulator would complete the pattern”. Rhythmic Stability was defined as “to what extent was the juggler able to juggle with the intended rhythm for the pattern, was there a good ‘flow’ to the pattern? Was the juggling pattern ‘smooth’”? For overall performance, the external raters were asked to “rate the overall quality of the juggling”. The ‘Rhythmic Stability’ and ‘Technical Accuracy’ scales were intended to index temporal and spatial components of the juggling, but were presented in a way that was more meaningful to the expert raters.

3.2 Procedure

Potential participants at the Juggling Convention received an information sheet about the experiment. If they wanted to participate they selected a time from a list provided and returned to the test area at that time. When participants arrived for testing they were asked to read and sign an informed consent form. Experiment instructions were read to participants who also received written instructions. The video camera was then turned on. Participants completed two minutes of juggling for their first condition followed by a thirty second break. Participants then juggled their second condition followed by a thirty second break, and so on for the four conditions. At the end of the four juggling trials, participants completed a debriefing questionnaire and the subjective juggling ratings.

3.3 Results and Discussion

Self Report Ratings. Participants rated their performance as being better in the Inaudible than Audible conditions but only for the 3BC, $t(23)=1.87$, $p=0.081$; for NT, $t(23)=2.652$, $p=0.015$.

Expert Ratings. Expert Jugglers rated performance of the participants as being better in the Audible Condition than in the Inaudible condition but only on Rhythmic Stability (Table 1).

Table 1. Differences between Ratings of performance in Audible (A) and Inaudible (I) conditions

	3BC	NT
Rhythmic Stability	Better in A than I, $t(45)=2.60$, $p<0.01$	Better in A than I, $t(45)=2.15$, $p<0.05$
Technical Accuracy	No difference, $t(45)=0.72$, $p=0.48$	No difference, $t(45)=0.19$, $p=0.85$
Overall Performance	No difference, $t(45)=1.77$, $p=0.08$	No difference, $t(45)=1.06$, $p=0.29$

Inter-rater reliability. There was significant correlation between expert raters for all comparisons. The minimum correlation was 0.597 ($p<0.003$) for Overall Performance on the Newest Trick Audible Condition and the maximum was 0.816 for Rhythmic Stability 3BC Inaudible condition ($p<0.001$).

The Pilot data suggested that jugglers could hear the sounds of the balls hitting their hands. Furthermore, the spectrogram suggests that the distribution of sound is sufficient for identifying information about where the balls are landing. In other words, the auditory feedback generated from balls hitting hands provides information about rhythm and technical accuracy. Using data from the Pilot, the hypothesis for Experiment 2 was that when

self-generated auditory feedback was removed, juggling would be worse on both spatial and temporal components. This hypothesis was partially supported. From the external ratings, it was clear that without auditory feedback juggling rhythm was impaired but there was no change in spatial accuracy. Additionally, and contrary to expectations, the effect of the auditory manipulation was found even for the 3BC, suggesting that auditory feedback is a strong feedback source for maintaining rhythmic stability. The divergence between subjective and objective ratings highlights the importance of collecting behavioral data for examining training and skill acquisition.

4. Conclusion

The results from the Pilot and Experiment suggest that for the SKILLS juggling training simulator, rendering the sound of the virtual balls hitting the virtual hands will likely support the trainee's acquisition of rhythmic stability. However, it appears less important to render these sounds in three dimensions. The results suggest that auditory feedback provides an important cue for maintaining rhythmic stability but these cues are of little benefit for supporting spatial accuracy.

On the basis of the data collected in the current study, and in contrast to [5] it can be argued that jugglers do appear to have an internally represented temporal clock. Previously, it was found that using a metronome was not sufficient for assisting learning. In light of the current data, one reason for the limited value of a metronome may be that the internally represented rhythm of the intended beat is relatively simple and that a metronome (which provides an index of the intended behavior) does not provide any new information to the juggler. The results from the current study would suggest that rather than use an auditory signal to represent the intended pattern, it would be preferred to present and amplify actual performance—thus making it easier for the trainee to mentally compare actual and intended performance.

Furthermore, while [5] stated that it may be more beneficial to provide spatial feedback than temporal feedback the results from the current study highlight the value of self-generated audio feedback for maintaining stable juggling performance. The use of self-generated auditory feedback, as compared to information from other modalities, comes almost naturally when performing rhythmically organized motor behavior [8]. For example, auditory-motor feedback interactions are particularly relevant in the accurate control of timing and rhythm during music performance [9]. When playing an instrument, the musician must listen to each note produced and implement appropriately timed motor adjustments. Pfordresher and Palmer showed that introducing asynchronies to self-generated auditory feedback can

significantly disrupt the timing of events during music performance [10]. Results suggest that disruptions occur because both percepts and actions depend on a single mental representation [10]. On this basis, one could hypothesize that if disrupting the auditory percept causes motor disruptions, then enhancing the percept can improve the action.

One factor not examined in the current study was the informal classification of patterns as technical or artistic. It is possible that the value of auditory feedback is relative to task demands. Indeed research suggests that limiting auditory feedback prevents the artistic, “expressive aspects” of music performance [9]. However, the extent to which this applies to gross motor tasks remains unknown and should be examined.

The wider implications of the current research is that for training behaviors involving rhythmic motor behavior, audio feedback generated directly from the trainees behaviors may be a more effective training aid than external cues. By extrapolating the current results, it is possible to argue that for all activities involving a rhythmic component (e.g. rowing, cycling, cross country skiing, running...) skill acquisition can be enhanced by providing amplified self-generated audio cues. Trainees may benefit from these cues because they can better compare their actual rhythmic performance with the intended rhythmic pattern.

5. References

- [1] Weller, J. M. (2004). Simulation in undergraduate medical education: bridging the gap between theory and practice. *Medical Education*, 38, 32-38.
- [2] Chopra, V., Gesink, B. J., De Jong, J., Bovill, J. G., Spierdijk, J., & Brand, R. (1994). Does training on an anesthesia simulator lead to improvement in performance? *British Journal of Anesthesia*, 73, 293-297.
- [3] Gopher, D., Krupenia, S., & Gavish, N. (2011). Skill training in multimodal virtual environments. Taylor and Francis. In D. Kaber, & G. Boy (Eds.). *Advances in Cognitive Ergonomics*. CRC Press.
- [4] Lippi, V., Avizzano, C.A., Mottet, D., & Ruffaldi, E. (2010). Effect of delay on dynamic targets tracking performance in behavior in virtual environment. In *Proceedings of the 19th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 446-451). Viareggio, Italy: IEEE.
- [5] van Santvoord, A. A. M., & Beek, P. J. (1996). Spatiotemporal variability in cascade juggling. *Acta Psychologica*, 91, 131-151.
- [6] Zelaznik, H.N., & Lantero, D. (1996). The role of vision in repetitive circle drawing. *Acta Psychologica*, 92(1), 105-118.
- [7] Beek, P. J., & van Santvoord, A. A. M. (1992). Learning the cascade juggle: A dynamical systems analysis. *Journal of Motor Behavior*, 24, 85-94.
- [8] Repp, B. H. & Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychological Research*, 68, 252-270.
- [9] Zatorre, R. J., Chen, J. L. & Penhune, V. B. (2007). When the brain plays music: auditory-motor interactions in music perception and production. *Neuroscience*, 8, 547-558.
- [10] Pfordresher, P. Q. & Palmer, C. (2006). Effects of hearing the past, present, or future during music performance. *Perception & Psychophysics*, 68(3), 362-376.