Research Article | Open Access

The Use of Audio Stimulation to Affect Sensorimotor Learning

Gregory Ranky[*] and Sergei Adamovich

Department of Biomedical Engineering, New Jersey Institute of Technology, NJ 07102, USA

[*]**Corresponding author:** Gregory Ranky, Department of Biomedical Engineering, New Jersey Institute of Technology, Fenster Hall, University Heights, Newark, NJ 07102, USA, Tel: 9735965268; E-mail: gregoryranky@gmail.com, sergei.adamovich@njit.edu

Abstract

Sensorimotor learning for the hand and fingers can be conducted using both hardware and software components, but the training regime is also important. Using repetitive sequence tapping allows measurement of defined metrics in a controlled, safe environment, and therefore statistical indications for subject improvement. The process of entrainment, when a subject's own movements synchronize to an external signal, has been tested in prior studies for memorization and recognition, but has not been investigated for correlation with sensorimotor learning. This is tested with selected custom isochronic audio tones, combined with sequential finger tapping on a standard computer keyboard. Whilst there were no significant differences between specific frequencies, testing blocks done during tone conditions show subject improvement in reduced mean sequence times compared to pre-stimulation, with no significant change in subsequent post-stimulation blocks.

Keywords: Isochronic; Audio; Tones; Unimanual; Upper extremity; Sensorimotor; Tapping

Introduction

Because human working memory is limited, multiple cultures across the world have independently developed physical and behavioral tools to aid in memorization. Amongst the oldest and most widespread of these is the use of music and rhythm to encode information for future use. As both music and dance contain repeating sound and motion patterns, learning to play or keep time with music requires the practitioner to maintain rhythms by using internal synchronization and body movements, whether in dance or playing instruments.

Entrainment is the matching of brain activity frequencies to external rhythms. This can occur through sensory means using audio, visual or tactile input, or applied means such as tACS (Transcranial Alternating Current Stimulation). Whilst prior work has found that using entrainment in combination with additional tasks has led to changes in recognition and memorization, the use of specific patterns, tones and frequencies and their effects on brain activity and sensorimotor learning however have had comparatively little study.

The brain itself has the ability to change its active wave frequencies to match external audio, visual or tactile rhythms and is defined as entrainment [1]. And it has been found that audiovisual entrainment in humans has a correspondence with the stimulation frequency. One example of this is the presence of entrainment at 2x the stimulation rate [2], with entrainment occurring at 1.33 Hz as measured via Electroencephalography (EEG), when stimulation was presented in the sub-delta band at 0.67 Hz. Another study [3] varied light flash frequencies, and found a positive correlation with cerebral blood flow (CBF) in the striate cortex from 0 to 7.8 Hz, with a decrease in CBF above this frequency and a 20-30% increase in CBF at a 7.8 Hz frequency of stimulation.

In addition to audio and visual stimulation, tACS has been utilized in recent work, on the M1 region of the brains of 15 right-handed subjects [4]. It was discovered that 10 Hz stimulation increased movement variability over 30 min post-stimulation, and 20 Hz stimulation caused movement slowing directly after stimulation. The conclusion drawn from this was that a 10 Hz neural oscillation interferes with inhibitory circuits, therefore increasing movement variability, an undesired result which would decrease subjects' precision.

Most significantly, regional cerebral blood flow has been measured using Positron Emission Tomography (PET) during a pattern-flash visual stimulation at frequencies of 0,1,4,7, and 14 Hz [5].The results showed increase in striate cortex activity at 7 Hz, with a decline at 14 Hz. Likewise, a study centered on a word recognition memory task with elderly participants [6], discovered optimal visual frequencies from 9.5-11 Hz for increased word recognition, with optimal recognition close to 10.2 Hz. Whilst these results point to the effects of optimal frequencies using visual entrainment, there has been no verification for the use of audio only at these same frequencies.

The presence of optimal or default frequencies for body movements as well as brain activity has also been investigated historically. This has led to discoveries such as a resonance for human walking at approximately 120 beats per minute, or 2 Hz, though the discoverers concluded that the biomechanics of the arm and hand may give different frequencies [7]. Other work has found an optimal tapping frequency for single finger tapping of ~600 ms or 1.667 Hz; as this has not been examined for sequences, there may be additional correlations with multiples of this frequency [8]. However, there is a difficulty here due to this being a repeating decimal, and it is unknown if there will be effects on tone creation due to rounding errors.

As the possible frequencies for optimal repetitive upper extremity and finger motion may not necessarily be identical to those for walking, they may span a range of values, or be a multiple of existing optimal frequencies. The discovery and use of optimal frequencies for repetitive movements has implications for such fields as sports medicine, workplace ergonomics and rehabilitation. Movements conducted at an optimal frequency would lead to greater efficiency of movement, fewer repetitive strain injuries, and reduced fatigue.

It is important to note that the frequencies that proved most effective in assisting recognition and memorization may not be optimal for upper extremity sensorimotor activity. Nevertheless, it is important to prioritize subject safety, and focus on finding as direct a correlation as possible between audio tones and upper extremity sensorimotor activity.

In addition to potential effects on human performance, there is also the issue of safety and comfort, whether the workplace is an office, a mine or a factory floor. Workplace noise has been found to be an explanatory factor in fatal industrial accidents between 1990 and 2005 [9], especially if worker communication was involved. In a less obvious example, worker performance and comfort can suffer with the presence of disruptive or unwanted noise; a survey conducted within eight European nations with 7441 participants revealed 'noise' as the variable with the highest association with occupants' comfort [10]. Finally, noise has been found to affect recognition memory, which can lead to diminished performance or accidents by diminishing vigilance [11].

Whether there exists a correlation-positive or negative-between audio tones and upper extremity activity, the result is significant in either case. We are surrounded by a multitude of sounds-both repetitive and random, and if there are any effects on our behavior or physiology, then it is necessary to determine what these are. It is also possible that there are related effects with other subject factors, such as handedness, gender or formal musical training.

In the event of a positive correlation, then the presence of isochronic tones may enhance concentration, improve sensorimotor activity or learning, or reduce fatigue, whilst a negative correlation may do the reverse. If there is no correlation, then the results are significant nonetheless, as it permits the presence of a range of ambient sound frequencies in our working environment. Many of these may be unavoidable or costly to reduce, and so their presence would not detract from the quality of work or living by those in the proximity.

Materials and Methods

A total of 30 subjects were tested for this specific procedure, and each was compensated $20 for their time. All subjects were college students with no history of neurological disorders and were right-hand dominant. Out of the 14 subjects who answered 'Yes' to Formal Musical Training, two replied that their training was vocal only. The remaining 12 had either a mixture of instruments and vocal training or instruments only. All the instruments listed required the use of the fingers of both hands.

Each subject performed a sequence of 3 blocks of continuous typing of the '[a f s d a d f]' sequence with their non-dominant, left hand. During the second block, auditory stimulation of either 4 Hz or 10 Hz was provided. Then the sequence of three blocks was repeated, and the auditory stimulation was once again provided during the second block, with a frequency of either 10 Hz or 4 Hz. The presentation order for the two stimulation frequencies were randomized across subjects. Each 370 s testing block had two recorded files-the keystroke sequence and corresponding time codes-from which MATLAB extracted several metrics: the Accuracy, defined as the total number of correctly performed sequences divided by the total number of performed sequences, then multiplied by 100; the Mean Sequence Time, defined as the time in seconds to perform a correct sequence; the Total # of Error States, a unitless value where each Error State is defined from the

beginning point of an incorrect sequence to the beginning of the next correct sequence.

In order to gain familiarity with the testing sequence, subjects performed a single testing block prior to recording any of the metrics. Subjects were allowed to see the typing sequence to perform before and after this practice block, but at no point after this. They were also instructed to prioritize accuracy before speed. The key sequence utilized involved all four fingers of the subjects' non-dominant hand, instead of the index finger of the dominant hand, and involved a fixed, repeated sequence instead of a single-finger self-paced rhythm. In order to reduce visual distractions, subjects also used their non-dominant hand both to increase challenge, and provide a more easily observed change in performance before and after testing.

The majority of surveyed prior research for unimanual sequential finger tapping focuses on five-digit sequences using the four fingers of the left hand: the left index finger corresponds to '1', the left middle finger to '2', the left ring finger to '3' and the left pinky to '4'. This gives the most common sequence found as '[4-1-3-2-4]'. This sequence has the advantage of providing challenge by avoiding more than 2 adjoining keys in a sequence, and difficulty can be increased further to a 7-digit sequence. In order to avoid two adjoining keys once again both in sequence and for successive sequences, the modified sequence is [4-1-3-2-4-2-1]. To use this on a standard keyboard, and as the majority of subjects are right-handed, the keys used for the non-dominant/left hand were the 'a s d f' location, which is taught as the starting position for the left hand when typing. The resulting sequence, when translated to these keys is '[a f s d a d f]', and the two tones (4 Hz and 10 Hz) were randomized across the subjects using Microsoft Excel's RAND function.

The tones used in this experiment were made using Audacity® software, this allowed the creation of tones where not only the isochronic frequency, but the carrier frequency, duration, and overall shape can be customized for this study. The carrier frequency was chosen to be 256 Hz, which represents the note of middle C in Scientific pitch. This differs from middle C at 261.62 Hz as used by concert orchestras, as 256 Hz is a whole number in the binary system, and allows all the octaves of C (an octave is 50% or 200% of a note's frequency) to remain whole numbers in both binary and decimals down to 1 Hz. Middle C was also chosen because it exists within the average human hearing range, and can also be sung. Unlike the commercially available tone used previously, the resulting custom tone did not have tapering in volume at the beginning or end of the block. Whilst it is possible to include tapering in custom tones, it introduces additional variables to adjust, such as the time from silent to full volume, and the shape of the volume increase. As this is experimental, it is prudent to minimize testing variables until a clearer model can be established.

With the carrier frequency chosen as a constant, the next step was to determine a selection of isochronic frequencies to vary. As shown earlier, prior art gives a number of choices for frequencies to use, and for this round of testing, having frequencies with even spacing between them will allow the recording of changes in a more even distribution. Therefore, the frequencies tested here were 4 and 10 Hz was also included, as it has been used as a visual frequency in prior art, and falls within the specified frequency range. In addition, this is close to a multiple of 6x the listed maximum for human finger tapping, or 10.002 Hz.

In order to ensure subject safety, it is necessary to exclude potential subjects who have a history of seizures, but also to minimize this risk within the testing procedure itself; this is one of the primary reasons why the tones are presented in audio as opposed to visual. The key frequency to avoid is 15 Hz, which is given in prior art as the frequency with the greatest risk of seizures, with the risk decreasing linearly on either side. Further seizure risks not covered in this study include the color red, stripes, and alternating light and dark patterns [12]. Whilst this has been primarily reported in visual stimulation only so far, it is important not to undergo unnecessary risks in experimental work. Specifically, reflex epilepsy can be trigger by environmental stimuli, not only visual, or photosensitive epilepsy, but also audio from music or human voices.

Separate four-way ANOVA was performed on each of the three outcome measures (Accuracy, Mean Sequence Times, Total # of Error States) with two between factors Gender (Male, Female) and Formal Musical Education (Training, No Training), and two repeated measures factors Repetition (First, Second) and Block (Pre, Stimulation, Post).

Subsequently, to investigate the effects of different frequencies of stimulation, data from the conditions where auditory stimulation was present were analyzed. Separate three way repeated measures ANOVAs were performed on each of the three Responses (Accuracy, Mean Sequence Times, Total # of Error States), with two between factors of Gender (Male, Female) and Formal Musical Education (Training, No Training) and the repeated measure factor of Tone Frequency (4 Hz, 10 Hz).

All variance analyses performed on the data used $p < 0.05$ as the probability level to accept statistical significance. Post hoc comparisons using both Bonferroni and the less conservative Tukey multiple comparisons tests showed very similar outcomes, so it was decided to report Bonferroni. For Accuracy and Mean Sequence Times, n=30, for Total # of Error States, n=27, as three Subjects had to be excluded due to technical reasons.

Results

Three separate ANOVAs with two repeated measures (Time (Pre, Stimulation, Post) and Repetition (First, Second)) were used to investigate the effects of auditory stimulation and motor learning on three outcome measures: Mean Sequence Time, Accuracy and Number of Error States.

Both main effects of Time and Repetition on Mean Sequence Time were significant ($F_{(2,52)}=37.94$, $p<0.0001$ and $F_{(1,26)}=54.83$, $p<0.0001$, respectively). Speed of typing increased during the second half of the experiment, with the mean (SD) Sequence Time reduced (Table 1) from 2.51 (0.08 s) during the first three blocks (First Repetition) to 2.19 (0.66 s) during the last three blocks of trials (Second Repetition).

Effect: Repetition	Count	Mean	STD
First	90	2.51	0.74
Second	90	2.19	0.66

Table 1: Means table for sequence time effect: Repetition.

For the factor Time, post hoc comparisons showed that Mean Sequence Time averaged across the two repetitions (Table 2) was significantly shorter in the two Stimulation blocks of trials with auditory stimulation (mean (SD) of 2.27 (0.72) s) than during the preceding Pre blocks of trials without the stimulation (mean (SD) of 2.53 (0.74) s). However, Mean Sequence Time was not different in the Post trial blocks (2.25 (0.67)) when compared to the Stimulation blocks.

Effect: Time	Count	Mean	STD
Pre	60	2.53	0.74
Stimulation	60	2.27	0.72
Post	60	2.25	0.67

Table 2: Means table for sequence time effect: Time.

Finally, there was a significant Repetition by Time interaction ($F_{(2,52)}=6.66$, $p=0.003$, Table 3; Figures 1 and 2). Post hoc comparisons show that the decrease in sequence time in the Stimulation block (when compared to the Pre block) was more pronounced during the first Repetition than during the second Repetition.

Repetition × Time	Count	Mean	STD
First, Pre	30	2.77	0.71
First, Stimulation	30	2.4	0.75
First, Post	30	2.36	0.71
Second, Pre	30	2.3	0.7
Second, Stimulation	30	2.14	0.67
Second, Post	30	2.15	0.62

Table 3: Means table for sequence time effect: Repetition × Time.
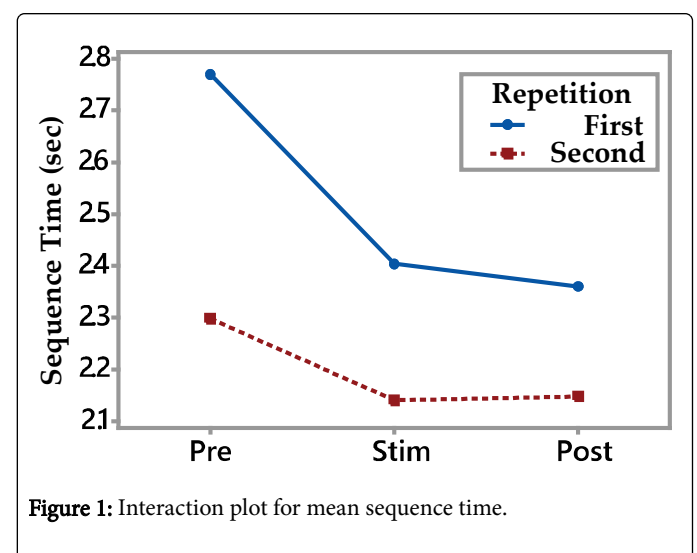

Figure 1: Interaction plot for mean sequence time.

Post hoc analysis using the Bonferroni multiple comparison procedure demonstrated that Mean Sequence Time was shorter in both blocks of trials where auditory stimulation was present when compared to the preceding blocks with no auditory stimulation (blocks 1 and 4, respectively). At the same time, Mean Sequence Time was not different between these two blocks with stimulation and the two

subsequent blocks, as shown in Figure 2, where there is not a significant difference between Stimulation and Post blocks.
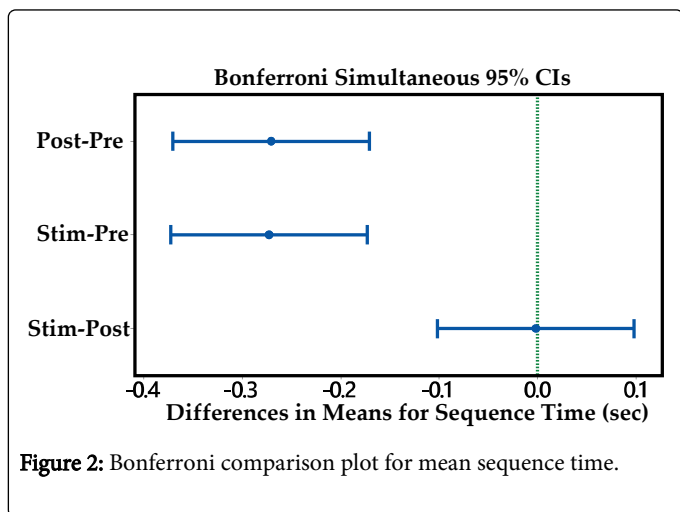

Figure 2: Bonferroni comparison plot for mean sequence time.

There were no significant main or interaction effects on Accuracy except for the Gender by Musical Education interaction effect ($F(1,26)=5.43$, $P=0.03$, Table 4).

Gender × Musical Education	Count	Mean	STD
Female, No	30	94.15	4.74
Female, Yes	36	85.54	17.94
Male, No	72	86.08	11.65
Male, Yes	42	92.24	2.8

Table 4: Means table for accuracy effect: Gender × Music.

The effect of Time on Number of error States was significant ($F(2,46)=5.71$, $p=0.006$) (Table 5). Post hoc comparisons revealed that the number of error states was not different between the Pre and the Stimulation blocks of trials. The same was true for the Stimulation versus Post comparison. However, the difference between the Pre and the Post blocks reached significance, probably because of the overall increased number of typing sequences due to faster typing at the end of the experiment.

Time	Count	Mean	STD
Pre	54	20.15	17.22
Stimulation	54	22.56	14.82
Post	54	25.06	17.56

Table 5: Means table for error states effect: Time.

In a subsequent analysis of only the trials where auditory stimulation was present, we investigated the potential differential effects of stimulation frequency (4 Hz vs. 10 Hz) on the three main responses. The three-way ANOVA with factors Stimulation Frequency, Gender and Formal Musical Training did not reveal any significant main or interaction effects of frequency stimulation (Figure 3).
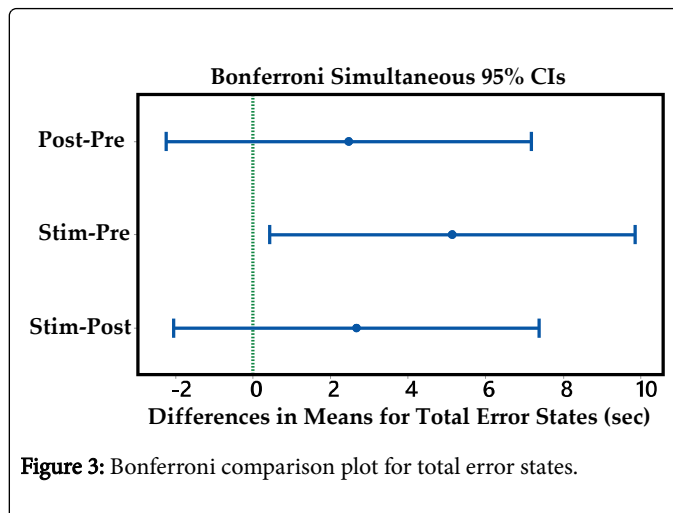

Figure 3: Bonferroni comparison plot for total error states.

Discussion

Although the use of isochronic tones on upper extremity sensorimotor learning and activity were not as pronounced as expected, the results were nonetheless significant.

Whilst Tone Frequency did not have any significant effects on the 3 metrics chosen for this protocol, Tone vs. Pre-stimulation conditions displayed a positive effect for Mean Sequence Times, giving shorter times regardless of gender or Formal Musical Training. The lack of significant increase or decrease in this metric for both Post-stimulation conditions indicates a degree of retention of the effects of applied Tones. Also, whilst there was an overall decrease in Mean Sequence Times from the first to the last block for each subject, the Tone block effects were more pronounced during the first Tone block, regardless of frequency. As for an explanation for the reduction in sequence time during Tone blocks, the cause is unlikely due to the Tone being a distraction, as Post-Tone blocks are not significantly lower or higher than the immediately preceding Tone blocks, and the overall decrease of Mean Sequence Time from Block 1 to Block 6 for each Subject occurs regardless. This supports the explanation that the presence of a Tone enhances a process that is already present in each Subject. A likely explanation is that Tones affect or enhance Subject vigilance, and keep them alert to allow them to acclimate faster to performing a repetitive upper extremity unimanual task.

Despite the lack of significant effects on Accuracy due to Tones, the interaction effect of Formal Musical Training and Gender did reveal within each gender, higher Accuracy and a smaller Standard Deviation for Female Subjects without training, whilst Male Subjects demonstrated lower Accuracy, and a larger Standard Deviation. Though this result alone cannot universally support the assertion that Formal Musical Training assists with repeating sequences and doing so with reduced variation, there is room in possible future studies to examine this further.

The metric of Total Error States was not affected significantly by Tone vs. No-Tone conditions. The only noteworthy finding here is that it is significantly affected by fatigue, with Post-testing blocks on average giving more Error States than Pre-testing, though whether the length or distribution of these Error States within each block is

significantly affected is a topic for future research. Furthermore, whilst all the tones in this experiment utilized sine waves as components, it may be the case that square or triangular waves have different effects on subject metrics if all other tone parameters remain unchanged.

Finally, in examining Tone testing blocks only, the lack of significant effects on any of the three metrics shows that the frequencies chosen for this study are not significantly different, though this does not preclude other frequencies from being significantly different, especially those closer to the maximum human tapping rate.

It is also useful to note that in prior work by Mentis et al., the use of audio tones was presented for longer durations than the 6 min used in this study, such as 10 min or 30 min; therefore, it may be necessary to have longer stimulation duration to achieve more statistically significant effects. The difficulty here is that including the before, during and after conditions for each additional specific frequency adds time spent on a session, and to avoid subject fatigue or acclimation to the sequence it is necessary to avoid a total testing time longer than approximately an hour. Having a subject test on successive days to try different sequences is a possibility, but the effects of acclimation are greater, as is the difficulty of fitting testing sessions to conform to subjects' schedules. There may be combinations of varying the durations of each of the a-b-c conditions such as maintaining the b/during condition for 6 min and reducing the a/before and c/after each to under 6 min, but at this stage it is unclear which combinations to aim for.

Also, as mentioned previously, given the maximum human finger tapping speed of 1.667 Hz, none of the audio frequencies used thus far overlap directly with this tapping range. It may be that for repetitive motor actions, audio frequencies that overlap with motion range have a significantly more pronounced effect on successive motions than on frequency ranges for brain activity.

The use of neurotypical subjects may not display significant effects in all metrics, and instead those with deficits, such as chronic post-stroke subjects, may display greater changes. Future work may also need to include an isochronic tone closer to the measured optimal human finger tapping rate of 1.667 Hz as determined by Keele et al., to determine if entrainment effects occur at a lower frequency range for the fingers than what has been used thus far. Alternatively, using tones with each of the isochronic frequencies chosen from multiples of 1.667 Hz may have noticeable effects.

Though not as strenuous as walking, typing requires integrating audio, visual and tactile information. As this was a self-paced activity, subjects' attention was split between perception of external stimuli and internal rhythm generation [13].

It has been found in prior work that closed-loop auditory feedback on walking with Parkinson's subjects, results in improved walking speed and stride length; and compared to open-loop it has residual effects-suggesting that it could be integrated into existing therapy programs [14]. In contrast, because the activity performed in this study gave tactile feedback in a closed-loop fashion, whilst the auditory Tone stimulation was open-loop, it is plausible that their sequence tapping negated a portion of the effects due to the Tones.

Using custom isochronic tones allows precision over the testing materials in not only frequency but in duration and waveform shape. Combining this with keyboard-derived unimanual sequential finger tapping metrics allows the measurement of significant changes across subjects.

Conclusion

Whilst there were no statistically significant effects of Tone frequency on Accuracy, Mean Sequence Time or Total Error States, Tone conditions did result in shorter Mean Sequence Times compared to Pre-Stimulation conditions.

In summary, it is highly unlikely that there is entrainment occurring during audio tones with isochronic frequencies above the maximum human finger tapping speed, as the pace is simply too fast for the human hand to match and stay in time to. However, if testing is performed in 6 min blocks accompanied by isochronic frequencies above the maximum human finger tapping speed, a self-paced unimanual task will not be affected by tone frequency, but the use of audio tones will result in improved Mean Sequence Time.

Potential future studies can include using one or more isochronic tones below that of maximum human finger tapping rates, or combining these with more complicated upper extremity sensorimotor tasks. In addition, whilst retention was not tested directly during this study, it remains a potential attribute to test for future work. Ultimately, if there exist larger effects under more specific circumstances, then it is necessary to clarify and study these for potential benefits to workplace activity; and the results found here can provide a foundation for future work.

Acknowledgement

References

1. Siever D (2003). Applying audio-visual entrainment technology for attention and learning. Biofeedback 31: 1-15.

2. Gomez-Ramirez M, Kelly SP, Molholm S, Sehatpour P, Schwartz TH, et al. (2011) Oscillatory Sensory Selection Mechanisms during Intersensory Attention to Rhythmic Auditory and Visual Inputs: A Human Electro-Corticographic Investigation. J Neurosci 31: 18556-18567.

3. Fox PT, Raichle ME (1984) Stimulus rate dependence of regional cerebral blood flow in human striate cortex, demonstrated by positron emission tomography. J Neurophysiol 51: 1109-1120.

4. Wach C, Krause V, Moliadze V, Paulus W, Schnitzler A, et al. (2013) Effects of 10 Hz and 20 Hz transcranial alternating current stimulation (tACS) on motor functions and motor cortical excitability. Behav Brain Res 241: 1-6.

5. Mentis MJ, Alexander GE, Grady CL, Horwitz B, Krasuski J, et al. (1997) Frequency Variation of a Pattern-Flash Visual Stimulus during PET Differentially Activates Brain from Striate through Frontal Cortex. Neuroimage 5: 116-128.

6. Williams JH, Ramaswamy D, Oulhaj A (2006) 10 Hz flicker improves recognition memory in older people, Neuroscience 7: 21.

7. Van Noorden L, Moelants D (1999) Resonance in the perception of musical pulse. J New Music Res 28: 43-66.

8. Keele SW, Ivry RI, Pokorny RA (1987) Force control and its relation to timing. J Motor Behav 19: 96-114.

9. Deshaies P, Martin R, Belzile D, Fortier P, Laroche C, et al. (2015) Noise as an explanatory factor in work-related fatality reports. Noise and Health 17: 294-299.

10. Sakellaris IA, Saraga DE, Mandin C, Roda C, Fossati S, et al. (2016) Perceived Indoor Environment and Occupants' Comfort in European "Modern" Office Buildings: The OFFICAIR Study. Int J Environ Res Public Health 13: 5.

11. Molesworth BR, Burgess M, Zhou A (2015) The effects of noise on key workplace skills. J Acoust Soc Am 138: 2054-2061.

12. Fisher RS, Harding G, Erba G, Barkley GL, Wilkins A (2005) Photic- and Pattern-induced Seizures: A Review for the Epilepsy Foundation of America Working Group, Epilepsia, Blackwell Publishing Inc 46: 1426-1441.

13. Hao Q, Ogata T, Ogawa K, Kwon J, Miyake Y (2015) The simultaneous perception of auditory-tactile stimuli in voluntary movement. Front Psychol 6:1429.

14. Baram Y, Aharon-Peretz J, Badarny S, Susel Z, Schlesinger I (2016) Closed-loop auditory feedback for the improvement of gait in patients with Parkinson's disease. J Neurolog Sci 363: 104-106.