

Lack of efficacy of music to improve sleep: A polysomnographic and quantitative EEG analysis

Stanley E. Lazic^{*}, Robert D. Ogilvie¹

Department of Psychology, Brock University, St. Catharines, Ontario, Canada L2S 3A1

Received 12 August 2006; received in revised form 13 October 2006; accepted 13 October 2006

Available online 22 November 2006

Abstract

An increasing number of studies have been examining non-pharmacological methods to improve the quality of sleep, including the use of music and other types of auditory stimulation. While many of these studies have found significant results, they suffer from a combination of subjective self-report measures as the primary outcome, a lack of proper controls, often combine music with some type of relaxation therapy, or do not randomise subjects to control and treatment conditions. It is therefore difficult to assess the efficacy of music to induce or improve sleep. The present study therefore examined the effects of music using standard polysomnographic measures and quantitative analysis of the electroencephalogram, along with subjective ratings of sleep quality. In addition, a tones condition was used to compare any effects of music with the effects of general auditory stimulation. Using a counter-balanced within-subjects design, the music was not significantly better than the tones or control conditions in improving sleep onset latency, sleep efficiency, wake time after sleep onset, or percent slow wave sleep, as determined by objective physiological criteria.

© 2006 Elsevier B.V. All rights reserved.

Keywords: EEG; FFT; Music therapy; Power spectral analysis; PSG; Sleep; SWS

1. Introduction

There are a number of published studies examining the use of music to promote relaxation or induce sleep (Fried, 1990a,b; Mornhinweg and Voignier, 1995; Bonebreak, 1996; Levin, 1998; Johnson, 2003; Tan, 2004; Loewy et al., 2005; Lai and Good, 2005). However, several of these studies use subjective self-report measures as the primary outcome index (Mornhinweg and Voignier, 1995; Tan, 2004; Lai and Good, 2005), which is not ideal as self-reports can be subject to biases and do not always correlate with objective physiological measures of sleep (Armitage et al., 1997; Baker et al., 1999; Chervin and Aldrich, 1999). In addition, since it is not possible to blind subjects to the experimental condition (music vs. control), it is necessary to

control expectation or placebo effects which might influence polysomnographic parameters (Fratello et al., 2005). Some studies have also combined music with relaxation therapy or some other intervention (Fried, 1990a,b; Lai and Good, 2005), making it difficult to draw conclusions on the efficacy of music to facilitate sleep onset, or to improve the duration or quality of sleep. One study which had both a music and a relaxation therapy condition found that relaxation therapy had a greater effect on EEG spectral activity than music (Jacobs and Friedman, 2004). It is therefore not clear if music can improve sleep.

There is an enormous variety of music available, and it is not possible to test the efficacy of each genre or individual recording. The music that was used for this experiment was the “Delta Sleep System”, which was created to be used at bedtime to promote sleep (Thompson, 1999). The manufacturer claims that this music induces sleep by entraining EEG in the delta (0.5–3.5 Hz) range, thus facilitating deep sleep, similar to the way in which occipital frequencies can be entrained with flashing lights in the phenomenon known as photic driving (Freedman, 1963; Stough et al., 2001). This commercially available CD was chosen over others because (1) it was

^{*} Corresponding author. Current address: Centre for Brain Repair, University of Cambridge, Forvie Site, Robinson Way, CB2 2PY, UK. Tel.: +44 1223 331160; fax: +44 1223 331174.

E-mail addresses: stan.lazic@cantab.net (S.E. Lazic), bmogilvie@sympatico.ca (R.D. Ogilvie).

¹ Tel.: +1 613 756 3476.

specifically designed to improve sleep, (2) subjects had no previous exposure and thus possible positive or negative associations to it, and (3) testable claims were made by the manufacturer that overall sleep quality would improve and that slow wave sleep (SWS) would increase.

In addition to the music and control conditions, a tone condition was used, as it was necessary to compare any effects of music with the effects of general auditory stimulation. A previous study found that a greater number of subjects fell asleep in a monotonous auditory stimulation condition (in the form of a 1000 Hz tone at 80 dB every 4 s), compared to a no-stimulation control group (Bohlin, 1971). Similarly, Webb and Agnew (1979) found that sleep onset latencies were shorter in a group that listened to intermittent tones, compared to a silent control condition.

The number of CDs that claim to increase creativity, induce relaxation, reduce stress, promote sleep, etc. is proliferating, and there is a need to empirically assess the efficacy of these approaches. The present study examined music as a non-pharmacological alternative for enhancing sleep quality, using standard polysomnographic measures of sleep quality, and quantitative analysis of the amount of delta power during the sleep onset period and the first SWS period as dependent measures.

2. Materials and methods

2.1. Subjects

Ten female students between the ages of 17 and 24 (mean = 19.9, S.D. = 1.91) participated in this study. They were drawn primarily from a first-year undergraduate psychology class, and all those that volunteered happened to be female. Subjects were excluded from the study (1) if they were taking any prescriptive medication, (2) had a history of epilepsy, (3) had any diagnosed sleep disorders, or (4) if they were hearing impaired.

During the initial interview – which was approximately 2 weeks prior to the over-night recordings – participants were given a tour of the sleep lab and the procedure for over-night monitoring was fully explained. Participants were given a Letter of Information, Informed Consent Form, and the Pittsburgh Sleep Quality Index (PSQI; Buysse et al., 1989), and Personality Assessment Inventory (PAI; Morey, 1991) questionnaires to fill out. A sleep log was also given to the participants to take home in order to obtain information on usual sleep patterns (i.e. bedtime, wake time, etc.) for a 2-week period, which was used to plan the bedtime and wake time for the experimental nights. This protocol was assessed and approved by the Brock University Research Ethics Board.

2.2. Questionnaires

Several questionnaires were used to gather information about the participants. The PSQI is a 19-item questionnaire that assesses sleep quality for the preceding month and it was used to obtain descriptive information about sleep. It yields seven subscales which include subjective sleep quality, sleep onset latency, sleep duration, habitual sleep efficiency, sleep distur-

bance, use of sleep medication, and daytime dysfunction. Each subscale can have a score from zero to three, and the total score can range from zero to twenty-one, with higher scores indicating worse sleep quality. Buysse et al. (1989) obtained 0.85 test–retest reliability for global sleep quality scores and range of 0.65 to 0.84 for the component scores. Validity was established by comparing scores to a physician's diagnoses as well as polysomnographic assessment. The PAI was used to acquire indices of anxiety and depression and is a standardized 344-item questionnaire in multiple choice format (Morey, 1991). Scores are standardised to have a mean fifty and a standard deviation of ten, with higher scores indicating greater anxiety or depression. Test–retest reliability is greater than 0.80 for the various subscales and convergent validity was established by comparing the PAI with clinical judgements and existing measures of personality (Morey, 1991).

Subjects completed a Bedtime Questionnaire before going to sleep each night in the laboratory, and upon waking up in the morning, subjects completed a Morning Questionnaire. These two questionnaires were used to gather information about the day (e.g. exercise, medications taken) and night (e.g. how typical or atypical the night was) in order to document any events or activities that might affect the interpretation of the results for that night. These questionnaires do not have associated reliability and validity measurements and were not formally analysed.

The Stimulus Evaluation Questionnaire (SEQ; adapted from Walker, 1977) was completed in the morning and assessed subjects' response to the music or tones that they had listened to the previous night. This questionnaire used a visual analogue scale (potential scores ranging from zero to eighty-eight) and subjects were required to put a slash through a line that had a dichotomized variable at either end, which included helped me sleep–kept me awake, relaxed–tense, pleasant–unpleasant, comforting–disturbing, depressing–exciting, happy–sad, irritating–soothing, familiar–unfamiliar, and listened attentively–did not pay attention. This questionnaire does not have published reliability or validity values.

2.3. Sleep Lab and polysomnographic data acquisition

The Brock University sleep lab was used for the collection of data. Facilities include two bedrooms, a monitoring room, and a bathroom. Each bedroom is electrically shielded, sound-attenuated, temperature- and light-controlled, and approximately 3 m × 3 m in size. The temperature in the bedrooms ranged from 19.6 °C to 20.9 °C during experimental nights. In order to make subjects feel more comfortable, the bedrooms were equipped with a dresser, mirror, and drapes were hung on the wall. A video camera (which operates in low light) recorded the participants continuously throughout the night, and the information was displayed in the monitoring room. Two-way communication was available via microphones and speakers placed in each room.

Polysomnographic data, including electroencephalogram (EEG), electrooculogram (EOG), electromyogram (EMG) and electrocardiogram (ECG) were collected from each participant for four consecutive nights. Mark-easy 10–20 caps (Optimit,

Tucson, AZ) were used to mark the standard electrode sites, with the following EEG electrode sites used: C3, C4, O1, O2, T3, T4, T5, T6, A1, and A2. Each of these sites was referenced bilaterally (A1+A2), and a ground was placed on the back of the shoulder. The two EMG electrodes were placed over the mentalis muscle, and the EOG electrodes were placed on the outer right and left canthi. Any electrodes with impedances greater than 10 Ω were reapplied. Electrodes placed on the skin were held in place with surgical tape, while electrodes on the top of the head were affixed with collodion glue.

Data were sampled at 200 Hz with a band-pass filter of 0.78 to 30 Hz and a notch filter of 60 Hz. Harmonie software (version 4.0, Stellate Systems, Montreal, Canada) was used to record, score, and analyse the data, which were recorded on the hard drives of two Pentium III 450 Hz computers. PSG records were scored in 30-s epochs according to standard criteria (Rechtschaffen and Kales, 1968) by a trained researcher. The records were scored blindly, without knowledge of which condition or which night was being scored.

2.4. Auditory stimuli

Both the tones and music were recorded onto a cassette player that was located outside of the bedrooms. The music and tones were transmitted into the rooms via a 4-in. Sony speaker placed on a shelf approximately 1 m above the head of the bed. The tones were played at a frequency of 300 Hz for 1 s, followed by silence for 1.5 s. The music was obtained from the first of two CDs of the Delta Sleep System (Thompson, 1999).

The information that accompanied the Delta Sleep System did not specify an optimal volume. Therefore, both the music and tones were played at approximately 15 dB above background levels. Room A had background noise levels of 38–40 dB, and Room B had levels of 30–32 dB. The slightly higher background decibel levels in Room A were due to ventilation noise differences between the rooms. The stimuli were loud enough to be heard clearly, yet they were quiet enough so as not to be intrusive.

Subjects were not given any specific instructions to either attend to or ignore the music and tones, and the subjects were simply told that the purpose of the study was “to examine the effects of music and tones on sleep neurophysiology”. None of the subjects were familiar with the Delta Sleep System prior to taking part in this study.

2.5. Over-night procedure

Subjects were scheduled two at a time for four consecutive nights. The first served as an adaptation night in order to account for a possible first night effect (Agnew et al., 1966). The other three nights served as the experimental nights. Subjects listened to music, tones, or neither, and the conditions were counterbalanced. A within-subjects design with counterbalanced treatments was used in order to control for factors such as sex (Goel et al., 2005), age (Pandi-Perumal et al., 2002), depression (Argyropoulos and Wilson, 2005), anxiety (Papadimitriou and Linkowski, 2005), and other factors which could

differ between groups and thus affect the interpretation of the results. Subjects were asked to arrive between 1 and 1.5 h before their regular bedtime (as indicated by the sleep log), to allow time for preparation for polysomnographic recording.

Once the subjects were in bed, a bio-calibration procedure was conducted. Subjects were asked to blink ten times, look left, look right, look up, look down, cough, swallow and grind their teeth. They were then given instructions to complete the bedtime questionnaire and go to sleep. Once lights were out, the music or tones began playing (on experimental nights) and were turned off after 5 min of continuous sleep, with sleep onset defined as beginning with the first spindle or K-complex. Subjects were then allowed to sleep uninterrupted for the remainder of the night.

2.6. Statistical analysis

To test for differences between experimental nights, a linear mixed-effects model was used (Pinheiro and Bates, 2000; Kristensen and Hansen, 2004). The FFT data were positively skewed and were therefore \log_{10} transformed to normalise the residuals; graphs display the untransformed data. In addition, data for sleep onset latency and wake time after sleep onset (WASO) were positively skewed, and thus \log_{10} transformed. A Pearson correlation was used to test for linear relationships between physiological measurements and questionnaire data, and a paired samples *t*-test (two-tailed) was used to test for differences on the SEQ. Analysis was conducted with the open-



Fig. 1. Results of the stimulus evaluation questionnaire. Subjects rated the music to be significantly more relaxing, comforting, pleasant, and soothing compared to the tones condition. There were no significant differences for the other variables. * $p < 0.05$.

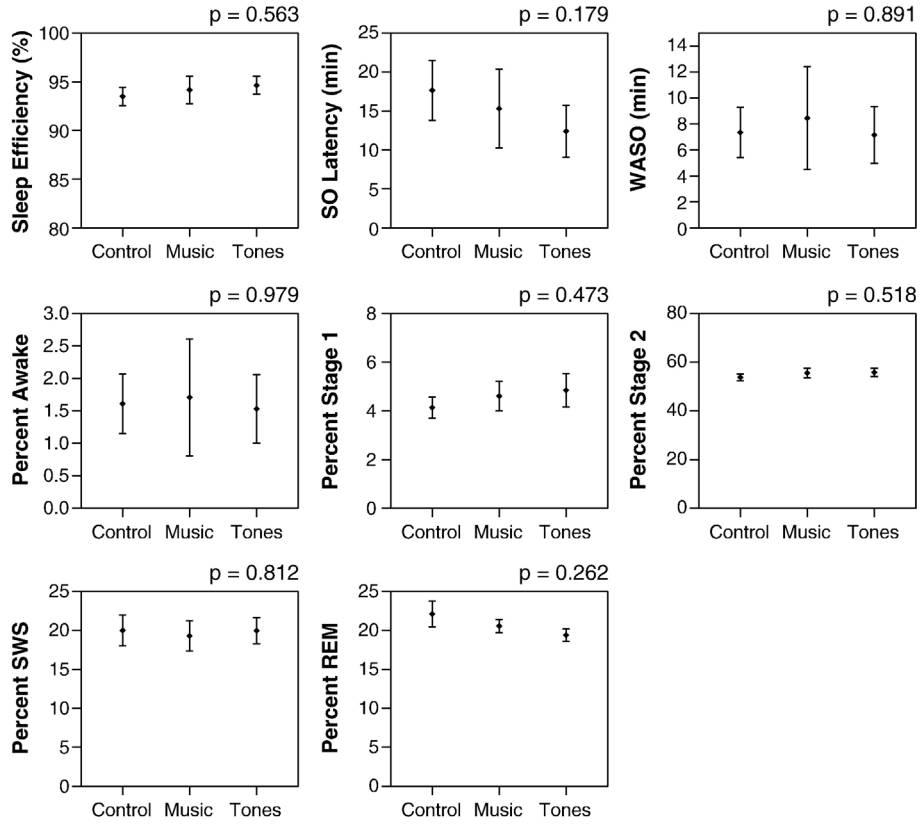


Fig. 2. Polysomnographic data. There were no significant differences between experimental conditions on the PSG variables. *p*-values are for overall differences between conditions and evaluated on (2,18) degrees of freedom. No post-hoc tests were significant, and data represent means ± S.E.M.

source statistical program R (version 2.2.1; Ihaka and Gentleman, 1996; R Development Core Team, 2003), available at www.r-project.org, and for all tests the Type I error rate (α) was set at 0.05.

3. Results

3.1. Pittsburgh Sleep Quality Index

The PSQI global score has a possible range of 0–21 points. Buysse et al. (1989) reported that 88.5% of patients have a global score above 5. Patient groups consisted of (1) depression, (2) disorders of initiating and maintaining sleep, and (3) disorders of excessive somnolence. Scores in the present study ranged from 2 to 11 points, with a global mean of 6.40 and a standard deviation of 2.80. Thus, subjects in the present study either were good sleepers or had scores towards the lower (normal) end of the scale.

3.2. Personality assessment inventory

The PAI was not used as a diagnostic measure to exclude participants. Its primary purpose in this study was to gather data on anxiety and depression, which can influence sleep parameters (reviewed in Papadimitriou and Linkowski, 2005; Argyropoulos and Wilson, 2005). The raw scores have been transformed to *t*-scores in order to compare them with a standardized sample of 1000 adults from a population with a

mean of 50 and a standard deviation of 10 (Morey, 1991). The global anxiety *t*-scores ranged from 40 to 65, with a mean of 49.70 and a standard deviation of 7.39. The global depression *t*-scores ranged from 39 to 65 with a mean of 48.70 and a standard deviation of 9.57. Thus, the mean scores for both

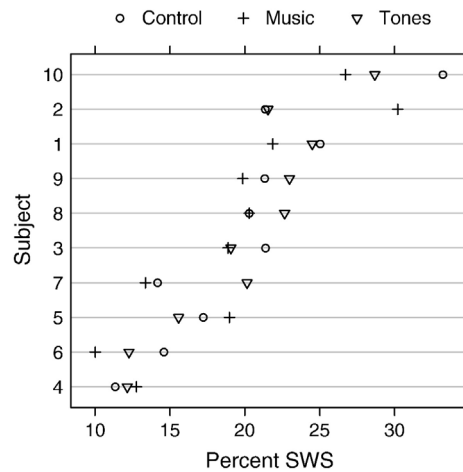


Fig. 3. Percent SWS for each subject by experimental condition; subjects are ordered from bottom to top with increasing mean percent SWS. The amount of time spent in SWS on different experimental nights was fairly consistent for each subject, there was much greater variability between different subjects than within subjects (i.e. between experimental conditions). It is also evident that there is no consistent pattern of individuals having more SWS in the music condition. Results were similar for the other PSG variables.

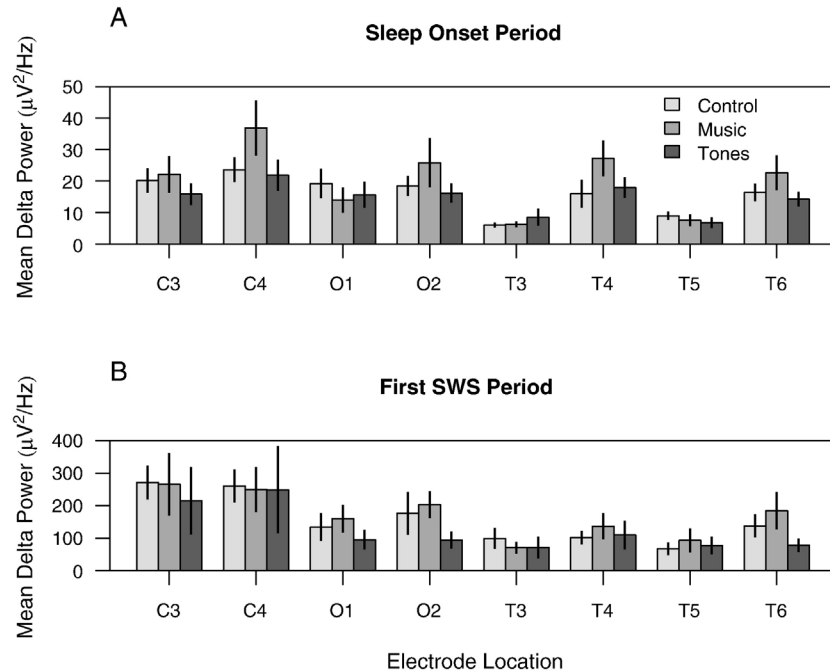


Fig. 4. Power spectral analysis. Amount of delta power at each electrode site between conditions during the sleep onset period and the first slow wave sleep period. There was a significant difference between conditions at T4 ($p=0.048$) during the sleep onset period and at T6 ($p=0.033$) during the first SWS period. Pair-wise post-hoc tests with a Bonferroni correction at these sites were not significant. Data represent means \pm S.E.M.

anxiety and depression in this study are quite close to the population means for these variables.

3.3. Stimulus evaluation questionnaire

The music differed significantly from the tones on four of the nine scales that were rated by subjects (Fig. 1). The music was rated as being more relaxing than the tones ($t_{(8)}=3.13$, $p=0.012$), more pleasant ($t_{(8)}=4.59$, $p=0.001$), more soothing ($t_{(8)}=4.27$, $p<0.001$), and more comforting ($t_{(8)}=2.77$, $p=0.022$). Participants also tended to think that the music helped them fall asleep faster compared to the tones, but this was not significant ($t_{(8)}=2.11$, $p=0.064$). There was no significant difference between conditions on the other four scales (happy, familiar, exciting, attentive listening).

3.4. Polysomnographic data

There were no significant differences in the primary outcome variables between the control, music and tones conditions (Fig. 2). There was no improvement in sleep efficiency, (defined as [time asleep/time in bed] \times 100; $F_{(2,18)}=0.59$, $p=0.563$), sleep onset latency ($F_{(2,18)}=1.89$, $p=0.179$), wake time after sleep onset ($F_{(2,18)}=0.12$, $p=0.891$), percent awake ($F_{(2,18)}=0.02$, $p=0.979$), or percent SWS, which was defined as stages 3 and 4 combined ($F_{(2,18)}=0.21$, $p=0.812$). Furthermore, there were no significant differences in percent stage 1 ($F_{(2,18)}=0.78$, $p=0.473$), percent stage 2 ($F_{(2,18)}=0.68$, $p=0.518$), or percent REM ($F_{(2,18)}=1.44$, $p=0.264$). Effect sizes and 95% confidence intervals were also calculated for differences between the control and music conditions using

Hedges's g (Kline, 2004) for sleep efficiency ($g=0.18$, 95% CI: $-0.26, 0.61$), sleep onset latency ($g=0.38$, 95% CI: $-0.08, 0.84$) and percent slow wave sleep ($g=0.11$, 95% CI: $-0.31, 0.54$). All three confidence intervals contained zero.

Subjects had very similar polysomnographic parameters, irrespective of experimental condition, with much greater variability between subjects than within subjects. For example, 79% of the variance for percent SWS was between different subjects, while only 21% was between experimental nights (i.e. within subjects). This is most clearly displayed by plotting each data point separately by subject and condition, and the data for percent SWS are shown (Fig. 3).

3.5. Power spectral (FFT) analysis

Two sample sections of digitized data were taken from each experimental night and analysed separately. The first was a 4-min section that consisted of all EEG data 2 min before and 2 min after the first spindle or K-complex. This was the sleep onset section, and the music or tones were playing continuously throughout this period. This section was used to test the hypothesis that the music can entrain EEG in the delta frequency.

The second section was comprised of 5 min of EEG data starting at the first SWS period. By this time, the stimulation had ceased. This section was examined to see if there were any lingering effects of the music on EEG, that is, would there be an elevated amount of delta power in the music condition after the music had stopped playing?

Any artefacts (such as body movements) within these intervals were eliminated manually and were not included in the analysis. The spectral record length was set to 5.12 s. Within

each spectral record, two 2.56-s FFT analyses were performed and averaged together. A Hanning tapering window with 50% overlap was used.

A linear mixed-effects model was used to test for differences in spectral power in the delta bandwidth (0.78–3.5 Hz) across the different nights, at each electrode site separately. The music condition tended to have greater delta power at electrode sites in the right hemisphere during the sleep onset period, but this was only significant at T4 ($p=0.048$). During the first SWS period there was only one site that had a significant difference between conditions (T6; $p=0.033$; Fig. 4), and this was partly due to the tones condition having lower values. Pair-wise post-hoc tests between conditions using a Bonferroni correction were not significant at either T4 during the sleep onset period or T6 during the first SWS period. In addition, there was no significant difference between conditions during the sleep onset period or first SWS period when collapsed across electrode site. Results of the statistical tests are displayed in Table 1.

3.6. Correlation of anxiety and depression with sleep onset latency

The relationship between depression and anxiety subscales on the Personality Assessment Inventory and PSG data was examined. Sleep onset latencies were averaged for the three experimental nights for each subject, and then correlated with anxiety. There was a moderate correlation between these variables, but it was not significant ($r=0.543$, $p=0.105$, 95% CI= -0.132 , 0.874). In addition, the raw anxiety subscale scores (physiological, cognitive, and affective) did not correlate significantly with average sleep onset latency. There was, however, a significant direct correlation between depression and sleep onset latency averaged across the three nights ($r=0.681$, $p=.030$, 95% CI= 0.089 , 0.917), with higher scores of depression associated with longer sleep onset latencies.

In order to test if those with higher anxiety benefited more from listening to music than those with lower anxiety, the difference in sleep onset latency between the control and music condition was calculated by subtracting the music condition

latencies from the control condition latencies. Therefore, a positive value indicates a shorter sleep onset latency in the music condition, and thus an improvement over the control condition. This was then correlated with the anxiety score. There was virtually no correlation between change in sleep onset latency and anxiety ($r=-0.099$; $p=0.785$; 95% CI= -0.686 , 0.566). In addition, the change in sleep onset latency was correlated with depression scores on the PAI in order to determine if those with greater depression benefited more from listening to music than those with lower depression. Similar to the anxiety data, there was little evidence for a relationship between the variables ($r=0.044$; $p=0.903$; 95% CI= -0.602 , 0.656).

4. Discussion

Subjects rated the music to be significantly more relaxing, comforting, pleasant, and soothing compared to the tones condition. In addition, subjects tended to think that the music helped them fall asleep faster compared to the tones, but the opposite tendency was actually observed from the PSG measures, where participants fell asleep 3 min sooner while listening to the tones compared to the music (this was not significant). This underlies the need for objective physiological measures of sleep parameters when testing the efficacy of music therapy. If only the subjective measures were used in this study, the results could be interpreted as providing some support for the efficacy of the music.

Based on the PSG data however, there is no support for the hypothesis that music can induce or improve the quality of sleep. Subjects in the music condition did not have a shorter sleep onset latency, greater sleep efficiency, or less WASO. There were only ten subjects and thus the power of the statistical tests was low, however, the means were quite similar across conditions, for example, the mean sleep efficiency for the control, music, and tones conditions was 93.5%, 94.2%, and 94.6%, respectively. In addition, the (Hedges's g) effect sizes were small, based on Cohen's (1988) convention (small effect size=0.2, medium=0.5, large=0.8). The effect size is "the degree to which the phenomenon exists" (Cohen, 1988, p. 4) and in this case, it refers to the difference between the means of the control and music condition, divided by the pooled within-condition standard deviation. Based on these results, it would appear that the experimental manipulation had little effect on overall sleep architecture. Similar non-significant results were obtained in another study which used PSG measures to determine if classical Indian Karnatic music could induce sleep (Gitanjali, 1998). In addition, another study using PSG measures which compared a back massage condition with a music/muscle relaxation/mental imagery and a control condition, in critically ill patients, did not find a significant improvement in sleep efficiency—the primary outcome measure (Richards, 1998). Therefore, this study and two others using PSG measures have not found significant effects of music. In addition, tones did not seem to improve sleep, unlike a previous study which showed a decrease in sleep onset latencies (Webb and Agnew, 1979). However, this study tested sleep onset latencies in the

Table 1
Results of FFT analyses for delta power

Electrode site	Sleep onset period	First SWS period
C3	$F=0.74$, $p=0.491$	$F=2.18$, $p=0.142$
C4	$F=1.38$, $p=0.277$	$F=2.00$, $p=0.164$
O1	$F=0.77$, $p=0.477$	$F=0.64$, $p=0.540$
O2	$F=0.30$, $p=0.743$	$F=3.34$, $p=0.059$
T3	$F=0.71$, $p=0.506$	$F=0.93$, $p=0.414$
T4	$F=3.61$, $p=0.048$	$F=1.48$, $p=0.255$
T5	$F=1.00$, $p=0.368$	$F=0.23$, $p=0.800$
T6	$F=1.10$, $p=0.354$	$F=4.14$, $p=0.033$
Average	$F=0.63$, $p=0.544$	$F=2.18$, $p=0.142$

F - and p -values are displayed during the sleep onset period and first slow wave sleep period at each electrode site, and for all sites combined. F -tests evaluated on (2, 18) degrees of freedom and significant values are in bold. Bonferroni's post-hoc tests were not significant for T4 during the SOP and T6 during the first SWS period.

morning, after a full night of sleep – under a condition of “artificial insomnia” – and therefore the experimental design of that study differed substantially from the present one.

Non-significant results between the music, tones and control conditions were obtained from the FFT analysis on fourteen of sixteen tests (without a correction for experiment-wise error rate). During the sleep onset period, the mean delta power was highest in the music condition for six of the eight EEG sites, and there was a tendency for the music condition to have the highest delta power in the right hemisphere (sites C4, O2, T4, T6), with T4 being statistically significant. This may represent the possibility of an effect of music (but not tones) on EEG entrainment, and might indicate that effects on delta activity are more likely to be seen while the music is playing. Therefore, music may have had an effect on the PSG parameters if it was played continuously throughout the night, but further studies are required to demonstrate this. During the first SWS period, the only significant difference in delta power was at T6, but this was due to the tones condition having lower values and therefore does not support the hypothesis that music increases delta activity.

Numerous studies have reported using music to induce relaxation or improve sleep (Fried, 1990a,b; Mornhinweg and Voignier, 1995; Bonebreak, 1996; Levin, 1998; Johnson, 2003; Tan, 2004; Loewy et al., 2005; Lai and Good, 2005), and thus the current study would seem at odds with some previous research. However, most of these studies used participants' subjective self-reports as outcomes, and self-reports can be subject to bias. Indeed, in the present study, subjects thought that the music helped them fall asleep better than the tones, but sleep onset latencies tended to be shorter in the tones condition. Furthermore, it is not possible to blind subjects to the experimental condition and therefore one must take into account any placebo effect, which can influence both subjective and PSG sleep parameters (Fratello et al., 2005). In addition, some studies used relaxation therapy combined with music (e.g. Lai and Good, 2005), which makes it impossible to determine the effect of the music alone. Yet these limitations and caveats are rarely mentioned when such studies are reproduced in media reports (Campbell, 2006).

This study used young females with depression and anxiety scores in the normal range, and low scores on the PSQI. Therefore the results may not generalise to older adults, males, or clinical populations, in particular those that suffer from insomnia. However, some previous studies have used subjects without sleep disorders to test the efficacy of music (Johnson, 2003; Lai and Good, 2005), making this study suitably comparable. These results also do not rule out the efficacy of other types of music, and this study did not test whether music is effective when used to block out obtrusive noise such as traffic. In addition, it is possible that subjects may have needed to listen to the music for several days or weeks for an effect to be detected. Nevertheless, under the experimental conditions used in this study, it is clear that music from the Delta Sleep System did not have a significant effect on the standard polysomnographic measures of sleep quality. Replications with a bigger sample size, using in-home PSG, clinical populations, and more diverse subjects might produce different results.

In summary, there is little evidence that music has a beneficial effect on sleep parameters in normal young subjects, although the efficacy of music to improve sleep in clinical populations remains to be determined. Complementary and alternative therapies constitute a large part of medical expenditure in Western countries, with an estimated £1.6 billion spent annually in the UK (Ernst and White, 2000) and \$27.0 billion in out-of-pocket expenditures in the US (Eisenberg et al., 1998). The use of music to induce sleep is one such alternative approach, and it is assumed by the general population that music might be useful for this purpose (Urponen et al., 1988). Future studies must use polysomnographic measures of sleep parameters if they are to convincingly demonstrate the efficacy of such treatments.

Acknowledgements

S.E.L. was supported by a Natural Sciences and Engineering Research Council (NSERC) Undergraduate Student Research Award. This project was supported by an NSERC grant to R.D.O.

References

- Agnew, A.W., Webb, W.B., Williams, R.L., 1966. The first night effect: an EEG study. *Psychophysiology* 2, 263–266.
- Argyropoulos, S.V., Wilson, S.J., 2005. Sleep disturbances in depression and the effects of antidepressants. *Int. Rev. Psychiatry* 7, 237–245.
- Armitage, R., Trivedi, M., Hoffmann, R., Rush, A.J., 1997. Relationship between objective and subjective sleep measures in depressed patients and healthy controls. *Depress. Anxiety* 5, 97–102.
- Baker, F.C., Maloney, S., Driver, H.S., 1999. A comparison of subjective estimates of sleep with objective polysomnographic data in healthy men and women. *J. Psychosom. Res.* 47, 335–341.
- Bohlin, G., 1971. Monotonous stimulation, sleep onset and habituation of the orienting reaction. *Electroencephalogr. Clin. Neurophysiol.* 31, 593–601.
- Bonebreak, K.J., 1996. A sound way to induce relaxation and natural sleep—a safe alternative to sedation. *Am. J. Electroneurodiagn. Technol.* 36, 264–268.
- Buysse, D.J., Reynolds, C.F., Monk, T.H., Berman, S.R., Kupfer, D.J., 1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry Res.* 28, 193–213.
- Campbell, G., 2006. 30th anniversary commentary on Lai H.-L. and Good M. (2005) Music improves sleep quality in older adults. *J. Adv. Nurs.* 49 (3), 234–244. *J. Adv. Nurs.* 53, 145–146.
- Chervin, R.D., Aldrich, M.S., 1999. The Epworth Sleepiness Scale may not reflect objective measures of sleepiness or sleep apnea. *Neurology* 52, 125–131.
- Cohen, J., 1988. *Statistical Power for the Behavioural Sciences*, 2nd ed. Lawrence Erlbaum Associates, New Jersey, pp. 25–26.
- Eisenberg, D.M., Davis, R.B., Ettner, S.L., Appel, S., Wilkey, S., Van Rompay, M., Kessler, R.C., 1998. Trends in alternative medicine use in the United States, 1990–1997: results of a follow-up national survey. *JAMA* 280, 1569–1575.
- Ernst, E., White, A., 2000. The BBC survey of complementary medicine use in the UK. *Complement. Ther. Med.* 8, 32–36.
- Fratello, F., Curcio, G., Ferrara, M., Marzano, C., Couyoumdjian, A., Petrillo, G., Bertini, M., De Gennaro, L., 2005. Can an inert sleeping pill affect sleep? Effects on polysomnographic, behavioral and subjective measures. *Psychopharmacology* 181, 761–770.
- Freedman, N.L., 1963. Bilateral differences in the human occipital electroencephalogram with unilateral photic driving. *Science* 142, 598–599.
- Fried, R., 1990a. Integrating music in breathing training and relaxation: I. Background, rational, and relevant elements. *Biofeedback Self-Regul.* 15, 161–169.

- Fried, R., 1990b. Integrating music in breathing training and relaxation: II. Applications. *Biofeedback Self-Regul.* 15, 171–177.
- Gitanjali, B., 1998. Effect of the Karnatic music raga “Neelambari” on sleep architecture. *Indian J. Physiol. Pharmacol.* 42, 119–122.
- Goel, N., Kim, H., Lao, R.P., 2005. Gender differences in polysomnographic sleep in young healthy sleepers. *Chronobiol. Int.* 22, 905–915.
- Ihaka, R., Gentleman, R., 1996. R: a language for data analysis and graphics. *J. Comput. Graph. Stat.* 5, 299–314.
- Jacobs, G.D., Friedman, R., 2004. EEG spectral analysis of relaxation techniques. *Appl. Psychophysiol. Biofeedback* 29, 245–254.
- Johnson, J.E., 2003. The use of music to promote sleep in older women. *J. Commun. Health Nurs.* 20, 27–35.
- Kline, R.B., 2004. *Beyond Significance Testing: Reforming Data Analysis Methods in Behavioral Research*. American Psychological Association, Washington, pp. 104–107.
- Kristensen, M., Hansen, T., 2004. Statistical analyses of repeated measures in physiological research: a tutorial. *Adv. Physiol. Educ.* 28, 2–14.
- Lai, H.-L., Good, M., 2005. Music improves sleep quality in older adults. *J. Adv. Nurs.* 49, 234–244.
- Levin, Y.A., 1998. “Brain Music” in the treatment of patients with insomnia. *Neurosci. Behav. Physiol.* 28, 330–335.
- Loewy, J., Hallan, C., Friedman, E., Martinez, C., 2005. Sleep/Sedation in children undergoing EEG testing: a comparison of chloral hydrate and music therapy. *J. Perianesth. Nurs.* 20, 323–331.
- Morey, L.C., 1991. *Personality Assessment Inventory Professional Manual*. Psychological Assessment Resources Inc, Odessa, FL.
- Mornhinweg, G.C., Voignier, R.R., 1995. Music for sleep disturbance in the elderly. *J. Holist. Nurs.* 13, 248–254.
- Pandi-Perumal, S.R., Seils, L.K., Kayumov, L., Ralph, M.R., Lowe, A., Moller, H., Swaab, D.F., 2002. Senescence, sleep, and circadian rhythms. *Ageing Res. Rev.* 1, 559–604.
- Papadimitriou, G.N., Linkowski, P., 2005. Sleep disturbance in anxiety disorders. *Int. Rev. Psychiatry* 17, 229–236.
- Pinheiro, J.C., Bates, D.M., 2000. *Mixed-Effects Models in S and S-PLUS*. Springer, London.
- R Development Core Team, 2003. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
- Rechtschaffen, A., Kales, A., 1968. *A Manual of Standardized Terminology, Techniques and Scoring System of Sleep Stages of Human Subjects*. US Government Printing Office, Washington, DC.
- Richards, K.C., 1998. Effect of a back massage and relaxation intervention on sleep in critically ill patients. *Am. J. Crit. Care* 7, 288–299.
- Stough, C., Donaldson, C., Scarlata, B., Ciorciari, J., 2001. Psychophysiological correlates of the NEO PI-R openness, agreeableness and conscientiousness: preliminary results. *Int. J. Psychophysiol.* 41, 87–91.
- Tan, L.P., 2004. The effects of background music on quality of sleep in elementary school children. *J. Music Ther.* 41, 128–150.
- Thompson, J., 1999. *Delta Sleep System [CD]*. The Relaxation Company, New York.
- Urponen, H., Vuori, I., Hasan, J., Partinen, M., 1988. Self-evaluations of factors promoting and disturbing sleep: an epidemiological survey in Finland. *Soc. Sci. Med.* 26, 443–450.
- Walker, J.L., 1977. Subjective reactions to music and brainwave rhythms. *Physiol. Psychol.* 5, 483–489.
- Webb, W.B., Agnew, H.W., 1979. Sleep onset facilitation by tones. *Sleep* 1, 281–286.