Electrophysiological Activity Associated With a Cross-Modal Anapaest Rhythm: Evidence for the Vestibular Syncopation Hypothesis

Neil P M Todd  
unswresearch.net.au

University of New South Wales  
https://orcid.org/0000-0003-0336-6684

Peter E Keller  
Aarhus University  
https://orcid.org/0000-0001-7579-6515

Sendhil Govender  
Neuroscience Australia  
https://orcid.org/0000-0003-0847-0974

James G Colebatch  
University of New South Wales  
https://orcid.org/0000-0002-2099-4976

Research Article

Keywords: rhythm and movement, cerebellum, vestibular system, syncopation

Posted Date: May 3rd, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2875207/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations:  
Declarations: there were no competing interests.

Version of Record: A version of this preprint was published at Timing & Time Perception on May 6th, 2024. See the published version at https://doi.org/10.1163/22134468-bja10108.
Abstract

We report an experiment that tested the vestibular syncopation rhythm hypothesis, which holds that the rhythmic effect of syncopation is a form of vestibular reflexive/automated response to a postural perturbation, for example during locomotion. Electrophysiological signals were recorded from the cerebral cortex and cerebellum during processing of rhythmic sequences in a sample of experienced participants. Recordings were made using four different stimulus modalities, auditory, axial, vestibular and visual, under different rhythmic timing conditions, irregular, regular and syncopated/uncertain. Brain current activity was measured using a 10 dipole source regions of interest model in each of the participants, each modality, each timing condition, and for each beat within the bar of the rhythm. The cross-modal spectral power in frontal EEG and cerebellar ECoG was also analysed. The results show that the brain activity increases from the irregular to the regular and then from the regular to the uncertain timing conditions. However, the vestibular modality induces the greatest total brain activity across the regions of interest, and exhibits the highest sensitivity to the interaction of beat structure with the timing conditions in both source currents and spectral power. These data provide further evidence to support the primal role of the vestibular system in rhythm perception.

Introduction

Music is a multisensory phenomenon. While the auditory modality is typically considered to be primary in musical communication, information received via other sensory modalities can affect the perception of musical sounds and the overall experience of musical performances. Observing a performer’s body movements influences judgements of musical expressivity and performance quality (Broughton & Stevens, 2009, Tsay, 2013, Vines, Krumhansl, Wanderley, & Levitin, 2006; Vuoskoski, Thompson, Spence, & Clarke, 2016), and movement observation and execution induce illusory percepts of basic features including pitch and duration (Manning & Schutz, 2013; Repp & Knoblich, 2009; Schutz & Lipscomb, 2007). Such biases demonstrate the relevance of multisensory integration in music perception, and more specifically highlight cross-modal interactions at the level of fundamental processing mechanisms in the brain (see Paraskevopoulos & Herholz, 2013; Walsh, 2003). The current article addresses the contribution of different sensory systems to the perception of musical rhythm, with particular focus on the vestibular system.

The neural mechanisms underlying perception of rhythmic stimuli have long been of interest to cognitive scientists, neurologists/neuroscientists and music researchers alike. A common theme has been the link between rhythm and movement (Todd, 1985; 1999), which can be traced to the earliest theoretical writings on the nature of rhythm (see Todd & Lee 2015a for review). This link has been exemplified in dance since antiquity and, more recently, in relation to the concept of groove—a multidimensional acoustic quality that induces the pleasurable urge to move in time with music (Janata, Tomic, & Haberman, 2012; Madison, 2006; Stupacher, Hove, & Janata, 2016; Todd & Lee, 2015b; Witek, Clarke, Wallentin, Kringelbach, & Vuust, 2014).

Musical features associated with groove include a strong beat at moderately fast tempi (100–120 BPM), low-frequency acoustic energy, and repetitive rhythms with moderate degrees of syncopation (Burger, Thompson, Luck, Saarikallio, & Toviainen, 2013; Etani, Marui, Kawase, & Keller, 2018; Hove, Martinez, & Stupacher, 2020; Madison, Gouyon, Ullén, & HöÖrstrÖm, 2011; Pressing, 2002; Sioros, Miron, Davies, Gouyon, & Madison, 2014; Stupacher, Hove, November, Schütz-Bosbach, & Keller, 2013; Witek et al., 2014). The level of groove in music has been found to influence the dynamics of postural body sway, the periodicity and quantity of motion in different body parts while dancing, and the force of finger taps in sensorimotor synchronization tasks (see Hove et al., 2020). A more general acoustic property that encourages movement in response to music is high sound intensity, as is common in dance clubs (Todd & Cody, 2000; van Dyck et al., 2013). While rhythms can influence movement not only in the auditory modality, as demonstrated in studies of visual and tactile stimulation for instance (e.g., Ammirante, Patel, & Russo, 2016; Hove, Fairhurst, Kotz, & Keller, 2013; Hove et al., 2020; Varlet, Marin, Issartel, Schmidt, & Bardy, 2012), questions remain about the relative strength of such effects. Moreover, the brain structures and functional processes that underpin links between rhythm and movement across sensory modalities are not fully resolved.

There is now a substantial body of evidence which supports the view that brain areas activated by rhythmic stimuli overlap considerably with those involved in the planning and generation of movement (see Todd & Lee 2015a for review). It is generally agreed that there are two distinct brain sub-systems involved, involving both the cerebellum and basal ganglia and associated subcortical and cortical structures (Ivry & Keele, 1989), but these are described in different terms and in different combinations of brain areas by different theorists. Lewis & Miall (2003) refer to “automatically” and “cognitively” controlled systems, with the cerebellum largely associated with “automatic” motor timing. Grahn & Rowe (2012) associate the basal ganglia with beat based rhythms, and thus a “beat prediction” system, while the cerebellum was more activated with irregular rhythms. Teki, Grube, Kumar, and Griffiths (2011) similarly associate the basal ganglia with “beat-based” auditory timing, while the cerebellum is more associated with “duration-based” timing in their scheme. Todd and Lee (2015a, 2015b) in an alternative perspective on rhythm perception, described two systems for “externally” vs “internally” guided action involving the respectively the cerebellum and basal ganglia.

THE VESTIBULAR SYNCOPEATION HYPOTHESIS
In addition to the above work on motor involvement in rhythm and beat perception, the influence vestibular system to rhythm perception has also been demonstrated (Phillips-Silver & Trainor, 2005, 2007, 2008; Trainor, Gao, Lei, Lehtovaara, & Harris, 2009). For example, it has been shown that activation of the vestibular apparatus is necessary, in this case by head movement, to observe effects on rhythm perception (Phillips-Silver & Trainor, 2007), but that vestibular influence on rhythm perception can be achieved directly by using galvanic vestibular stimulation without head movement (Trainor et al. 2009). These and other data led Todd and Lee (2015a) to propose that not only is the central vestibular system primal to rhythm perception, but that the central vestibular system is one and the same as the system which mediates rhythm perception. In other words, rhythm perception is a form of vestibular perception, even without overt head movement. However, the matter remains controversial, with some authors suggesting that the vestibular influence is not direct (Riggle, 2009; Phillips-Silver & Trainor, 2007; Trainor & Unrau, 2009). A further unresolved issue concerns what specific properties of the vestibular system, compared with other sensory systems, account for its role in rhythm perception.

One possibility is that the vestibular system may exert its influence on rhythm perception via its unique connectivity with the cerebellum (Ivry & Keele, 1989), especially the nodular/uvular and floccular/paraflloccular divisions of the cerebellum, which receive mossy fiber primary-afferent input from the vestibular apparatus and secondary projections from the vestibular nuclei (Barmack, 2016; Büttner-Ennever, 1999; Voogd, Nieuwenhuys, van Dongen, & ten Donkelaar, 1998; Walker et al., 2010). The nodulus/uvula and floccular/paraflloccular lobes in turn exert their influence by means of Purkinje cell inhibition of the deep cerebellar nuclei. The strength of this influence can be modulated by vestibular and other sensorimotor inputs to the cerebellum via the inferior olive and climbing fiber inputs to the nodulus/uvula and floccular paraflloccular lobes (Barmack, 2016; Voogd et al., 1998). The deep cerebellar nuclei project sub-cortically and also widely to the cerebral cortex, especially frontal areas involved in movement planning and preparation. Thus, the vestibular system could influence rhythm and beat perception via vestibulocerebellar projections to frontal cortex (Watson, Becker, Apps, & Jones, 2014; Todd, Govender, Lemieux, & Colebatch, 2021a; Todd, Keller, Govender, & Colebatch, 2021c).

In our own recent work we have shown that short latency vestibular cerebellar evoked potentials (VsCEPs) and an associated spontaneous electro-cerebellogram (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 and impulsive head accelerations (Govender, Todd, & Colebatch, 2020; Todd, Govender, & Colebatch, 2017, 2018a, 2018b, 2019, 2021b; Todd et al., 2021a, c). For air-conducted sound vestibular thresholds are typically around 80 dB sensation level (Todd et al. 2014b). These responses show considerable plasticity and can be modulated by eye gaze, head position and optokinetic stimuli.

Most recently we have shown, with the use of a custom 10% cerebellar extension montage, source localisation of generators of both short latency and long latency potentials within cerebellar and frontal cortical regions in response to vestibular and axial stimulation. These are achieved by means of short impulses delivered to the mastoid behind the ear for vestibular stimulation and at the C7 vertebra at the back of the neck for axial stimulation (Todd et al., 2021a). In the present paper we report an extension of this study to include auditory and visual stimuli using an experimental paradigm originally described by Todd and Seiss (2004), later reanalysed by Todd and Lee (2015b). This used irregular, regular anapaest (“Three blind mice”) and syncopated anapaest auditory rhythms. Syncopation was created in this context by omitting the tone on the third (metrically strong) beat (“mice”) of the four-beat metric cycle, leading to a situation where the tone on the second (weak) beat (“blind”) was followed by a silent strong beat.

We hypothesised that the vestibular stimulus would elicit both a higher degree of rhythmicity in brain activity and sensitivity to rhythmic syncopation than the other modalities. The vestibular syncopation hypothesis was also explicitly stated by Todd and Lee (2015b), who theorized that the power of syncopation and its link to movement was mediated through vestibular brainstem generated reflexes. In other words, a response to syncopation was equivalent to a reflexive/automated response to an unexpected postural perturbation, or unexpected absence of an expected postural perturbation, based on the prior temporal conditioning of a regular sequence with fixed temporal interval, such as may occur during locomotion. Given that the axial stimulus also produces a postural perturbation we would expect that axially evoked brain activity would also exhibit some of the above properties. Responses to syncopation across different sensory modalities were assessed using two complementary EEG analysis methods—event-related potentials and time-frequency analysis—which have proven informative in past studies on the brain bases of rhythm perception.

**EVENT-RELATED POTENTIALS**

The use of event related potentials (ERPs) from averaged stimulus or movement locked EEG is a powerful tool for investigating brain activity associated with sensory and motor behavior. For rhythm and beat induction, three categories of ERPs have been studied, namely stimulus following potentials, movement/stimulus preceding potentials and reafference potentials. Of the stimulus following category the cognitive ERP component waves, and in particular the P300 and mis-match negativity (MMN), have proven to be the most revealing in the analysis of rhythm perception and beat induction (Snyder & Large, 2005). The application of the P300 in a rhythmic context has been demonstrated by a number of researchers, including for subjective accents and omitted events (Brochard, Abecasis, Potter, Ragot, & Drake, 2003; Jongsm, Desain, & Honing, 2004; Jongsm et al., 2005). Several studies have shown sensitivity of the MMN to metrical structure (Vuust et al., 2005; Pablo Martín...
et al., 2007; Bendixen, Schroger, & Winkler, 2009; Geiser, Ziegler, Jancke, & Meyer, 2009; Schwartz, Rothermich, Schmidt-Kassow, & Kotz, 2011; Bouwer, van Zuilen, & Honing, 2014; Hove, Marie, Bruce, & Trainor, 2014). The classic movement preceding potential is the readiness potential or Bereitschaftspotential (or BP; Kornhuber & Deecke, 1990; Shibasaki & Hallett, 2006). It has been well established that this potential has its origin in the frontal areas of the brain and in particular in the supplementary motor area (SMA) and cingulate motor areas (CMA), as well as in the motor cortex (MI) in its latter stages. There are some differences depending on whether the movement is self-paced or cued (Jankelowitz & Colebatch, 2002). The related contingent negative variation (CNV) has been proposed as an index of brain processes underlying time estimation and is generally observed over frontal regions during the temporal gap between two events, usually a warning and imperative stimulus (Macar & Vidal, 2004; Praamstra, Kourtis, Kwok, & Oostenveld, 2006). There is now a considerable body of literature on the role of the CNV and other slow potentials in time interval estimation, but the application of the method specifically to beat induction in a metrical context has been limited (Zanto, Snyder, & Large, 2006). The third set of electrophysiological components relevant to the sensorimotor approach to beat induction is the reafference potentials (RAPs), of somatosensory origin, which are a consequence of movement. Such potentials can be observed following the BP, usually in the form of post-motion potentials of both negative and positive polarity (Kornhuber & Deecke, 1990; Shibasaki & Hallett, 2006).

All three of these species of potentials come together within the realm of sensorimotor synchronization (Müller et al., 2000). A study of error correction in sensorimotor synchronization by Praamstra, Turgeon, Hesse, Wing, and Perryer et al. (2003) using EEG demonstrated examples of stimulus and movement related potentials. Essentially there are two morphologies, depending on whether the averaging is done relative to the stimulus during passive listening or the movement during synchronization: (a) stimulus related potentials, consisting of the P1, N1, P2, and N2 component waves, as above, and (b) movement related potentials, consisting of a PMN, equivalent to the BP, a re-afference negativity (RAN) and a post-motion positivity (PMP; Jankelowitz & Colebatch, 2002). A third hybrid morphology is obtained when locking the averaging to the stimulus during active synchronization. Whether these potentials are functionally relevant to similar degrees for rhythm perception across different sensory modalities is not known. Comparing ERPs across the vestibular and other modalities provides an avenue for examining modality specific versus amodal neural processes underpinning rhythm perception.

**TIME-FREQUENCY ANALYSIS**

A complementary method to ERPs is use of time-frequency analysis of EEG spectral power. The application of time-frequency analysis methods is one which has become central to the field of rhythm perception and production, both in terms of acoustic signal processing and analysis of electroencephalographic signals. Spectral analysis of rhythmical signals was pioneered by Todd and colleagues in the 1990s (Todd, 1993, 1994, 1995, 1999; Todd & Brown, 1996), who went on to develop this into a sensory-motor theory of how such signals are processed in the brain (Todd, O'Boyle, & Lee, 1999; Todd, Lee, & O’Boyle, 2002). The essential idea was that sensory cortex represents the temporal properties of signals in the form of a spectro-temporal (for auditory) or spatio-temporal (for visual) power spectrum in a range of approximately 0.5–20 Hz. Thus, a regular isochronous sequence would be represented by a harmonic series, with the fundamental frequency corresponding to the inverse of the interval, but even single intervals, filled or empty, would be represented spectrally. More complex and syncopated rhythms would be represented by a more complex spectral profile, with the higher harmonics referred to as *metrical harmonics*. When viewed through the window of the motor system, the sensory evoked spectra would be weighted by what Todd and colleagues termed a sensory-motor filter which shaped the spectra to have dominant harmonics close to the natural or eigen-frequencies of human bodily movement, including locomotion, which has been established as coinciding with the existence region of the beat, i.e. the relatively narrow range of tempi at which a tactus is felt (Fraisse, 1982; Parnicutt, 1994; van Noorden & Moelants 1999; Styns, Van Noorden, Moelants, & Leman, 2007). This would explain why with a change of tempo the character of a rhythm would shift when its dominant harmonics flipped up or down and the beat shifted up or down a metrical level, e.g. how an anapaest pattern may change from a “Three blind mice” to a “Jingle bells” to a “William Tell” rhythm. The theory also suggested a role for the cerebellum as a feed-forward model in making temporal predictions based on the weighted cortical modulation spectrum (Todd et al., 1999; 2002), and further, that the central vestibular system was critical as it provided the essential link between sensory and motor representations, including via the vestibular cerebellum (Todd & Lee, 2015a,b; Todd, Keller, Govender, & Colebatch, 2021c).

While there have been several experimental studies which support the importance of the vestibular system to rhythm and beat perception (Philips-Silver & Trainor 2005, 2007, 2008; Trainor, Gao, Lei, Lehtovaara, & Harris, 2009), several outstanding questions remain. Specifically, does the vestibular system contribute to a spectral sensory-motor representation? How do these relate to the natural rhythms of the EEG/ECeG? Is this mediated by the vestibular cerebellum? Is metrical structure, i.e. the sense of strong-weak beats, vestibular/locomotive in origin? Can this explain the effect of syncopation, or metrical surprise? This last question was formulated as the vestibular syncopation hypothesis (Todd & Lee 2015a,b; Todd et al., 2021d). The current study aims to go some way to answering these questions by applying a cross-modal power analysis of EEG/ECeG in response to a simple rhythmic input. Time-frequency analysis is a technique that has long been applied to the study of rhythm in EEG signals but the non-invasive ECeG, which as its own natural rhythms (Todd, Govender, & Colebatch, 2018; 2019), has only just been discovered and to date there are no studies on the effect of rhythmic signals.
Within the EEG domain when time-frequency analysis has been applied to specifically the role of temporal prediction in auditory rhythm perception it has been suggested that induced gamma (20–60 Hz) band activity could be an index of anticipation (Pantev, 1995; Bertrand & Tallon-Baudry, 2000). Snyder and Large (2005) found evidence of anticipation in gamma band activity (GBA) using a simple binary rhythm with random omissions on loud or soft tones such that the timing of the GBA on omitted beats was close to the onset of the beat. Zanto, Snyder, and Large (2006) suggested that the generators of GBA might be similar to those of anticipatory slow potentials in frontal cortex, i.e., the BP and CNV, and were thus distinct from the beat following potentials, i.e., MMN and P300. Iversen, Repp, and Patel (2009) further investigated both gamma and beta (10–30 Hz) MEG activity in passive and imagined beat conditions for simple rhythms formed from two tones with a gap. They reported evidence that the beta band in particular, the lower end of the Snyder and Large (2005) gamma range (20–30 Hz), was active in subjectively imagined beats. The beta oscillations, it was suggested, allowed the synchronized coupling of the various sensory and motor areas. Fujioka, Trainor, Large, and Ross (2009) examined both beta (15–30 Hz) and gamma oscillations (>30 Hz) in the auditory cortex to rhythms composed of loud and soft tones. They reported that beta activity synchronized with an isochronous sequence, decreasing after tones, but increased “excessively” after an omitted tone. In contrast the gamma activity peaked 80 ms after a present tone, but 110 ms after an omitted tone. They suggested that the two oscillation bands subserved two distinct functions. Whereas the gamma band in their view reflected an endogenous process of anticipatory entrainment, the beta band reflected an exogenous audio-motor coupling process (Fujioka, Zendel, & Ross, 2010). Subsequently, Fujioka, Trainor, Large, and Ross (2012) applied a beta band analysis to isochronous sequences of different rates suggesting the beta “rebound” between stimuli related to an interval timing process.

More recent EEG/MEG studies, inspired by the “neural entrainment” theory of rhythm perception (Large & Kolen 1994; Large, 2008), have given emphasis on lower frequency bands which cover the natural frequencies of the rhythmic signals at the level of the beat, with evidence to show coupling of delta and beta activity during sensorimotor synchronisation (e.g. Morillon, Hackett, Kajikawa, & Schroeder, 2015; Rimmele, Morillon, Poeppel, & Amal, 2018; Zalta, Petkoski, & Morillon, 2020). Related work has examined how such activity supports auditory processing at multiple timescales in signals with regular or irregular timing, including speech (Kotz & Schwarte, 2010; Stockert, Schwartz, Poeppel, Anwander, & Kotz, 2021; Teng & Poeppel, 2019). A spectral approach using narrow-band Fourier analysis to focus on changes to specific metrical harmonic components represented in the EEG, referred to as “frequency tagging”, has been proposed as evidence for the existence of coupled “neural oscillations” underlying rhythm perception and production (Nozaradan, Zerouali, Peretz, & Mouraux, 2015). Frequency tagging studies have found that beat-related peaks in the EEG frequency spectrum are larger in amplitude in syncopated rhythms than in unsyncopated rhythms (e.g., Nozaradan, Peretz, & Keller, 2016), especially when the rhythms are conveyed by low-pitched sounds (Lenc, Keller, Varlet, & Large, 2019). Thus, these recent interests in the field converge on spectral representations of rhythmic stimuli that vary in regularity, albeit from a radically different perspective to the sensory-motor theory as outlined above (for a review of the differences of these perspectives see Todd & Lee (2015a)).

**THE PRESENT STUDY**

In the present study, we used both ERP and time-frequency as complementary methods to investigate the effects of four modalities in differing timing conditions (regular, and syncopated/uncertain), as previously described (Todd & Seiss 2004; Todd & Lee 2015b). For the ERP analysis like Todd & Lee (2015b) we also used a source current model to measure the changing brain activity in regions of interest, including cortical and subcortical areas, but generalized to allow for cross-modal comparison. Further, the use of a cerebellar extension to the standard 10–20 EEG montage provided proof-of-concept that non-invasive recording from cerebellar sources relevant to rhythmic processing is feasible. The resulting source current waveforms were then segmented according to the global field power associated with cross-modal equivalents of the AEP N1, P2, N2 and P300 ERP component waves for each beat of the rhythm. The segmented current sources then served as the dependent variable for the cross-modal analysis. The vestibular syncopation hypothesis would be testable as an interaction in an analysis of variance. Specifically, support for this hypothesis would manifest in greater sensitivity to syncopation and beat, along with an interaction of syncopation by beat, for the vestibular modality, but given the shared postural aspect (and convergence at the level of the brainstem) we would also expect the axial modality to exhibit a similar sensitivity. For the time-frequency analysis, the resulting EEG/ECG power were segmented using the same time periods as for the currents, which then served as the dependent variable for the cross-modal analysis. The vestibular syncopation hypothesis would, as with the ERP analysis, again be testable as an interaction in an analysis of variance of power.

**Materials and Method**

**PARTICIPANTS**

Five adult healthy participants (4 males and 1 female) were recruited from staff and students at the Prince of Wales Hospital and Western Sydney University. The participants were selected as being skilled musicians and/or trained experimental participants in order to ensure the highest quality artefact free recordings. We opted to test a small sample of experienced participants who generated high quality data rather than a large mixed sample, given our focus on basic sensorimotor processes that were not expected to display large inter-individual variation. All participants gave written informed consent before experimentation and the study was approved by the Human Research Ethics Committee.
Detailed analysis was performed on two electrodes – FCz and Iz as being representative of frontal and posterior fossa responses, respectively. This was followed by analysis of the EMG and accelerometry recordings. Global field power was computed by the RMS of all electrodes combined. After recording EEG/ECG we performed source localization using the whole 63 channels, spectral power analyses of selected channels, and calculated the interaction of EEG and EMG. The triggers for the stimuli (either one, two or three per trial) were generated from digital outputs with Signal software. Markers for epoch zero were used to synchronise the parallel EEG and EMG recording systems controlled by custom software driving a second Power1401 digital output.

Mice rhythm (46–50). Participants were asked to report the number of times they heard the “Three blind mice” rhythm (36–40) and for the uncertain condition to count the number of times they heard the complete “Three blind mice” rhythm (46–50). The visual stimuli were delivered via a colour display unit (Samsung, model C24F390FHE) at a viewing distance of about 30 cm and consisted of alternating vertical black/white bars generated by a custom software. The arrangement of stimuli in each modality into sequences in the three timing conditions (one regular, and two syncopated or uncertain) is described below.

EEG/EcEG RECORDING

FIGURE 1

63 channels of EEG/ECG were recorded from over the scalp and neck using a novel 10% cerebellar-extended 10–20 system cap (EASYCAP GmbH, Germany) (see Fig. 1). The cerebellar extension was designed by completing the population of the Iz row, labelled using the standard nomenclature, from P9 to P10, and adding two additional 10% rows inferiorly, from P11 to P12 and from P13 to P14. Subsequent to the cap design, Heine, Dobrota, Schomer, Wigton, and Herman (2020) published an extended electrode placement nomenclature which we have adopted here. The electrodes were Ag/AgCl type and maintained at 10 kOhms or less. A ground electrode was placed at AFz with reference at Nz. Signals were amplified using a combination of amplifiers, 32-channel ActiChamp and 32-channel Digitimer D360/D120 (Digitimer Co, UK), filtered at 0.5 Hz to 3 kHz and sampled at 10 kHz. The 32 ActiChamp channels were recorded using BrainVision software (version 1.22, Brain Products GmbH, Germany) and the 32 Digitimer amplified channels were sampled using a CED Power 1401 and recorded using Signal software (version 6.02, Cambridge Electronic Design, UK).

EXPERIMENTAL PROCEDURE

FIGURE 2

Recordings were made using the four stimulus modalities under three timing conditions (see Fig. 2): a regular condition, in which the stimuli were presented in the form of an anapaest (“Three blind mice”) rhythm, with inter stimulus intervals of 600 ms and 1200 ms (inter-trial interval 2400 ms), and a recording epoch of 2100 ms; and two uncertain/syncopation conditions, where the third beat of the anapaest rhythm was randomly absent on 50% of trials, thus forming two trial types uncertain (present) and uncertain (absent), i.e. three in total. An irregular condition, in which the stimuli were presented with a random inter-trial interval (800 ms to 1400 ms), with a recording epoch of 700 ms, was also included in the recordings. The stimulus modalities and timing conditions were scheduled pseudo-randomly. The recordings took place with the participants in the standing position; for the axial stimulus they were asked to lean forward, but look towards the horizon. Between recordings participants were allowed to rest sitting down. For the regular condition subjects were asked to count the number of times they heard the “Three blind mice” rhythm (36–40) and for the uncertain condition to count the number of times they heard the complete “Three blind mice” rhythm (46–50). Participants were asked to report the number and this was recorded. The timing of the trials, stimulus delivery and synchronisation of the parallel EEG and EMG recording systems was controlled by custom software driving a second Power1401 digital output. The triggers for the stimuli (either one, two or three per trial) were generated from digital outputs with Signal software. Markers for epoch zero point and trial type were also recorded. The whole procedure lasted about 90 minutes, including setting up.

DATA ANALYSIS

After recording EEG/EcEG we performed source localization using the whole 63 channels, spectral power analyses of selected channels followed by analysis of the EMG and accelerometry recordings. Global field power was computed by the RMS of all electrodes combined. Detailed analysis was performed on two electrodes – FCz and Iz as being representative of frontal and posterior fossa responses, respectively. For the vestibular condition, PO10 was used rather than Iz as we have previously shown a lateralised maximum (Govender et al., 2020).
EEG/ECeG analysis. All EEG/ECeG recordings were screened for blinks and other artefacts (about 5–10% trials) and merged together using the Scan software (version 4.5, Compumedics Ltd, Australia) and BESA software (version 6.1, MEGIS Software GmbH, Germany). For the electrical source analysis the data were averaged across all timing conditions in order to improve the signal-noise ratio for each modality.

Brain Electrical Source Analyses (BESA). The standard four-shell ellipsoidal head model was employed with radial thicknesses of 85, 6, 7 and 1 mm for respectively the head, scalp, bone and CSF, with conductivities of 0.33, 0.33, 0.0042 and 1.0, respectively. Both cerebrum and cerebellum fall within the CFS volume conductor and are not discriminated by the BESA algorithm. The fitting was carried out using the BESA genetic algorithm with default parameter settings after remontaging to an average reference. A modelling strategy was adopted to run a 10 dipole (maximum allowed by degrees of freedom) genetic algorithm fit 10 times to test its reproducibility using different starting points. The resultant 100 locations for each modality were then subject to a hierarchical cluster analysis, using the between-groups linkage method with squared Euclidian distance measure, in order to eliminate the non-viable and very weak sources. A 5 mm³ standard error was imposed on the cluster volumes and any isolated single dipole sources which resulted from that constraint were eliminated. In addition to the mean Talairach-Tournoux coordinates of the final surviving clusters, a weight was attributed to the clusters derived from the number of dipoles making up the cluster divided by 10 (runs). Thus if the same source appeared for every run its weight would be 1.0. For cerebellar coordinates the Schmahmann et al. (2000) atlas was used to determine the anatomical locations, while for other locations the Talairach Client application (version 2.4.3) was employed with a +/- 5mm cube search (see Todd et al., 2021a).

10 Regions of interest source current analyses. In order to make a comparison of brain activity across all four modalities and timing conditions, a generic 10 dipole source model with fixed locations was designed to be representative of the areas common to all modalities. The specific source coordinates were derived from those obtained from the individual modalities. This generic model was then optimised for each modality, by allowing the dipole orientations to fit over the whole epoch, and then applied to each individual participant. The resultant 10 current profiles for each of the five participants, four modalities and three timing conditions were further segmented into five “waves”, a baseline (BL) epoch, a short latency (SL) epoch (selected to exclude electrical stimulus artefact), and three long-latency N1, P2 and N2/P300 epochs and three “beats”. The P300 has been lumped with the N2 as they occur during the same epoch mutually exclusively. Within each segment the mean absolute source current was taken to be the dependent variable in units of nano-Amperes (nA).

Analysis of variance (ANOVA). The regular/uncertain data were combined together in a single repeated measures analysis of variance (ANOVA) which was applied to the currents with within-subjects factors of “Timing Condition” (regular, uncertain present & uncertain absent), “Beat” (1, 2 & 3), “Modality” (auditory, axial, vestibular and visual), “Source” (1–10), and ERP component “Wave” (BL, SL, N1, P2, N2/P300). To test for the effect of regularity we also ran an ANOVA on combined irregular/regular data but treating the irregular a “beat zero”. In general a 5% level was taken as the threshold for statistical significance in the ANOVA main effects. However, in order to further ensure against type I error, in the interpretation of main effects this was increased to 1% for post-hoc pair-wise comparisons if uncorrected. The 1% level was also strictly applied to ANOVA interactions, apart from when there was a prior hypothesis, in which case a 5% level was applied. Subsequently, as we had made a specific modality dependent hypothesis the data was subsequently split by modality and ANOVAs run independently. In addition we also provided within-subjects contrasts to assist in the interpretation of non-linear contributions to the effects when there were more than two levels of the within-subjects factors, with again a 1% level strictly applied to interactions. In all cases the Greenhouse-Geisser correction for sphericity violations was applied to p-values.

Post-hoc analyses using the General Linear Mixed Model (GENLINMIX). As the repeated measured ANOVA is limited to main effects in pair-wise comparisons, for post-hoc analysis of interaction effects we followed up on targeted fixed effects on data subsets, specifically the N2/P300 epoch, using the GENLINMIX approach. For all such analyses, degrees of freedom were derived using the Satterthwaite method with Bonferroni correction on all pair-wise comparisons.

Spectral power analyses of spontaneous cerebellar activity. After recording EEG/ECeG we performed spectral power analyses of selected channels. In order to measure any high-frequency pausing or bursting, characteristic of post-climbing fibre responses, a spectral power analysis was also conducted on electrode sites at which the CEPs had previously been identified the scalp (Iz for the axial, PO10 for the vestibular) using the continuous wavelet transform (CWT) as implemented in the MATLAB toolbox (R2019b, Mathworks, Natick, CA). In the present analysis a Morlet wavelet was employed at a density of 24 voices per octave over 9 octaves. The CWTs were further transformed to scaleograms (time-frequency images) from the absolute value of the CWT and rescaled to be in dB per Hz re 1 µV². Scaleograms were computed for all trials, then further split into eight frequency bands; delta (δ: 1.8 Hz – 4 Hz), theta (θ: 4–7.8 Hz), alpha (α: 7.8–12.5 Hz), beta (β: 13–30 Hz), gamma (γ: 30–80 Hz), ultra-gamma (u-γ: 80–160 Hz), very high frequency (VHF: 160–320 Hz) and ultra-high frequency (UHF: 320-1 kHz). These were then further segmented using the same time boundaries as for the current source analyses and submitted to ANOVA.

Results

Our results are divided into three parts dealing with, respectively, the evoked potentials, the source currents, and the EEG/ECeG spectral power. In the first part we consider the nature of the evoked potentials generated from averaged EEG/ECeG for each of the four modalities. The global...
field power of the potentials from each of the sensory modalities was used to define the segmentation epochs for each of five sub-epochs, baseline (BL), short-latency (SL), N1, P2 and N2. In the second part we conducted a brain electrical source analysis (BESA) for each the modalities, which enabled us to define a region of interest 10 source model as a cross-modality measuring tool. We then applied this tool to statistically compare current activity across modality as a function of timing condition and beat position, effects of which could indicate sensitivity to syncopation and metrical structure. For the time-frequency analyses we computed the cross-modal spectral power along with statistical analysis of both frontal EEG and the cerebellar ECoG. Ideally, the spectral range of this computation should approximately span three decades, from about 0.5 Hz to 500 Hz, to cover those frequencies associated with the driving signal, including the beat interval of 600 ms (corresponds to 100 BPM or 1.67 Hz), and up to the UHF band of the ECoG. Unfortunately, this full span requires a long epoch of four repetitions of the anapaest rhythm which was only achievable for the frontal set of electrodes because the CED system recorded each epoch online. We nevertheless, report the full range initially for FCz in order to have a clear interpretation of the metrical components of spectra. Subsequently, we employed the shorter epoch which allows cross comparison of the two sites.

Results I: Potentials

GRAND MEANS OF EEG/ECEG

FIGURES 3 HERE

Figure 3 shows the grand means for responses to the auditory (Fig. 3A) and visual (Fig. 3B) modalities averaged over all timing conditions. The axial and vestibular means have been previously published (Todd et al., 2021a). Unlike the axial and vestibular response, the auditory and visual responses do not show a prominent short latency (SL) response, but display the classical long latency responses, N1 and P2 potentials, which are common to all four modalities. A subsequent N2 potential was absent for the visual stimulus.

FIGURE 4 HERE

Figure 4 shows the grand means for selected electrodes FCz and Iz/PO10 for all four modalities and for the three anapaest conditions. Both axial and vestibular modalities show very prominent SL responses, especially over the cerebellum, but also frontally. For the uncertain, beat-missing condition a P300 appears to be clearly present in the axial and vestibular cases, less so for the auditory and not obviously present for the visual modality. These visual impressions are tested below.

GLOBAL FIELD POWER

FIGURE 5 HERE

Table 1 HERE

Figure 5 shows the grand means for FCz and Iz/PO10 for all four modalities collapsed across beats and for all the timing conditions, along with a butterfly plot and global field power (GFP). The lobulated structure of the GFP allowed us to define the epoch boundaries which are given in Table 1. Although there is some variation in the boundaries between lobes, there is broad agreement, with the SL-N1 boundary ranging from 55–77 ms, the N1-P2 boundary ranging from 126–150 ms and the P2-N2 boundary 253–273 ms. Figure 5 also shows top meridian maps for the three long latency potentials for each modality. The N1 potentials are quite modality specific, with the auditory and visual cases showing their classical distributions reflecting activation in bilateral temporal for the auditory and midline occipital for the visual. The N1 potentials for the axial and vestibular are though more complex indicating multiple foci. In contrast the P2/N2 potentials appear to be less modality specific and centrally focussed.

Results II: Currents

BRAIN ELECTRICAL SOURCE ANALYSES (BESA)

Source analyses for each modality

TABLES II AND III HERE

The outcome of the 10 dipole source modelling and cluster analyses are shown in Tables II and III for the auditory and visual modalities. The source/cluster analyses for the axial and vestibular modalities have been previously published (Todd et al., 2021a). All four modalities show wide-spread cortical and sub-cortical activations, including midline frontal, basal ganglia and cerebellar sources, in addition to their primary areas.

The generic regions of interest model
The regions of interest source model locations are given in Table 4 and their orientation optimised fit illustrated in Fig. 6. Over the whole epoch the four optimised models give residual variances of respectively 5%, 8%, 13% and 4% for the auditory, axial, vestibular and visual modalities. Both the axial and the vestibular modalities show large short-latency cerebellar and frontal activations, and more widely in all locations. For all four modalities, the mid-cingulate source is a dominant contributor to the long-latency potentials, although for the visual case the occipital P100 coincides with the cingulate N1 and for the auditory cases the temporal N1/P2 overlaps with the cingulate N1/P2 generator.

**STATISTICAL ANALYSES OF THE CURRENT SOURCES**

**FIGURE 7 HERE**

**TABLE V HERE**

Table 5 shows the results of the ANOVA for the combined data for the regular/uncertain conditions, and then split by modality. All of the factors, apart from "beat", showed significant main effects, and also a large number of within-subjects factor interactions and contrasts. The irregular condition is included in Fig. 7 for comparison but tested separately below against the regular condition.

**Effects of uncertainty**

The effect of "timing condition" (Fig. 7B) shows that the uncertain rhythm increases brain activity across the 10 regions compared to the regular anapaest condition. The uncertain 3rd stimulus present condition yielded the highest activity compared to the other two (pair-wise p < .005 in both cases). The main effect of "modality" (Fig. 7C) indicates that the vestibular stimulus tends to produce higher overall activity across the 10 regions of interest than the others followed by axial, then visual (pair-wise p < .01), and last, auditory (pair-wise p = .01). The main effect of "source" (Fig. 7D) indicates that the midline cingulate and midline cerebellar sources (sources 4 and 7) are dominant. Finally, the main effect of "wave" (Fig. 7E) is consistent with overall activity being higher during the SL epoch compared to the baseline (pair-wise p < .005) and later epochs.

Three significant, at 1% or less, two-way interactions were yielded by the combined ANOVA (Table 5). The first, "timing condition" with "beat" (p < .001), appears to indicate that the effect of "beat" changes over the timing conditions, with the divergence especially on (absent) third beat. The second, "wave" with "beat" (p = .01), indicates a change in the beat profile across the wave conditions, with a tendency for the baseline activity to increase with beat position (from relatively weak (W) to relatively strong (S)), an apparent tendency for the N1 in particular to decrease with beat position (from relatively strong (S) to relatively weak (W)) and for the activity during the N2/P300 ERP component wave to remain flat, or even increase again from beat two to beat three (a relatively strong(S)-weak(W)-strong (S) pattern, tested explicitly in the following section). The third illustrated two-way interaction, "wave" with "modality" (p = .001) supports the ERP component wave effect being particularly prominent for the vestibular and axial modalities, with the short-latency (SL) and N1 activity outstanding compared to the other two modalities. There was also one significant (at 1% or less) three-way interaction, i.e. "wave" by "beat" by "timing condition". This is consistent with the change in beat profile for the different ERP component waves, apparent in the two-way "wave" by "beat" interaction, being enhanced or exaggerated by syncopation, but with the missing stimulus likely making the major contribution. It also revealed that the activity during the SL wave remains largely unchanged by beat or timing condition, other than after the omitted stimulus, so that the strong-weak beat pattern is carried primarily by the later ERP component waves.

**Effects of regularity**

To test for effects of regularity we ran a separate ANOVA including the irregular condition against the regular condition, but treating the irregular as a beat "zero" (Fig. 7A). This yielded main effects as in the regular/uncertain comparison but with the main effect of "beat" reaching significance at the 5% level (F(3,12) = 8.6, p < .05). Pair-wise comparison indicated that this was primarily due to the difference between the irregular 0th beat and the 1st beat of the regular rhythm. However, this analysis also produced the same two-way interactions of "modality" by "wave" (F(12,48) = 17.0, p < .0001) and "beat" by "wave" (F(12,48) = 8.2, p < .01) as in the regular/uncertain comparison. This confirmed that the "beat" effect is primarily carried by the later ERP component waves rather than SL components, but in this case extending to the effect of regularity over irregularity in the late components and in addition to the baseline epoch which would include anticipatory activity in the regular condition. Thus, regularity had the effect of increasing activity, but, for this choice of intervals, primarily in the late response components produced by the initial strong beat and in its anticipation. Overall this revealed an activity hierarchy across the timing conditions from irregular to regular, and then from regular to syncopated.

**Separate modalities**
As we had at the outset hypothesised specific modality interaction effects, and both three-way interactions of "modality" with "timing condition" by "beat" and four-way with "wave" by "beat" by "timing condition" were significant at the 5% level, the data was split by modality (Table 5). All of the four modalities show main effects and trends consistent with the combined data, albeit with lower significance levels due to the reduction in power. Both axial and visual showed the two-way interaction of "beat" by "timing condition". However, only the vestibular modality independently produced main effects of both "timing condition" and "beat" at the 5% level, with a pair-wise comparison confirming the missing stimulus as the major source of the effect. Further, only the vestibular modality exhibited at the 1% level or less the three-way interaction of "wave" by "beat" by "timing condition" described above for the combined data. This is illustrated in Fig. 8 where we compare the non-significant auditory versus the vestibular three-way interactions. This confirms that the auditory contribution (Fig. 8A) to the beat effects in the combined data interaction is small. In contrast the vestibular modality (Fig. 8B) exhibits all of the features present in the combined interaction. Thus, evidence of an emerging metrical structure, i.e. a S-W-S beat pattern, with syncopation appearing to be present for the vestibular modality, but not the auditory.

**Results III: Power**

EEG VERSUS ECeG SPECTRAL POWER

**Continuous wavelet transform scaleograms and power spectra of the EEG at FCz: long epoch**

Figure 10 illustrates scaleograms and associated power spectra for the four modalities at FCz for the regular anapaest rhythm. All four exhibit α-, β- and γ-band activity, although the α-band is partially suppressed in the visual case, with spectral peak frequencies at about 8.5–9.0 Hz, 18–21 Hz and 81–84 Hz respectively (Fig. 10D). As well as any tonic activity in these bands, the axial and vestibular modalities also exhibit stimulus evoked short-latency β- and γ-band activity, as well as well-defined stimulus related α-band modulation (Fig. 10B, C). The auditory modality also shows stimulus related α-band modulation to a lesser degree (Fig. 10A).

All modalities, to differing degrees, also exhibit well-defined low-frequency exogenous spectral components which are related to the temporal properties of anapaest driving rhythm, with peaks corresponding to the fundamental and harmonics associated with the beat frequency – an interval of 600 ms corresponds to a tempo of 100 BPM or 1.67 Hz. We refer to these as metrical harmonics (as in Todd, 1994), with a metrical fundamental, or $M_0$ of 1.67 Hz, and 1st and 2nd metrical harmonics, or $M_1$ and $M_2$, respectively of 3.3 and 5.0 Hz. For this tempo, the $M_0$ falls within high β-band, while the $M_1$ and $M_2$ fall within the θ-band. The $M_2$ harmonic also exhibits stimulus related modulation in phase with the α-band modulation. Although metrical harmonics can be discerned for all the modalities, they are noticeably more prominent, at least qualitatively, for the axial and vestibular modalities, and for the visual modality the $M_0$ is not well-defined. The axial and vestibular modalities also include a sub-harmonic below $M_0$.

**Continuous wavelet transform scaleograms and power spectra of the EEG at FCz versus PO10: short epoch**

Figures 11 and 12 illustrate short epoch (2.1 s) scaleograms averaged across trials and subjects for each of the four modalities and three timing conditions at respectively the electrode positions FCz and PO10, along with their associated spectra. Clearly, there is a dramatic difference between the spectral profiles of the two electrode sites. Whereas the frontal site has an overall low-pass profile dominated by the low-frequency bands (Fig. 11), the posterior site has an overall band-pass profile with a high-frequency peak around 100 Hz in the ultra-γ-band (Fig. 12). Unlike the frontal profile, the posterior profile does not exhibit clearly distinct β- and γ-bands but rather a single broad high-frequency band. The tonic level of high-frequency power is both modality and timing condition dependent. The auditory, axial and vestibular modalities do show a clear α-peak in the posterior profile (Fig. 12D, H L), in contrast to the visual posterior profile (Fig. 12P).
Regarding the metrical components, all four modalities show evidence of a low-frequency peak corresponding to the upper shoulder of the metrical fundamental $M_2$ in the posterior profile, although this is not easily discernible in the frontal profile due to the short epoch truncation and general low-pass characteristic. The metrical harmonics $M_1$ and $M_2$ are also variously present, but also apparently dependent on the timing conditions, especially for the vestibular frontal profile at FCz, as illustrated in Fig. 11 (R, L), consistent with the results reported above.

In addition to their static spectral properties, the scaleograms also exhibit dynamic properties in time. On visual inspection at least, the $\beta$- and $\gamma$-bands are relatively invariant to timing condition for both locations, but the $\theta$- and $\alpha$-bands in contrast show a greater range of responsiveness, with evidence of both beat and timing condition sensitivity.

**STATISTICAL ANALYSES OF THE EEG VERSUS ECeG SPECTRAL POWER**

In order to statistically evaluate the effects apparent in the scaleograms, in the first instance power at the two electrode sites was broken down by “band”, “timing condition”, “beat” and “wave” and combined across modalities and electrode for a single omnibus ANOVA. Due to the problem of end-effects, caused by the shortness of the total epoch, a baseline segment and the $\delta$-band were excluded. Further, given the absence of significant EEG power in the highest bands, the statistical analysis was further confined to just five spectral bands, $\theta$, $\alpha$, $\beta$, $\gamma$ and $\mu$-bands. As this might reduce the power of the test we relaxed the strict 1% criterion for interactions to 5%.

**TABLE VI HERE**

**FIGURE 13 HERE**

The outcome of this omnibus analysis is given in Table 6, where all main effects are presented, but only the most significant interactions, also illustrated in Fig. 13, along with the hypothesised modality interactions. As would be expected, there was a highly significant main effect of “electrode” (Fig. 13A, p < .005) and associated interaction of “electrode” by “band” (Fig. 13B, p < .001), supporting the qualitative observation that there is a significantly higher amount of high-frequency power in the ECeG than EEG. This is further supported when broken down by electrode where both EEG and ECeG profiles have main effects of “band”, but whereas the EEG shows a typical low-pass spectral profile, the ECeG shows a high-pass spectral profile over the range of frequencies considered. The omnibus analysis also showed a highly significant main effect of “modality” (Fig. 13C, p < .005), indicating that both axial and vestibular modalities produce greater total power than the auditory or visual modalities. When broken down by electrode, the modality effect is seen to be predominantly in the ECeG (p < .001), rather than the EEG, and then predominantly a high-frequency effect when broken down by band, especially in the $\gamma$-band (p < .005), i.e. it is predominantly the $\gamma$-band ECeG that discriminates the axial and vestibular from auditory and visual modalities, as can be seen qualitatively in Fig. 13.

Effects of “beat” (Fig. 13D, p < .05) and “wave” (Fig. 13E, p < .05) were also observed, which indicated a tendency for greater power to occur on the 1st beat and during the SL segment compared to later segments. When broken down by electrode, these effects were generally equally distributed in the EEG and ECeG, but when broken down by band both effects were distributed in the $\beta$, $\gamma$- and $\mu$-bands. The two effects also interacted significantly (Fig. 13F, p < .01), indicating that the tendency for greater power to occur on the 1st beat compared to the 2nd and 3rd was confined to the SL, N1 and P2 segments, but not the N2/P300. This interaction tended to be more significant for the EEG compared to ECeG, occurred across the bands, more evidently for the $\beta$- and $\gamma$-bands (p < .01), compared to the $\theta$- and $\alpha$-bands (p < .05), and across the modalities, except the auditory modality.

The differential quality of the low-frequency ($\theta$/$\alpha$-bands) and high-frequency ($\beta$/$\gamma$-bands) “beat” by “wave” interaction, which is described above, is revealed in the three-way interactions of “timing condition” by “beat” by “wave” (p = .005) and “band” by “beat” by “wave” (p < .05) and their breakdowns. The tendency for greater power to occur on the 1st beat compared to the 2nd beat is primarily a low-frequency N1/P2 effect (p < .05), whereas the high-frequency aspect (p < .1) is due to greater SL sensitivity to the missing stimulus on the 3rd beat. Thus the low-frequency component of the “timing condition” by “beat” by “wave” interaction is an accent effect due to the timing context, whereas its high-frequency counterpart is primarily a missing stimulus effect.

**FIGURE 14 HERE**

The three-way interaction of “timing condition” by “beat” by “wave” is especially evident for the vestibular (Fig. 14B, p < .005), ECeG (p < .005) and $\gamma$-band (p < .001) cases in the breakdowns. The modality specificity of this interaction is further illustrated by the fact that the auditory modality in contrast, although showing the same general trends in EEG versus ECeG effects and their interaction with “band”, does not appear to exhibit any significant interactions associated with the rhythmic structure of the stimuli (Fig. 14A). The axial and visual modalities exhibit the interactions to varying degrees.

**Discussion**
In the present study we have collected evoked electrophysiological responses to rhythmic stimuli of four different modalities using a 63 channel cerebellar extension montage and conducted both ERP and time-frequency analyses. We discuss the outcomes of these analyses in two parts respectively.

**ERP FINDINGS**

Source analyses of the grand means of these responses indicated widespread brain activity. A generic 10 dipole regions of interest model was used to conduct a quantitative cross-modal analysis of this differential activity. This analysis revealed some cross-modal commonalities, most notably the presence of centrally focussed cross-modal long latency N1/P2/N2 event related responses, although the N1 was more complex for the axial and vestibular modalities, and the N2 was not prominent for the visual modality. Another was the presence of dominant midline sources in cerebellar and cingulate locations. A third commonality was that the regular rhythm produced higher brain activity than the irregular (at least for the present choice of intervals), and then, in turn, the uncertain rhythms produced higher brain activity compared to regular rhythms. There were, however, also very clear differences. The most notable of these differences was that irrespective of the timing condition, the vestibular stimulation gave rise to higher overall activity. This result is unlikely to be explained by the montage, which was widely distributed, or source locations, which were chosen to be representative of all modalities and then independently optimised for each modality. The primary explanation is likely that while the axial and vestibular modalities exhibited large short-latency responses, which were generated by both cerebellar sources and also short latency frontal activity (Todd et al., 2021a), the auditory and visual modalities showed relatively weak sub-cortical activity and little evidence of associated early frontal activity. Further, sensitivity to beat effects and the interaction of beat differentiation with timing condition were not uniformly distributed across the modalities. When split by modality, the greater vestibular sensitivity to the interaction of timing condition and beat in the ANOVA was confirmed in a post-hoc analysis. This was especially apparent in the vestibular interaction of “timing condition”, “beat” and “wave”, which indicated the emergence of a metrical structure, especially in the N2/P300 ERP wave components with syncopation. The auditory modality by contrast, while showing sensitivity to timing condition, did not yield strong evidence of metrical structure in the post-hoc analysis for the N2/P300. The axial modality did show evidence of the ANOVA interaction of timing condition with beat structure, but to a lesser degree than the vestibular. The vestibular and axial modalities also showed evidence of greater sensitivity to stimulus absence/presence in the SL and N1 waves.

These results we believe provide further evidence to support the vestibular syncopation hypothesis articulated in the introduction. Not only did vestibular stimulation produce greater overall brain activity, but it also showed greater sensitivity to beat and the interaction of beat structure with timing condition, with additional evidence of the emergence of metrical structure (i.e. a S-W-S beat pattern) in the N2/P300 component. Although it is likely that factors other than syncopation contribute to the main effect of timing condition, such as having a longer inter-stimulus interval after missing 3rd beat, we believe that the timing condition and beat interactions effects are genuine syncopation effects for the vestibular modality as both the uncertain conditions produced the S-W-S pattern and thus likely driven by the omission, i.e. the syncopation, on the 2nd strong beat of the anapaest rhythm.

However, a number of issues are raised by these results. It is important to rule out that the modality differences are merely differences in sensation level or properties of peripheral adaptation. If sensation level differences can be ruled out what other possible mechanisms could account for the modality differences? Despite the modality differences there are some important commonalities, notably the dominance of the mid-cingulate source in contributing to the N1/P2 component waves (see Fig. 6). Is this evidence of a universal sensorimotor process? What are the neural generators underlying this process and how do they relate to beat perception and movement induction (as in groove)?

Taking the first of these, prior cVEMP threshold estimates for impulsive head accelerations (Paillard, Kluk, & Todd, 2014; Colebatch, Dennis, Govender, Chen, & Todd, 2014) give estimates of between ~ 45 and ~ 40 dB re 1g, values within an order of magnitude for the detection of whole body motion (Benson, Spencer, & Stott, 1989; Sokya, Robuffo, Beykirch, & Bulthoff, 2011). This would give as estimate the intensity of our vestibular stimulus at about 40 dB above threshold. Based on an earlier estimate of otolithic receptor physiological thresholds to 100 Hz vibration (Todd et al., 2008), the estimate would be significantly higher at about 70 dB above threshold, but this may be an over estimate, however, given the differences in the stimulus frequencies being compared. For the auditory modality, assuming a minimum audible pressure of 15 dB SPL at 2 kHz (Killion, 1977), the sensation level of our auditory stimulus would be somewhere between 75 and 95 dB above threshold. For the axial stimulus, which is likely mediated by proprioreceptors (Todd et al. 2022), the threshold is very close to the tactile somatosensory threshold, and thus once again, the effective sensation level below or close to that of the vestibular modality. Thus, even taking a conservative approach we can rule out sensation level as accounting for the differences between the auditory, axial and vestibular effects on brain activity. Regarding peripheral adaptation, vestibular and auditory modalities are both mediated by mechanoreceptive hair-cells which have common properties of adaptation. For the visual modality estimates of threshold are more difficult, in part due to the visual stimuli involving a complex of factors such as luminosity, spatial and temporal frequency sensitivity. However, although there is for the visual stimuli at least some uncertainty over the modality main effects, modality specific interactions should be independent of modality main effects.
If sensation level or adaptation differences, either due to physical stimulus levels and/or peripheral sensitivity, are unlikely to account for the modality effects that we observed, then these must lie in differences in the neuroanatomy and efficacy of central projections for the different modalities. Both the vestibular and axial stimuli have been demonstrated to produce large short-latency activations of cerebellar and frontal sources, which may mediate rapid associated automated/habitual or voluntary movements related to their preceding brain-stem/spinal reflexes (Todd et al., 2021a). As noted above for the vestibular modality, the associated short-latency frontal activity and consequent long-latency activity are in part likely to be a function of powerful cerebellar-frontal projections from both the vestibulo- and spino-cerebellar regions via the deep cerebellar nuclei (Middleton & Strick, 2001; Watson et al., 2014). It is noteworthy, therefore, that in the post-hoc analysis the predominant source contributing to the vestibular N2/P300 interaction effect was from the left frontal source, with additional contributions from right frontal, mid cingulate, middle and right cerebellar sources.

Within the rhythm perception literature, frontal, cingulate and SMA are strongly implicated particularly for beat based rhythms (Grahn & Rowe, 2012). Todd and Seiss (2004) (also Todd & Lee, 2015b) in their source analysis of the N2 associated with beat induction suggested vestibular dependent frontal sources. Indeed, this was the starting point for their hypothesis that the N2 represents a readiness for action cognitive reflex which may become entrained during beat induction to form a pre-movement negativity. For regular rhythms, according to this proposal, the reflexive/automated readiness response facilitates the rapid transfer of a timed conditioned response to single intervals from an “external” guidance pre-motor-cerebellar pathway and to an entrained premotion “internal” predictive activity in CMA and basal ganglia (Todd & Lee, 2015a; Grahn & Rowe, 2012).

This interpretation is consistent with the well-established role that the cerebellum plays in classical conditioning involving the generation of timed conditioned responses to paired conditional and unconditional stimuli (Marr, 1969; Albus, 1970; Todd et al., 2021b, 2023). In the context of a regular sequence we can consider the cerebellum to play an analogous role in temporal conditioning to classical conditioning whereby each successive stimulus acts as both conditional and unconditional to its following and preceding stimuli. Once the temporally conditioned response had been acquired by the cerebellum, then the transfer of a forward prediction would be available via cerebellar projections to basal ganglia and frontal cortical areas which mediate learned, highly automated/habitual voluntary movements and which are involved in the generation of the movement related potentials (Watson et al., 2014; Chen, Fremont, Arteaga-Bracho, & Khodakhah, 2014; Todd et al., 2021a). This interpretation is also consistent with current views that the cerebellum and basal ganglia should not be considered as the two separate timing/learning systems alluded to in the introduction, but rather “nodes” of a single network (Bostan & Strick, 2018).

Although the initial short-latency cerebellar activations involved in the modulation of the brain-stem/spinal reflexes are likely invariant to posture, the later automated/habitual or voluntary movements and their cingulate generators are almost certainly gated by posture. In the present experiments our participants were standing in order to measure the vestibular and axial reflexes in leg muscles (Todd et al., 2021a). It is an interesting question whether rhythm and beat perception could also be posture dependent. By definition rhythms which have a high groove rating induce a compulsion to move and dance, which require a standing posture thus likely gating on the reticular mechanisms. Further, most rhythmic sounds which are designed for dancing to are loud with a high bass contribution (Todd & Cody, 2000), likely to induce head and body vibration (Sundberg & Verrillo, 1992; Todd, 1993), thus additionally contributing to vestibular and proprioceptive sensory activations (Todd et al., 2008). While research on links between musical rhythm and movement has focused primarily on how acoustic parameters—notably those associated with groove—might directly induce illusory and real body motion, our findings underscore the importance of considering the role of the vestibular system, its trans-cerebellar and trans-reticular projections to basal ganglia and frontal cortex in modulating this link.

**TIME-FREQUENCY FINDINGS**

The time-frequency analyses confirmed, as previously demonstrated (Todd et al., 2018), that when the spectral power of EEG and ECeG are compared, while the EEG shows a typical low-pass characteristic, the ECeG, in contrast shows a band-pass spectral profile with a peak in the α and β-band. The EEG at FCz exhibits the expected distinct α, β and γ-band spectral peaks. In addition, low-frequency δ and θ-band metrically related features in the form of a harmonic series were present corresponding to the stimulus timing characteristics, including a metrical fundamental at the beat frequency of 1.7 Hz. The ECeG, while showing a distinct α peak, is characterised by a dominant broad α-band spectral peak, with less distinct β and γ-band sub-peaks as shoulders on the dominant band. δ- and θ-band metrical structure is also present, although due to a shorter epoch the metrical fundamental could not be fully resolved in this data. Further, due to the dominant high-frequency activity, the total power in the ECeG was significantly higher than that of the EEG. Aspects of these spectral features were present for all four modalities, although the low-frequency peaks, including the α peak, were not easily discernible for the visual modality, while at FCz the metrical fundamental appeared to be most strongly represented in the vestibular and axial modalities, but this would need to be confirmed in a future study. Although this last specific feature was not tested here, there was a highly significant difference in the total power across the modalities with the axial and vestibular standing out from auditory and visual, especially for the high-frequency ECeG. These observations provide evidence to support a sensory-motor view of the spectral representation of rhythmic sequences as both vestibular and axial stimuli strongly evoke frontal activity (Todd, et al., 2021b).
Although effects at a lower significance threshold were obtained for “beat” and “wave”, there was a highly reliable interaction of “beat” and “wave” with “timing condition”, but which turned out to be highly specific for the vestibular modality, for the ECoG and concentrated in the γ-band, with less potent α and β-band contributions. This interaction, we suggest, provides further evidence for the primal role of vestibular for inputs in rhythm perception (Phillips-Silver & Trainor, 2005, 2007, 2008; Trainor et al., 2009). The vestibular modality, especially in its ECoG activation, showed greater power, as well as current activity, and sensitivity to the interaction of beat and timing condition. In contrast, the auditory modality was practically inert to these interactions, as it was for the current analysis. On its own, however, the α and β-band interaction cannot be taken as evidence for the vestibular syncopation hypothesis, unlike the N2/P300 results in the current analysis. It is important in the interpretation to distinguish between high-frequency SL contributions and between timing condition effects which produce the S-W-S pattern in the N1/P2 waves those which produced the S-W-S pattern in the N2/P300 waves. The α-band contribution could be a genuine syncopation effect because the introduction of uncertainty in the presence or absence of a stimulus increased the tendency for a strong-weak relationship for the 1st to 2nd beat, especially for the N1 and P2 waves, i.e. syncopation had the effect of enhancing the metrical structure of the bar in the neural representation compared to the regular unsyncopated stimulus (cf. Lenc et al., 2018). However, it could also be interpreted as an accent effect on the 1st beat of the rhythm due to the presence of longer inter-stimulus intervals preceding the 1st beat in the uncertain conditions. The high-frequency, primarily a SL γ-band effect, was essentially sensitivity to missing stimuli on the 3rd beat, because there was less evidence of metrical structure. A further experiment with more elaborate rhythm alternatives and with longer epochs to capture fully the δ- and θ-band effects, will be needed to attempt to tease out these various contributions. The high sensitivity of vestibular cerebellar γ-band activity could still play an important role in rhythm perception, even if not a syncopation effect, by contributing to a heightening of rhythmic discrimination due to fine amplitude changes. As reviewed in the introduction, γ-band activity has been proposed as an index of anticipatory entrainment to rhythm (Snyder & Large, 2005; Zanto et al., 2006), although our data did not show significant evidence of induced gamma activity on missing stimuli, even though our method for computation of power is equivalent to that for “induced” activity, i.e. by rectification before averaging. This may in part be due to the fact that, as this was a cross-modal study, we did not attempt to apply the analyses to estimates of auditory cortical activity (Zanto et al., 2005), but rather from a frontal amodal common generator.

The contrast between the significant vestibular interaction and the absence of an auditory interaction, also replicated from the source current results, is striking. This raises a number of important issues. We went to some effort to make our auditory stimulus purely auditory without any contamination from other modalities, especially vestibular. In the “real world”, however, pure auditory stimulation is in fact rare to the extent that listeners actively sense the world around them (Schroeder, Wilson, Radman, Scharffman, & Lakatos, 2010). Everyday instances of music listening can be multisensory experiences involving visual, and possibly other, modalities of stimulation in addition to the auditory channel (Russo, Ammirante, & Fels, 2012; Thompson, Graham, & Russo, 2012). Research studies with live musical performances or video recordings have shown that visual perception of performers’ bodies can influence the auditory perception of a range of features, including rhythm and meter (Davidson & Broughton, 2016; Lee, Barrett, Kim, Lim, & Lee, 2015). Under certain circumstances, notably with loud sounds, vibrotactile and vestibular stimulation can contribute to the experience of musical beat (Hove, Martinez, & Stupacher, 2020; Todd & Cody, 2000; Todd et al., 2008). Even in the context of ‘passive’ listening to audio recordings, a vestibular contribution to rhythm perception is possible to the extent that body movement is difficult to inhibit. For example, when participants are instructed to stand still, involuntary micromotion occurs in response to certain types of music, characterised by moderately fast tempo and a clear beat (Gonzalez-Sanchez, Zelechowska, & Jensenius, 2018). Listening through headphones further increases the quantity of such micromotion compared to loudspeakers (Zelechowska, Gonzalez-Sanchez, Laeng, & Jensenius, 2020). Furthermore, body movement induced by music is not merely a receptive process, but can produce feedforward signals that influence auditory perception. Actively moving along with the beat can thus enhance auditory rhythm processing, as reflected in perceptual judgements, behavioral responses, and brain activity (e.g., Chemin, Mouraux, & Nozaradan, 2014; Manning & Schutz, 2013; Su & Pöppel, 2012). The rich literature on such effects of overt movement on auditory perception is, moreover, supplemented by growing evidence that covert motion can produce related effects. For instance, the process of motor preparation—that is, planning a movement without actually executing it—can influence temporal perception in a range of contexts (e.g., Hagura, Kanai, Orgs, & Haggard, 2012; Tomassini, Gori, Baud-Bovy, Sandini, & Morrone, 2014).

If this evidence indicates that the direct auditory contribution to rhythm perception is weak, it raises the further question of what exactly does the auditory system contribute? Quite obviously, the cochlea as a spectacular spectral analyser provides the sensori-neural basis for pitch and thus harmony perception and cognition (Helmholtz, 1885; Langner & Ochs, 2006; Tramo, Cariani, Delgutte, & Braida, 2003), as well as scene analysis (Bregman, 1994), and along with binaural contributions, to spatial analysis (Grothe, Pecka, & McAlpine, 2010; Moore, 1991). How then are these cochlear specific features combined with the rhythmic and motoric aspects of listening in which the vestibular system is dominant? One possible mechanism is a direct one with established hard-wired vestibular projections to auditory cortex. Imaging studies which have used acoustic, galvanic and caloric stimuli agree that the temporal lobe is a significant node within the network vestibular areas. These are supported by electrophysiology which indicates vestibular contributions to auditory evoked potentials when air conducted sounds exceed the vestibular threshold (Todd, Paillard, Kluk, Whittle, & Colebatch, 2014a,b). Another possible mechanism is an indirect “soft-wired” learned one, e.g. by conditioning. By the repeated pairing of auditory conditional stimuli with vestibular unconditional stimuli, especially in a temporally
predictive rhythmic context, auditory stimuli which would otherwise have no or only weak motoric or metrical attributes could acquire them by association, an area of ongoing research (Todd, Govender, & Colebatch, 2021a, 2023a/b).

Conclusion

To conclude, we have provided new evidence for the role of the vestibular modality and associated cerebellar and frontal cerebral cortical activity, in the motion inducing effects of beat and meter, especially with syncopated rhythms. Support for the role of the vestibular system in rhythm perception was observed in neural responses spanning a range of complementary electrophysiological measures, including ERPs, source currents, spectral power of EEG/ECeG, and EEG-ECeG coherence across multiple frequency bands for rhythmic stimuli from different sensory modalities. These results are thus consistent, we suggest, with the prior observations of Philips-Silver and Trainor (2005; 2007; 2008) and also that of Trainor et al. (2009) which demonstrated the primal role of the vestibular system in determining rhythm. In addition, our analyses of N2/P300 effects under syncopation provide further evidence to support the vestibular syncopation hypothesis. Detecting and localizing neural activity attributable to specific subcortical sources was enabled through the use of a novel cerebellar extension to the standard 10–20 EEG montage that allowed non-invasive recording from brain regions that were previously challenging to access reliably. Observing stronger effects for rhythms conveyed by vestibular stimulation than for auditory stimuli raises both methodological issues that inform the investigation of rhythm perception and production, and conceptual issues related to viewing music as a primarily auditory versus a cross-modal phenomenon.

Declarations

Acknowledgements.

The work reported in this paper was supported by a grant from the Australian Research Council DP210100552 “Imaging the human cerebellum during motor learning and timing”.

References


https://doi.org/10.21203/rs.3.rs-2839188/v1


### Tables

**TABLE I: Epochs, latencies and amplitudes of grand mean global field power**

<table>
<thead>
<tr>
<th>Lobe</th>
<th>Modality</th>
<th>SL</th>
<th>N1</th>
<th>P2</th>
<th>N2/P300</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Auditory</strong></td>
<td>Epoch (ms)</td>
<td>7 - 55</td>
<td>55 - 133</td>
<td>133 - 273</td>
<td>273 - 500</td>
</tr>
<tr>
<td>Peak latency (ms)</td>
<td>85</td>
<td>175</td>
<td>360</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amplitude (mV)</td>
<td>1.9</td>
<td>1.5</td>
<td>0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Axial</strong></td>
<td>Epoch (ms)</td>
<td>7 - 74</td>
<td>74 - 126</td>
<td>126 - 253</td>
<td>253 - 500</td>
</tr>
<tr>
<td>Peak latency (ms)</td>
<td>85, 115</td>
<td>180</td>
<td>370</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amplitude (mV)</td>
<td>1.1, 1.2</td>
<td>2.4</td>
<td>1.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Vestibular</strong></td>
<td>Epoch (ms)</td>
<td>7 - 74</td>
<td>74 - 133</td>
<td>133 - 254</td>
<td>254 - 500</td>
</tr>
<tr>
<td>Peak latency (ms)</td>
<td>85, 110</td>
<td>170</td>
<td>390</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amplitude (mV)</td>
<td>2.1, 2.1</td>
<td>3.1</td>
<td>1.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Visual</strong></td>
<td>Epoch (ms)</td>
<td>7 - 77</td>
<td>77 - 150</td>
<td>150 - 270</td>
<td>270 - 500</td>
</tr>
<tr>
<td>Peak latency (ms)</td>
<td>110</td>
<td>200</td>
<td>350</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amplitude (mV)</td>
<td>3.8</td>
<td>2.0</td>
<td>0.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**TABLE II: Summary of sources using a 10 dipole model for auditory stimulation**

<table>
<thead>
<tr>
<th>Location</th>
<th>Side</th>
<th>Area</th>
<th>Weight</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CEREBRAL CORTEX</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal cortex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ant Ins (10mm)</td>
<td>R</td>
<td>13</td>
<td>0.5</td>
<td>29</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>IFG, PrCeG</td>
<td>R</td>
<td>44,45</td>
<td>0.3</td>
<td>49</td>
<td>10</td>
<td>18</td>
</tr>
<tr>
<td>CG</td>
<td>L</td>
<td>24,32</td>
<td>0.2</td>
<td>-13</td>
<td>12</td>
<td>28</td>
</tr>
<tr>
<td>MiFG, CG</td>
<td>R</td>
<td>31,24</td>
<td>0.4</td>
<td>14</td>
<td>-22</td>
<td>34</td>
</tr>
<tr>
<td>Fronto-parietal cortex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG, PoCeG</td>
<td>L+R</td>
<td>23</td>
<td>0.2</td>
<td>-2</td>
<td>-23</td>
<td>27</td>
</tr>
<tr>
<td><strong>Temporal cortex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncus, ITG</td>
<td>L</td>
<td>20</td>
<td>0.2</td>
<td>-32</td>
<td>-2</td>
<td>-33</td>
</tr>
<tr>
<td>STG, MTG, Ins</td>
<td>L</td>
<td>22, 21,13</td>
<td>0.4</td>
<td>-49</td>
<td>-25</td>
<td>-3</td>
</tr>
<tr>
<td>STG, TTG, Ins</td>
<td>L</td>
<td>41,13</td>
<td>0.2</td>
<td>-44</td>
<td>-33</td>
<td>14</td>
</tr>
<tr>
<td>FG</td>
<td>L</td>
<td>37</td>
<td>0.3</td>
<td>-54</td>
<td>-58</td>
<td>-22</td>
</tr>
<tr>
<td><strong>Parieto-temporal cortex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ins, PrCeG / PoCeG, TTG</td>
<td>R</td>
<td>13,43,41</td>
<td>0.6</td>
<td>48</td>
<td>-15</td>
<td>17</td>
</tr>
<tr>
<td>IPL, PoCeG, STG</td>
<td>L</td>
<td>40,2,43</td>
<td>0.7</td>
<td>-62</td>
<td>-24</td>
<td>24</td>
</tr>
<tr>
<td>IPL, Ins, STG</td>
<td>R</td>
<td>40</td>
<td>0.6</td>
<td>54</td>
<td>-29</td>
<td>23</td>
</tr>
<tr>
<td><strong>Parietal cortex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPL, SMG</td>
<td>R</td>
<td>40</td>
<td>0.2</td>
<td>60</td>
<td>-36</td>
<td>36</td>
</tr>
<tr>
<td><strong>SUBCORTEX</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eyes</td>
<td>L</td>
<td>0.2</td>
<td>28</td>
<td>36</td>
<td>-48</td>
<td></td>
</tr>
<tr>
<td><strong>Basal Ganglia</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caudate (10mm)</td>
<td>L</td>
<td>0.3</td>
<td>-25</td>
<td>-21</td>
<td>29</td>
<td></td>
</tr>
<tr>
<td>R</td>
<td>0.2</td>
<td>18</td>
<td>-39</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Brainstem</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medulla</td>
<td>R+L</td>
<td>0.4</td>
<td>1</td>
<td>-34</td>
<td>-55</td>
<td></td>
</tr>
<tr>
<td><strong>Cerebellum</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lobule VIII A</td>
<td>R</td>
<td>0.4</td>
<td>36</td>
<td>-53</td>
<td>-54</td>
<td></td>
</tr>
<tr>
<td>Lobule VIII B</td>
<td>R</td>
<td>0.4</td>
<td>16</td>
<td>-51</td>
<td>-56</td>
<td></td>
</tr>
<tr>
<td>L</td>
<td>0.8</td>
<td>18</td>
<td>-60</td>
<td>-56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crus II/VIIB</td>
<td>R</td>
<td>0.7</td>
<td>-54</td>
<td>-58</td>
<td>-22</td>
<td></td>
</tr>
<tr>
<td>Lobule VI</td>
<td>L</td>
<td>0.3</td>
<td>-18</td>
<td>-74</td>
<td>-21</td>
<td></td>
</tr>
</tbody>
</table>
**Abbreviations:** Anterior insula (Ant Ins), cingulate gyrus (CG), inferior frontal gyrus (IFG), inferior parietal lobule (IPL), inferior temporal gyrus (ITG), insula (Ins), fusiform gyrus (FG), middle frontal gyrus (MiFG), middle occipital gyrus (MOcG), middle temporal gyrus (MTG), postcentral gyrus (PoCeG), precentral gyrus (PrCeG), supramarginal gyrus (SMG), transverse temporal gyrus (TTG).

**TABLE III:** Summary of sources using a 10 dipole model for visual stimulation

<table>
<thead>
<tr>
<th>Location</th>
<th>Side</th>
<th>Area</th>
<th>Weight</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CEREBRAL CORTEX</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontal cortex</td>
<td>R</td>
<td>MeFG, SuCaG, RG, IFG</td>
<td>0.5</td>
<td>10</td>
<td>12</td>
<td>-18</td>
</tr>
<tr>
<td><strong>Temporal cortex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STG, Unc (12 mm)</td>
<td>R</td>
<td>38</td>
<td>0.3</td>
<td>13</td>
<td>9</td>
<td>-50</td>
</tr>
<tr>
<td>STG, Unc, MTG (10 mm)</td>
<td>R</td>
<td>38,20</td>
<td>0.2</td>
<td>29</td>
<td>6</td>
<td>-51</td>
</tr>
<tr>
<td>STG, Ins</td>
<td>R</td>
<td>38,22,13,21</td>
<td>0.2</td>
<td>43</td>
<td>1</td>
<td>-9</td>
</tr>
<tr>
<td>ITG, MTG (10 mm)</td>
<td>R</td>
<td>20</td>
<td>0.2</td>
<td>40</td>
<td>-9</td>
<td>-52</td>
</tr>
<tr>
<td>FG, ITG, PHG</td>
<td>R</td>
<td>20,36</td>
<td>0.3</td>
<td>44</td>
<td>-26</td>
<td>-20</td>
</tr>
<tr>
<td>Ins, TTG</td>
<td>R</td>
<td>13,41</td>
<td>0.3</td>
<td>30</td>
<td>-31</td>
<td>17</td>
</tr>
<tr>
<td>PHG, STG</td>
<td>L</td>
<td>19</td>
<td>0.2</td>
<td>-34</td>
<td>-43</td>
<td>2</td>
</tr>
<tr>
<td><strong>Fronto-parietal cortex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PoCeG, PrCeG, IPL</td>
<td>L</td>
<td>2,3,4,40</td>
<td>0.2</td>
<td>-43</td>
<td>-23</td>
<td>37</td>
</tr>
<tr>
<td><strong>Temporo-parietal cortex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MTG, STG, AG, IPL</td>
<td>R</td>
<td>39</td>
<td>0.2</td>
<td>41</td>
<td>-51</td>
<td>27</td>
</tr>
<tr>
<td><strong>TPO cortex (10 mm)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCun, MTG, STG, PoCeG, Cun</td>
<td>R</td>
<td>31,39,7,18,30</td>
<td>0.3</td>
<td>27</td>
<td>-63</td>
<td>23</td>
</tr>
<tr>
<td><strong>Occipito-temporal cortex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOcG, MTG</td>
<td>R</td>
<td>19,18</td>
<td>0.2</td>
<td>37</td>
<td>-83</td>
<td>16</td>
</tr>
<tr>
<td><strong>Occipital cortex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOcG, MOcG, LG</td>
<td>R</td>
<td>18,19</td>
<td>0.8</td>
<td>33</td>
<td>-83</td>
<td>-1</td>
</tr>
<tr>
<td>MOcG</td>
<td>L</td>
<td>18,19</td>
<td>0.8</td>
<td>-32</td>
<td>-85</td>
<td>7</td>
</tr>
<tr>
<td>Cun, MOcG</td>
<td>R</td>
<td>18,17</td>
<td>1.1</td>
<td>8</td>
<td>-86</td>
<td>12</td>
</tr>
<tr>
<td>Cun, LG</td>
<td>L</td>
<td>17,18</td>
<td>0.9</td>
<td>-5</td>
<td>-86</td>
<td>6</td>
</tr>
<tr>
<td><strong>SUBCORTEX</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basal Ganglia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caudate, CG</td>
<td>L</td>
<td>31</td>
<td>0.2</td>
<td>-17</td>
<td>-37</td>
<td>21</td>
</tr>
<tr>
<td><strong>Cerebellum</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crus I</td>
<td>L</td>
<td>0.3</td>
<td>-47</td>
<td>-48</td>
<td>-32</td>
<td></td>
</tr>
<tr>
<td>Lobule IX (tonsil)</td>
<td>L</td>
<td>0.7</td>
<td>-16</td>
<td>-53</td>
<td>-51</td>
<td></td>
</tr>
<tr>
<td>Lobule VI</td>
<td>L</td>
<td>0.4</td>
<td>-39</td>
<td>-61</td>
<td>-23</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** Angular gyrus (AG), anterior cingulate gyrus (ACG), cingulate gyrus (CG), cuneus (Cun), inferior frontal gyrus (IFG), inferior occipital gyrus (IOcG), inferior parietal lobule (IPL), inferior temporal gyrus (ITG), insula (Ins), lingual gyrus (LG), fusiform gyrus (FG), medial frontal gyrus (MeFG), middle occipital gyrus (MOcG), middle temporal gyrus (MTG), parahippocampal gyrus (PHG), postcentral gyrus (PoCeG), posterior cingulate gyrus (PCG), precentral gyrus (PrCeG), precuneus (PCun), rectal gyrus (RG), subcallosal gyrus (SuCaG), transverse temporal gyrus (TTG), uncus (Unc).

**TABLE IV: 10 dipole regions of interest model for source current measurements**

<table>
<thead>
<tr>
<th>Source</th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>Lobe</th>
<th>Location</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>30</td>
<td>10</td>
<td>Frontal</td>
<td>Ant Cingulate</td>
<td>24,32</td>
</tr>
<tr>
<td>2,3</td>
<td>+50</td>
<td>15</td>
<td>10</td>
<td>Frontal</td>
<td>FFL/Ant Ins</td>
<td>44,45,13</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>-25</td>
<td>35</td>
<td>Frontal</td>
<td>Mid Cingulate</td>
<td>31,23,24</td>
</tr>
<tr>
<td>5,6</td>
<td>+50</td>
<td>-30</td>
<td>15</td>
<td>Temporal/Parietal</td>
<td>STG/IPL/TTG/PostCG</td>
<td>41,40,42</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>-50</td>
<td>-40</td>
<td>Cerebellum</td>
<td>Vermis/Brainstem</td>
<td></td>
</tr>
<tr>
<td>8,9</td>
<td>+30</td>
<td>-60</td>
<td>-55</td>
<td>Cerebellum</td>
<td>Lobule VIII</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>-86</td>
<td>10</td>
<td>Occipital</td>
<td>Cuneus/Ling G</td>
<td>17,18,19</td>
</tr>
</tbody>
</table>

**TABLE V: ANOVA of current sources, combined and broken down by modality.**
### TABLE VI:
ANOVA of EEG/ECeG power over five bands, broken down by modality, band and electrode

<table>
<thead>
<tr>
<th>OMNIBUS</th>
<th>Effect</th>
<th>MOD</th>
<th>BAND</th>
<th>ELE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>df</td>
<td>F</td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>MODALITY (M)</td>
<td>3,12</td>
<td>12.5</td>
<td>&lt;.005</td>
<td>ns</td>
</tr>
<tr>
<td>TIMING COND (TC)</td>
<td>2,8</td>
<td>56.4</td>
<td>&lt;.001</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>BEAT (B)</td>
<td>2,8</td>
<td>6.4</td>
<td>&lt;.053</td>
<td>ns</td>
</tr>
<tr>
<td>SOURCE (S)</td>
<td>9,36</td>
<td>9.1</td>
<td>&lt;.01</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>WAVE (W)</td>
<td>4,16</td>
<td>27.0</td>
<td>&lt;.001</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>

| TC*B      | 4,16  | 37.0 | <.001 | ns   |
| B*W       | 8,32  | 9.9  | =.01  | ns   |
| TC*B*W    | 16,64 | 13.2 | <.001 | ns   |

### MOD interactions

<table>
<thead>
<tr>
<th>MOD interactions</th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>M*TC</td>
<td>6.24</td>
<td>2.6</td>
<td>=.13</td>
<td>ns</td>
</tr>
<tr>
<td>M*B</td>
<td>6.24</td>
<td>6.4</td>
<td>&lt;.05</td>
<td>ns</td>
</tr>
<tr>
<td>M*W</td>
<td>12.48</td>
<td>12.6</td>
<td>&lt;.001</td>
<td>ns</td>
</tr>
<tr>
<td>M<em>TC</em>B</td>
<td>12.48</td>
<td>4.4</td>
<td>&lt;.05</td>
<td>ns</td>
</tr>
<tr>
<td>M<em>TC</em>W</td>
<td>24.96</td>
<td>4.2</td>
<td>&lt;.05</td>
<td>ns</td>
</tr>
<tr>
<td>M<em>TC</em>B*W</td>
<td>48.192</td>
<td>4.9</td>
<td>&lt;.05</td>
<td>ns</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>EEG</th>
<th>ECeG</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;.001</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

| Ba*E | 6,32 | 33.7 | <.001 | <.05 | <.001 | <.001 | <.005 | <.05 | <.005 |
| Be*W | 24.96 | 12.1 | <.01 | <.05 | <.05 | <.05 | <.05 | <.05 | <.05 |
| TC*Be*W | 56,224 | 13.4 | <.005 | <.05 | <.005 | <.05 | <.05 | <.005 | <.05 |
| Ba*Be*W | 24.96 | 5.6  | <.02 | <.05 | <.08 | <.08 | <.08 | <.08 | <.08 |

### Figures
Figure 1

The layout of the custom 10% cerebellar extension montage with select 63 channels indicated (light grey electrodes). The reference (dark grey) and earth electrodes (black) were at Nz and AFz respectively.
A Regular anapaest

ITI = 2400 ms, ISI = 600 & 1200 ms, E = 2100 ms

B Uncertain anapaest (3rd stimulus present)

ITI = 2400 ms, ISI = 600/1200/1800 ms, E = 2100 ms

C Uncertain anapaest (3rd stimulus absent)

ITI = 2400 ms, ISI = 600/1200/1800 ms, E = 2100 ms

D Irregular

ITI/ISI = 800 - 1400 ms, E = 700 ms

Figure 2

An illustration of the stimulus categories and EEG recording epochs. (A) A regular timing condition in the form of an anapaest rhythm with an ITI of 2400 ms, ISI of 600 and 1200 ms and recording epoch of 2100 ms. (B, C) An uncertain timing condition in form of an anapaest rhythm where the 3rd stimulus was randomly absent in 50% of trials. (D) The irregular timing condition which shows equal ITI and ISI intervals from 800 to 1400 ms and a recording epoch of 700 ms. ITI = inter-trial interval, ISI = inter-stimulus interval (ISI) and E = epoch.
Figure 3

Grand means of auditory (A) and visual (B) evoked responses averaged across timing conditions. Auditory evoked longer latency N1/P2/N2 ERP component waves can be observed frontally which are largest at FCz. The visually evoked potential (VEP; P100) can be seen posteriorly over Oz, with the other end of the dipole distributed over the frontal electrodes. For illustrative purposes, only a subset of the 63 electrodes is shown. *The AFz waveform reflects the average of recordings from the AF3 and AF4 locations. Dashed vertical line on traces indicate stimulus onset.
Figure 4

Cross-modal comparison of grand mean recordings to each of the three anapaest conditions at select electrodes: FCz (all four modalities), Iz (auditory, axial and visual modalities) and PO10 (vestibular modality). While all four stimuli show long latency potentials at FCz, the axial (B) and vestibular (C) modalities demonstrate additional short latency potentials at FCz, Iz and PO10 electrodes. For the uncertain anapaest timing condition with 3rd stimulus absent, a P300 response can be observed at FCz for the axial and vestibular modalities, less so for the auditory modality (A) and no such response during visual stimulation (D).
Figure 5

Cross-modal comparison of grand mean recordings collapsed across beats and associated butterfly plots, global field power (GFP) and top meridian maps. Grand means are shown at FCz and Iz for auditory (A), axial (B) and visual (D) modalities, while for vestibular stimulation FCz and PO10 are used. All four modalities produce long-latency N1, P2 and N2 waves. Vertical dashed lines on GFP traces illustrate the boundaries used to define the SL, N1, P2 and N2 epochs (refer to Table 1). Top meridian maps (E-H) for the long latency potentials show the modality specific distributions for the N1 potential, whereas P2/N2 distributions appear amodal.
Figure 6

Cross-modal comparison of dipole orientations and current sources for the generic 10 dipole source model applied to the grand mean for the auditory (A), axial (B), vestibular (C) and visual modalities (D). The axial and vestibular modalities show large short-latency cerebellar and frontal activations while all modalities demonstrate a dominate mid-cingulate source for the long-latency potentials. AC = anterior cingulate, LF = left frontal, RF = right frontal, MC = midcingulate, LT = left temporal, RT = right temporal, Vr = vermis, LCb = left cerebellar, RCb = right cerebellar, Oc = occipital.
Marginal means for the main effects for the current sources responses combined across timing conditions. Comparing the irregular with the regular anapaest rhythm (A), where the irregular condition was treated as “beat zero”, the greatest difference in activity was between the irregular and 1st beat of the regular condition. Comparing the regular with the uncertain condition, overall, the highest activity was during the uncertain present condition, with the uncertain absent condition yielding higher activity than the regular anapaest condition (B). Activity was highest during vestibular stimulation (C). Midline cingulate and cerebellar locations were the most predominant sources (D; red, purple and green bars). Activity increased from baseline levels with the short latency wave producing the highest levels overall (E). Irreg = irregular, Reg = regular, Unc = uncertain, P = present 3rd stimulus, A = absent 3rd stimulus, Aud = auditory, Axial = axial, Vest = vestibular, Vis = visual, BL = baseline, SL = short-latency. *** p < 0.005, ** p < 0.01, * p < 0.05
Marginal means for three-way interactions of “beat” by “wave” by “timing condition” for the current sources responses combined across timing conditions for the auditory (A) and vestibular (B) modalities. The auditory modality shows very little contribution to the beat effects observed in the combined analysis, whereas the vestibular modality is the primary contributor to the overall pattern of beat effects across waves and timing conditions (see Fig 9D).
Figure 9

Marginal means for post-hoc analyses of auditory (left column) versus vestibular (right column) N2/P300 components showing “timing condition” and “beat” main effects and “beat” by “timing condition” interactions. Both modalities showed a main effect of “timing condition” (A & D) on the N2/P300, whereas the main effect of “beat” was significant only for the vestibular (E) but not the auditory modality (B). Both modalities showed a two-way interaction of “beat” by “timing condition” (C & F), although only the vestibular modality exhibited the S-W-S beat pattern in the pairwise comparisons.
Figure 10

Scaleograms (left column) and power spectra (right column) for all four stimulus modalities at FCz over the longer epoch. The auditory (A), axial (B) and vestibular (C) modalities all show α-, β- and γ-band activity, with partial suppression in α-band activity for the visual modality (D). The power spectra also reveals low frequency spectral components with metrical harmonics ($M_0, M_1, M_2$) being more distinct for the axial and vestibular modalities.
Figure 11

Scalograms (A-C, E-G, I-K, M-O) and spectral power at FCz over the shorter epoch. Spectral power is shown averaged for each modality (D, H, L, P) and timing condition (Q, R, S) and reveals a predominately low-pass profile. The metrical harmonics ($M_1$ and $M_2$) showed a dependency on timing condition.
Figure 12

Scalograms (A-C, E-G, I-K, M-O) and spectral power at PO10 over the shorter epoch. Spectral power is shown averaged for each modality (D, H, L, P) and timing condition (Q, R, S). In contrast to the spectral profile demonstrated at FCz, PO10 demonstrated a band-pass profile dominated by a broad high-frequency (100 Hz) peak. An α-peak is clearly present for the auditory, axial and vestibular modalities.
ANOVA effects show higher power in the ECeG (PO10) than EEG (FCz) electrode (A). The higher ECeG power also occurred in the higher frequency bands (B). Overall, power was greater for the axial and vestibular modalities (C), for the 1st beat (D) and during the SL segment (E). The wave by beat interaction (F) shows that the greater power in the 1st beat was mainly confined to the SL, N1 and P2 waves.
Figure 14

ANOVA effects for the three-way wave by beat by timing condition interaction shown for the auditory (A) and vestibular (B) modalities.