Effects of Stimulus Intensity and Frequency on the Force and Timing of Sensorimotor Synchronisation

Neil P. M. Todd1,*, Peter E. Keller2,**, Sendhil Govender3,4 and James G. Colebatch3,4,***

1Department of Psychology, University of Exeter, Exeter EX4 4QC, UK
2MARCS Institute for Brain, Behaviour and Development, Western Sydney University, Penrith, Sydney, NSW, 2751, Australia
3Prince of Wales Clinical School and
4Neuroscience Research Australia, UNSW, Sydney, NSW 2052, Australia

Received 4 August 2020; accepted 18 May 2021

Abstract
We report an experiment to investigate possible vestibular effects on finger tapping to an auditory anapaest rhythm. In a sample of 10 subjects, index finger acceleration and tapping force were recorded along with extensor/flexor activity and the associated electroencephalographic activity measured at central and cerebellar surface electrodes. In a prior session with a standard short air-conducted 500-Hz pip, vestibular evoked myogenic potential thresholds were measured and subsequently used to set the acoustic intensity. During the main experiment subjects were asked to synchronise tapping to the pips arranged in the anapaest at two different frequencies, 500 Hz vs 5 kHz, so that only the low-frequency high-intensity condition was a vestibular, as well as an auditory stimulus. We hypothesised that a vestibular effect would manifest in an interaction between the frequency and intensity factors for a range of dependent measures of tapping performance. No clear evidence was found for vestibular effects, but this was likely due to the confounding effects of an independent effect of intensity and the relative weakness of the acoustic vestibular stimulus. However, the data did show novel evidence for two distinct timing processes for the flexion and extension stages of a tap cycle and two distinct timing strategies, which we refer to as ‘staccato’ and ‘legato’, characterised by different profiles of force and extension.

Keywords
Rhythm, timing, kinematics, sensorimotor synchronisation, vestibular, cerebellum

*To whom correspondence should be addressed. E-mail: n.todd@exeter.ac.uk
** ORCID: 0000-0001-7579-6515
*** ORCID: 0000-0002-2099-4976

© Neil P. M. Todd, et al., 2021
DOI: 10.1163/22134468-bja10041
This is an open access article distributed under the terms of the cc by 4.0 license.
1. Introduction

Sensorimotor synchronisation in the form of finger tapping to a regular auditory metronome is a well-established paradigm to investigate mechanisms of human motor timing (Repp, 2005; Repp & Su, 2013). It is generally agreed that there are two distinct brain subsystems involved, involving respectively the cerebellum and basal ganglia and associated structures (Ivry & Keele, 1989), but these are described in different terms by different theorists. Lewis and Miall (2003) refer to ‘automatically’ and ‘cognitively’ controlled systems, while Teki et al. (2011) distinguish ‘duration-based’ and ‘beat-based’ auditory timing, Grahn and Rowe (2013) ‘beat detection’ and ‘beat prediction’ systems, and Todd and Lee (2015a, 2015b) two systems for ‘externally’ vs ‘internally’ guided action involving the cerebellum and basal ganglia. It has also been suggested that the cerebellar (variously ‘automatic’, ‘duration-based’, ‘beat detection’, or ‘externally focussed’) system is primarily activated for error correction once a clear beat has been learned (Repp & Su, 2013; Teki et al., 2011).

There has in the last decade or so been an increasing awareness of a contribution of the vestibular system to timing, with suggestions that it may play an important role in synchronisation, especially with regards to the beat (Phillips-Silver & Trainor, 2005, 2007, 2008; Trainor et al., 2009). It has been demonstrated that head movement and therefore activation of the vestibular apparatus is necessary to observe effects on beat perception (Phillips-Silver & Trainor, 2007), and further that vestibular influence on beat perception can be achieved by using galvanic vestibular stimulation (Trainor et al., 2009). However, the matter remains controversial, with some authors suggesting that the vestibular influence is not direct (Riggle, 2009; Trainor, 2007; Trainor & Unrau, 2009).

In our own recent work we have shown that short-latency vestibular cerebellar evoked potentials (VsCEPs) and an associated spontaneous electrocerebellogram (ECeG) can be recorded over the posterior fossa in response to stimuli that activate vestibular evoked myogenic potentials (cervical and ocular VEMPs), including 500 Hz air- and bone-conducted sound as well as impulsive head accelerations (Govender et al., 2020; Todd et al., 2017, 2018a, b, 2019). These responses show considerable plasticity and can be modulated by eye gaze, head position and optokinetic stimuli. These responses are likely therefore to be generated by cerebellar regions which are involved in both vestibular and ocular–motor control, including the nodulus/uvula and flocculus/parafloccular lobes (Kheradmand & Zee, 2011). VsCEPs thus provide an informative dependent measure, as they indicate conditions under which vestibular receptors have been activated sufficiently to produce recordable changes in central neural activity. Conversely, the absence of such potentials indicates that the stimuli may not be strongly vestibular.
A question arises then as to whether it may be possible to use such acoustic/inertial vestibular stimuli to influence sensorimotor synchronisation (Todd & Lee, 2015b). The aim of the experiment reported here, therefore, was to investigate possible vestibular effects on the kinematics, dynamics and electromyography of finger tapping to a simple anapaest (“Three blind mice”) rhythm using air-conducted sound. The use of the anapaest allows assessment of continuation tapping as the fourth beat of the rhythm does not contain a stimulus and also facilitates metre induction, and hence beat effects on timing. As air-conducted sound also, of course, activates the auditory system it is necessary to make use of control stimuli, and here we made use of both a low-intensity and high-frequency control using short tone pips of 500 Hz (within the range of vestibular sensitivity) and 5 kHz (outside the range of sensitivity) at two intensities below and above the cervical VEMP threshold. We hypothesised that a vestibular effect would manifest in an interaction between the frequency and intensity factors for any dependent measures of tapping performance and in the appearance of VsCEPs in the vestibular (low-frequency, high-intensity) condition.

Previous research has yielded mixed results with regard to the effects of sound intensity on movement timing. One study found that, while simple reaction times decreased as the intensity of auditory stimuli increased (see also Jaśkowski, 1996), tap timing was not affected by the same intensity manipulation in a sensorimotor synchronisation task with auditory pacing sequences (Białuńska et al., 2011). However, studies on the effects of sound accentuation have found that movement timing is more stable for sequences containing regular variations in acoustic intensity (e.g., every second, third, or fourth sound is louder than intervening sounds) than with irregular or no accents (Bouvet et al., 2020; Etani et al., 2019; Keller & Repp, 2005). Sound intensity in these previous studies was well below the cervical VEMP threshold, leaving open the question of how air-conducted vestibular stimuli affect tap timing.

A further open question concerns whether effects of frequency and intensity vary across the time course of individual tap movement cycles. Previous work on relations between movement kinematics and timing has segmented tap cycles into stages, including downward flexion and upward extension (Balasubramaniam et al., 2004), as well as dwell at the target surface (Hove et al., 2014a). Flexion is typically relatively high in velocity, short in duration, and similar across tempi, while extension is lower in velocity, longer in duration, and scales with tempo (Hove et al., 2014a; Pabst & Balasubramaniam, 2018), suggesting that these two tap stages are independent to some degree. The duration of each stage also varies with the style of tapping – for example, flexion times are shorter for sharp ‘staccato’ than smooth ‘legato’ taps (Hove et al., 2014a) – which can be influenced by musical experience (Krause et al., 2010). We, therefore, divided tap cycles into stages and took musical training into account in our analyses.
2. Methods

2.1. Subjects

Ten young adults (six female, four male) were recruited from staff and students at the Prince of Wales Hospital and from Western Sydney University. The sample size was motivated by prior studies of both VsCEPs (e.g., Todd et al., 2017, 2018a, b, 2019) and sensorimotor synchronisation (e.g., Hove et al., 2014a; Keller & Repp, 2005). If the subjects had had musical training and played regularly either as an amateur or professional they were classified as musicians. None of the subjects had a history of vestibular or hearing impairment. The experiment was conducted in two parts and prior to testing all subjects gave written consent according to the Declaration of Helsinki.

2.2. Acoustic Stimuli

Auditory stimuli were generated using custom software and a CED laboratory interface (1401plus, Cambridge Electronic Design, Cambridge, UK), and signal amplification was achieved using a custom amplifier. The sounds consisted of sinusoidal 2-ms tone bursts (0-ms rise and fall) at two frequencies, 500 Hz and 5 kHz, and were delivered using audiometric headphones (TDH 49, Telephonics Corp., Farmingdale, NY, USA). The outputs were calibrated using a type 4192 pressure field microphone with a 4153 artificial ear and a 2260 sound level meter (Bruel & Kjær, Naerum, Denmark). Sound intensity was controlled by the applied drive voltage in dB re 5 V peak and for the same voltage the two stimuli calibrated to within 2 dBA. In the main experiment tone pips were arranged in an anapaest rhythm consisting of three pips with an interval of 600 ms and a gap of 1200 ms.

2.3. Cervical VEMP Recording

Surface electromyography (EMG) was recorded from the sternocleidomastoid ipsilateral to the stimulated ear using self-adhesive Ag/AgCl electrodes. Active surface electrodes were placed over the middle of the sternocleidomastoid belly and were referred to electrodes placed on the medial clavicular head. EMG was amplified (by a factor of 2,000), bandpass filtered (5 Hz–1 kHz) and sampled at 5 kHz using a Power1401 interface (Cambridge Electronic Design). The EMG recording began 10 ms before and finished 90 ms following stimulus onset.

2.4. EMG/EEG/Finger Movement Recording

During the main experiment scalp electroencephalography (EEG)/EMG was recorded in eight channels with Ag/AgCl electrodes secured with electrode paste at bilateral infra-ocular positions (at IO1/IO2 with an ocular VEMP montage), over the motor cortex (C3/C4–A1/A2), over the posterior fossa (CB1/CB2–A1/A2) and on the neck over the splenius (SP1/PS2–A1/A2) referred to ipsilateral ear lobe (A1/A2) using a biological amplifier (D360, Digitimer Ltd., Welwyn Garden City, Hertfordshire, UK) set to a gain of 20,000. In parallel, EMG was recorded
with a bipolar montage over the extensor indicis and flexor digitorum muscles of the index finger with a separate D360 amplifier set to a gain of 20,000. Also recorded was finger acceleration by means of an accelerometer (Model 3026-010-S, IC Sensors, Milpitas, CA, USA) secured to the top of the index finger and finger tap force by means of a 2 kg load cell (Model 632–736, RS Components Pty Ltd., NSW, Australia). All 12 channels were recorded at a sampling rate of 10 kHz using a Power1401 interface (Cambridge Electronic Design) with the finger EMG rectified. The recording epoch was 2,200 ms with 200 ms recorded prior to the first stimulus of the anapaest sequence. After recording, EEG was averaged over the whole duration of the anapaest rhythm, including for the last long interval, which acoustically contains no stimulus but does perceptually contain a weak beat.

2.5. Experimental Procedure

The experiment was conducted in two sessions on two separate occasions. The first was to measure VEMP thresholds as a reference level for the auditory stimuli, the second to carry out the main procedure.

2.5.1. Threshold Determination

In the first session subject cervical VEMP thresholds were determined for the auditory stimuli. Subjects were tested lying supine on a couch, with the backrest tilted to 30–45° from the horizontal, and required to lift their heads against gravity to activate the sternocleidomastoid muscles. Up to 200 stimuli were presented at a rate of 5 Hz. The presence or absence of a cervical VEMP was determined by visual inspection. Threshold was determined to within 3 dB and this was used to set the intensity of the stimuli, at −3 dB vs +18 dB re cervical VEMP thresholds, for the main experimental session.

2.5.2. Main Experiment

In the main experiment index finger acceleration and tapping force apparatus were set up along with the electrode arrangement for extensor/flexor activity. The EEG/ECeG was recorded using surface electrodes to allow measurement of both evoked and spontaneous activity (recording details above). After this preparation, subjects were positioned comfortably in a supine position and were given instructions. They were told that they would be presented with a single rhythmic figure (a bit like “Three blind mice”) consisting of three pips and a gap in a number of blocks lasting a few minutes each. In half of the blocks they would be required just to listen passively but to be as relaxed as possible and avoid unnecessary movements with gaze fixed while maintaining attention on the rhythm. These passive blocks would alternate with blocks where subjects would be required to tap a regular beat to the rhythm, including in the gap, on the force plate which was positioned and secured to be comfortably under the index finger of their dominant hand. They were shown an example of tapping by one of the experimenters in which the style of tapping required them to have a period of tonic extension between taps. Subjects were also informed that although the rhythm would remain constant,
the pips would vary in both frequency (low vs high), intensity (low vs high) and in the ear they were presented to (left vs right). They were also informed that at the start of each block they would be given a verbal warning when it would start, that between blocks they could relax and that they would be allowed two practice blocks before starting the main experiment. In order to circumvent order effects the 16 experimental blocks (each containing 50 trials) were presented pseudo-randomly for each subject. Data from passive blocks will not be presented here but will be reported in a follow-on paper that compares neural data for active tapping and passive listening (as in Todd & Lee, 2015b).

2.6. Data Analysis

2.6.1. Measurement of EMG/Movement Data
Measurements of the EMG/movement data were made for each beat, for each condition, using Gaussian or logistic curve-fitting methods on individual subject blocks after averaging across trials. The curve fitting allowed the estimate of within-block tap amplitude, time location and spread.

2.6.2. Processing of EEG/ECeG Data
After recording grand means were made of EEG/ECeG data across subjects for each of the conditions. A subsequent sub-average was made for two timing groups which emerged in the analysis of the movement data. No further analyses were made of EEG/ECeG data in the present paper.

2.6.3. Statistical Analysis
The data thus obtained were concatenated into a single matrix using the SPSS software (Statistical Product and Service Solutions, IBM, Australia), which could be then used for a principal components analysis (PCA), and cluster/regression analysis. Subsets of this data matrix were then further transformed into a within-subjects format, to allow for repeated-measures analyses of variance (ANOVA). Missing values were replaced by averages of the remaining subjects. Data can be made available on reasonable request to the corresponding author.

3. Results

3.1. Cervical VEMP Thresholds
Mean cervical VEMP thresholds were $-24 \pm 5$ dB and $-25 \pm 7$ for the right and left ears respectively and did not differ between the sides. While thresholds varied between individuals (range: $-15$ dB to $-33$ dB), they were within the reported normal limits (Rosengren et al., 2011).

3.2. Grand Means
Figure 1A shows the grand mean for a single supra-threshold condition of selected EEG/ECeG channels, infra-ocular EMG, along with finger extensor, flexor,
Figure 1. (A) The grand mean for a supra-threshold condition of infra-ocular EMG, EGeG at CB2, EEG at C3, extensor and flexor EMG, acceleration, contact force and reconstructed kinematics over a single rhythmic bar of four beats. Finger velocity and displacement were reconstructed by integration and double integration of the acceleration. The time base for the grey-shaded area is further expanded and used in part (B) to illustrate the (1) extensor relaxation, (2) extensor tonic relaxation, (3) extensor re-contraction and (4) extensor tonic contraction tap stages for the third beat.
acceleration and contact force over a single rhythmic bar of four beats, including on beat 4 which has a missing stimulus. Finger velocity and displacement was reconstructed by integration and double integration of the acceleration. Given the regularity of the movement, for the purpose of analysis each tap cycle was divided into four stages defined by extensor activity (Fig. 1B). These four stages were: (1) extensor relaxation, during which the flexion movement is initiated; (2) extensor tonic relaxation, during which the flexor contracts and reaches a peak, in parallel with the downwards acceleration reaching its peak; (3) extensor re-contraction, during which the contact force reaches its peak, the flexion movement ends and the extension movement starts; and (4) extensor tonic contraction, during which the extension movement reaches its maximum.

Within each tap cycle, in addition to regular movement patterns, it is also possible to see in the averaged EEG well-defined potentials at electrode C3 (approximately located over the left-hand sensorimotor cortex S1/M1). These potentials are likely to be both stimulus-related, auditory-evoked potentials (AEPs) and also movement-related potentials (MRPs) preceding the stimulus. Possible AEPs are shown as N1, P2 and N2 waves. The probable MRPs include a pre-movement negativity (PMN), which here is assumed to overlap with AEP N2, and what appears to be motor potentials (MPs) preceding the flexor and extensor activity. The traces for the averaged contralateral infra-ocular (at IO2) and ECG (at CB2) recordings did not show any clear evidence of the presence of ocular VEMPs and VsCEPs, even though the stimulus intensity was well above the cervical VEMP threshold.

3.3. Principal Components Analysis of the Tap Stage Parameters

The boundaries of the tap stages were not always easy to determine. However, within each stage reliable measurements could be made of (1) the extensor relaxation with a sigmoid curve fit within stage 1, (2) the down acceleration with a Gaussian curve fit within stage 2, (3) the contact force with a Gaussian curve fit within stage 3 and (4) the extensor re-contraction with a Gaussian or sigmoid curve fit at the end of stage 3. From each of these, an estimate of amplitude, phase and spread were measured for each subject. These estimates were made within an anapaest frame, and across the frequency and intensity conditions, as well as for the four individual beats within the anapaest. The phase was defined relative to the stimulus onset, i.e., 0 (or 2π) radians was aligned with the start of each inter-onset interval.

Given the large volume of data that resulted, a PCA was first conducted on all of the measures combined (i.e., all four stages and three parameters, for each of the within-subject conditions for all 10 subjects, forming a 12 by 320 between-subjects coded matrix). Four resultant principal components (PCs) could explain about 80% of the variance. Table 1 shows the PCA rotated matrix components, indicating in bold how each of the 12 stage parameters are correlated to the four
Tables: (1) the amplitude/phase/spread of the contact force, the phase/spread of extensor contraction and spread of the down acceleration, (2) the amplitudes of the extensor relaxation, the flexion/acceleration and the extensor contraction, (3) the phases of the extensor relaxation and flexion/acceleration and (4) the spread of the extensor relaxation. PC 1 and 3 thus represent the independent timing of respectively finger flexion (extensor relaxation/flexor contraction) and extension (flexor relaxation/extensor contraction). The corresponding phase parameters for the flexion and extension are further indicated in Table 1 with the phase parameters underlined. The first three PCs are illustrated in Fig. 2A which shows how the loadings in bold in Table 1 correspond to three well-defined clusters.

3.4. Regression and Cluster Analyses

Given the centrality of phase to the two timing PCs, a regression analysis was further conducted on the phase parameters for the flexion/acceleration vs extensor relaxation and plate force vs extensor contraction (Fig. 2B and 2C). Both linear regressions were highly significant (Fig. 2B, $R = 0.71, F_{1,286} = 65, p < 0.001$ and Fig. 2C, $R = 0.91, F_{1,286} = 1,381, p < 0.001$). For extension timing there appeared to be two distinct subject groups as the force phase, in particular, was clearly bimodal in

Table 1. Rotated component matrix using Varimax and Kaiser Normalization.

<table>
<thead>
<tr>
<th>Displacement</th>
<th>Stage</th>
<th>Parameter</th>
<th>PC1</th>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flexion</td>
<td>Relaxation</td>
<td>Amplitude</td>
<td>.033</td>
<td>.944</td>
<td>-.108</td>
<td>-.040</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phase</td>
<td>.048</td>
<td>.033</td>
<td>.904</td>
<td>-.092</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sigma</td>
<td>.222</td>
<td>-.168</td>
<td>-.182</td>
<td>.804</td>
</tr>
<tr>
<td></td>
<td>Acceleration</td>
<td>Amplitude</td>
<td>-.579</td>
<td>-.570</td>
<td>.308</td>
<td>-.039</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phase</td>
<td>.011</td>
<td>-.092</td>
<td>.891</td>
<td>.029</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sigma</td>
<td>.453</td>
<td>-.126</td>
<td>-.230</td>
<td>-.461</td>
</tr>
<tr>
<td>Extension</td>
<td>Force</td>
<td>Amplitude</td>
<td>.628</td>
<td>.314</td>
<td>.170</td>
<td>.340</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phase</td>
<td>.888</td>
<td>-.202</td>
<td>.271</td>
<td>.152</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sigma</td>
<td>.932</td>
<td>-.179</td>
<td>-.188</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>Contraction</td>
<td>Amplitude</td>
<td>-.241</td>
<td>.925</td>
<td>.041</td>
<td>.047</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Phase</td>
<td>.921</td>
<td>-.277</td>
<td>.068</td>
<td>.084</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Sigma</td>
<td>-.748</td>
<td>.108</td>
<td>-.378</td>
<td>-.165</td>
</tr>
</tbody>
</table>

PC = Principal component. Symbols in parentheses correspond to those in Fig 2A. Numbers in bold indicate those indicate the most significant rotated factor loadings, with the loadings for the phase parameters further indicated by underlining. Thus, PC1 is predominantly related to extension timing (force and contraction phase), while PC3 is predominantly related to flexion timing (relaxation and acceleration phase).
Figure 2. (A) The loadings plot for the first three components from a principal components analysis (PCA) of the measurements of amplitude, phase and sigma for each of the four tap stages. Three well-defined clusters can be observed; PC1, PC2 and PC3 – refer to Table 1 for corresponding values. (B) Scatterplots of acceleration vs extensor relaxation phase. (C) Scatterplot of force vs extensor contraction phase. (D) Box plots of force phase clusters corresponding to the two groups – ‘staccato’ and ‘legato’ timers. (E) Histogram demonstrating a clear bimodal distribution for the force phase and cluster analysis showing the two partitioned distributions for the staccato (F) and legato (G) groups which closely align to the original distribution.
distribution (Fig. 2E). In order to confirm this statistically, we first tested for non-normality (Shapiro–Wilk statistic = 0.97, df = 320, p < 0.001) and then employed a hierarchical cluster analysis to separate out the two distributions so that the means of the two clusters aligned closely with the two modes of the original distribution (Figs 2D, F and G). A cross-tabulation of subjects by average linkage of clusters confirmed the natural separation (Table 2A, $\chi^2_{9} = 210, p < 0.001$) and subjects were thus assigned to the groups based on the number of cases which fell into the two clusters. One group ($n = 5$) was tightly clustered around the force phase of approximately zero radians (Fig. 2D), which we refer to hereafter as the ‘staccato’ timers. The other ($n = 5$) was clustered around an extension phase of about $\pi/2$ radians (Fig. 2C), which we refer to hereafter as the ‘legato’ timers. The

### Table 2.

(A) Cross-tabulation of subject by average linkage (between groups), $\chi^2_{9} = 210, p < 0.001$.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Cluster 1 count</th>
<th>Cluster 2 count</th>
<th>Total</th>
<th>Group</th>
<th>Musician</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>6</td>
<td>26</td>
<td>32</td>
<td>LEG (2)</td>
<td>NON</td>
</tr>
<tr>
<td>S2</td>
<td>32</td>
<td>0</td>
<td>32</td>
<td>STA (1)</td>
<td>MUS</td>
</tr>
<tr>
<td>S3</td>
<td>32</td>
<td>0</td>
<td>32</td>
<td>STA (1)</td>
<td>NON</td>
</tr>
<tr>
<td>S4</td>
<td>32</td>
<td>0</td>
<td>32</td>
<td>STA (1)</td>
<td>MUS</td>
</tr>
<tr>
<td>S5</td>
<td>32</td>
<td>0</td>
<td>32</td>
<td>STA (1)</td>
<td>MUS</td>
</tr>
<tr>
<td>S6</td>
<td>2</td>
<td>30</td>
<td>32</td>
<td>LEG (2)</td>
<td>NON</td>
</tr>
<tr>
<td>S7</td>
<td>8</td>
<td>24</td>
<td>32</td>
<td>LEG (2)</td>
<td>MUS</td>
</tr>
<tr>
<td>S8</td>
<td>30</td>
<td>2</td>
<td>32</td>
<td>STA (1)</td>
<td>MUS</td>
</tr>
<tr>
<td>S9</td>
<td>5</td>
<td>27</td>
<td>32</td>
<td>LEG (2)</td>
<td>NON</td>
</tr>
<tr>
<td>S10</td>
<td>12</td>
<td>20</td>
<td>32</td>
<td>LEG (2)</td>
<td>MUS</td>
</tr>
<tr>
<td>Total</td>
<td>191</td>
<td>129</td>
<td>320</td>
<td>5 L, 5 S</td>
<td>4 N, 6 M</td>
</tr>
</tbody>
</table>

Subjects were assigned to groups (clusters) based on the number of cases which fell into each cluster. Thus, if the majority of cases fell into cluster 1, they were assigned to the "staccato" group, and vice versa.

(B) Cross-tabulation of subject by timing group, $\chi^2_{1} = 53.3, p < 0.001$.

<table>
<thead>
<tr>
<th>Musician</th>
<th>Group</th>
<th>LEG</th>
<th>STA</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>MUS</td>
<td>64 (2)</td>
<td>128 (4)</td>
<td>192 (6)</td>
<td></td>
</tr>
<tr>
<td>NON</td>
<td>96 (3)</td>
<td>32 (1)</td>
<td>128 (4)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>160 (5)</td>
<td>160 (5)</td>
<td>320 (10)</td>
<td></td>
</tr>
</tbody>
</table>

LEG, legato; MUS, musician; NON, non-musician; STA, staccato.
groups are closely, but not exactly, correlated with musicianship – of the ‘staccato’
group four were musicians, while three of the ‘legato’ group were non-musicians.
Although the correlation was not complete, a $\chi^2$ test showed that it was still statistically highly significant ($\chi^2 = 53.3, p < 0.001$, Table 2B).

3.5. Dynamics and Kinematics of the Tapping Style of the Two Timing Groups

Given the existence of the two timing groups it is relevant to consider the differences in the dynamics and kinematics of their tapping style. The acceleration, velocity, displacement, force and extension profiles of the two groups are illustrated as superimposed grand means (Fig. 3A) and separately (Fig. 3B). These show that essentially the ‘staccato’ group have a much more asymmetric flexion/extension ratio than the ‘legato’ tappers. This difference is associated with a higher flexion acceleration and deceleration in the ‘staccato’ group (see also Hove et al., 2014a; Repp & Su, 2013). Subsequent analyses thus included timer group as a between-subjects factor, as defined by their distinct extension timing strategy.

The group differences in the averaged movement data were also reflected in the averaged EEG data. Corresponding with the sharper movement timing, the

Figure 3. (A) The EEG, EMG and kinematic grand means over a single rhythmic bar of four beats for the ‘legato’ vs ‘staccato’ timer groups plotted together. The time base of the grey-shaded area (before and after the second beat) is shown expanded in part (B) and illustrates the grand means for the staccato (top) and legato timers (bottom) divided up into the four tap stages.
‘staccato’ group show larger and more well-defined MRPs, with what appear to be distinct flexor and extensor MPs (Fig. 3A, B). In both cases the ‘motor delay’ between peak MPs and peak EMG at about 40–50 ms, close to the established values (Colebatch, 2007). As in the grand average, neither group average shows clear ocular VEMPs nor VsCEPs. Cervical VEMP thresholds did not differ between ‘staccato’ and ‘legato’ groups ($t_{8} = 0.9, p = 0.400$).

3.6. ANOVAs

In order to make the analyses more efficient, ‘tap stage’ (1–4) was treated as an additional factor, rather than conducting four separate ANOVAs for each stage. For each of the three measurement parameters (amplitude, phase and spread) repeated-measures ANOVAs were then carried out for tap force amplitude with within-subjects factors of ‘ear’ (left vs right), ‘frequency’ (500 Hz vs 5 kHz), ‘intensity’ (high vs low), ‘beat’ (1, 2, 3 & 4) and ‘stage’ (1, 2, 3 & 4). ‘Timing group’ (‘staccato’ vs ‘legato’) was then used as a between-subjects factor.

3.6.1. Tap Stage Amplitudes

Given the lack of homogeneity in the amplitude parameters for EMG, force and acceleration ($\mu$V, N and g), the amplitudes for the different measures across stages were first converted to Z-scores based upon their means and variance for the entire data set. The outcomes of the ANOVA for amplitudes using Z-scores are illustrated in Figs 4 and 5, with numerical values given in Table 3. Main effects of ‘frequency’ (Fig. 4A), ‘intensity’ (Fig. 4B) and ‘beat’ (Fig. 4D) were obtained, indicating that tap amplitudes were greater for low-frequency stimuli compared to high-frequency, for high intensity compared to low intensity, and for beat three compared to the preceding and following beats. The interaction of ‘frequency’ by ‘intensity’ (Fig. 4C) indicated that the vestibular condition (low-frequency/high-intensity) exhibited the highest amplitude, as hypothesised, but this was not statistically significant (at least in a multiplicative sense, see Discussion). The ‘intensity’ factor also interacted with ‘beat’ (Fig. 4E), consistent with the third beat being especially emphasised for high-intensity stimuli.

In addition to the above, the ‘stage’ factor also interacted with ‘frequency’ (Fig. 5A), ‘intensity’ (Fig. 5B) and ‘beat’ (Fig. 5C). The figures show that the main factors differing between the conditions are acceleration and force. There was no ‘timing group’ effect, but ‘group’ did also interact with ‘stage’ (Fig. 5D), consistent with the ‘staccato’ timers emphasising the extension acceleration, while for the ‘legato’ timers the emphasis appeared to be on the contact force.

3.6.2. Tap Stage Phases

As above for amplitudes, a repeated-measures ANOVA was conducted on tap stage phase with the same five within-subjects factors and a single between-subjects factor. As all phases were measured using the same units, the ‘stage’ factor could be applied directly to phase rather than Z-scores. There were no main effects of
There were, however, highly significant within-subjects effects of ‘beat’ (Fig. 6A, Table 4) and ‘stage’ (Fig. 6B), along with a highly significant between-subjects ‘group’ effect (Fig. 6C), and an interaction of ‘stage by group’ (Fig. 6D). The ‘stage’ and ‘group’ effects and their interaction follow from the definitions of these two factors. The ‘beat’ effect indicates that the phase across stages is more advanced for the middle two beats, compared to beats 1 and 4.

‘frequency’ nor ‘intensity’, nor was there any interaction between these factors. Although not significant, amplitudes were greatest for the vestibular condition (high intensity 500 Hz stimuli) as hypothesised. Amplitudes were particularly large on the third beat during high intensity stimulation (E)*, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$.

Figure 4. ANOVA effects for Z-scores of overall amplitudes for force, acceleration and EMG. Amplitudes were greater for 500 Hz stimuli (A), using the higher intensities (B) and for the third beat (D). Although not significant, amplitudes were greatest for the vestibular condition (high intensity 500 Hz stimuli) as hypothesised. Amplitudes were particularly large on the third beat during high intensity stimulation (E).* $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$. 

Downloaded from Brill.com07/14/2022 02:39:31AM via free access
3.6.3. Tap Stage Variability

The repeated-measures ANOVA applied to tap stage timing variability, as with phase, yielded no main effects of ‘frequency’, ‘intensity’ nor ‘beat’, nor was there an interaction between these factors. Of the within-subjects factors, only ‘stage’ yielded a significant main effect (Fig. 7A, Table 5), indicating that the timing variability increased as the stages progressed. There was also a highly significant between-subjects ‘group’ effect (Fig. 7B), indicating that the ‘staccato’ timers’ tapping performance was overall less variable than that of the ‘legato’ timers. The two-way interaction of ‘stage’ by ‘timer group’ (Fig. 7C) showed the superior (lower) timing variability was the case for the force and extensor contraction measures.

4. Discussion

In a sample of 10 subjects the flexor/extensor activity and associated acceleration and contact force were recorded during synchronised finger tapping to an anapaest auditory rhythm along with S1/M1 EEG, EGeG and extra-ocular EMG. The tapping performance could be separated into four stages and, for each of these, three parameters of amplitude, phase and variability were measured for force, acceleration and EMG. A PCA showed two independent timing processes of extension (stages 1 and 2, PC3) and flexion (stages 3 and 4, PC1). Examination of the flexion stages further revealed that the subjects naturally fell into two distinct
groups which we have labelled as 'staccato' and 'legato', in line with the prior literature (Repp & Su, 2013). The two groups also showed distinct MRPs, consistent with differences in movement timing accuracy.

Regarding our original hypothesis of possible vestibular effects, the data showed main effects of frequency and intensity for amplitude, such that the amplitude of movement was greater for the low-frequency and high-intensity stimuli. Although there was no significant interaction to support a specific vestibular effect, the combination of the two main effects did produce the highest overall amplitude in the vestibular condition, while amplitude was lowest in the high-frequency/low-intensity condition. Observing an additive effect of frequency and intensity in the predicted direction suggests that the vestibular hypothesis should not yet be abandoned. It might be the case that low frequency and high intensity are both necessary to elicit a vestibular response, even if the effects of frequency and intensity are independent of one another.

One possible reason why we did not observe a multiplicative interaction may have been that it was confounded by the fact that there was an intensity effect in the amplitude measures. Intensity effects have been observed in reaction time.
and other experiments (e.g., Miller et al., 1999) and in retrospect it was a design limitation of the experiment because the relative vestibular input was not strong. AC activation of the saccule gives good cervical VEMP responses but less effective for the ocular VEMP, due to the higher threshold (Rosengren et al., 2011). Although we did not conduct a detailed analysis of the neural data here, not all subjects showed a clear ocular VEMP and nor were there clear VsCEPs, consistent with the averages (Figs 1 and 3). The latter are correlated with ocular VEMP

Table 4.
ANOVA effects for phase.

<table>
<thead>
<tr>
<th>Factor</th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>ε</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within-subjects effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EAR (E)</td>
<td>1,8</td>
<td>0.7</td>
<td>ns</td>
<td>1.0</td>
</tr>
<tr>
<td>FREQUENCY (F)</td>
<td>1,8</td>
<td>4.3</td>
<td>ns</td>
<td>1.0</td>
</tr>
<tr>
<td>INTENSITY (I)</td>
<td>1,8</td>
<td>0.0</td>
<td>ns</td>
<td>1.0</td>
</tr>
<tr>
<td>BEAT (B)</td>
<td>3,24</td>
<td>25.8</td>
<td>&lt; 0.001</td>
<td>0.3</td>
</tr>
<tr>
<td>STAGE (S)</td>
<td>3,24</td>
<td>518</td>
<td>&lt; 0.001</td>
<td>0.3</td>
</tr>
<tr>
<td>BEAT*I</td>
<td>1,8</td>
<td>0.7</td>
<td>ns</td>
<td>1.0</td>
</tr>
<tr>
<td>Between-subjects effects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GROUP (G)</td>
<td>1,8</td>
<td>45.7</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>S*G</td>
<td>3,24</td>
<td>8.8</td>
<td>= 0.001</td>
<td></td>
</tr>
<tr>
<td>BEAT<em>S</em>G</td>
<td>3,24</td>
<td>4.7</td>
<td>&lt; 0.05</td>
<td></td>
</tr>
</tbody>
</table>

Figure 6. ANOVA tap stage phase effects. Significant main effects of beat (A) and stage (B) and timer group (C) and an interaction of timer group with stage (D). *** P ≤ 0.001.
amplitudes, indicating that they are likely to share the same afferent input (Todd et al., 2017, 2018a). Therefore, it is possible that activation of the cerebellum via saccular acoustic-evoked projections was weak. The saccular acoustic-activated projections to the cerebellum are also asymmetric, with a strong bias to the right ear/left hemisphere on average (Govender et al., 2020).

Figure 7. ANOVA tap stage variability effects. Variability in timing increased across the tap stages (A). Overall, staccato timers were less variable than legato timers (B) and this was mainly in the force and extension re-contraction stages (C). ***, P ≤ 0.001.
Both the confound of the intensity effect and the relative weakness of the acoustic-activated projections to the cerebellum could explain the lack of clear vestibular effects in the present data. We did, nevertheless, observe a frequency effect, consistent with previous research showing enhanced behavioural and neural responses to rhythmic low-frequency sounds (Hove et al., 2007, 2014b; Lenc et al., 2018; Varlet et al., 2020). In our study this cannot be explained as a proxy intensity effect due to the outer/middle ear transfer function as a careful calibration of the low-frequency (500 Hz) and high-frequency (5 kHz) stimuli showed that there was at most a 2 dBA difference, and in favour of the 5 kHz stimulus. A further experiment using a stronger vestibular input without a change of acoustic intensity, to avoid this confound, will be needed to resolve this matter.

Finally, on the matter of possible vestibular effects, it should be noted that the repetitive and unchanging nature of the timing of the stimulus could itself have contributed to the absence of a strong vestibular cerebellar effect. As noted in the Introduction it is generally accepted that once a sensory-motor coupling has been learned and the perceptual system is not being challenged then the role of the cerebellum is diminished and it is likely that SMA and basal ganglia take the main role in timing (Teki et al., 2011; Todd & Lee, 2015b). In the present case vestibular cerebellar responses may have been adapted or habituated after the first few presentations. An alternative way of generating cerebellar involvement would be to introduce perturbations of phase so as to force error correction. The hypothesis would be that a vestibular timing stimulus should produce better error correction (Repp & Su, 2013; Teki et al., 2011). Although we did not observe the hypothesised vestibular effects, we did observe independent flexion/extension timing and two distinct timing groups. It is of interest to compare these observations with those of prior studies. In the present case we used an EMG measure to define the tap cycle stage stages whereas most prior studies used a kinematic one based on finger vertical position (e.g., Balasubramaniam et al., 2004; Hove et al., 2014a; Krause

### Table 5.
ANOVA effects for variability.

<table>
<thead>
<tr>
<th>Factor</th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>ε</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within-subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EAR (E)</td>
<td>1,8</td>
<td>0.2</td>
<td>ns</td>
<td>1.0</td>
</tr>
<tr>
<td>FREQUENCY (F)</td>
<td>1,8</td>
<td>0.2</td>
<td>ns</td>
<td>1.0</td>
</tr>
<tr>
<td>INTENSITY (I)</td>
<td>1,8</td>
<td>3.1</td>
<td>ns</td>
<td>1.0</td>
</tr>
<tr>
<td>BEAT (B)</td>
<td>3,24</td>
<td>2.2</td>
<td>ns</td>
<td>0.3</td>
</tr>
<tr>
<td>STAGE (S)</td>
<td>3,24</td>
<td>59</td>
<td>&lt; 0.001</td>
<td>0.3</td>
</tr>
<tr>
<td>F*I</td>
<td>1,8</td>
<td>0.7</td>
<td>ns</td>
<td>1.0</td>
</tr>
<tr>
<td>Between-subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GROUP (G)</td>
<td>1,8</td>
<td>17.1</td>
<td>&lt; 0.005</td>
<td></td>
</tr>
<tr>
<td>S*G</td>
<td>3,24</td>
<td>18.5</td>
<td>&lt; 0.001</td>
<td></td>
</tr>
<tr>
<td>B<em>S</em>G</td>
<td>9,72</td>
<td>3.5</td>
<td>&lt; 0.05</td>
<td></td>
</tr>
</tbody>
</table>
et al., 2010). However, reconstruction of finger position by double integration of acceleration allows a mapping between our dynamical measures (EMG/force/acceleration) and the conventional kinematic representation. This indicates that the transition between flexion and extension movement occurs close to the tap peak force, while the transition from extension to flexion occurs during extensor relaxation prior to flexor muscle contraction and close to our measure of the 50% extensor relaxation phase. When carried over to the two groups there is a clear difference in the degree of asymmetry between the extension and flexion stages, consistent with the prior kinematic studies.

It is possible that there are some inconsistencies in phase, as our displacement measure was estimated indirectly from acceleration rather than directly. In our measures the phase of the displacement minimum (and peak tap force) of the whole group was very close to zero (slightly negative for the staccato tappers and positive for the legato tappers). A prominent feature of the kinematic approaches and prior simpler measures based on tap contact is a tendency of taps to precede target tones, referred to as the ‘negative mean asynchrony’ (or NMA) (Aschersleben, 2002; Repp, 2005; Repp & Su, 2013). The properties of the NMA are more closely aligned with the phase of our peak down acceleration, which corresponds to the contact with the force plate, and is consistently negative for both groups and closer in magnitude to the normal range of NMA values. As the finger briefly continues downwards movement after contact, due to the natural compliance of the finger pad and elasticity of the force plate, the phase of the peak contact downwards force and assumed minimum of displacement occur after contact. Whether this may provide an explanation for the NMA as a consequence of measurement technique will require further experimentation.

For the ‘staccato’ group their timing (phase) was focussed on a force peak of approximately zero radians, whereas the ‘legato’ group had a peak re-extension phase of about \( \pi/2 \) radians behind the beat. Analyses of the each of the three parameters for effects of frequency, intensity and beat confirmed the group distinctions with the ‘staccato’ group having amplitude emphasis on the acceleration stage, with much more rhythmic emphasis of the beat pattern and with much lower overall timing variability. The ‘legato’ timers in contrast had more amplitude emphasis on the tap force, with less of a rhythmic beat effect and less overall accuracy of timing. ‘Staccato’ tapping may benefit timing to the extent that its kinematic profile effectively heightens sensory feedback and minimises motor noise, thereby providing a reliable source of salient temporal information to guide movement control (see Elliott et al., 2009; Hove et al., 2014a; Krause et al., 2010).

Finally, as noted above, there was a strong relationship between the two timing groups and musicianship. The greater rhythmicity and accuracy of style and timing of the ‘staccato’ group could therefore be one of musical training or aptitude. There is indeed a substantial literature which demonstrates a difference between
musicians and non-musicians (e.g., Franěk et al., 1991; Kincaid et al., 2002; Repp, 2010) whereby musicians generally tap with less tap contact time (more staccato), with more accuracy (less variability) and with a reduced NMA (which approximately corresponds to the phase of peak acceleration here).

In sum, we found no compelling evidence for vestibular effects on the timing and kinematics of movements executed in synchrony with auditory rhythms that varied in sound frequency and intensity. Nevertheless, despite failing the strict test of an interaction between frequency and intensity, the observation that low-frequency, high-intensity sounds were associated with greatest movement amplitude suggests that the vestibular hypothesis deserves further testing under more tightly controlled conditions (Todd et al., unpublished results). Our finding that subjects spontaneously adopted two different modes of tapping that were correlated with musical experience and timing accuracy highlights the importance of taking movement kinematics into account in such future work and in studies of sensorimotor synchronisation more generally.

Acknowledgement

The first author was supported by a grant from the Prince of Wales Hospital Foundation.

References


