International Journal of Behavioral Medicine (2021) 28:250–258
https://doi.org/10.1007/s12529-021-09973-1

SPECIAL ISSUE: SLEEP SCIENCE

Test Anxiety and Poor Sleep: A Vicious Cycle

Nancy Hamilton[1] · Ronald Freche[1] · Yichi Zhang[1] · Gabriella Zeller[1] · Ian Carroll[1]

Accepted: 19 February 2021 / Published online: 17 March 2021

Abstract

Background Test anxiety may be better thought of as a biopsychosocial process affecting academic performance during the days leading up to an exam, rather than a static appraisal of attitudes related to test taking. This was a passive observational study following students 2 days before a midterm exam and was designed to test the Sleep Anxiety Performance Process (SAPP) model in the context of a psychology statistics exam.

Methods Undergraduates (N = 167) enrolled in a statistics class, January–November 2015. Participants completed an electronic battery of measures and Sleep Mood Study Diaries (SMS) during the mornings, 2 days before a statistics exam. Instructors confirmed exam scores.

Results A path model showed a reciprocal bi-directional relationship between Sleep Quality and restfulness (Q&R) and test anxiety 2 days before a scheduled exam, with test anxiety measured in the morning, before the exam predicting exam performance. Prior exam performance, being a non-native English speaker (ESL), and class performance motivation also predicted exam performance.

Conclusions These data support the SAPP model's premise that that sleep and anxiety feed one another, as a reciprocal process, that collectively impairs academic performance, with direct effects on academic performance, but with implications for overall student health.

Keywords College students · Anxiety · Sleep · Academic performance

Introduction

Test anxiety has been defined as the emotional, physiological, and behavioral responses in the anticipation of negative evaluation from an upcoming test or exam [1]. Between 10 and 40% of students experience test anxiety and students with higher levels of test anxiety perform more poorly on tests and have lower Grade Point Averages (GPA: the accumulated average of point values assigned to final letter grades, typically ranging from 1 to 4, with an $A = 4$ points, a $B = 3$, etc.) [2]. However, consequences of test anxiety extend beyond the classroom and are correlated with poor health behaviors, including dysregulated sleep patterns and poor sleep quality [3]. Given the well-known bi-directional relationship between sleep and anxiety [4], it would be important to understand the specific pathways by which anxiety and sleep predicts exam performance and potentially unfold over the course of the several days prior to an exam.

Characteristics of the relationship between test-taking and anxiety have been well documented. Although trait anxiety is related to test anxiety and GPA [5], state and test-specific anxiety measured immediately before an exam have been shown to have stronger relationships to performance than do more distal measures of state anxiety [6, 7]. In one particularly well-designed study, researchers assessed trait anxiety, state anxiety, and statistics specific anxiety. Measured 2 weeks before the exam, trait anxiety was related to statistics specific anxiety and state anxiety. However, performance was predicted by attitudes about math and interest in the class, statistics specific anxiety, and state anxiety measured proximal to the test [6]. Test anxiety varies across time [7], across classes, and attitudes about subject matter are important. In particular, statistics classes are associated with high levels of anxiety [1, 8], especially for female students [9], and class performance has been shown to be driven both by attitudes about the specific class and anxiety [6].

✉ Nancy Hamilton
nancyh@ku.edu

[1] Department of Psychology, University of Kansas, KS 66044 Lawrence, USA

Although most test anxiety studies have only examined the relationship between test anxiety and performance, extant research suggests that sleep is related to both academic performance and anxiety. For instance, sleep deprivation has been found to compromise a broad array of memory processes [10]. Average sleep duration, sleep habits, and chronotype have been found to predict GPA [11–14] and medical students' exam scores [15]. In an illustrative study, end of the semester undergraduate course performance was predicted by prior academic performance, study habits, chronotype, sleep duration, and sleep quality [16]. In any study of exam anxiety, it would be important to measure both anxiety and sleep as sleep problems such as insomnia and anxiety are highly correlated [17]. State and trait anxiety have documented relationships to shorter total sleep time and disrupted sleep [18, 19], longer sleep onset latencies, more Wake Time after Sleep Onset (WASO [20]), and there is evidence of a bidirectional relationship between poor sleep and anxiety [4].

A few studies have examined exam-related anxiety, sleep, and exam performance. Among college students, test anxiety was related to sleep duration and quality the night before an exam, but only test anxiety predicted actual exam performance [18]. Stress (a construct related to anxiety) was found to increase, and sleep quality (but not duration) was found to decrease in the week before an exam period [21]. Stress and sleep duration were related to academic performance among medical students [22, 23]. During the month before a high-stakes exam, stress increased and sleep duration decreased; however, the individual relationship of stress and sleep disruption were not examined [22]. Collectively, these studies indicate that anxiety (and stress) and poor sleep (shorter duration and poor quality) go hand in hand. However, the temporal relationship between sleep disruption and anxiety has not yet been explicitly examined.

The Present Study

Test anxiety is a serious academic problem for college students. However, the bidirectional relationship between sleep and anxiety suggests that it would be important to examine the temporal relationship between sleep (duration and quality) and test anxiety as predictors of academic performance. Because there is an abundance of research documenting a reciprocal relationship between sleep problems such as insomnia and anxiety, we proposed that the effects of test anxiety on academic performance would unfold as a reciprocal process, across several days, as represented by the Sleep Anxiety Performance Process (SAPP) model. Moreover, given the relationship of sleep problems to decrements in memory, we predicted that sleep would mediate the relationship between anxiety and academic performance. In order to understand the temporal relationship between test-taking anxiety and sleep (duration and quality), we used electronic diaries to examine the trajectory of anxiety and sleep during two days before a statistics exam. However, we were also mindful that there were other relevant predictors of achievement and potential confounds of the relationships in question; thus, we controlled for prior exam performance, and individual difference variables including trait anxiety and symptoms of depression, enduring sleep patterns (chronotype and insomnia), and baseline sleep quality, as well as class performance motivation and study habits. After controlling for these individual differences, the SAPP model predicted that the effect of anxiety on exam performance would be mediated by sleep quality and sleep duration.

Method

Design and Participants

Undergraduate students were recruited from seven sections of introductory statistics (STATS) classes at a large university in the Midwestern USA. Data were gathered from STATS classes taught by five different instructors, gathered during a single calendar year, two spring sections ($n = 40$), one summer section ($n = 9$), and five fall sections ($n = 118$).

A total of 314 students were offered participation in the study ($N = 167$) consented to participate. Power was analyzed via simulation, where relationships between the core variables of varying strength were evaluated.

Consistent with demographics of psychology majors at this university, 75% of participants were female; the majority were Caucasian (74%); 8% identified as African-American, 5% Hispanic, 8% Asian, and 5% identified themselves as having multiple ethnic affiliations. English was a second language for 9% of participants. STATS is a class for second year college students and ages ranged from 18 to 40 ($M = 20.65$, $SD = 2.83$).

Procedure

A research assistant who explained the nature of the study recruited students from classrooms. Recruitment and baseline assessments varied according to the semester schedule, with study entry following the first exam. Interested STATS students linked to an online baseline questionnaire (Qualtrics survey software), signed an electronic consent form, provided their email addresses, and were sent Sleep, Mood, Study daily diaries (SMS diaries) at 6:00 a.m. each morning for the 6 days leading up to a target exam. Participants were asked to respond to time-stamped diaries as quickly as possible after waking. In an effort to minimize response burden and capture the strongest anxiety-performance relationship [6], test anxiety was measured only twice: the morning before the exam and the morning of the exam. After completing the final SMS diary, instructors confirmed students' self-reported exam scores. At that time students were also provided with a debriefing statement. Each student could receive up to 15 total extra credit points. The protocol for this study was approved by the University of Kansas Institutional Review Board (IRB).

Measures

Although all questionnaires were administered in an electronic format, psychometric properties (i.e., internal consistency and construct validity) remained similar to the paper and pencil versions of the questionnaires.

Baseline Battery

Student characteristics in the battery included items to identify participants' age, sex, ethnicity, year in school, employment, school credits, previous exam score, and class performance motivation ("How important is it for you to do well in this class?") (4-point Likert items).

The Pittsburgh Sleep Quality Inventory (PSQI) [24] was used as an omnibus measure of sleep quality (Cronbach's $\alpha = 0.66$). PSQI Total Score range: 0–21. A score of 5 or greater indicates clinically relevant sleep disturbance.

The Morning-Eveningness Questionnaire (MEQ) was used to measure chronotype. MEQ contains 19 Likert items (Cronbach's α = 0.76-0.79 [25]:). Higher scores indicate greater "morningness" (70–86: definite; 59–69: moderate), lower scores indicate "eveningness" (16–20: definite; 31–41: moderate), and central scores indicate "intermediate" tendencies (42–58).

The Insomnia Severity Index (ISI) was used to measure symptoms of insomnia (Cronbach's α = 0.90–0.91 [26]. The ISI consists of 7 Likert items that measure clinical insomnia. Higher total scores denote higher severity of clinical insomnia, clinical cut-offs: sub-threshold insomnia (8–14), moderately severe insomnia (15–21), and severe clinical insomnia (22–28).

The Beck Depression Inventory-II (BDI) was used to measure symptoms of depression (Chronbach's α = 0.86–0.95; [27]). Scores exceeding 10–12 indicate risk for depression.

The State Trait Anxiety Inventory, Trait form (STAI-T) was used to measure symptom of anxiety [28]. The STAI-T uses 20, 4-point Likert items, with higher scores indicating greater anxiety (Cronbach's α = 0.86–0.95). Recent administrations of the STAI-T with college students show average scores ranging from 37 to 41.51[29–31].

Sleep Mood Study Diaries

Sleep was measured each morning using a sleep diary adapted from the Consensus Sleep Diary (Qs 1–10) [32]. Questions asked for bedtime, time it took to fall asleep, number of awakenings, number of awakenings after sleep onset (WASO), time out of bed, sleep duration. Sleep Quality (Q), and Restfulness (R) (7-point Likert items) were highly correlated (0.81, $p < 0.01$) and thus were averaged (Q&R). From this measure, the following variables were calculated: sleep onset latency (SOL), total sleep time (TST), and sleep efficiency (SE).

Study behavior was measured by asking the following questions: "How many minutes did you study for your statistics class yesterday?" "Did you go to (scheduled) class"?

Test anxiety was assessed on the mornings before (day 5) and of the exam (day 6). Students rated test-taking anxiety using a version of the State Trait Anxiety Inventory-State Version [28], with an altered instruction set to capture anxiety about the upcoming exam (STAI-test). The STAI-test was selected because it is brief (20 4-point Likert items), reactive to changes in anxiety across successive days, but with high observed internal consistency using this instruction set (Cronbach's α = 0.85) that is similar to established instruction sets (Cronbach's α = 0.86–0.95). Although sleep diary and study behaviors were gathered for six mornings, STAI-Test was measured only twice, in close proximity to the test in order to minimize response burden.

Procedure

A research assistant who explained the nature of the study recruited students from classrooms. Recruitment and baseline assessments varied according to the semester schedule, with study entry following the first exam. Interested STATS students linked to an online baseline questionnaire (Qualtrics survey software), signed an electronic consent form, provided their email addresses, and were sent Sleep, Mood, Study daily diaries (SMS diaries) at 6:00 a.m. each morning for the 6 days leading up to a target exam. Participants were asked to respond to time-stamped diaries as quickly as possible after waking. In an effort to minimize response burden and capture the strongest anxiety-performance relationship [6], test anxiety was measured only twice: the morning before the exam and the morning of the exam. After completing the final SMS diary, instructors confirmed students' self-reported exam scores. At that time, students were also provided with a debriefing statement. Each student could receive up to 15 total extra credit points. The protocol for this study was approved by the University of Kansas Institutional Review Board (IRB).

Table 1 Descriptive data

	M/%	(SD)	1	2	3	4	5	6	7	8	9
Age	20.7	(2.94)									
Female	75%										
ESL	9%										
Variable name			1	2	3	4	5	6	7	8	9
Grade	79.36	(17.97)	– 0.14	0.04	0.17	– 0.18	– 0.24**	– 0.02	0.06	0.04	– 0.04
1. BDI	9.03	(7.37)		0.59**	– 0.12	0.41**	0.36**	– 0.16	– 0.16	– 0.19*	0.16
2. STAI-T	42.81	(8.77)			– 0.15	0.36**	0.35**	– 0.11	– 0.12	– 0.16	0.13
3. MEQ	44.41	(8.55)				– 0.27**	– 0.14	– 0.14	0.08	0.12	– 0.12
4. ISI	16.45	(6.07)					0.68**	– 0.10	– 0.37**	– 0.34**	0.30**
5. PSQI	6.62	(3.08)						– 0.05	– 0.38**	– 0.27**	0.25**
6. WASO	1.56	(0.32)							0.01	0.11	– 0.13
7. TST	415.52	(57.17)								0.25**	– 0.11
8. SE	91%	(0.10)									– 0.60**
9. SOL	34.27	(36.28)									

ESL English as a second language, *BDI* Beck Depression Inventory, *STAI-T* State Trait Anxiety Inventory, trait version, *MEQ* Morningness-Eveningness Questionnaire, *ISI* Insomnia Severity Inventory, *PSQI* Pittsburgh Sleep Quality Inventory, *Q&R* Average of Sleep Quality and Restfulness, *WASO* number of awakenings after sleep onset, *SE* sleep efficiency; *SOL* sleep onset latency

*Correlation is significant at the 0.05 level. **Correlation is significant at the 0.01 level

Method of Analysis

Data were analyzed in R using the “lavaan” package. The small amount of missing data was addressed using a saturated correlates model for full-information maximum likelihood (FIML). Because students were nested within classroom and semester, we investigated whether there it would be necessary to use a multilevel SEM model. In no instance did analyses reveal evidence of systemic between-cluster differences (composed either of instructor or semester).

Simulation data were used to estimate power [33]. Given the plausible range of effect sizes and range of probable covariates; $N = 150$ conferred sufficient power to detect a moderate sized simple effect of the primary predictors (i.e., Test Anxiety, Sleep Q&R, and TST). The total fraction of missing information in the data used for the final analysis model was 7.8%, with 11–14% for individual SMS diary items. This fraction of missing data is small and within the range where FIML can recover lost information and point estimates on which to base unbiased parameters estimates [34].

Results

Descriptive Data

Table 1 includes descriptive data and correlations among individual difference and demographic data. BDI scores were below clinical range, and STAI-T scores only slightly higher than recently reported college student samples [30, 35, 36]. However, there were indications of sleep problems, with ISI scores in the moderately severe range and PSQI total scores that exceeded the clinical cut off for sleep disturbance. On average, students fell into the “intermediate” range in terms of morning-evening tendencies. As shown in Table 1, correlations among individual difference variables were in expected directions (e.g., STAI-T, BDI, PSQI, and ISI were all positively correlated).

Primary Hypotheses

Analyses were constructed in a sequential manner, meaning that a simpler model was constructed first, that explained variability in final exam scores using individual difference variables (i.e., BDI, STAI-T, PSQI, ISI, MEQ chronotype, age, sex, race/ethnicity, and English as a Second Language (ESL), class performance motivation, first exam score). Non-significant predictors were pruned from the model based on Wald tests for predictor significance (seeTable 2 for detailed output).[1] A model constraining non-significant

[1] It should be noted that race/ethnicity was a significant predictor of exam score. However, the effect was entirely dependent upon a single student’s outlier (poor) test score. Because of the low probability for replication, this effect is not pictured in the model or reported in Table 2.

Table 2 Baseline model

Variable name	$\Delta\chi^2$	*Df*	*p* value
Significant predictors of academic performance			
ESL	14.511	1	< 0.001
First exam grade	20.228	1	< 0.001
STAI-test Day 6	17.195	1	< 0.001
Class performance motivation	4.756	1	0.029
Non-significant baseline predictors of academic performance			
BDI	0.009	1	0.925
PSQI	0.477	1	0.450
Trait anxiety	1.969	1	0.161
ISI	0.073	1	0.787
MEQ	0.638	1	0.424
Age	0.256	1	0.613
Sex	0.438	1	0.507

ESL English is a Second Language, *STAI* State Anxiety Inventory, *BDI* Beck Depression Inventory II, *PSQI* Pittsburgh Sleep Quality Inventory, *MEQ* Morningness, Eveningness Questionnaire (Chronotype), *ISI* Insomnia Severity Inventory

predictor slopes to zero was then analyzed and was determined to have excellent fit (CFI = 0.996, TLI = 0.994).

The modified baseline model with individual differences was then incorporated into the hypothesized path model, such that theorized sleep-anxiety-performance processes could be evaluated while simultaneously controlling for relevant individual difference variables. Variables pruned from the individual differences model that nevertheless could potentially impact steps in the sleep-anxiety-performance process were also included. The final model included the following individual difference variables: Age, Sex, Race/Ethnicity, ESL, First Exam Grade, and Class Performance Motivation; Class time (AM/PM); and ISI and Chronotype. To this, we added reports from the SMS diaries reflecting reports from night four (exam-2 days) and five (exam-1 day): Minutes studied, Sleep Q&R, WASO, TST, and STAI-Test (exam-1 only), and on the day of the exam STAI-test and the Final Exam Grade.

The final model largely supports the hypothesized conceptualization of the SAPP model and fit the data well (CFI = 0.963, TLI = 0.935, RMSEA 0.029). Figure 1 can be broken down into its component features: Higher Q&R preceded lower levels of STAI-test for both the night prior to the exam (χ^2 [1] = 7.17, $p = 0.007$) as well as two nights before (χ^2 [1] = 6.789, $p = 0.009$). Increasing levels of STAI-test had an adverse effect on Q&R (for STAI-test, the day prior to the exam, χ^2 [1] = 5.172, $p = 0.023$) as well as exam performance itself (for STAI-test on exam day, χ^2 [1] = 16.15524, $p \leq 0.001$). In fact, after controlling for prior exam performance and other predictors, test anxiety explained approximately 9.8% of variability in test performance, increasing the multiple *R*-squared for test performance from ~ 30 to ~ 40% significantly improving the fit of the model, $p < 0.001$. Although Sleep Q&R predicted test

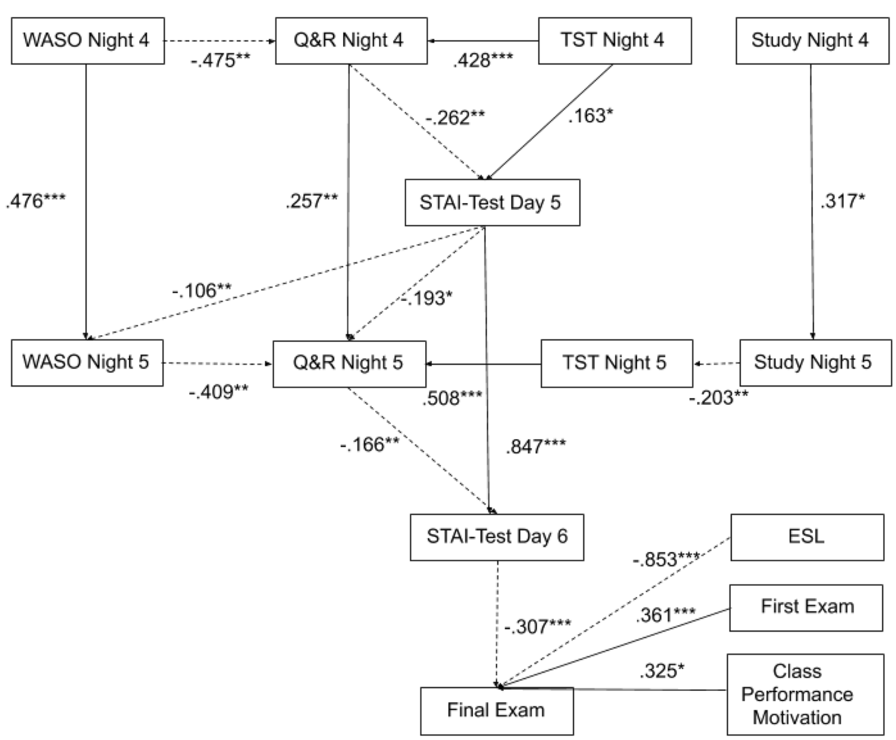

Fig. 1 The Sleep Anxiety Performance Process Model (SAPP). WASO number of awakenings after sleep onset, Q&R Sleep Quality and Restfulness, TST Total Sleep Time, STAI-Test, State Test Anxiety, ESL English as a Second Language, solid lines = positive path; dashed lines = negative path. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

anxiety, our results did not indicate that sleep was directly related exam performance. In fact, none of the sleep parameters (e.g. Q&R, TST, SOL, SE) had significant direct or lagged effects on the final exam score.

Few individual difference variables predicted exam scores. Neither symptoms of psychopathology, nor general demographic differences, nor general sleep dysfunction as measured by the ISI, chronotype, nor baseline sleep quality as measured by the PSQI mediated the relationship between sleep quality and anxiety, nor had any direct effects on exam performance. Individual difference variables that did predict exam score include first exam score, class performance motivation, and ESL. Each of these variables significantly predicted final exam scores in expected ways: Class performance motivation and first exam scores were positive indicators of Final Exam scores, whereas ESL students had significantly lower final exam scores.

Risk of Failure

As a follow-up analysis, we examined whether significant predictors of exam scores could be used to predict whether students passed or failed the exam. Using logistic regression, first exam grade, STAI-Test, ESL status, and Sleep Q&R were entered into an equation to predict whether students passed (score $\geq$ 70) or failed the exam (score < 70). Nearly one-third of the students failed their exam (27.86%). Single-unit increases in first exam scores were associated with greater odds of passing the exam (1.067, Wald 11.956, $p < 0.05$) and higher STAI-test scores reduced the chance of a passing score on the exam (0.943, Wald = 7.932, $p < 0.05$), as did speaking a first language other than English (14.017, Wald = 4.043, $p < 0.05$). Although this model explained a significant portion of the variance, Nagelkerke $R^2 = 0.33$, and correctly classified 77% of students, it proved less likely to identify true failures 44.1% than true passers 89%.

Discussion

Universities are struggling with the high rate of drop-out. Almost 40% of freshmen fail to return to their original college or university after the first year, many for academic reasons [31]. One significant contributor to academic failure is anxiety about academic performance, specifically test-taking anxiety. Students who are at increased risk for ineffective or maladaptive coping with the stressors of college may be even more vulnerable to the negative effects of anxiety on their academic performance and health-behaviors including sleep [3]. Consistent with this observation, this study found that students who experienced even moderate amounts of anxiety had lower test scores. Rather than having separate effects on exam performance, over the course of two days and nights, poor sleep quality predicted increased anxiety, which in turn perpetuated disrupted sleep, which then predicted further increases in anxiety and ultimately poor academic performance. Although the snapshot of the sleep-anxiety relationship captured here was not consistent with the specific a-priori predicted relationship (poor sleep mediating the relationship between test anxiety and exam performance), the model is consistent with the overall precept of the SAPP model, that sleep and anxiety feed one another, creating a vicious cycle that collectively impairs academic performance.

This was not a clinical sample. Nevertheless, test anxiety had a significant effect on test performance. A one-unit increase of test anxiety on the day of the test was associated with a 0.388-point lower test score. For students who were one standard deviation above the mean for anxiety, the effect size equates to a nearly 5-point lower (4.96) exam score. This could mean the difference between earning a D vs. a C-. Although the relationship between test anxiety and performance is fairly well understood, what was new in this study was documentation of the reciprocal relationship between sleep and anxiety. In particular, test anxiety and poor sleep quality appeared to operate as a positive feedback loop, with poor sleep disinhibiting anxiety, and those feelings and thoughts, in turn, disrupting sleep. Instead of poor sleep mediating the relationship between test anxiety and exam performance, anxiety, and sleep appeared to form a feed forward process, with test anxiety ultimately predicting exam performance.

Although poor sleep predicted test anxiety, none of the other sleep parameters (e.g., SOL, WASO, TST, SE) were either directly related to academic performance or test anxiety. Instead, in this sample of young adults, the effects of WASO and TST appeared to have been captured in the appraisal of poor sleep quality and restorativeness. Had we chosen a different focus of this model, we could have modeled WASO and TST as mediators of the relationship of poor sleep quality to test anxiety.

Other predictors of exam performance were primarily related to academic history. Prior exam performance was a strong predictor of future exam performance. However, this is not a "clean" academic covariate. Similar to what we observed during the 2 days prior to the statistics exam, we would expect that previous exams were also affected by the cyclical relationship between sleep and anxiety. Dispositional measures (e.g., Trait anxiety, depression, insomnia symptoms, circadian rhythms) were largely unrelated to final exam performance, underscoring the fact that this was a largely mentally healthy sample and that the effects of anxiety and sleep on performance were situational and not driven by general response biases or trait level effects. Non-native speakers had significantly lower scores than native speakers and were 14 times more likely to fail than a native speaker.

It should be noted that this was a not a planned analysis, only a small proportion of our sample (roughly 10%) were non-native speakers, and the student's level of English fluency was not measured. These caveats aside, the results of this study suggest the need for additional interventions to support non-native English speakers.

It is worth commenting on the relationship of motivation and study habits to academic performance. Perhaps it is not surprising that class performance motivation to do well in a class was related to actual performance, but was surprising that study habits did not relate to exam score. In fact, the only relationship of study time was to shorter sleep time the night before the exam. Follow-up exploratory analyses on study time showed that academic performance was not correlated with total minutes studied, average or variability, minutes of study during any day during the week, nor a curvilinear function. Examination of the data revealed no obvious sources of bias in reporting of study times. Although surprising, this finding is consistent with a body of work that shows little relationship between homework and achievement in elementary and secondary education [29]. As this is a single study and this was not a hypothesis driven finding, replication would be critical.

Test Anxiety and Broader Implications for Health

It is tempting to conceptualize test anxiety as a psychological state that has effects that are limited to academic performance and success. However, in other studies, highly test anxious students were found to have higher levels of stress, increased levels of maladaptive health behaviors such as increased caffeine intake, smoking, and lower levels of adaptive behaviors such as healthy eating, physical activity, and hygienic sleep [3]. The SAPP model may help to explain Oaten's findings by documenting that test anxious students may have poor sleep quality multiple days before an exam. Thus, it seems plausible that at least some of the effects of test anxiety on health behaviors (e.g., caffeine intake, healthy eating) observed by Oaten, were attempts to mitigate feelings of fatigue associated with poor sleep quality.

Limitations

We attempted to comprehensively measure the relationships among sleep, anxiety, and academic performance by using a combination of trait level measures and daily diary measures of sleep and anxiety. However, there were several notable limitations. First, non-invasive wrist worn actigraphs would have provided a more accurate measure of TST, WASO, and SOL. A wealth of literature has documented that sleep diary reports of sleep parameters do not correlate well with objective sleep measured via polysomnography [37]. Self-reported sleep often is more strongly related to measures of mood and anxiety and has a more modest association with biological and performance measures [38]. In addition, test anxiety and sleep were measured concurrently in order to minimize response burden for non-financially compensated research participants. However, morning and evening reports would reduce commonly recognized sources of bias (e.g., mood-dependent recall) [39]. Moreover, these were observational data, weakening our ability to make causal statements about the relationships observed in this study. And finally, generalizability is limited by class type (statistics) and should be made with care in universities with more racially and ethnically diverse undergraduate populations.

Summary and Clinical Application

The SAPP model offers a comprehensive way to understand the relationship between sleep and anxiety on academic performance. Test anxiety appears separate from trait anxiety and is thus likely to be more variable. One study that tracked the ebb and flow of test anxiety across the semester found that anxiety peaked at the beginning of the semester and dissipated as students become more familiar with the subject matter and test formats [7]. However, unremitting anxiety was among the most commonly identified mental health concern named by students leaving college [40]. Test anxiety has typically been thought of as an emotion that arises as students face a test. However, the SAPP model proposes that test anxiety does not start and stop on the day of an exam. Instead, the non-trait nature of test anxiety and the temporal relationship between sleep and text anxiety span at least 2 days leading up to the exam, which suggests a clear opportunity for intervention.

As universities and high schools struggle to retain students, interventions to address anxiety-related performance deficits may be part of a broader strategy to prevent drop out because of academic failure. Research on test anxiety interventions has primarily focused on the cognitive–behavioral methods, skill-building, and relaxation interventions to reduce test anxiety and improve test performance (e.g., [41]. However, individual or even group therapy for all test anxious students is not likely to be a practical solution to this problem. Fortunately, lower-burden interventions show promise. Unlimited exam-time for math and statistics exams [9] and a seven-minute expressive writing intervention directing students to write about "thoughts and feelings" about an upcoming exam have been shown to alleviate or minimize the effect of anxiety on a quantitative reasoning test [42]. However, the good news for test anxious students and for universities who wish to retain test anxious students is that interventions such as these are effective, easily delivered via instructional change, and come at no additional cost to the university.

Declarations

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

Conflict of Interest The authors declare that they have no conflict of interest.

References

1. Zeidner M. Does test anxiety bias scholastic aptitude test performance by gender and sociocultural group? J Pers Assess. 1990;55:145–60.
2. Gregor A. Examination anxiety: live with It, control it or make it work for you? Sch Psychol Int. 2005;26(5):617–35.
3. Oaten M, Cheng K. Academic examination stress impairs self-control. Journal of social and clinical psychology. J Soc Clin Psychol. 2005;24(2):254–70.
4. Gillespie NA, Gehrman P, Byrne EM, Kendler KS, Heath AC, Martin NG. Modeling the direction of causation between cross-sectional measures of disrupted sleep, anxiety and depression in a sample of male and female Australian twins. J Sleep Res. 2012;21(6):675–83.
5. Chapell MS, Blanding ZB, Silverstein ME, Takahashi M, Newman B, Gubi A, McCann N. Test anxiety and academic performance in undergraduate and graduate students. J Educ Psychol. 2005;97:268–74.
6. Macher D, Paechter M, Papousek I, Ruggeri K, Freudenthaler HH, Arendasy M. Statistics anxiety, state anxiety during an examination, and academic achievement. Br J Educ Psychol. 2013;83:535–49.
7. Hunsley J. Test anxiety, academic performance, and cognitive appraisals. J Educ Psychol. 1985;77(6):678–82. https://doi.org/10.1037/0022-0663.77.6.678
8. Onwuegbuzie AJ, Wilson VA. Statistics anxiety: nature, etiology, antecedents, effects, and treatments—a comprehensive review of the literature. Teach High Educ. 2003:195–209.
9. Onwuegbuze AJ. Satistics test anxiety and female students. Psychol Women Q. 1995;19:413–8.
10. Kopasz M, Loessl B, Hornyak, M, et al. Sleep and memory in healthy children and adolescents—a critical review. Sleep Med Rev. 2010;14:167–77.
11. Dimitriou D, Knight FLC, Milton P. The role of environmental factors on sleep patterns and school performance in adolescents. Front Psychol. 2015;6.
12. Trockel MT, Barnes MD, Egget DL. Health-related variables and academic performance among first year college students: implications for sleep and other behaviors. J Am Coll Health. 2000;49:125–31.
13. Borisenkov MF, Perminova EV, Kosova AL. Chronotype, sleep length, and school achievement of 11-to 23-year-old students in northern European Russia. Chronobiol Int. 2010;27(6):1259–70.
14. Taylor DJ, Vatthauer KE, Bramoweth AD, Roggero C, Roane B. The role of sleep in predicting college academic performance: is it a unique preictor? Behav Sleep Med. 2013;11(3):159–72.
15. Genzel L, Ahrberg K, Roselli C, al. e. Sleep timing is more important than sleep length or quality for medical school performance. Chronobiol Int. 2013;30(6):766–71.
16. Gomes AA, Tavares J, de Azevedo MH. Sleep and academic performance in undergraduates: a multi-measure, multi-predictor approach. Chronobio Int. 2011;28(9):786–801.
17. Uhde TW, Cortese BM, Vedeniapin A. Anxiety and sleep problems: emerging concepts and theoretical treatment implications. Curr Psychiatry Rep. 2009;11(4):269–76.
18. Blankstein KR, Flett GL, Watson MS, Koledin S. Test anxiety, self-evaluative worry, and sleep disturbance in college students. Anxiety Res. 1990;3(3):193–204.
19. Rosa RR, Bonnet MH. Reported chronic insomnia is independent of poor sleep as measured by electroencephalography. Psychosom Med. 1983;62:474–82.
20. Tsypes A, Aldao A, Mennin DS. Emotion dysregulation and sleep difficulties in generalized anxiety disorder. J Anxiety Dis. 2013;27(2):197–203.
21. Campbell R, Soenens B, Beyers W, Vansteenkiste M. University students' sleep during an exam period: the role of basic psychological dneeds and stress. Motiv Emot. 2018;42(5):671–81.
22. Ahrberg K, Dresler M, Niedermaier S, Steiger A, Genzel L. The interaction between sleep quality and academic performance. J Psychiatry Res. 2012;46(12):1618–22.
23. Alsaggaf MA, Wali SO, Merdad RA, Merdad LA. Sleep quantity, quality, and insomnia symptoms of medical students during clinical years: relationship with stress and academic performance. Saudi Med J. 2016;37(2):173–80.
24. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res. 1989;28(2):193–213.
25. Horne JA, Ostberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiol. 1976;4(2):97–110.
26. Morin CM, Belleville G, Bélanger L, Ivers H. The Insomnia Severity Index: psychometric indicators to detect insomnia cases and evaluate treatment response. Sleep. 2011;34(5):601–8.
27. Beck AT, Steer RA, Garbin MG. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. Clin Psychol Rev. 1988;8:77–100.
28. Spielberger CA, Gorsuch RL, Lushene RE. Manuel for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press; 1970.
29. Trautwein U, Köller O. The relationship between homework and achievement—still much of a mystery. Educ Psychol Rev. 2003;15(2):115–45.
30. Sadigh MR, Himmanen SA, Scepansky JA. An investigation of the prevalence of insomnia in college students and its relationship to trait anxiety. Coll Stud J. 2014;48:397–406.
31. Center NSCR. Snapshot report: Persistence: National Research Center 2016.
32. Carney C, E., Edinger JD, Manber R, Garson C, Segal ZV. Beliefs about sleep in disorders characterized by sleep and mood disturbance. J Psychosomatic Res. 2007;62:179–88.
33. Wolf EJ, Harrington KM, Clark SL, Miller MW. Sample size requirements for structural equation models: an evaluation of power. Bias Solut Propriety Educ Psychol Meas. 2013;76(6):913–34. https://doi.org/10.1177/0013164413495237.
34. Enders CK, Bandalos DL. The relative performance of full information maximum likelihood estimation for missing data structural equation models. Struct Equ Modeling. 2001;8(3):430–57.
35. Lindsey C. Trait anxiety in college students: the role of the approval seeking schema and separation individuation. Coll Stud J. 2014;48:407–18.

36. Clements AD, Bailey BA. The relationship between temperament and anxiety. J Health Psychol. 2010;15:515–25.
37. Matthews KA, Patel SJ, Pantesco EJ, et al. Similarities and differences in estimates of sleep duration by polysomnography, actigraphy, and self-reported habitual sleep in a community sample. Sleep Health. 2018;4(1):96–103.
38. Bernstein JPK, DeVito A, Calamia M. Subjectively and objectively measured sleep predict differing aspects of cognitive functioning in adults. Arch Clin Neuropsychol. 2019;34:1127–37.
39. Bower GH. Mood and memory. Am Psychol. 1981;36:129–48.
40. Grassgreen A. Students rate mental health services. Inside High Educ 2012.
41. Ergene T. Effective interventions on test anxiety reduction. Sch Psychol Int. 2003(24):313 28.
42. Park CL, Cohen LH, Murch RL. Assessment and prediction of stress-related growth. J Personal. 1996;64:71–105.