Neurobiology of Learning and Memory 171 (2020) 107206

Contents lists available at ScienceDirect

Neurobiology of Learning and Memory

journal homepage: www.elsevier.com/locate/ynlme

Classical music, educational learning, and slow wave sleep: A targeted memory reactivation experiment

Chenlu Gao[a], Paul Fillmore[b], Michael K. Scullin[a,*]

[a] Baylor University, Department of Psychology and Neuroscience, Waco, TX, United States
[b] Baylor University, Department of Communication Sciences and Disorders, Waco, TX, United States

ARTICLE INFO

Keywords:
STEM learning
Achievement gap
Sleep disparity
Knowledge transfer
Memory consolidation

ABSTRACT

Poor sleep in college students compromises the memory consolidation processes necessary to retain course materials. A solution may lie in targeting reactivation of memories during sleep (TMR). Fifty undergraduate students completed a college-level microeconomics lecture (mathematics-based) while listening to distinctive classical music (Chopin, Beethoven, and Vivaldi). After they fell asleep, we re-played the classical music songs (TMR) or a control noise during slow wave sleep. Relative to the control condition, the TMR condition showed an 18% improvement for knowledge transfer items that measured concept integration ($d = 0.63$), increasing the probability of "passing" the test with a grade of 70 or above ($OR = 4.68$, 95%CI: 1.21, 18.04). The benefits of TMR did not extend to a 9-month follow-up test when performance dropped to floor levels, demonstrating that long-term-forgetting curves are largely resistant to experimentally-consolidated memories. Spectral analyses revealed greater frontal theta activity during slow wave sleep in the TMR condition than the control condition ($d = 0.87$), and greater frontal theta activity across conditions was associated with protection against long-term-forgetting at the next-day and 9-month follow-up tests ($rs = 0.42$), at least in female students. Thus, students can leverage instrumental music—which they already commonly pair with studying—to help prepare for academic tests, an approach that may promote course success and persistence.

1. Introduction

The pervasiveness of poor sleep in students is well chronicled. Sixty percent of college students are habitually poor sleepers, and students sleep fewer than the recommended 7 h on 50–65% of nights (Hirshkowitz et al., 2015; Lund, Reider, Whiting, & Prichard, 2010; Scullin, 2019). Such statistics should be of wide concern to students, teachers, and administrators because insufficient sleep compromises immune functioning, exacerbates stress reactivity, and impairs numerous cognitive functions. Without sleep, students suffer from impairments to attention (Doran, Van Dongen, & Dinges, 2001), creativity (King, Daunis, Tami, & Scullin, 2017), and memory consolidation (Rasch & Born, 2013).

Sleep restriction is detrimental to people of all demographic groups. However, there is longstanding evidence that women are more likely to suffer from insomnia than men (Zhang & Wing, 2006). More importantly, there is emerging evidence that females are more vulnerable than males to the consequences of sleep loss (Goldstein-Piekarski et al., 2018; Gao, Terlizzese, & Scullin, 2019; Prather, Epel, Cohen, Neylan, & Whooley, 2013; but see Okano, Kaczmarzyk, Dave, Gabrieli, & Grossman, 2019). In rats, 72 h of paradoxical sleep deprivation significantly impaired spatial learning and short-term memory in female rats, but not male rats (Hajali, Sheibani, Esmaeili-Mahani, & Shabani, 2012). In humans, Rångtell et al. (2019) found that one night of total sleep deprivation impaired working memory performance in women, but not in men. Gender-related sleep disparities are particularly provocative when considered within the broader educational context of achievement gaps for females in science, technology, engineering, and mathematics (STEM) disciplines (especially prevalent in mathematics-based disciplines; Ballard & Johnson, 2005; Fryer & Levitt, 2010).

To combat sleep-loss-related cognitive impairments for all students, some educators have implemented sleep education programs (Hershner & O'Brien, 2018). Unfortunately, education programs tend to only impact students' knowledge/awareness of sleep, rather than motivate

Abbreviations: TMR, targeted memory reactivation; STEM, science, technology, engineering, and mathematics; SWS, slow wave sleep; REM, rapid eye movement; PSG, polysomnography; RAPM, Raven's Advanced Progressive Matrices; MEQ, Morningness-Eveningness Questionnaire; PSQI, Pittsburgh Sleep Quality Index; ESS, Epworth Sleepiness Scale; WASO, wake after sleep onset

* Corresponding author at: Baylor University, One Bear Place 97334, Waco, TX 76798, United States.
E-mail address: Michael_Scullin@Baylor.edu (M.K. Scullin).

https://doi.org/10.1016/j.nlm.2020.107206
Received 14 October 2019; Received in revised form 20 January 2020; Accepted 2 March 2020
Available online 04 March 2020
1074-7427/

them to change their sleep habits/behaviors (van Rijn et al., 2019; Wing et al., 2015). Another approach is to take a system-wide effort to delay school start times (e.g., Wahlstrom, 2000). Doing so improves sleep duration and academic outcomes, but system-wide changes are often resisted by politicians, administrators, and other stakeholders (American Academy of Pediatrics Adolescent Sleep Working Group, 2014). A third approach is to incentivize individual students to go to bed earlier (Scullin, 2019), but this approach is also limited because sleeping longer does not necessarily change sleep microarchitecture or increase the probability that classroom/textbook memories will be consolidated.

A cross-disciplinary challenge for educational, cognitive, and sleep sciences is to devise methods that increase sleep-dependent memory consolidation of educational concepts in students. Declarative memories, such as those learned in the classroom, are theorized to initially be encoded into a temporary store in the hippocampus (Born & Wilhelm, 2012), and during slow-wave sleep (SWS), the memories are reactivated and redistributed to more permanent stores in the neocortex (for an alternative view, see Yonelinas, Ranganath, Ekstrom, & Wiltgen, 2019). Critically, memory consolidation is theorized to be a *selective* process (Cairney, Durrant, Hulleman, & Lewis, 2014; Saletin, Goldstein, & Walker, 2011). Only the daily experiences that are perceived as most important, emotionally-salient, or future-relevant are those that are *spontaneously* reactivated and consolidated (Bennion, Payne, & Kensinger, 2015). Therein lies the problem for students. Though educators might perceive that classroom learning is highly important, classroom/textbook memories compete against everyday social, entertainment, and extracurricular memories.

A potential solution is to bias which daily experiences an individual consolidates during SWS via targeted memory reactivation (TMR; Rudoy, Voss, Westerberg, & Paller, 2009). In a seminal experiment, Rasch, Büchel, Gais, and Born (2007) had participants learn spatial card locations while smelling a rose odor. Subsequently, researchers presented the same odor (or control) during SWS, rapid eye movement (REM) sleep, or wakefulness prior to testing memory again the next day. Only odor re-exposure during SWS significantly improved memory consolidation. This TMR finding has now been replicated in more than 30 published studies (for review, see Schouten, Pereira, Tops, & Louzada, 2017), most commonly using laboratory-based learning materials (e.g., cat picture location) paired with discrete sound auditory stimuli (e.g., cat's meow). Such experiments have been useful in identifying that TMR can trigger changes in SWS theta power, spindles, or slow oscillations en route to enhancing retention (Cellini & Capuozzo, 2018; Schreiner & Rasch, 2014).

Despite TMR's robustness in the laboratory, there are three issues that currently limit its translational value to addressing educational challenges. The first issue is that student success in higher education is often determined by conceptual learning and integration. Such learning processes are not well-captured by the research tradition of having participants learn word lists, images, or spatial locations. This distinction between rote learning and conceptual learning is quite critical, as underscored by recent findings in the "testing effect" literature (Karpicke & Roediger, 2008). The testing effect literature has shown that retrieving a memory while awake reliably improves rote learning (item memory), but that simple memory retrieval while awake does not normally benefit conceptual learning or integration (Miyatsu, Nguyen, & McDaniel, 2018; Pan & Rickard, 2018). Thus, if TMR is simply and only reactivating item memories, then it should not benefit conceptual learning.

A second issue for translating TMR from the laboratory to educational settings is that most students study in dormitory or library settings. In such settings, it is doubtful that most students would attempt to pair their studying with odors or to attempt to match exact study content to exact discrete sounds. A more translatable approach would be to capitalize on existing habits of listening to music while studying, a "sensory-study pairing" habit shared by more than half of college students (Danhauer et al., 2009). Listening to *lyrical* music might impede studying, but empirical work shows that listening to instrumental *classical* music does not harm encoding (Jäncke, Brügger, Brummer, Scherrer, & Alahmadi, 2014) or affect nighttime sleep (Harmat, Takács, & Bodizs, 2008). Thus, if instrumental music can be demonstrated as an effective TMR tool, then instrumental music would be an optimal candidate for broadly implementing TMR into educational settings.

A third issue for understanding TMR's translational value concerns whether TMR effects are acute or sustained. Whereas most students are primarily concerned with strategies that immediately enhance performance (i.e. acute effects; Miyatsu et al., 2018), most educators are interested in techniques that prevent the steep forgetting rates that occur across a semester or academic year (i.e. sustained effects; Conway, Cohen, & Stanhope, 1991; Ebbinghaus, 1885). TMR studies generally only test memory the next morning (a few studies used a 1-week delay; e.g., Hu et al., 2015). Thus, there is a need for data on whether TMR buffers against typical long-term-forgetting curves, or conversely, if TMR should only be applied the night before select tests.

The current work was a double-blind, placebo-controlled study in which college students took a virtual lecture on microeconomics while listening to classical music. Later that night, when participants entered stable SWS, they were re-exposed to the classical music (or a white noise control). The next morning, as well as approximately 9-months later (akin to the length of one academic year), participants took a microeconomics test that included concepts they were explicitly trained to solve as well as problems that required novel conceptual integration (hereafter, *knowledge transfer*).

One remaining note is that the current study considered gender differences in TMR, as motivated by two literatures. First, a recent meta-analysis found divergent trends of the benefits of TMR across males and females (Hu, Cheng, Chiu, & Paller, 2020). Second, research at the intersection of music arts and brain sciences (Cheever et al., 2018) has indicated that females process music more efficiently than males, with better association and recognition of familiar music (Fancourt, Burton, & Williamon, 2016; Feizpour, Parkington, & Mansouri, 2018; Koelsch, Maess, Grossmann, & Friederici, 2003; Miles, Miranda, & Ullman, 2016). Musical cues might therefore be especially strong retrieval cues for females. Based on these literatures, we supplemented our primary analyses with gender-stratified analyses to investigate whether classical-music TMR affected males and females similarly.

2. Materials and methods

2.1. Participants

Fifty college students between the ages of 18 and 33 (M_{age} = 21.16, SD_{age} = 2.77, 70% female) were recruited via campus flyers to complete a two-night study protocol consisting of educational tasks and polysomnography procedures. Participants were not eligible if they had taken a college-level Economics course, had a history of psychiatric, neurological, or sleep disorders, or were younger than 18 years old. The Baylor University Institutional Review Board approved this study and all participants provided written informed consent.

2.2. Material and measurements

2.2.1. Educational materials

The educational learning task resembled course material from an undergraduate microeconomics class, with the virtual lecture being developed based on topics from a leading undergraduate textbook (Baumol & Blinder, 2007). During the learning phase, participants navigated through a self-paced, computer-based interactive lecture for up to 30 min. The lecture included a section on supply and a section on demand in a counterbalanced order. As participants navigated through the lecture, they learned to identify and calculate changes in supply or demand in a given scenario, followed by 6 practice questions (Gao et al., 2019).

During the initial testing phase, participants completed a microeconomics test with 18 trained questions (9 supply and 9 demand) and 9 knowledge-transfer questions. The 18 supply/demand questions were similar to lecture example questions where participants needed to shift only the supply curve or only the demand curve to calculate the correct answer. The 9 knowledge-transfer questions required integration of supply and demand concepts, which was not explicitly taught in the lecture. Participants needed to shift both the supply curve and the demand curve and identify that the correct answer was the intersection of the two curves. All questions on the test were open-ended questions scored using a rubric.

To evaluate metacognition (i.e., self-awareness of learning), participants reported from 0% to 100% their predictions for learning (start of lecture sections) and postdictions for learning (conclusion of supply and demand sections). They repeated these prediction and postdiction estimates before and after taking the next-day test and the follow-up test. Table S1 summarizes the meta-cognitive assessment outcomes, which were consistent in pattern with the objective test outcomes.

2.2.2. Questionnaires

Participants kept a sleep diary for a week prior to the first session (Carney et al., 2012). During the study, participants completed questionnaires and intelligence scales. Questionnaires and tests that were relevant to sleep and cognitive functioning included the Pittsburgh Sleep Quality Index (Buysse, Reynolds, Monk, Berman, & Kupfer, 1989), Epworth Sleepiness Scale (Johns, 1991), Morningness-Eveningness Questionnaire (Horne & Östberg, 1976), a vocabulary test, and the Raven's Advanced Progressive Matrices test (Raven, 1938). Participant responses are summarized in Table 1.

2.2.3. Polysomnography

We used the Grass Comet XL Plus system to objectively record sleep in a sound-attenuated sleep laboratory. For the experimental night, we recorded sleep from F3, F4, FZ, C3, C4, CZ, P3, P4, PZ, T3, T4, T5, T6, O1, O2, and OZ sites (200 Hz), left and right EOG, mentalis EMG, forehead cerebral oximetry, and fingertip pulse oximetry. A certified polysomnography technician scored the sleep stages in 30 s epochs, and according to AASM guidelines (Iber, Ancoli-Israel, Chesson, & Quan, 2007).

2.2.4. Spectral power analysis

Spectral analysis characterizes the power of EEG waves and has been used to investigate the mechanisms of TMR (Schouten et al., 2017). We used BrainVision Analyzer 2.0 software to conduct spectral analyses. First, trained research personnel visually inspected each epoch and excluded epochs containing artifacts. Second, we filtered EEG data with high- and low-pass cutoffs of 0.3 Hz and 35 Hz. We then modified the sampling rate to 128 Hz. Next, we segmented each epoch into four-second segments with 50% overlap. We applied a symmetric Hanning window and performed Fast Fourier Transformation with 0.25 Hz resolution. Last, we computed spectral power values for each sleep stage for the relevant frequency ranges: 0.5–1 Hz for slow oscillation; 1–4 Hz for delta; 4–8 Hz for theta; 8–12 Hz for alpha; 12–16 Hz for sigma; and 16–32 Hz for beta. We analyzed spectral power in the slow oscillation, delta, and theta bands averaged across all frontal sites during SWS, averaged across all sites during SWS, and averaged across all sites during all sleep epochs. For visual comparison across groups, we normalized power across the whole night within each frequency band to highlight the relative spatial power distribution across the head.

2.2.5. Spindle detection

Using Matlab 9.0, we implemented a previously-validated, wavelet-based algorithm shown to have high agreement with expert visual identification (Wamsley et al., 2012; Warby et al., 2014). First, we modified the sampling rate to 100 Hz. Second, we performed a time–-frequency transformation on the EEG data using Morlet wavelet. Third, we classified spindles as EEG events that occurred in the 10–16 Hz frequency range, that exceeded 4.5 times the mean amplitude of all artifact-free epochs, and that lasted between 300 and 3000 ms (Scullin et al., 2019).

2.3. Procedure

Participants completed an overnight polysomnography-adaptation visit prior to returning for the experimental night. During the experimental night, research personnel set the laboratory environment brightness to 45 lx to simulate evening studying with a desk lamp. The room temperature was set to 68 degrees Fahrenheit (20 degrees Celsius), with blankets provided as needed. Participants arrived at approximately 8:45 pm. Both TMR and control condition participants completed the economics virtual lecture while listening to background classical music played from the computer. The music was played to all participants in loop at 40 dB, which is the volume equivalent of soft background noise in a library setting. Decibel level was confirmed each evening prior to participant arrival using an Extech 407735 Digital

Table 1
Demographic Characteristics and Learning Outcomes.

	TMR (n = 19)	Control (n = 22)	Comparison between TMR and control groups (p-value)
Age	20.89 (2.18)	21.05 (2.08)	0.82
Gender (female)	68.42%	77.27%	0.52
Race/Ethnicity (White)	63.16%	63.64%	0.98
Participants with music experience	73.68%	81.82%	0.75
Fluid intelligence (RAPM) (%)	65.12 (14.36)	59.85 (12.36)	0.22
Lecture Duration (min)	16.18 (3.99)	17.80 (5.79)	0.31
Vocabulary test (%)	72.89 (8.22)	70.45 (9.69)	0.40
MEQ score	47.05 (10.53)	49.32 (6.43)	0.42
PSQI	5.22 (2.44)	6.61 (2.69)	0.10
ESS	9.53 (4.64)	10.14 (3.71)	0.64
Sleep diary total sleep time (min)	426.20 (44.44)	443.71 (56.68)	0.30
Habitual bedtime (PSQI)	11:42 pm	11:31 pm	0.56
Difference between habitual bedtime and experimental night bedtime (min)	87.05 (56.23)	70.82 (55.70)	0.37
Completed the 9-month follow-up session	63.16%	68.18%	0.74

Note. Data presented as percentage or mean (standard deviation).
RAPM = Raven's Advanced Progressive Matrices.
MEQ = Morningness-Eveningness Questionnaire.
PSQI = Pittsburgh Sleep Quality Index.
ESS = Epworth Sleepiness Scale.

Sound Meter.

The classical music consisted of three pieces: Moonlight Sonata 1st movement (duration: 6:04; Beethoven, 1802), Spring movement 1: Allegro (duration: 3:13; Vivaldi, 1725), and Nocturne in E-Flat Major, Op.9, No.2 (duration: 4:28; Chopin, 2005). These pieces were selected for their distinctive melodies, because previous TMR work found that sensory stimuli must be distinctive to become associated with learning materials (cf. ocean wave sensory stimuli; Donohue & Spencer, 2011). Using Audacity software (version 2.1.3; https://audacityteam.org/), we digitally smoothed the volume and applied fade-in and fade-out to each song to remove sudden volume fluctuations (Sanchez & Bootzin, 1985). Music and economics learning materials are publicly available at https://osf.io/ems7z/. After completing the lecture, trained personnel applied polysomnography (PSG) electrodes while participants completed questionnaires. Lights out was at approximately 10:30 pm.

To minimize expectation effects and other potential sources of bias, the participants and experimenters were masked to the experimental conditions. Condition assignments were determined via blocked randomization methodology (block sizes of 2 and 4). Once participants reached their first bout of stable slow wave sleep, the overnight technician opened a sealed envelope to determine whether to play the classical music or white noise (matched in decibel level). Classical music or white noise was played through speakers at the bedside (Bose Companion 2 Series III). The overnight technician was not involved in administering any lecture or test materials.

Each classical music piece was played once (total duration = 13.75 min); and white noise was played to the control group for the same duration. If a participant showed an EEG arousal, the technician paused the music/white noise until after the participant returned to stable SWS. Music/white noise started to play 33.81 (SD = 29.66) minutes after sleep onset in the TMR condition and 34.24 (SD = 31.26) minutes after sleep onset in the control condition, which did not differ between conditions, $t(39) = 0.05$, $p = .96$. Music/white noise presentation time also did not differ across males (M = 32.72, SD = 34.08) and females (M = 34.53, SD = 29.19), $t(39) = 0.17$, $p = .87$.

The next day, lights on was at approximately 7:30 am, with the exact timing determined by spontaneous awakenings or by waiting until the participant reached stage 1 sleep. Participants reported whether they heard any sounds during the night and then started the microeconomics test at approximately 7:45 am. They were given up to 45 min to complete the test. Approximately nine months after the initial session (M_{TMR} = 263.58, SD_{TMR} = 84.88 days; M_{Control} = 292.93, SD_{Control} = 116.37 days; $t(25) = 0.73$, $p = .47$), 27 participants completed a second microeconomics test that included different test items, but was similar in difficulty and structure (i.e., consisting of 9 supply and 9 demand questions that participants were trained to solve and 9 knowledge-transfer questions that required application of learned concepts; all questions were open-ended). Fourteen participants did not complete the follow-up test because they had enrolled in economics courses (n = 2), were not available to participate (n = 4), or did not respond to requests for follow-up (n = 8). The probability of completing the follow-up session did not differ across TMR and control conditions (12 TMR participants, $\chi^2(1) = 0.11$, $p = .74$) or across gender groups (20 female participants, $\chi^2(1) = 0.03$, $p = .86$).

All microeconomics questions were scored independently by two raters who were blinded to the experimental conditions. Discrepancies were resolved by discussion between raters. Overall performance was calculated as the percentage of correctly answered questions.

2.4. Statistical analysis

Participants were excluded from statistical analyses if there was a protocol deviation (n = 5 had the music/control noise played prematurely during light sleep; n = 1 had the speakers turned off during the initial lecture). Two other participants did not complete the study and one wrote responses that were ungradable by independent, blinded experimenters. Exclusions did not differ across conditions ($\chi^2(1) = 1.22$, $p = .27$).

We conducted statistical analyses using SPSS (version 25) and JASP (version 0.10.2.0) software. We used Pearson's correlations to investigate associations between test performance and only a small group of sleep variables that were previously implicated in TMR studies (frontal theta, delta, spindle, and slow oscillation activity; see Cellini & Capuozzo, 2018). We conducted t-tests and ANOVAs to compare test performance between the TMR and control conditions, first collapsed across genders and then stratified by males and females (following the approach taken by Gao et al., 2019). EEG data were analyzed and presented in aggregate form (averaged across relevant channels), and displayed as individual channels for visualization of spatial patterns. We conducted additional Bayesian analyses to illustrate the effects on EEG data. Though all TMR effects were hypothesized to occur in one direction (TMR > Control), we report the statistics as two-tailed tests with results of $p \leq 0.05$ considered to be statistically significant. Bayesian analyses results with Bayes factors greater than 3 were considered as moderate evidence that the data favored the alternative hypothesis (Dienes, 2014; Lee & Wagenmakers, 2014).

3. Results

3.1. Baseline Measures

Table 1 presents the demographic characteristics of the sample. The TMR and control conditions were similar in age, gender, race/ethnicity, music-related experience (i.e., instrument playing and choir singing experience), fluid intelligence, chronotype, sleep quality in the past month, daytime sleepiness, and total sleep time from the sleep diary (all $ps > 0.05$). Table 2 shows that PSG sleep architecture variables and

Table 2
Polysomnography Measures.

	TMR (n = 19)	Control (n = 22)	Comparison between TMR and control groups (p-value)
Stage 1 sleep (min)	25.34 (16.61)	23.29 (14.62)	0.68
Stage 2 sleep (min)	289.11 (37.42)	282.96 (40.99)	0.62
Slow wave sleep (min)	88.53 (27.14)	92.09 (28.50)	0.69
REM sleep (min)	106.32 (31.39)	111.36 (32.42)	0.62
Total sleep time (min)	509.28 (33.97)	509.87 (37.79)	0.96
Sleep latency (min)	11.50 (10.35)	13.86 (12.25)	0.51
Sleep efficiency (%)	92.14 (5.09)	93.02 (5.35)	0.59
WASO (min)	32.66 (26.74)	24.57 (25.75)	0.33
Number of wake epochs when auditory stimuli were presented	1.16 (2.67)	0.14 (0.35)	0.12
Number of arousals when auditory stimuli were presented	2.47 (4.51)	1.09 (1.74)	0.19
Reported hearing sounds at night	36.84%	36.36%	0.98

Note. Data presented as percentage or mean (standard deviation).
WASO = Wake After Sleep Onset.

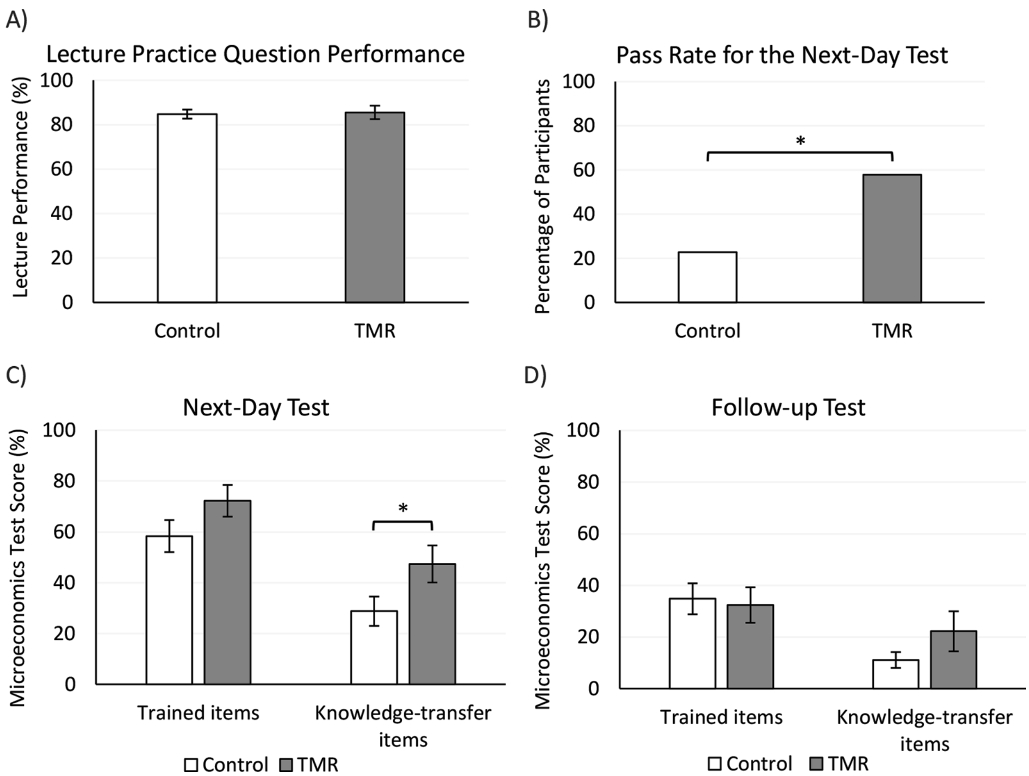

Fig. 1. (a) Performance on lecture practice questions was not different between TMR and control groups; (b) Proportion of participants who passed the test with a total score ≥70% was significantly higher in the TMR group; TMR benefited performance on (c) the next-day microeconomics test, but performance on (d) the follow-up microeconomics test dropped to floor levels. Error bars: standard error. * $p \leq 0.05$.

reports of hearing sounds during the night did not differ significantly between TMR and control groups (all $ps > 0.10$). The TMR and control groups took equally long to complete the microeconomics lecture ($t(39) = 1.03$, $p = .31$; Table 1), and they showed equal levels of initial learning on practice questions during the lecture, $t(39) = 0.19$, $p = .85$ (Fig. 1a; Creery, Oudiette, Antony, & Paller, 2015). Male and female students spent similar amounts of time on the lecture ($t(39) = 0.70$, $p = .49$) and showed equal levels of initial learning ($t(39) = 0.29$, $p = .78$).

3.2. Next-day test

At most colleges in the United States, a grade of 70 is considered the cutoff for passing a test, and thus we used this criterion to define a passing score, representative of successful learning. As shown in Fig. 1b, TMR increased the likelihood of "passing" the microeconomics test (57.89% of participants) relative to the control condition (22.73% of participants; Odds Ratio = 4.68, 95%CI: 1.21, 18.04, $\chi^2(1) = 5.30$, $p = .02$). Specifically, this was driven by the TMR group showing higher performance than the control group on the challenging, integrative knowledge-transfer questions, by 18.58 percentage points ($t(39) = 2.03$, $p = .050$, $d = 0.63$; Fig. 1c). Fig. 2a illustrates how the benefit of TMR on knowledge transfer questions seemed to be more pronounced in female participants (Fig. 2a; $t(28) = 3.05$, $p = .005$, $d = 1.13$) than in male participants ($t(9) = 0.40$, $p = .70$, $d = 0.24$), though the direct comparison across genders was only marginally significant, $F(1,37) = 3.46$, $p = .07$, partial $\eta^2 = 0.09$. Performance on basic/trained questions showed similar patterns but was less sensitive to condition ($t(39) = 1.56$, $p = .13$, $d = 0.50$; Fig. 2b). By the 9-month follow-up visit, test performance dropped to floor levels in both the TMR and control conditions (Fig. 1d: basic/trained questions: $t(25) = 0.27$, $p = .79$, $d = 0.10$; knowledge-transfer questions: $t(25) = 1.34$, $p = .20$, $d = 0.56$), with no evidence for a gender by condition interaction, $F(1,23) = 0.18$, $p = .68$, partial $\eta^2 = 0.01$.

3.3. Spectral analysis results

Table 3 shows that TMR affected sleep microarchitecture. TMR specifically increased frontal theta activity averaged across SWS epochs throughout the night, $t(39) = 2.67$, $p = .01$, $d = 0.87$ (the next largest effect was for spindle density, $t(39) = 1.89$, $p = .07$, $d = 0.55$). To restrict the potential inflation in family-wise error rates for conducting multiple null hypothesis significance tests, we complemented our analyses with Bayesian analyses. Note that because corrections for family-wise type 1 error rates are adjustments of p-values, and the Bayesian analytical approach does not depend on this limitation of p-values, corrections for family-wise type I error rates are unnecessary in Bayesian analyses (Bender & Lange, 2001). Table 3 illustrates that Bayesian analyses yielded similar findings.

Gender stratified analyses indicated similar trends in both gender groups for theta activity during SWS at frontal sites (Fig. 3a; Females: $t(28) = 1.78$, $p = .09$, $d = 0.69$; Males: $t(9) = 1.99$, $p = .08$, $d = 1.14$; Gender by Condition interaction: $F(1,37) = 2.07$, $p = .16$, partial $\eta^2 = 0.05$). Next, we tested whether frontal theta activity during SWS predicted test scores, with the scatterplots shown in Fig. 4. Although frontal theta activity was not associated with levels of learning during the lecture in females ($r(28) = -0.01$, $p = .95$), Fig. 3b shows that greater frontal theta activity significantly predicted higher performance on knowledge transfer questions on the next-day test ($r(28) = 0.42$, 95% CI: 0.07, 0.68, $p = .02$). We observed a similar effect size at the 9-month follow-up test, albeit with marginal significance given the sample size attrition ($r(18) = 0.42$, 95% CI: -0.02, 0.73, $p = .06$, Fig. 4). There were too few male participants to conduct definitive correlational tests in this subgroup, and therefore we only report the correlation values for archival purposes (lecture performance: $r(9) = -0.34$, $p = .30$; next-day test: $r(9) = -0.01$, 95%CI: -0.61, 0.59, $p = .98$; follow-up test: $r(5) = 0.11$, 95% CI: -0.70, 0.80, $p = .82$; Fisher's r-to-z-transformation test across the gender groups showed $Z = 1.14$, $p = .25$ for next-day test; $Z = 0.61$, $p = .54$ for follow-up test; Fisher, 1915).

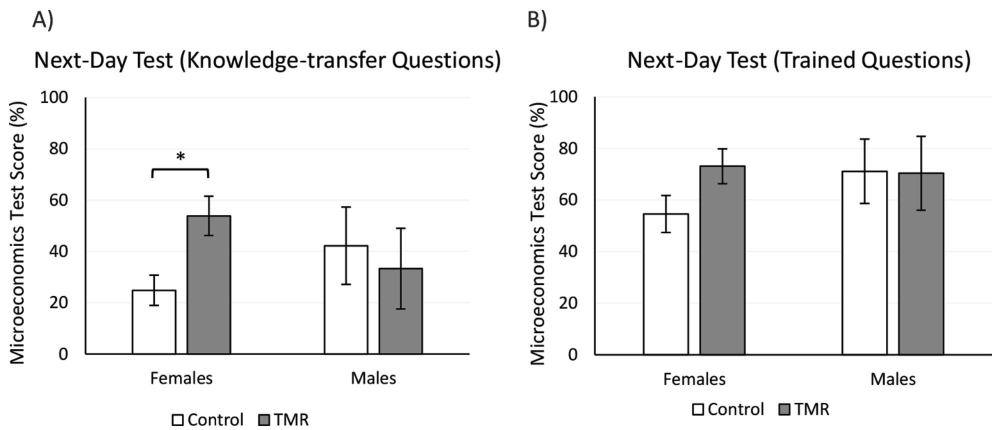

Fig. 2. Performance on the next-day test in male and female students, separated by (a) knowledge-transfer questions and (b) trained questions. Error bars: standard error. * $p \leq 0.05$.

Table 3
Spectral Power and Spindle Density in TMR and Control Conditions.

	TMR (n = 19)	Control (n = 22)	Condition Effect (p-value)	Condition Effect (Bayes Factor$_{10}$)
Slow oscillation activity				
SWS (frontal)	34.58 (12.87)	33.19 (15.16)	0.76	0.32
SWS (all sites)	20.39 (6.40)	20.38 (7.61)	> 0.99	0.31
All sleep epochs (all sites)	7.40 (1.99)	8.29 (5.26)	0.47	0.37
Delta activity				
SWS (frontal)	87.38 (38.24)	75.99 (37.33)	0.34	0.44
SWS (all sites)	41.55 (14.32)	38.86 (16.12)	0.58	0.35
All sleep epochs (all sites)	14.37 (4.50)	13.65 (5.42)	0.65	0.33
Theta activity				
SWS (frontal)	6.51 (2.58)	4.72 (1.48)	0.01*	5.64†
SWS (all sites)	3.97 (1.44)	3.28 (0.91)	0.08	1.19
All sleep epochs (all sites)	2.29 (0.82)	1.95 (0.64)	0.14	0.74
Spindle density (N2, frontal)	2.90 (0.32)	2.57 (0.75)	0.07	1.07
Spindle density (SWS, frontal)	1.54 (0.74)	1.35 (0.76)	0.43	0.39

Note. Data presented as mean (standard deviation). Slow oscillation is 0.5–1 Hz; μV^2; Delta activity is 1–4 Hz; μV^2; Theta activity is 4–8 Hz; μV^2.
* $p \leq 0.05$.
† Bayes factor > 3 indicates that the data favor the alternative hypothesis over the null hypothesis.

4. Discussion

This double-blinded TMR study showed that naturalistic sensory stimuli (classical music) can be leveraged to promote integration of college-level educational concepts. TMR increased theta activity during SWS, and greater theta activity across conditions was associated with better subsequent test performance. Interestingly, some results suggested that classical music TMR might particularly benefit females, a finding that converges with the literature on music neuroscience and that has implications for addressing achievement gaps in economics or other mathematics-based fields.

4.1. TMR mechanisms

Most previous TMR work has focused on the mechanisms, or EEG signature, of memory reactivation during SWS. For example, studies of odor and discrete-sound TMR cues indicated that TMR increased spindle density (Creery et al., 2015; Oyarzún, Morís, Luque, de Diego-Balaguer, & Fuentemilla, 2017), delta/slow oscillation activity (Creery et al., 2015), or frontal theta activity (Farthouat, Gilson, & Peigneux, 2016; Oyarzún et al., 2017), and that such increases predicted subsequent performance. In the current study, using more continuous auditory stimuli, classical music TMR had no discernible impact on delta/slow oscillations, and there was only a marginal trend for an increase in spindle activity. However, classical music TMR was associated with significantly more frontal theta activity than the white-noise control group, and frontal theta activity predicted later test performance.

Increased theta activity during SWS can indicate successful processing of auditory cues (Cox, Korjoukov, de Boer, & Talamini, 2014; Farthouat et al., 2016). By some views, theta activity is also theorized to contribute to memory formation through hippocampal-neocortical long-term potentiation and synaptic plasticity mechanisms (Axmacher, Mormann, Fernández, Elger, & Fell, 2006; Cantero et al., 2003; Nyhus & Curran, 2010). Consistent with this latter view, the re-occurrence of theta rhythms during SWS was predictive of retention of specific memories in at least two studies (vocabulary/translation words and paired associative learning in Schreiner & Rasch, 2014; Farthouat et al., 2016, respectively). The current study indicates that similar theta-dependent mechanisms can be triggered across bouts of SWS by classical music, with the outcome being the integration of learned educational concepts rather than solely the consolidation of specific discrete memories.

An alternative view is that classical music during SWS benefits test performance not because of frontal theta activity or other memory reactivation mechanisms (i.e. TMR), but instead because there is some intrinsic cognitive benefit to listening to classical music. This notion, which has been popularized as the "Mozart effect," is rooted in a few studies that found that listening to Mozart (or other energetic classical music) improved performance on the Stanford-Binet intelligence test (Rauscher, Shaw, & Ky, 1993). However, all reported benefits to performance have been temporary (15 min), meaning that this literature would not predict that listening to classical music while sleeping would help test performance 9 h later. In addition, since the original media-popularized studies in the 1990s, most researchers have concluded that

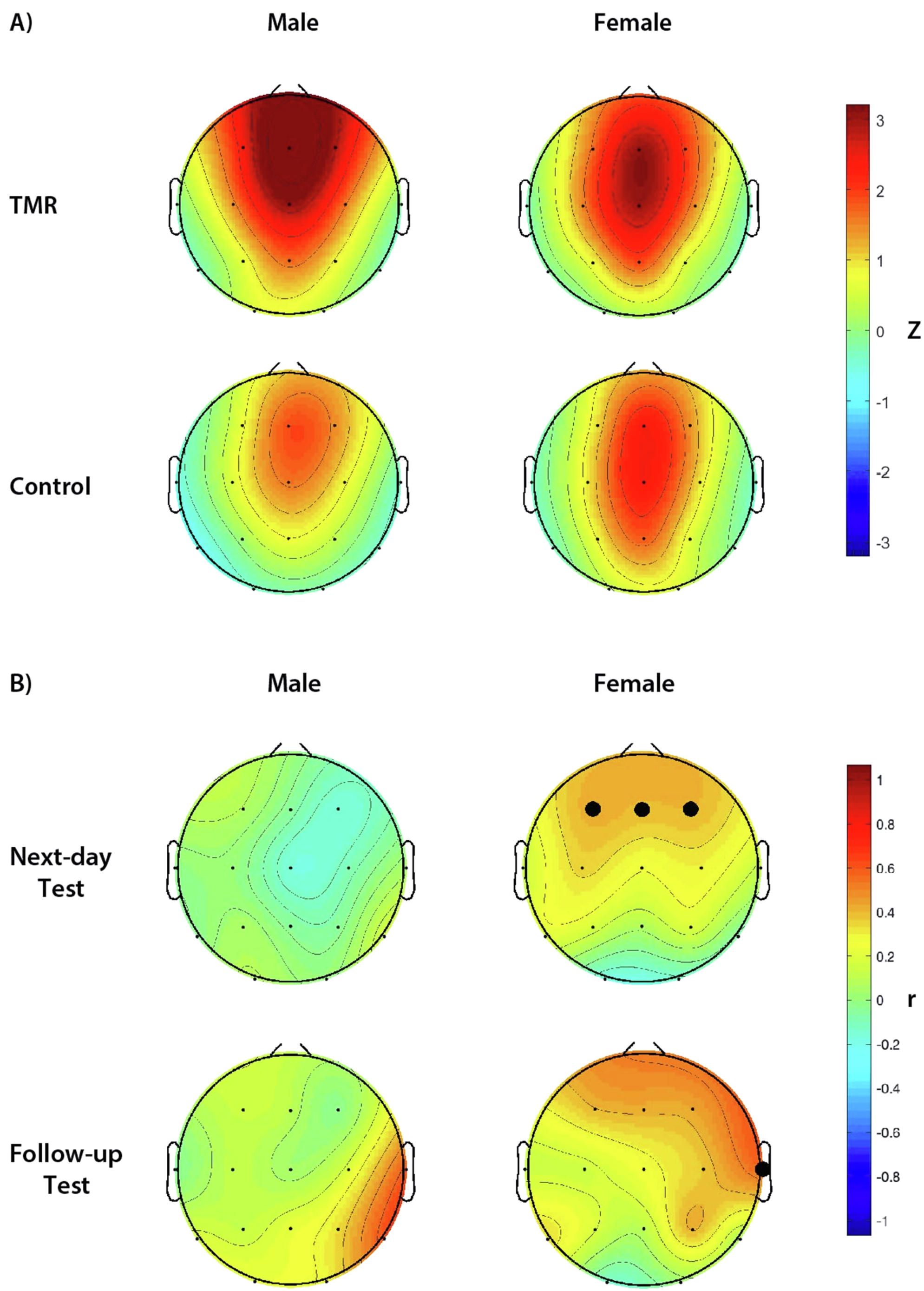

Fig. 3. (a) The TMR group showed increased frontal midline theta power during SWS for both males and females. Values shown are normalized to z-scores, to highlight relative spatial power distributions across groups. (b) Knowledge transfer test performance was positively correlated with theta power during SWS in females, but not males. Significant correlations are marked with dark dots.

the Mozart effect is an artifact of arousal and mood (Thompson, Schellenberg, & Husain, 2001), with one meta-analysis finding that evidence for the Mozart effect was largely restricted to one laboratory (Pietschnig, Voracek, & Formann, 2010). In other words, Mozart does not make memories; but, pairing music with studying *and then* listening to that music again during sleep can be memory-promoting via TMR mechanisms.

4.2. Gender differences in music processing and academic achievement

Research at the intersection of music and neuroscience is developing rapidly (Cheever et al., 2018), and there are several studies demonstrating that females process music differently than males (Gaab, Keenan, & Schlaug, 2003; Koelsch et al., 2003). There are at least two mechanisms by which gender differences in music processing would explain why classical music TMR effects seemed very strong in females. First, in multitasking scenarios, females show that they can maintain cognitive task performance levels while listening to music whereas males cannot (Fancourt et al., 2016; Feizpour et al., 2018), implying that females may be able to study more effectively than males while listening to music, a pre-requisite for TMR to be successful (Creery et al., 2015). Second, females outperform males at later recognizing familiar music (Miles et al., 2016), implying that when familiar classical music was played during SWS, females' ability to efficiently recognize

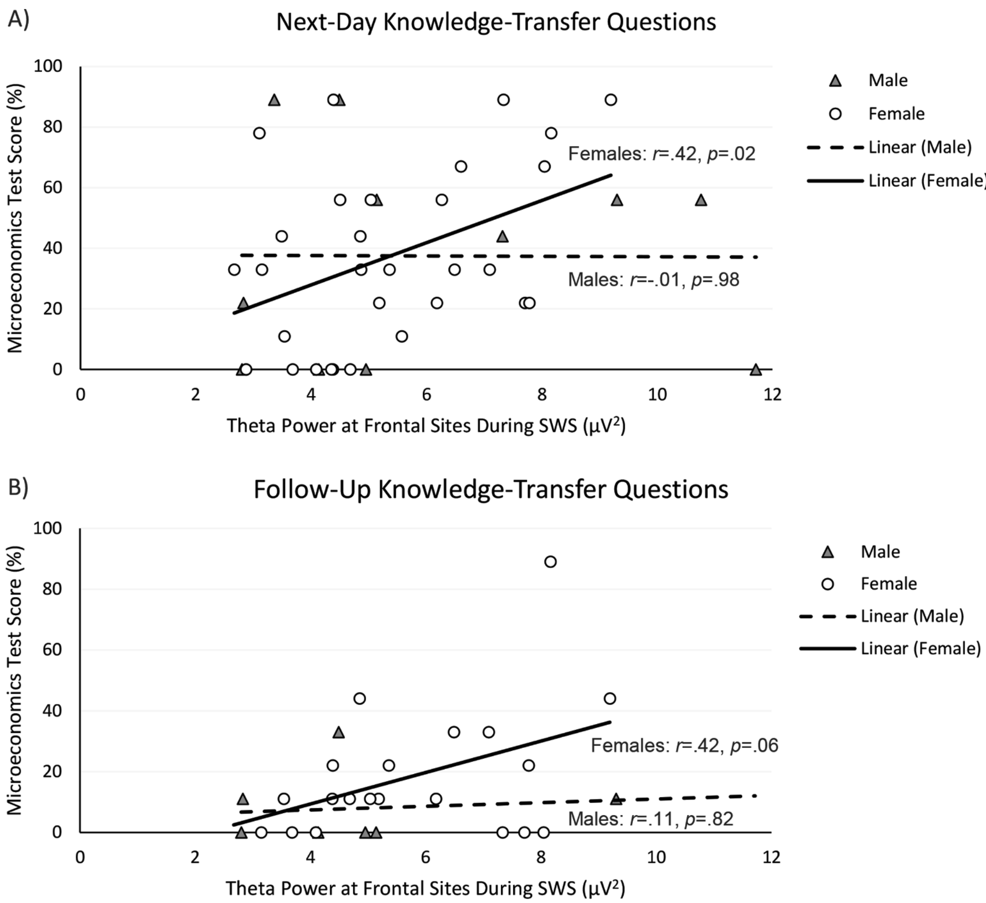

Fig. 4. (a) Theta power at frontal sites during slow wave sleep predicted performance on next-day knowledge-transfer questions in females, but not in males; (b) Theta power trended toward predicting knowledge-transfer question performance on the follow-up test in females, but not in males. Circles and solid lines represent females; triangles and dash lines represent males.

the music would facilitate reactivation of the associated educational content. Of course, future studies with considerable statistical power to detect gender interactions are required to confirm (or falsify) this hypothesis that gender differences in musical processing allow females to specifically benefit from classical-music TMR.

Advancing knowledge of TMR and music processing might inform pathways to bridging achievement gaps in STEM learning. Gender differences have long been observed in mathematics-based disciplines (Fryer & Levitt, 2010). There are numerous factors that contribute to the underrepresentation of women in some STEM fields (e.g., disposition, occupational interests); however, many of the factors known to influence STEM achievement gaps are themselves influenced by sleep and sleep restriction: cognitive abilities/relative strengths, field-specific ability beliefs, and gender-related stereotypes (Wang & Degol, 2017). Restricting sleep amplifies stress/emotional reactivity and activates stereotypes (Ghumman & Barnes, 2013; Goldstein-Piekarski et al., 2018), which in a cyclical fashion can contribute to more chronic sleep disturbances, thereby impairing attention and memory consolidation. If females are more susceptible to these detrimental effects than males (Gao et al., 2019; Hajali et al., 2012; Rångtell et al., 2019), then addressing sleep disparities may bridge STEM achievement gaps (Heissel, Levy, & Adam, 2017). Classical music TMR may be one sleep-based approach to reducing sleep disparities and achievement gaps.

4.3. Limitations and conclusions

Limitations of the current study include attrition, sample size inequality, and low performance at the nine-month follow-up. Although a single TMR session may help students to cram for the next-day test, for effects to persist throughout the academic year, TMR sessions may need to be repeated and progressively spaced apart (Kornell, 2009). Furthermore, while the current study used an ecological research paradigm, the research procedures were carried out in a controlled-laboratory setting. Future research that translates TMR procedures to home settings will be an important step towards remediating deficits in memory consolidation caused by insufficient sleep in college students (Goldi & Rasch, 2019).

It may also be fruitful for future studies to include multiple control groups, for example, a group that listens to classical music during sleep but not during the lecture. Adding such a condition would help to disentangle TMR-based explanations from brainwave-entraining explanations. According to the brainwave-entraining view, listening to classical music during sleep entrains brainwave oscillations in a manner that leads to memory enhancement (Obleser & Kayser, 2019). For this explanation to be considered viable, several methodological conditions must be met. First, when music entrains brainwaves, the elicited state evoked potentials will synchronize to the tempo of the music (Daly et al., 2014); but, in the current study, the classical music tempos never matched theta frequency, meaning that they should not entrain theta waves (110 to 147 beats per minute, corresponding to 1.83 to 2.45 Hz). Second, for auditory stimulation during sleep to entrain specific brainwave frequencies, the stimulation must consistently occur in-phase with ongoing oscillations (Ngo, Claussen, Born, & Mölle, 2013). However, the Chopin, Beethoven, and Vivaldi pieces used here lacked consistency in interstimulus intervals, again disfavoring an entrainment explanation. The practical recommendation, therefore, is that students only listen to classical music while sleeping if they first paired that music with studying.

The current work showed that TMR is achievable with continuous auditory cues (classical music) and integrative educational concepts. Classical music TMR can improve performance on knowledge-transfer questions on the next-day test. Some findings indicated that females were particularly benefited by classical music TMR, but definitive gender-based conclusions will require additional future testing. Future work at the intersection of music, memory theory, and neuroscience is needed to enhance educational outcomes and bridge achievement gaps in STEM learning.

Acknowledgements

The authors are grateful to Daniel Howard for developing and pilot testing the microeconomics materials. The authors are additionally grateful for the assistance of Hannah Ballard, Nikita Chapagain, Mary High, Taylor Luster, Stacy Nguyen, Zachary Pranske, Natalya Pruett, Cole Rowley, Kiersten Scott, Taylor Terlizzese, Ali Villagran, and Daniel Zeter.

Author contributions

MKS conceptualized and designed the study. All authors performed statistical analyses. CG conducted spectral analysis. PF performed spindle detection. All authors wrote the manuscript.

Funding sources

This work was supported by the National Science Foundation (1920730, 1943323) and National Institutes of Health (AG05316). The funding sources had no role in study design, data collection, data analysis and interpretation, writing of the manuscript, and decision to submit the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.nlm.2020.107206.

References

American Academy of Pediatrics Adolescent Sleep Working Group (2014). School start times for adolescents. *Pediatrics, 134*(3), 642–649. https://doi.org/10.1542/peds.2014-1697.

Axmacher, N., Mormann, F., Fernández, G., Elger, C. E., & Fell, J. (2006). Memory formation by neuronal synchronization. *Brain Research Reviews, 52*(1), 170–182. https://doi.org/10.1016/j.brainresrev.2006.01.007.

Ballard, C., & Johnson, M. (2005). Gender, expectations, and grades in introductory microeconomics at a US university. *Feminist Economics, 11*(1), 95–122. https://doi.org/10.1080/1354570042000332560.

Baumol, W. J., & Blinder, A. S. (2007). *Microeconomics: Principles and policy.* Florence, KY: Southwestern Learning.

Beethoven, L. (1802). *Piano Sonata No. 14 [Op. 27, No. 2, movement 1].* Vienna, Austria: Giovanni Cappie Comp.

Bender, R., & Lange, S. (2001). Adjusting for multiple testing—when and how? *Journal of Clinical Epidemiology, 54*(4), 343–349.

Bennion, K. A., Payne, J. D., & Kensinger, E. A. (2015). Selective effects of sleep on emotional memory: What mechanisms are responsible? *Translational Issues in Psychological Science, 1*(1), 79–88. https://doi.org/10.1037/tps0000019.

Born, J., & Wilhelm, I. (2012). System consolidation of memory during sleep. *Psychological Research, 76*(2), 192–203. https://doi.org/10.1007/s00426-011-0335-6.

Buysse, D. J., Reynolds, C. F., III, Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*(2), 193–213. https://doi.org/10.1016/0165-1781(89)90047-4.

Cairney, S. A., Durrant, S. J., Hulleman, J., & Lewis, P. A. (2014). Targeted memory reactivation during slow wave sleep facilitates emotional memory consolidation. *Sleep, 37*(4), 701–707. https://doi.org/10.5665/sleep.3572.

Cantero, J. L., Atienza, M., Stickgold, R., Kahana, M. J., Madsen, J. R., & Kocsis, B. (2003). Sleep-dependent θ oscillations in the human hippocampus and neocortex. *Journal of Neuroscience, 23*(34), 10897–10903. https://doi.org/10.1523/JNEUROSCI.23-34-10897.2003.

Carney, C. E., Buysse, D. J., Ancoli-Israel, S., Edinger, J. D., Krystal, A. D., Lichstein, K. L., et al. (2012). The consensus sleep diary: Standardizing prospective sleep self-monitoring. *Sleep, 35*(2), 287–302. https://doi.org/10.5665/sleep.1642.

Cellini, N., & Capuozzo, A. (2018). Shaping memory consolidation via targeted memory reactivation during sleep. *Annals of the New York Academy of Sciences, 13855*. https://doi.org/10.1111/nyas.13855.

Cheever, T., Taylor, A., Finkelstein, R., Edwards, E., Thomas, L., Bradt, J., et al. (2018). NIH/Kennedy Center Workshop on Music and the Brain: Finding harmony. *Neuron, 97*(6), 1214–1218. https://doi.org/10.1016/j.neuron.2018.02.004.

Chopin, F. (2005). Nocturne in E-Flat Major [Op.9, No.2]. New York City, NY: Alfred Music.

Conway, M. A., Cohen, G., & Stanhope, N. (1991). On the very long-term retention of knowledge acquired through formal education: Twelve years of cognitive psychology. *Journal of Experimental Psychology: General, 120*(4), 395–409. https://doi.org/10.1037/0096-3445.120.4.395.

Cox, R., Korjoukov, I., de Boer, M., & Talamini, L. M. (2014). Sound asleep: Processing and retention of slow oscillation phase-targeted stimuli. *PLoS ONE, 9*(7), e101567. https://doi.org/10.1371/journal.pone.0101567.

Creery, J. D., Oudiette, D., Antony, J. W., & Paller, K. A. (2015). Targeted memory reactivation during sleep depends on prior learning. *Sleep, 38*(5), 755–763. https://doi.org/10.5665/sleep.4670.

Daly, I., Hallowell, J., Hwang, F., Kirke, A., Malik, A., Roesch, E., et al. (2014). Changes in music tempo entrain movement related brain activity. *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 4595–4598). IEEE.

Danhauer, J. L., Johnson, C. E., Byrd, A., DeGood, L., Meuel, C., Pecile, A., et al. (2009). Survey of college students on iPod use and hearing health. *Journal of the American Academy of Audiology, 20*(1), 5–27. https://doi.org/10.3766/jaaa.20.1.2.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology, 5*, 781. https://doi.org/10.3389/fpsyg.2014.00781.

Donohue, K. C., & Spencer, R. M. (2011). Continuous re-exposure to environmental sound cues during sleep does not improve memory for semantically unrelated word pairs. *Journal of Cognitive Education and Psychology, 10*(2), 167–177. https://doi.org/10.1891/1945-8959.10.2.167.

Doran, S. M., Van Dongen, H. P. A., & Dinges, D. F. (2001). Sustained attention performance during sleep deprivation: Evidence of state instability. *Archives Italiennes de Biologie, 139*(3), 253–267.

Ebbinghaus, H. (1885). *Ueber das Gedächtnis: Untersuchungen zur experimentellen Psychologie.* Leipzig: Verlag Von Duncker & Humblot.

Fancourt, D., Burton, T. M., & Williamon, A. (2016). The razor's edge: Australian rock music impairs men's performance when pretending to be a surgeon. *Medical Journal of Australia, 205*(11), 515–518. https://doi.org/10.5694/mja16.01045.

Farthouat, J., Gilson, M., & Peigneux, P. (2016). New evidence for the necessity of a silent plastic period during sleep for a memory benefit of targeted memory reactivation. *Sleep Spindles Cortical Up States, 1*(1), 14–26. https://doi.org/10.1556/2053.1.2016.002.

Feizpour, A., Parkington, H. C., & Mansouri, F. A. (2018). Cognitive sex differences in effects of music in Wisconsin Card Sorting Test. *Psychology of Music, 030573561879503*. https://doi.org/10.1177/0305735618795030.

Fisher, R. A. (1915). Frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika, 10*(4), 507–521.

Fryer, R. G., Jr, & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *Am Economic J Applied Economics, 2*(2), 210–240. https://doi.org/10.1257/app.2.2.210.

Gaab, N., Keenan, J. P., & Schlaug, G. (2003). The effects of gender on the neural substrates of pitch memory. *Journal of Cognitive Neuroscience, 15*(6), 810–820. https://doi.org/10.1162/089892903322370735.

Gao, C., Terlizzese, T., & Scullin, M. K. (2019). Short sleep and late bedtimes are detrimental to educational learning and knowledge transfer: An investigation of individual differences in susceptibility. *Chronobiology International, 36*(3), 307–318. https://doi.org/10.1080/07420528.2018.1539401.

Ghumman, S., & Barnes, C. M. (2013). Sleep and prejudice: A resources recovery approach. *Journal of Applied Social Psychology, 43*, E166–E178. https://doi.org/10.1111/jasp.12045.

Goldi, M., & Rasch, B. (2019). Effects of targeted memory reactivation during sleep at home depend on sleep disturbances and habituation. *NPJ Science of Learning, 4*, 5. https://doi.org/10.1038/s41539-019-0044-2.

Goldstein-Piekarski, A. N., Greer, S. M., Saletin, J. M., Harvey, A. G., Williams, L. M., & Walker, M. P. (2018). Sex, sleep deprivation, and the anxious brain. *Journal of Cognitive Neuroscience, 30*(4), 565–578. https://doi.org/10.1162/jocn_a_01225.

Hajali, V., Sheibani, V., Esmaeili-Mahani, S., & Shabani, M. (2012). Female rats are more susceptible to the deleterious effects of paradoxical sleep deprivation on cognitive performance. *Behavioural Brain Research, 228*(2), 311–318. https://doi.org/10.1016/j.bbr.2011.12.008.

Harmat, L., Takács, J., & Bodizs, R. (2008). Music improves sleep quality in students. *Journal of Advanced Nursing, 62*(3), 327–335. https://doi.org/10.1111/j.1365-2648.2008.04602.x.

Heissel, J. A., Levy, D. J., & Adam, E. K. (2017). Stress, sleep, and performance on standardized tests: Understudied pathways to the achievement gap. 2332858417713488 *AERA Open, 3*(3), https://doi.org/10.1177/2332858417713488.

Hershner, S., & O'Brien, L. M. (2018). The impact of a randomized sleep education intervention for college students. *Journal of Clinical Sleep Medicine, 14*(03), 337–347. https://doi.org/10.5664/jcsm.6974.

Hirshkowitz, M., Whiton, K., Albert, S. M., Alessi, C., Bruni, O., DonCarlos, L., et al. (2015). National Sleep Foundation's sleep time duration recommendations: Methodology and results summary. *Sleep Health, 1*(1), 40–43. https://doi.org/10.

1016/j.sleh.2014.12.010.
Horne, J. A., & Östberg, O. (1976). A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. *International Journal of Chronobiology, 4*(2), 97–100.
Hu, X., Antony, J. W., Creery, J. D., Vargas, I. M., Bodenhausen, G. V., & Paller, K. A. (2015). Unlearning implicit social biases during sleep. *Science, 348*(6238), 1013–1015. https://doi.org/10.1126/science.aaa3841.
Hu, X., Cheng, L. Y., Chiu, M. H., & Paller, K. (2020). Promoting memory consolidation during sleep: A meta-analysis of targeted memory reactivation. *Psychological Bulletin, 146*(3), 218–244. https://doi.org/10.1037/bul0000223.
Iber, C., Ancoli-Israel, S., Chesson, A., & Quan, S. F. (2007). *The AASM manual for the scoring of sleep and associated events: Rules, terminology and technical specifications.* Westchester, IL: American Academy of Sleep Medicine.
Jäncke, L., Brügger, E., Brummer, M., Scherrer, S., & Alahmadi, N. (2014). Verbal learning in the context of background music: No influence of vocals and instrumentals on verbal learning. *Behavioral and Brain Functions, 10*(1), 10. https://doi.org/10.1186/1744-9081-10-10.
Johns, M. W. (1991). A new method for measuring daytime sleepiness: The Epworth sleepiness scale. *Sleep, 14*(6), 540–545. https://doi.org/10.1093/sleep/14.6.540.
Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*(5865), 966–968. https://doi.org/10.1126/science.1152408.
King, E., Daunis, M., Tami, C., & Scullin, M. K. (2017). Sleep in studio-based courses: Outcomes for creativity task performance. *Journal of Interior Design, 42*, 5–27. https://doi.org/10.1111/joid.12104.
Koelsch, S., Maess, B., Grossmann, T., & Friederici, A. D. (2003). Electric brain responses reveal gender differences in music processing. *NeuroReport, 14*(5), 709–713. https://doi.org/10.1097/00001756-200304150-00010.
Kornell, N. (2009). Optimising learning using flashcards: Spacing is more effective than cramming. *Applied Cognitive Psychology, 23*(9), 1297–1317. https://doi.org/10.1002/acp.1537.
Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian cognitive modeling: A practical course.* Cambridge: Cambridge University Press.
Lund, H. G., Reider, B. D., Whiting, A. B., & Prichard, J. R. (2010). Sleep patterns and predictors of disturbed sleep in a large population of college students. *Journal of Adolescent Health, 46*(2), 124–132.
Miles, S. A., Miranda, R. A., & Ullman, M. T. (2016). Sex differences in music: A female advantage at recognizing familiar melodies. *Frontiers in Psychology, 7*, 278. https://doi.org/10.3389/fpsyg.2016.00278.
Miyatsu, T., Nguyen, K., & McDaniel, M. A. (2018). Five popular study strategies: Their pitfalls and optimal implementations. *Perspectives on Psychological Science, 13*(3), 390–407. https://doi.org/10.1177/1745691617710510.
Ngo, H. V. V., Claussen, J. C., Born, J., & Mölle, M. (2013). Induction of slow oscillations by rhythmic acoustic stimulation. *Journal of Sleep Research, 22*(1), 22–31. https://doi.org/10.1111/j.1365-2869.2012.01039.x.
Nyhus, E., & Curran, T. (2010). Functional role of gamma and theta oscillations in episodic memory. *Neuroscience & Biobehavioral Reviews, 34*(7), 1023–1035. https://doi.org/10.1016/j.neubiorev.2009.12.014.
Obleser, J., & Kayser, C. (2019). Neural entrainment and attentional selection in the listening brain. *Trends in Cognitive Sciences, 23*(11), 913–926. https://doi.org/10.1016/j.tics.2019.08.004.
Okano, K., Kaczmarzyk, J. R., Dave, N., Gabrieli, J. D., & Grossman, J. C. (2019). Sleep quality, duration, and consistency are associated with better academic performance in college students. *npj Science of Learning, 4*(1), 16. https://doi.org/10.1038/s41539-019-0055-z.
Oyarzún, J. P., Morís, J., Luque, D., de Diego-Balaguer, R., & Fuentemilla, L. (2017). Targeted memory reactivation during sleep adaptively promotes the strengthening or weakening of overlapping memories. *Journal of Neuroscience, 37*(32), 7748–7758. https://doi.org/10.1523/JNEUROSCI.3537-16.2017.
Pan, S. C., & Rickard, T. C. (2018). Transfer of test-enhanced learning: Meta-analytic review and synthesis. *Psychological Bulletin, 144*(7), 710–756. https://doi.org/10.1037/bul0000151.
Pietschnig, J., Voracek, M., & Formann, A. K. (2010). Mozart effect–Shmozart effect: A meta-analysis. *Intelligence, 38*(3), 314–323. https://doi.org/10.1016/j.intell.2010.03.001.
Prather, A. A., Epel, E. S., Cohen, B. E., Neylan, T. C., & Whooley, M. A. (2013). Gender differences in the prospective associations of self-reported sleep quality with biomarkers of systemic inflammation and coagulation: Findings from the Heart and Soul Study. *Journal of Psychiatric Research, 47*(9), 1228–1235. https://doi.org/10.1016/j.jpsychires.2013.05.004.
Rångtell, F. H., Karamchedu, S., Andersson, P., Liethof, L., Olaya Búcaro, M., Lampola, L., et al. (2019). A single night of sleep loss impairs objective but not subjective working memory performance in a sex-dependent manner. *Journal of Sleep Research, 28*(1), e12651. https://doi.org/10.1111/jsr.12651.
Rasch, B., & Born, J. (2013). About sleep's role in memory. *Physiological Reviews, 93*, 681–766. https://doi.org/10.1152/physrev.00032.2012.
Rasch, B., Büchel, C., Gais, S., & Born, J. (2007). Odor cues during slow-wave sleep prompt declarative memory consolidation. *Science, 315*(5817), 1426–1429. https://doi.org/10.1126/science.1138581.
Rauscher, F. H., Shaw, G. L., & Ky, C. N. (1993). Music and spatial task performance. *Nature, 365*(6447), 611.
Raven, J. C. (1938). *Raven's progressive matrices.* Los Angeles, CA: Western Psychological Services.
Rudoy, J. D., Voss, J. L., Westerberg, C. E., & Paller, K. A. (2009). Strengthening individual memories by reactivating them during sleep. 1079 1079 *Science, 326*(5956), https://doi.org/10.1126/science.1179013.
Saletin, J. M., Goldstein, A. N., & Walker, M. P. (2011). The role of sleep in directed forgetting and remembering of human memories. *Cerebral Cortex, 21*(11), 2534–2541. https://doi.org/10.1093/cercor/bhr034.
Sanchez, R., & Bootzin, R. R. (1985). A comparison of white noise and music: Effects of predictable and unpredictable sounds on sleep. *Journal of Sleep Research, 14*, 121.
Schouten, D. I., Pereira, S. I., Tops, M., & Louzada, F. M. (2017). State of the art on targeted memory reactivation: Sleep your way to enhanced cognition. *Sleep Medicine Reviews, 32*, 123–131. https://doi.org/10.1016/j.smrv.2016.04.002.
Schreiner, T., & Rasch, B. (2014). Boosting vocabulary learning by verbal cueing during sleep. *Cerebral Cortex, 25*(11), 4169–4179. https://doi.org/10.1093/cercor/bhu139.
Scullin, M. K. (2019). The eight hour sleep challenge during final exams week. *Teaching of Psychology, 46*(1), 55–63. https://doi.org/10.1111/joid.12135.
Scullin, M. K., Gao, C., Fillmore, P., Roberts, R. L., Pruett, N., & Bliwise, D. L. (2019). REM sleep mediates age-related decline in prospective memory consolidation. *Sleep*. https://doi.org/10.1093/sleep/zsz055.
Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science, 12*(3), 248–251. https://doi.org/10.1111/1467-9280.00345.
van Rijn, E., Koh, S. Y., Ng, A. S., Vinogradova, K., Chee, N. I., Lee, S. M., et al. (2019). Evaluation of an interactive school-based sleep education program: A cluster-randomized controlled trial. *Sleep Health*. https://doi.org/10.1016/j.sleh.2019.10.006.
Vivaldi, A. (1725.). *The four seasons [Concerto No. 1. La primavera (Spring), movement 1: Allegro]. Amsterdam, the.* Netherlands: G. Schirmer Inc.
Wahlstrom, K. L. (2000). School start time and teen sleep. *High School Magazine, 7*(9), 40–41.
Wamsley, E. J., Tucker, M. A., Shinn, A. K., Ono, K. E., McKinley, S. K., Ely, A. V., et al. (2012). Reduced sleep spindles and spindle coherence in schizophrenia: Mechanisms of impaired memory consolidation? *Biological Psychiatry, 71*(2), 154–161. https://doi.org/10.1016/j.biopsych.2011.08.008.
Wang, M. T., & Degol, J. L. (2017). Gender gap in science, technology, engineering, and mathematics (STEM): Current knowledge, implications for practice, policy, and future directions. *Educational Psychology Review, 29*(1), 119–140. https://doi.org/10.1007/s10648-015-9355-x.
Warby, S. C., Wendt, S. L., Welinder, P., Munk, E. G., Carrillo, O., Sorensen, H. B., et al. (2014). Sleep-spindle detection: Crowdsourcing and evaluating performance of experts, non-experts and automated methods. *Nature Methods, 11*(4), 385. https://doi.org/10.1038/nmeth.2855.
Wing, Y. K., Chan, N. Y., Man Yu, M. W., Lam, S. P., Zhang, J., Li, S. X., et al. (2015). A school-based sleep education program for adolescents: A cluster randomized trial. *Pediatrics, 135*(3), e635–e643. https://doi.org/10.1542/peds.2014-2419.
Yonelinas, A. P., Ranganath, C., Ekstrom, A. D., & Wiltgen, B. J. (2019). A contextual binding theory of episodic memory: Systems consolidation reconsidered. *Nature Reviews Neuroscience, 20*, 364–375. https://doi.org/10.1038/s41583-019-0150-4.
Zhang, B., & Wing, Y. K. (2006). Sex differences in insomnia: A meta-analysis. *Sleep, 29*(1), 85–93.