Eye Tapping: How to Beat Out an Accurate Rhythm using Eye Movements

Rasmus Bååth*
Lund University
Cognitive Studies
Kungshuset, Lundagård
222 22 Lund, Sweden
rasmus.baath@lucs.lu.se

Thomas Strandberg
Lund University
Cognitive Studies
Kungshuset, Lundagård
222 22 Lund, Sweden
thomas.strandberg@lucs.lu.se

Christian Balkenius
Lund University
Cognitive Studies
Kungshuset, Lundagård
222 22 Lund, Sweden
christian.balkenius@lucs.lu.se

ABSTRACT

The aim of this study was to investigate how well subjects beat out a rhythm using eye movements and to establish the most accurate method of doing this. Eighteen subjects participated in an experiment were five different methods were evaluated. A fixation based method was found to be the most accurate. All subjects were able to synchronize their eye movements with a given beat but the accuracy was much lower than usually found in finger tapping studies. Many parts of the body are used to make music but so far, with a few exceptions, the eyes have been silent. The research presented here provides guidelines for implementing eye controlled musical interfaces. Such interfaces would enable performers and artists to use eye movement for musical expression and would open up new, exiting possibilities.

Keywords

Rhythm, Eye tracking, Sensorimotor synchronization, Eye tapping

1. INTRODUCTION

Humans have the ability to synchronize movements to an external rhythm. This is an unique ability not found among any other mammal, not even among the greater apes. Recently it has been shown that parrots and cockatoos have a limited ability to entrain to music [2]. Still it can not compare with the richness of human rhythmical expression whether it is dance or music making. It is not strange then that *sensorimotor synchronization*, the rhythmic coordination of perception and action, is an active field of research [11].

While most research is done on tasks such as finger tapping, eye movements and rhythm is an unexplored area. This lack of research might be because of the technical difficulties in measuring eye movements. Unobtrusive devices for measuring eye movements, so called *eye trackers,* have been available for many years. To accurately measure the timing of eye movements the temporal resolution of an eye tracker should be high however, and this is a problem as many eye trackers have had and still has relatively low temporal resolution (< 250 Hz). Nowadays there are commercial high-speed eye trackers available, some with a temporal resolution of 500 Hz or more, and resolution should not be a limitation anymore. There exists a few examples of using eye movements to control music [6, 4, 8] but none of them describe the use of eye movements to trigger sounds in a rhythmical fashion. A recent study [5] used eye movements to generate hand clap sounds in time with a metronome.

When finger tapping a rhythm, or when drumming, it is obvious when the actual strike to be synchronized with the beat is made. It occurs when the finger or drum stick hits a surface. When beating out a rhythm using the gaze and an eye tracker there is no surface to strike and it is not obvious when the actual strike is made.

When beating out a rhythm using gaze, henceforth called *eye tapping*, where is the "strike" felt? One obvious alternative is to use eye blinks as the triggering method but even if only the actual eye movements are used there are still many alternatives. Two types of eye movements that could be used for eye tapping are *fixations* and *saccades*. A fixation occurs when the gaze maintains the focus on a single location; saccades are the fast eye movements made between fixations. An example of how to trigger a sound by an eye movement would be to use the fixation onset, another example would be to trigger a sound in the middle of a saccade.

The aim of this study was twofold. The first aim was to investigate if, and how well, it would be possible to beat out a rhythm using eye movements. The second aim was to establish the most accurate method of eye tapping. There are four main reasons why it is interesting to investigate how to best control rhythm with the eyes and how well it can be done:

1. **To make it possible to use eye moments for musical composition.** Many parts of the body are used to make music; the mouth and lungs control wind instruments, fingers and arms control string instruments, and legs and feet are used when drumming but so far, with a few exceptions [6, 4], the eyes have been silent. To enable performers and artists to use eye movement for musical expression would open up new, exiting possibilities.

2. **To enable people with physical disabilities to make music.** Eye tracking is already used by people with physical disabilities to interact with computers. If more was known about how to control rhythms with eye movements it would make it possible for them to enjoy new opportunities for musical expression.

3. **To learn more about sensorimotor synchronization.** The literature on sensorimotor synchronization

*Corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
NIME'11, 30 May–1 June 2011, Oslo, Norway.
Copyright remains with the author(s).

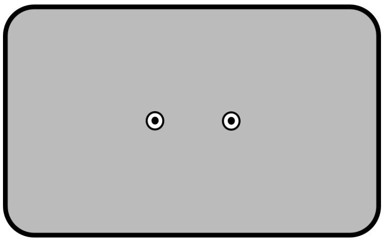

Figure 1: The display shown to the subjects during the experiment. The distance between the two fixation points was either 10 or 20 visual degrees.

has been preoccupied with studying finger tapping and few other means of striking a beat has been tried. Eye movements and finger movements might have different properties and our knowledge about sensorimotor synchronization in humans would not be complete without accounting for eye movements.

4. **To enable the use of rhythmic control in gaze controlled interfaces.** Knowledge about how to best control rhythm with eye movements could not only be used in interfaces for music composition but rhythm could more generally be used as a input variable in gaze controlled interfaces. The most obvious application being computer games as the game mechanics in many games are rhythm based.

A goal with the study was also to generate useful guidelines for what properties of rhythmic eye movements that needs to be considered when implementing an eye movement based instrument.

2. METHOD

18 subjects (11 male) were recruited from the student population of Lund, Sweden. All were volunteers and no payment was given. Their age ranged from 19 to 50 with a mean age of 26. Twelve reported to have had musical training and on average a subject had 7 years of musical training.

Subjects were seated in front of a tower-mounted, SMI Hi-Speed eye tracker with a temporal resolution of 500 Hz. A chin rest was used to constrain subjects' head movements and maintain an eye-to-screen distance of 67 cm. The screen used as stimuli display was a Samsung SyncMaster 245T, a 52×33 cm large LCD monitor with a resolution of 1920×1200 px.

The task given to the subjects was to let their gaze alternate between two fixation points in time with a beat (see fig. 1). The beat was given by a sequence of 50 msec square wave beeps of 440 Hz with a fixed inter-onset interval (IOI). The beeps were played to the subjects through a pair of full sized head phones (Philips SHP2500). A subject was given 16 session, where each session consisted of 20 sec. of beat following and 10 sec. of rest. The sessions were kept short in order to avoid fatigue of the subjects. Two factors were varied in the sessions, tempo and the span of the fixation points, and each session included one combination of factors. Tempo was either 60 or 120 bpm, corresponding to an IOI of 1.0 and 0.5 sec. The span of the fixation points was either 10 or 20 visual degrees. Each factor combination was used in four sessions but the order of the combinations was randomized for each subject. Both visual and auditory stimulus was presented using Matlab[1] with the Psychophysics Toolbox extensions [7] and eye tracking data was recorded using iView X 2.5[2].

After collecting the eye tracking data five different methods of eye tapping were implemented and applied on the gaze position data. This yielded the tap onsets that would have been present if the eye tapping methods had been used during the sessions. As the fixation points were placed on a horizontal line only the x-coordinates of the gaze position were used. The five methods to generate taps were:

1. When crossing the midline. This generates a tap every time the midline between the fixation points is crossed, but only after 100 msec have been spent on one side of the midline.

2. At the beginning of a saccade. This generates a tap when leaving a fixation point after having looked at it for more than 100 msec.

3. At the end of a saccade. This generates a tap when arriving at a fixation point after having crossed the midline.

4. At maximum saccade velocity. This generates a tap at the velocity peak when saccading between the two fixation points.

5. When fixating. This generates a tap every time a fixation is made after having crossed the midline. A gaze point is defined as a fixation if the point-to-point velocity is below $20°/s$ [12]. Note that this eye tapping method does not use the actual position of the fixation points.

Each beep onset was then paired with the closest tap onset generated by each of the five methods (see fig. 2). Beep onsets from the three first sec of each session were disregarded as no cue was given subjects when a session was about to start.

One variable of interest is beep-to-tap asynchronies, that is, the difference between the beep onsets and the closest taps of the methods. Here a negative value would indicate that the tap onset occurred before the beep onset and a positive value would indicate that the tap onset occured after the beep onset. This will show if subjects tend to eye tap before or after the beat. This is not a good measure of performance however. If the beep onsets are [1.0, 2.0, 3.0, 4.0] and the corresponding taps generated by one of the methods above are [0.8, 1.9, 3.0, 4.3] this gives the asynchronies [-0.2, -0.1, 0.0, 0.3] but a mean asynchrony of 0.0. A better performance measure is to take the absolute value of the asynchronies, [0.2, 0.1, 0.0, 0.3], which then gives a mean *absolute* asynchrony of 0.15. This is the main variable of interest and it shows how well the different methods, and the subjects, performed.

Processing of the eye tracking data was done using the Ruby programming language[3] and statistical analyses were conducted using the R environment [9]. The raw eye tracking data and the script used to generate eye taps are available for download at `http://www.sumsar.net/files/eye_tapping_experiment1.7z` .

3. RESULT

3.1 Performance of the Tapping Methods

Table 1 summarizes the result of the experiment. All tapping methods had a mean absolute asynchrony in the range of 130-190 msec. The mean direction of the asynchronies was negative for all methods, that is, all methods tended to

[1]http://www.mathworks.com/

[2]http://www.smivision.com/

[3]http://www.ruby-lang.org/

beeps	midline	beginning	end	velocity	fixation
3	2.74	2.72	2.77	2.72	2.96
4	3.93	3.91	3.97	3.93	3.99
5	4.77	4.75	4.80	4.77	4.83
6	5.93	5.92	5.96	5.92	6.00
7	6.83	6.81	6.86	6.83	6.89

Figure 2: An excerpt from session one of subject one with beep onsets (IOI = 1.0 sec) and the corresponding tap onsets generated by the five eye tapping methods.

	Mean	SD	Abs mean	Abs SD
(1) Midline	-71	185	163	112
(2) Saccade beg.	-58	203	173	121
(3) Saccade end	-22	261	185	185
(4) Max. velocity	-61	202	169	125
(5) Fixation	-49	175	137	119

Table 1: Mean and SD in msec. of the asynchrony and absolute asynchrony and of the five tapping methods for the 18 subjects.

yield taps that occurred before the beep onset. A repeated measures ANOVA was used to test for absolute asynchrony differences between the five tapping methods. Absolute asynchrony differed significantly across the five methods, $F(4, 17) = 5.76, p < 0.001$. Fisher's LSD test showed that method (5), to generate a tap when fixating, had a significantly lower mean absolute asynchrony than the other methods ($p < 0.05$). This is also visible in figure 3 .

3.2 Performance of the Subjects

As the fixation based tapping method was significantly better that the other four methods it was used in the subsequent analyses of the performance of the subjects. The mean asynchrony and the mean absolute asynchrony of the subjects are shown in figure 4. The mean absolute asynchrony over all subjects was 137 msec ($SD = 55$) and a one sample t-test showed that it differed significantly from zero ($p < 0.001$, 95% CI[110, 165]). All subjects, except one, had a negative mean asynchrony. A one sample t-test showed that mean asynchrony over all subjects differed significantly from zero ($M = -49.6$ msec, $p < 0.01$, 95% CI[-77, -22]). This is comparable to the negative mean

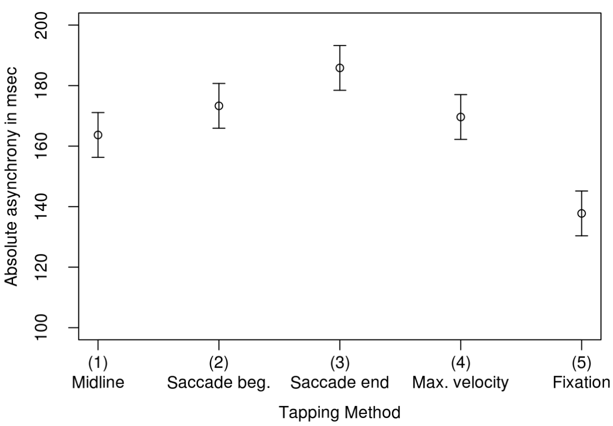

Figure 3: Mean absolute asynchrony, in msec., of the 18 subject as a function of tapping method. The error bars shows the standard error given by the ANOVA.

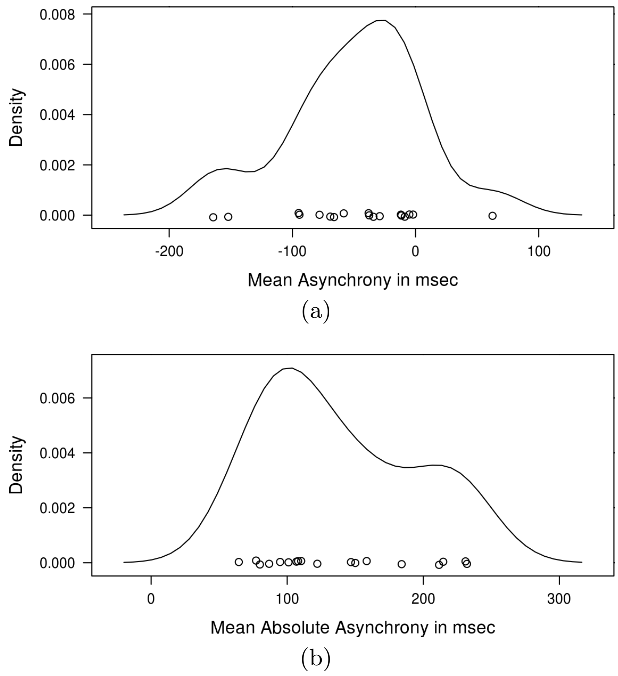

Figure 4: Mean asynchrony (a) and mean absolute asynchrony (b) in msec of the fixation tapping method for each subject. The points show the means of the individual subjects and the line shows the density created using a gaussian kernel.

asynchrony of 20–80 msec frequently found in finger tapping tasks when no auditory feedback is given [1]. Another finding in the finger tapping literature is that musicians perform better in finger tapping tasks than non-musicians [3, 11]. No significant correlation between reported number of years of musical training and mean absolute asynchrony was however found (Pearson's product-moment correlation, $r = -0.035$, $p = 0.89$). A negative correlation was found between mean asynchrony and mean absolute asynchrony ($r = -0.67$, $p < 0.01$), that is, bad performance is related to a tendency to eye tap to early. A positive correlation between the standard deviation of the mean asynchrony and the mean absolute asynchrony ($r = 0.89$, $p < 0.001$) indicate that low performing subjects are also less consistent in their timing.

4. DISCUSSION

The most accurate method of triggering rhythmical sounds from eye movements is by using fixation onsets (5) as the fixation based method had a significantly lower mean absolute asynchrony compared to the other methods. When building an eye controlled instrument accuracy might not be the only criteria when choosing an eye tapping method. Another criteria might be ease of implementation and if this is important the second best method, the midline based (1), might be a better choice.

All subjects managed the task of synchronizing their eye movements to a given beat and the conclusion is that it is possible to beat out a rhythm using eye movements. What is evident is that eye tapping differs from finger tapping and drumming in that subjects are more inaccurate [10]. The mean error in this study was 137 msec which is a quite noticeable error when beating out a rhythm. If implementing an eye controlled instrument one should be prepared for that the rhythmic accuracy of the users of your instrument might be quite low. Another difference from finger tapping studies is that no effect of musical training was found. It would be interesting to see if, and how much, training of eye tapping over an extended period can improve accuracy.

In this study there was a strong tendency of negative mean asynchrony. When finger tapping the negative mean asynchrony is known to decrease or even disappear when auditive feedback is given [11]. An interesting continuation of this study would then be to conduct an experiment where fixation based eye tapping is used to trigger taps and to compare the results with the finger tapping litterature. What also should be done is to use eye tapping to create new musical interfaces to allow performers and artists to use eye movements to create and perform music.

5. ACKNOWLEDGEMENT

This study was conducted at the Humanities Laboratory at Lund University[4]. The authors gratefully acknowledges support from the Linnaeus environment Thinking in Time: Cognition, Communication and Learning, financed by the Swedish Research Council, grant no. 349-2007-8695.

6. REFERENCES

[1] G. Aschersleben. Temporal Control of Movements in Sensorimotor Synchronization. *Brain and Cognition*, 48(1):66–79, 2002.

[2] W. T. Fitch. Biology of music: another one bites the dust. *Current biology : CB*, 19(10):R403–4, May 2009.

[3] M. Frank, J. Mates, T. Radil, K. Beck, and E. Pöppel. Finger tapping in musicians and nonmusicians. *International Journal of psychophysiology*, 11(3):277–279, 1991.

[4] A. Hornof and L. Sato. EyeMusic: making music with the eyes. *Proceedings of the 2004 Conference on New Interfaces for Musical Expression (NIME04)*, pages 3–5, 2004.

[5] A. Hornof and K. Vessey. The sound of one eye clapping: Tapping an accurate rhythm with eye movements. In *Proceedings of the 55nd Annual Meeting of the Human Factors and Ergonomics Society, to appear.*, 2011.

[6] J. Kim, G. Schiemer, and T. Narushima. Oculog: playing with eye movements. In *Proceedings of the 7th international conference on New interfaces for musical expression*, pages 50–55. ACM, 2007.

[7] M. Kleiner, D. Brainard, and D. Pelli. What's new in Psychtoolbox-3. In *Perception 36 ECVP Abstract Supplement.*, 2007.

[8] A. Polli. Active vision: Controlling sound with eye movements. *Leonardo*, 32(5):405–411, 1999.

[9] R Development Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2010.

[10] B. Repp. Rate limits in sensorimotor synchronization with auditory and visual sequences: The synchronization threshold and the benefits and costs of interval subdivision. *Journal of Motor Behavior*, 35(4):355–370, 2003.

[11] B. H. Repp. Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, 12(6):969, 2005.

[12] D. Salvucci and J. Goldberg. Identifying fixations and saccades in eye-tracking protocols. In *Proceedings of the 2000 symposium on Eye tracking research & applications*, pages 71–78. ACM, 2000.

[4] http://www.humlab.lu.se/