

RESEARCH ARTICLE

The association between depressive symptoms and self-reported sleep difficulties among college students: Truth or reporting bias?

Zhiyong Huang¹ , Fabrice Kämpfen² *

1 Southwestern University of Finance and Economics, Chengdu, China, **2** Population Studies Center, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America

 These authors contributed equally to this work.

* kampfenf@sas.upenn.edu



OPEN ACCESS

Citation: Huang Z, Kämpfen F (2021) The association between depressive symptoms and self-reported sleep difficulties among college students: Truth or reporting bias? PLoS ONE 16(2): e0246370. <https://doi.org/10.1371/journal.pone.0246370>

Editor: Jacob Resch, University of Virginia, UNITED STATES

Received: September 14, 2020

Accepted: January 18, 2021

Published: February 19, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0246370>

Copyright: © 2021 Huang, Kämpfen. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [Supporting information](#) files.

Abstract

The strong association between self-reported sleep difficulties and depressive symptoms is well documented. However, individuals who suffer from depressive symptoms could potentially interpret the values attached to a subjective scale differently from others, making comparisons of sleep difficulties across individuals with different depressive symptoms problematic. The objective of this study is to determine the existence and magnitude of reporting heterogeneity in subjective assessment of sleep difficulties by those who have depressive symptoms. We implement an online survey using Visual Analogue Scales and anchoring vignettes to study the comparability of subjective assessments of sleep difficulties among college students in Switzerland ($N = 1, 813$). Using multivariate linear regressions and double-index models, our analysis shows that reporting heterogeneity plays only a marginal role in moderating the association between sleep difficulties and depression, irrespective of the severity of the depressive symptoms of the individuals. This suggests that unadjusted comparisons of self-reported sleep difficulties between college students are meaningful, even among individuals with depressive symptoms.

Introduction

Sleep is associated with a wide range of mental health outcomes. Persons with higher sleep quality have greater well-being and better psychological functioning [1–9]. Insomnia and other sleep difficulties, on the other hand, are strong predictors of depression, anxiety, stress, and neuroticism, to name a few [1].

The study of the relationship between sleep difficulty and mental health is of particular relevance among college students as they transition into adulthood. College students, and more generally young adults (18–25 years of age), have been shown to be particularly prone to depression because this is a period in the life course that coincides with many important challenging transitions to adulthood such as separation from family, financial independence and

Funding: This study was funded by the HEC Research Fund 2017 - University of Lausanne, Switzerland (<https://www.unil.ch/hec/en/home.html>) awarded to FK and ZH. This study was also supported by the National Natural Science Foundation of China under the project "Quality of Life Research: Based on Capacity Approach: A Population Study in Southwestern China" (Project number 71804151) awarded to ZH. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

increased work-load [10–13]. At the same time, young adulthood is a period in which there are great shifts in circadian rhythm and sleep patterns [14, 15]. Problems with sleep length among college students have been associated with poor mental health such as depression and anxiety [16–21], suicidal thoughts and self-harm behaviors [22]. Given the increasing prevalence of depression among young adults in the recent years [23] and the fact that depression in young adults constitutes an important risk factor for depression and bipolar disorder during adulthood [24, 25], having a clear understanding of the association between sleep difficulties and depression is crucial if one wants to formulate lifestyle guidance and design policies to promote healthy development of college students and young adults.

Subjective sleep quality measures have been a central feature to assess insomnia, non-restorative sleep and sleep health in general [26]. Several tools and instruments have been implemented to assess the subjective sleep quality and corresponding sleep difficulties of individuals. Questionnaires such as the Pittsburgh Sleep Quality Index (PSQI), the Insomnia Severity Index, the Athens Insomnia Scale, Epworth Sleepiness Scale (ESS), the Leeds Sleep Evaluation Questionnaire (LSEQ), to name a few, are commonly used as effective means to aid diagnosis of sleep difficulty [27–31]. Moreover, single-item questions are commonly employed in large-scale population health surveys such as the Health and Retirement Study (HRS) in the USA [32], the Survey of Health, Aging and Retirement in Europe (SHARE) [33], the WHO Study on Global Aging and Adult Health (SAGE) [34], the China Health and Retirement Longitudinal Study (CHARLS) [35] to assess individual's sleep difficulties in non-clinical settings.

One particular characteristic of all these instruments is that they are based on self-reports and may therefore suffer from subjective bias. It has been shown that there exist significant differences between objective measures, such as polysomnography and actigraphy, and subjective measures of sleep difficulties [26, 36–40]. Buysse et al. (2008) [36] have attributed the weak association between subjective and objective measures of sleep to whether one assesses habitual patterns of sleep versus sleep on a discrete occasion. Baker et al. (1999) [39] have argued that the weak correlation between objective and subjective measures are due to between-individual variation while Jean-Louis and Kripke (2000) [40] have speculated that the subjective measures may be biased by the depressive mood of respondents.

As documented, inconsistencies between subjective and objective measures of sleep difficulties can arise because of several reasons, one of them being systematic between-individual reporting heterogeneity. Reporting heterogeneity, also known as differential item functioning (DIF) in the literature, arises when individuals with different characteristics interpret the possible values attached to a scale differently from one another, i.e., when they adopt different reporting scales to rate their own characteristics. For instance, one of the items of the PSQI, which is the most widely used sleep health assessment tool in both clinical and non-clinical settings [41], is phrased as "During the past month, how would you rate your sleep quality overall?", with possible answer being "very good", "fairly good", "fairly bad" and "very bad". In the case of individuals with depression, the subject of our study, one critical concern would be that respondents with depression might have very different interpretation of the words "good", "bad", "fairly" and "very". Individuals with depressive symptoms might therefore report poor sleep quality, either because they are indeed deprived of sleep as a result of depression, or simply because they are inclined to rate everything more negatively, making subjective sleep quality assessments between individuals with and without depressive symptoms hard to compare [42].

As another example, LSEQ, which is very commonly used to assess sleep difficulties [43], contains a Visual Analogue Scale (VAS) to measure sleep difficulties. VAS is a measurement instrument that is used to rate a given characteristic over a continuum of values (a horizontal or vertical line) with "anchors" at the two end-points, representing the best and worst

scenarios. The anchors in LSEQ are “more difficult than usual” as one endpoint and “easier than usual” as another. Both of these endpoints can be interpreted very differently by individuals with different depression level. And because respondents use these anchors to evaluate their sleep difficulty, the variations in the interpretations of these anchors across groups of respondents could potentially mean that they use different reporting scales, making comparisons between self-reported measures of sleep difficulty across these groups hard to interpret. Other factors than depression, such as sex and income, can also bias self-reported sleep difficulty if respondents interpret the scale of the sleep difficulty measure differently from others [44].

Because it is important to understand whether the association between a subjective measure such as sleep difficulty, and its important risk predictors –depressive symptoms in our case– still holds after controlling for reporting differences among individuals, researchers have designed techniques to account for reporting heterogeneity in self-assessed measures. One example of such techniques is anchoring vignettes [45–47], which are short descriptions of hypothetical individuals that are assessed by respondents alongside their self-assessment of the same domain. For example, an anchoring vignette on sleep patterns included in SAGE is profiled as: “[Mark] falls asleep every night within five minutes of going to bed. He sleeps soundly during the whole night and wakes up in the morning feeling well-rested.” Respondents are asked to evaluate the sleep quality of this hypothetical person alongside their own sleep quality.

As anchoring vignettes are predefined and invariant across individuals, their ratings are supposed to only reflect differences in reporting scales, which in turn allows to identify DIF among respondents. Anchoring vignettes can therefore be used to adjust subjective measures by purging reporting heterogeneity from the self-assessed measures, making the adjusted self-assessment better comparable across individuals. Anchoring vignettes have been used in many research areas such as health [47–54], healthcare [55, 56], subjective well-being and quality of life [46, 57–60], political efficacy [45, 61], job satisfaction [62], working disability [63] and social status [64].

The objective of this study is to determine the existence of reporting heterogeneity in subjective assessment of sleep difficulties among college students, and explore whether this reporting heterogeneity is explained by differences in depressive symptoms as measured by the standard and validated PHQ-9 instrument. We use data collected from a sample of college students in the region of Lausanne in Switzerland whom we asked to complete an online survey. Online surveys are particularly well-adapted for this kind of research question [63] and have been used in many other studies to assess reporting heterogeneity in self-reported evaluations [63, 65, 66]. We hypothesize that college students with and without depressive symptoms adopt different reporting scales when assessing their sleep difficulties, to an extent that is hard to evaluate a priori. To test our hypotheses, we first document the unadjusted associations between depressive symptoms and self-assessed sleep difficulties. We then exploit five sleep-related anchoring vignettes that are derived from the existing literature to adjust for reporting heterogeneity in the relationship between self-assessed sleep difficulties and depressive symptoms. The econometric specification we put in place will allow to (i) disentangle the “true” difference in sleep difficulties from reporting heterogeneity and (ii) uncover the individual characteristics that are driving this reporting heterogeneity. In addition to various socio-demographic information that we collected as part of our survey, our econometric models will also control for various measures of sleep patterns that could confound the relationship between sleep difficulties and depressive symptoms, as explained below.

Although some studies have examined the association between depressive symptoms and self-assessed sleep quality and difficulties [2, 67, 68], very few have attempted to assess the existence of reporting heterogeneity in subjective sleep assessment and to correct for it. Tareque

et al. (2016) [44] seems to be the only exception. However, Tareque et al. (2016) [44] do not look specifically at depression as the main source of reporting heterogeneity and they measure sleep difficulties on a Likert scale, making interpretation of the magnitude of reporting heterogeneity difficult. Using self-assessed sleep difficulties and anchoring vignettes measured on VAS, we can evaluate the magnitude, if any, of reporting heterogeneity in self-assessed sleep difficulties. In addition, our study is the first to investigate the existence of reporting heterogeneity in subjective assessment of sleep difficulties among college students, which, as discussed above, is a strata of the population that is particularly prone to both sleep difficulties and mental disorders.

Methods

Participants

We recruited our sample of college students from three schools of higher education in Canton de Vaud in Switzerland (the Swiss Federal Institute of Technology in Lausanne (EPFL), the University of Lausanne (UNIL) and Ecole hoteliere de Lausanne (EHL)). College students from these three schools can register to the Laboratory of Behavioral Experiments program at the University of Lausanne and participate in experiments and surveys. The online questionnaire we used was created on Qualtrics and sent out to all students who registered in the University of Lausanne's experiment program. The invitation email we sent to students and the introductory page of our online survey did not mention the specific aim of the study but referred to it as a study on their quality of life in general. 6,578 emails were sent, out of which 1,938 students completed our survey in the following two weeks (response rate = 29.5%). For their participation, students who completed our online survey were automatically entered into a lottery for which the highest prize was about USD 300. The experimental protocol and informed consent were approved by HEC Ethics Committee of the University of Lausanne in February 2017, and all subjects gave informed consent.

Out of the 1,938 college students who completed our online questionnaire, we kept those who were between 18 and 25 years old to capture "standard" college progression. This age restriction also corresponds to late adolescent/young adulthood [69], which represents a period in life with important transitions from adolescence to adulthood [70] that coincide with major changes both in mental health and sleep patterns/circadian rhythm [10–15]. Out of 1,938 students who answered our survey, only 3 respondents were 17 years old (as compared to 160 college students who were 18 years old), while 78 respondents were 25 years old, as opposed to 39 respondents who were 26. After discarding 125 observations that were out of the age range, our final sample consisted of 1,813 observations.

Survey design and measurements

After completing the consent forms, college students were first asked to evaluate our set of five sleep-related anchoring vignettes using a VAS ranging from 0 ('no difficulty') to 100 ('extreme difficulty'). We then asked them to evaluate their own sleep difficulties, using the same VAS. As discussed in Hopkins and King (2010) [61], asking individuals to evaluate anchoring vignette first allows to prime respondents into interpreting the self-assessment question in a similar light and defining the response scale in a common way. The self-assessment question was then followed by a series of more general questions on socio-demographic characteristics from which we derive information to create our set of control variables.

Measures on subjective sleep difficulties and anchoring vignettes. Our subjective measure of sleep difficulties was derived from the respondents' answer to the question "Overall in the last 30 days, how much of a problem did you have with sleeping, such as falling asleep,

waking up frequently during the night or waking up too early in the morning?”, which is directly taken from the SAGE/World Health Survey designed by the World Health Organization (WHO) [71] and very similar in essence to Tareque et al. (2016) [44] to facilitate comparisons. Respondents had to use a VAS that was operationalized online as a horizontal graphic slider on which they could select any point among the continuum set of possible values ranging from 0 (‘no difficulty’) to 100 (‘extreme difficulty’). These two anchors are the corresponding labels to the Likert scale used by the WHO in the World Health Survey (‘none’ and ‘extreme’) [71]. Only the minimum and maximum values at both ends of the horizontal graphic slider, 0 and 100, were displayed and no tick points in between were shown.

The five sleep-related anchoring vignettes we use in our study are taken from the World Health Survey [71] developed by the WHO as well. Widely used and validated, the original WHO vignettes were intended for adult populations and not for college students. Hence, to fit our college student population, we made a trivial change to one of the vignettes in which the expression “s/he is late to *work* 4 days out of 5” in the original text was modified to “s/he is late to *school* 4 days out of 5”. The vignettes that were developed by WHO and that we use in our study include several factors associated with sleep difficulties such as time it takes for individuals to fall asleep, the quality of the sleep, the difficulty the person has to wake up as well as whether the person feels well-rested in the morning. Moreover, the vignettes respondents have to evaluate are sex-specific, matching the sex of the respondent to the sex of the fictitious persons described in the hypothetical scenarios. As documented by Jürges and Winter (2013) [72], the sex of the persons described in the vignettes could potentially be a confounder when vignettes are evaluated. Therefore, survey including vignettes usually either assign gender-specific vignettes to respondents [45, 49, 73, 74] or control for sex in all estimations [63, 65]. We follow the literature and design sex-specific vignettes and control for sex in our analysis. The five vignettes we use in our study can be found in [S1 Appendix](#).

In addition to self-assessed sleep difficulties and the five vignette evaluations, college students were further asked questions about their sleep patterns, such as the number of effective hours of sleep they had per night on average over the past 30 days, at what time they usually went to bed, how long it took them to fall asleep and what time they woke up in the morning. All these factors are strongly correlated with both depression and sleep difficulties [75] and can therefore confound the association between these two variables. We will control for these factors in our econometric models and show that adding these sleep patterns in our specifications do not alter our results.

Measures on depressive symptoms. We assess the presence and severity of depressive symptoms in our sample of college students using the standard Patient Health Questionnaire-9 (PHQ-9). The PHQ-9 is a validated 9-item instrument that is frequently used to screen for depression in primary care and non-clinical settings [76–78], including the screening of depression among college students [79, 80]. It is worth noting that its computerized version has also been validated [81]. The PHQ-9 can be found in [S2 Appendix](#).

Its score ranges from zero to 27. Following Kroenke et al. (2001) [82], we categorize our respondents as having “no or minimum symptoms” if their score is between 0 and 9, “mild depressive symptoms” if between 10 and 14, “moderate depressive symptoms” if between 15 and 19 and “severe depressive symptoms” if above 19 points. The Cronbach value of the PHQ-9 in our sample equals to 0.791, indicating good reliability of the items to measure the same construct [83]. As a robustness check, we present results when PHQ-9 score instead of depressive symptoms is used as a measure of mental health. We show that our results when analyzing the presence of depressive symptoms are similar when analyzing depressive symptom severity.

Measures on socio-demographic characteristics. We obtained information on our respondents about their age, nationality, relationship status (in couple or single), quality of the

relationship with their parents as well as their parents' education level, family income, school performance, and the number of close friends. These are factors that can explain both sleep difficulties and depression [20, 84] and we hence include these characteristics as control variables in our econometric specification.

Statistical methods

To gain a preliminary understanding of the existence and strength of the correlation between sleep difficulties and depressive symptoms in our college student population, we first use Epanechnikov kernel densities to investigate the bivariate relationship between these two variables. We further investigate the relationship between sleep difficulties and depressive symptoms using multivariate linear regressions, including various socio-demographic characteristics and self-reported measures of sleep patterns that are not subject to reporting heterogeneity, e.g., hours of sleep, as control variables. The use of linear specifications is supported by various tests we implement and report in [S5 Appendix](#).

We then explore two different ways to determine the existence and influence, if any, of reporting heterogeneity in subjective sleep difficulties. Perhaps the first and most straightforward way to control for reporting heterogeneity in our setting is to include vignette evaluations as additional controls in the linear regressions. This assumes that respondents use the same reporting scale when evaluating their own sleep difficulties as the scale they use when evaluating the sleep difficulties of the person described in the anchoring vignettes. This strategy, however, does not allow us to determine what the characteristics of the individuals that drive reporting heterogeneity are. Another way to disentangle reporting heterogeneity and sleep difficulties is to perform a double-index model [60]. This model exploits respondents' evaluations of sleep difficulties for themselves and for vignettes, assuming that individuals use the same reporting scales. Under this assumption, one can control for reporting heterogeneity as one controls for individual fixed effects in panel data models (where several observations are recorded for the same individual over time). The advantage is that the model allows us to not only control for reporting heterogeneity but also estimate the characteristics of the individuals that drive reporting heterogeneity.

More formally, we assume y_{i0}^* , the latent level of sleep difficulties of respondent i , can be modeled as a linear function of a set of factors x_i' subject to an error term ϵ_{i0} :

$$y_{i0}^* = x_i' b_0 + \epsilon_{i0} \quad (1)$$

When the level of sleep difficulties y_{i0}^* is self-reported, the measurement can be potentially biased due to reporting heterogeneity, i.e., respondents with different characteristics report value of y_{i0}^* in systematically different ways. By assuming that reporting heterogeneity can take the form of an additive and unobserved individual effect c_i , the reported value of sleep difficulties of respondent i , y_{i0} , takes the form:

$$y_{i0} = y_{i0}^* + c_i \quad (2)$$

In addition to these assumptions, the use of anchoring vignettes to correct for reporting heterogeneity requires the very common vignette-related assumptions, namely, vignette equivalence and response consistency [45, 47, 55, 62, 63]. Vignette equivalence assumes that vignettes are perceived in the same way by all respondents up to an idiosyncratic error term. Under this assumption, the perception of vignette j is:

$$y_{ij}^* = b_j + \epsilon_{ij} \quad \text{for } j \geq 1 \quad (3)$$

where b_j is the location of vignette j . On the other hand, vignette consistency assumes that

respondents' evaluations of the anchoring vignettes are subject to the same reporting heterogeneity as their self-reported variable of interest (see Eq 2), i.e.,:

$$y_{ij} = y_{ij}^* + c_i \quad \text{for } j \geq 1 \quad (4)$$

By defining c_i as a linear function of x_i , i.e., $c_i = x_i'\gamma$, we can plug in y_{i0}^* into Eq 2 and y_{ij}^* into Eq 4 to obtain the following system of equations:

$$y_{i0} = x_i'b_0 + x_i'\gamma + \epsilon_{i0} \quad (5)$$

$$y_{ij} = b_j + x_i'\gamma + \epsilon_{ij} \quad \text{for } j \geq 1 \quad (6)$$

We can then simultaneously estimate this system of equations with OLS, where b_0 represents the “true” effect of x on subjective sleep difficulties and γ represents the effects of x on reporting heterogeneity c .

Results

Descriptive statistics

Table 1 reports the descriptive statistics of our study sample. The average score of sleep difficulty is 37.7, on a scale where 0 represents “no difficulty” and 100 “extreme difficulty”. Up to 16%, 6% and 1.8% of the college students in our sample have mild, moderate, or severe depressive symptoms, respectively, as reflected by their PHQ-9 score. Females account for 51% of our sample. More than half (52%) of our respondents are Swiss, and a significant portion of them are coming from Western Europe (35%) and Africa (7.3%). Most of the respondents are undergrad students (82%). 34% of our respondents are from families that are making between USD 7,000 and 12,000 per month, which is comparable to the average gross income per month per household in Switzerland in 2014 (about USD 10,0798). 51% of our respondents claim they are single at the time of the online survey, 2.4% admit to have no close friends and 7.7% have no siblings. Among our respondents, about 3% of them evaluate the relationship with their parents as being bad. More than half of them consider their school performance as average and another 13% think their performance to be below average.

Fig 1 shows the Epanechnikov kernel density of the self-assessed sleep difficulty of the college students in our sample. This distribution appears to be highly bimodal, with the middle inflection point at about 40. The majority of respondents in our sample appear to have mild sleep difficulties, where the density distribution is the highest at roughly 15. There is also a large share of our respondents who appear to have difficulties with their sleep, as evidenced by the large number of evaluations that are above 50.

Fig 2 reports the distributions of the five anchoring vignette evaluations. We can see that our five vignettes cover the range of possible sleep difficulty values very well. The first plot in the first row represents the evaluation by our respondents of the first vignette. There appears to be unanimity across our respondents: the sleep pattern of the fictitious person described in vignette 1 represents an ideal scenario (most of the density is located at 0). At the complete opposite, the middle plot in the second row represents the evaluation by our respondents of vignette 5. Although somewhat more skewed towards the left, most of our sample evaluated scenario 5 as representing the scenario of someone experiencing the most severe difficulty with one's sleep pattern. Vignettes 2, 3, and 4 represent mid-range scenarios. It is interesting to note here that the evaluations of vignette 2 appear to be quite discordant, as revealed by the density which, in addition to displaying bimodality, is relatively high from value 20 to 70.

Table 1. Descriptive statistics (N = 1, 813).

Variable	Mean	SD	Min	Max
Sleep difficulty (VAS score)	37.7	27.4	0	100
Depressive symptoms (based on PHQ-9 score)				
No or minimum symptoms (reference group)	0.76	0.43	0	1
Mild	0.16	0.37	0	1
Moderate	0.060	0.24	0	1
Severe	0.018	0.13	0	1
Female	0.51	0.50	0	1
Age (in years)	21.0	1.85	18	25
Place of origin				
Swiss (reference group)	0.52	0.50	0	1
Western Europe	0.35	0.48	0	1
Eastern Europe	0.030	0.17	0	1
Africa	0.073	0.26	0	1
Others	0.033	0.18	0	1
Education				
Bachelor (reference group)	0.82	0.38	0	1
Master	0.16	0.36	0	1
Others	0.020	0.14	0	1
Family income				
Less than USD 4000 (reference group)	0.081	0.27	0	1
(4000,7000]	0.21	0.41	0	1
(7000,12000]	0.34	0.47	0	1
More than 12000	0.26	0.44	0	1
Don't know	0.10	0.30	0	1
School performance				
Above average (reference group)	0.27	0.45	0	1
Average	0.60	0.49	0	1
Below average	0.13	0.34	0	1
Single (not in a relationship)	0.51	0.50	0	1
Close friends				
None (reference group)	0.024	0.15	0	1
One	0.041	0.20	0	1
Two	0.16	0.37	0	1
Three or more	0.78	0.42	0	1
Siblings				
None (reference group)	0.077	0.27	0	1
One	0.46	0.50	0	1
Two	0.31	0.46	0	1
Three or more	0.16	0.36	0	1
Relationship with parents				
Good (reference group)	0.66	0.47	0	1
Good, but not always	0.31	0.46	0	1
Bad	0.031	0.17	0	1

Note: Unweighted sample characteristics of the students who have completed our online survey in April 2017.

<https://doi.org/10.1371/journal.pone.0246370.t001>

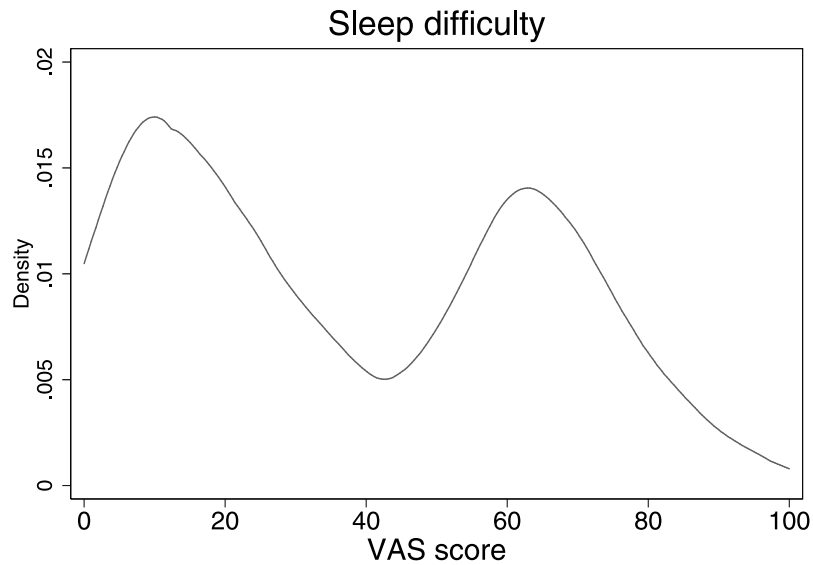


Fig 1. Distribution of sleep difficulty evaluations. Note: 0 means ‘no difficulty’ and 100 means ‘extreme difficulty’ with sleep.

<https://doi.org/10.1371/journal.pone.0246370.g001>

Relationship between depressive symptoms and sleep difficulties

We start our analysis by examining the bivariate relationship between self-assessed sleep difficulties and depressive symptoms. Fig 3 shows the levels of self-assessed sleep difficulties of respondents in our sample by depressive symptoms status. As can be seen from the densities, the bimodality of the distribution exists for all four depressive symptoms categories, although

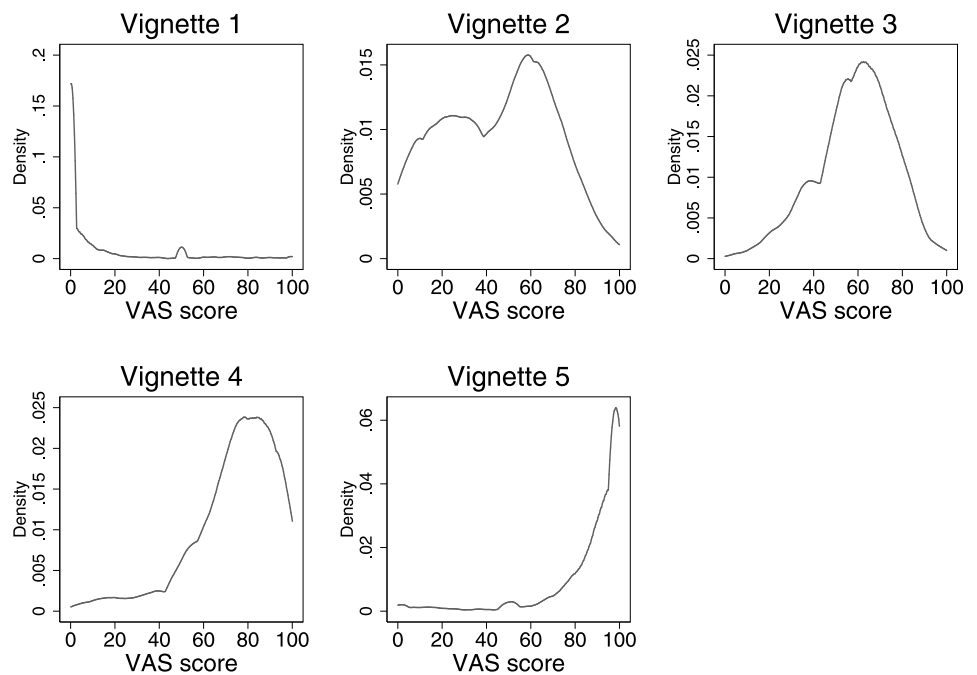


Fig 2. Distribution of vignette evaluations. Note: 0 means ‘no difficulty’ and 100 means ‘extreme difficulty’ with sleep.

<https://doi.org/10.1371/journal.pone.0246370.g002>

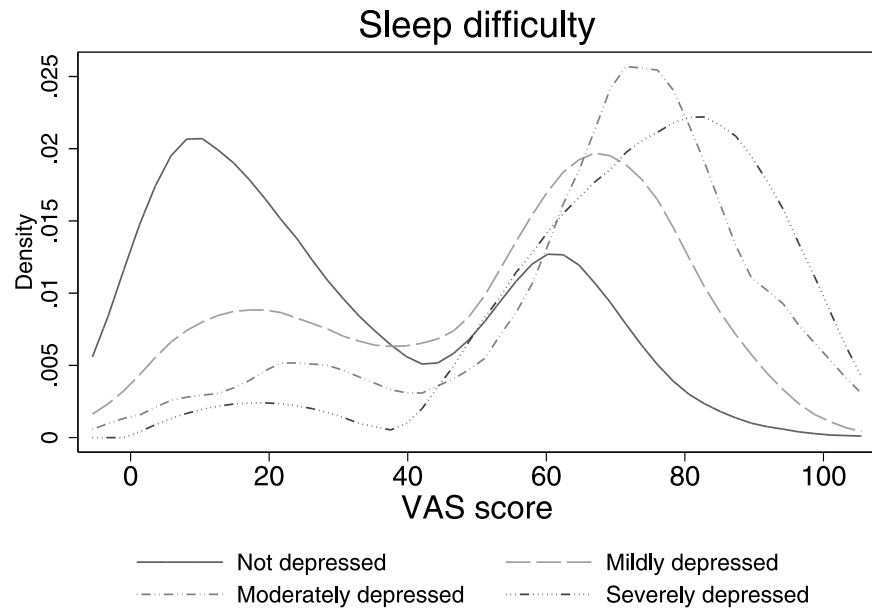


Fig 3. Distribution of sleep difficulty evaluations by depressive symptoms. Note: 0 means 'no difficulty' and 100 means 'extreme difficulty' with sleep.

<https://doi.org/10.1371/journal.pone.0246370.g003>

it is somewhat less pronounced for those who are moderately depressed and those who are severely depressed. The relationship between sleep difficulties and depressive symptoms is very strong: almost all individuals suffering from depressive symptoms, irrespective of the severity, have difficulties with their sleep, with the density shifting towards 100 the more severe the depressive symptoms are.

Results from a multivariate linear regression of self-assessed sleep difficulties measured from 0 to 100 on our set of control variables are reported in the first column of Table 2. Sex, age, income and relationship status appear to not affect sleep difficulties. Note that we report in the tables only the coefficients associated with sex, age, and depression symptoms status, as they are the main focus of this study. Details on the coefficient of the other variables are available upon request to the authors. Individuals suffering from 'mild', 'moderate' and 'severe' depressive symptoms however have experienced great difficulties with sleep as they have rated their sleep difficulties 20, 32, and 37 points higher, respectively, compared to college students with 'no or minimal' symptoms. These effects are all significant at the level of 1% ($p < 0.01$).

One may wonder whether the relationship between sleep difficulty and depressive symptoms could be confounded with sleep time and various patterns of sleep. Indeed, the association could be spurious if it turns out that sleep time or any other variables affecting both sleep difficulties and depressive symptoms are not controlled for in the analysis.

Columns 2, 3, 4, 5, and 6 of Table 2 however show that controlling for various measures of sleep patterns and efficiency that are commonly used in the literature does not critically affect the relationship between sleep difficulties and depressive symptoms. Descriptive statistics of these variables are presented in S1 Table. Controlling first for the reported effective hours of sleep of the respondents (column 2) [67, 75, 85, 86], which is derived from the question: *Over the last month, how many hours of sleep have you had on average every night?*, one can see that the association between depressive symptoms and sleep difficulties remains very large and significant. For instance, the increase in sleep difficulties for those suffering from severe depressive symptoms, compared to those with no depressive symptoms, only changed from 37.4

Table 2. Linear regressions of sleep difficulties on our set of control variables.

	(1) Sleep difficulty	(2) Sleep difficulty	(3) Sleep difficulty	(4) Sleep difficulty	(5) Sleep difficulty	(6) Sleep difficulty
Female	1.24 (1.20)	1.58 (1.18)	1.29 (1.21)	1.62 (1.20)	0.87 (1.14)	0.21 (1.09)
Age	0.20 (0.39)	0.078 (0.38)	0.16 (0.39)	0.20 (0.39)	0.13 (0.37)	0.14 (0.36)
Mild depressive symptoms	19.6*** (1.69)	18.0*** (1.67)	19.7*** (1.69)	19.2*** (1.68)	15.6*** (1.66)	14.1*** (1.57)
Moderate depressive symptoms	32.4*** (2.40)	30.8*** (2.19)	32.5*** (2.38)	31.2*** (2.34)	21.9*** (2.52)	24.7*** (2.27)
Severe depressive symptoms	37.4*** (3.69)	35.7*** (3.59)	37.6*** (3.66)	36.8*** (3.61)	25.9*** (4.28)	28.5*** (3.40)
Hours of sleep (reported)		-5.26*** (0.59)				
Hours of sleep (wake up time—bed time)			-0.77 (0.56)			
Hours of sleep, adjusted for time it takes to fall asleep				-2.95*** (0.54)		
Time to fall asleep (in hours)					25.8*** (2.71)	
Take more than 20 mins to fall asleep						17.7*** (1.17)
Sleep less than 7 hours						6.76*** (1.42)
Low sleep efficiency						9.07*** (1.48)
Constant	17.5* (9.93)	59.9*** (10.9)	24.6** (10.6)	39.7*** (10.5)	9.68 (9.88)	11.9 (9.18)

Note: Robust standard errors are in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). We also control for parent's income, education level, nationality, number of siblings, relationship status, relationship with parents, number of close friends and school performance. These coefficients are not reported in the table but are available upon request.

<https://doi.org/10.1371/journal.pone.0246370.t002>

points ($p < 0.01$) to 35.7 ($p < 0.01$). Controlling next for the hours of sleep defined as the difference between the time a person wakes up and the time he or she goes to bed [75], the effect of depressive symptoms again remains almost identical (column 3). In column 4, we control again for hours of sleep as defined in column 3, but this time correcting for the time it takes for respondents to fall asleep [75, 85]. Again the effect of depressive symptoms on sleep difficulties is very robust: for instance, the increase in sleep difficulties for those suffering from moderate depressive symptoms, compared to those with no depressive symptoms, only change from 32.4 points ($p < 0.01$) to 31.2 ($p < 0.01$), which corresponds to a change of only about 4%. The association also holds when we control only for the time it takes individuals to fall asleep instead (column 5) [67, 75, 87]. Two things are worth noting here. The first is that this variable captures some of the depressive symptoms-sleep difficulties gradients as the coefficients associated with depressive symptoms decreased by about 30%. Second, the time it takes individuals to fall asleep explains quite a lot of the sleep difficulties variation, “independently” of depressive symptoms status, given the large and statistically significant effect of the variable on sleep difficulties (about 25.8 points, $p < 0.01$).

Finally, in column 6, we control for whether it takes more than 20 minutes for individuals to fall asleep (the usual cutoff point for sleep latency [2]) and whether college students sleep less than 7 hours per night on average [2]. In the same regression, we also add an indicator variable that reflects the sleep inefficiency of an individual as defined as taking the value of 1 if the ratio of the number of effective hours of sleep over the difference between the waking up time and the time the individual goes to bed is lower than 0.85 [75, 88–90]. Results show that including these three indicator variables in our statistical specification, although they are individually all positive and highly significant ($p < 0.01$), do not appear to attenuate the associations between depressive symptoms and sleep difficulties.

Note that our results are robust to using PHQ-9 score as a continuous measure instead of descriptive depression categories. Indeed, as evidenced in [S2 Table](#), the effect of a marginal increase in PHQ-9 score on sleep difficulty is equal to about 3 points ($p < 0.01$) and remain positive and statistically significant even after including our various measures of sleep patterns and efficiency as control variables in the econometric specification.

Sleep patterns and hours of sleep, therefore, do not seem to explain the association between depressive symptoms and sleep difficulties as the association still appears to be strong and statistically significant after controlling for these different confounding variables. Could the association between depressive symptoms and sleep difficulties be explained, at least partially, by reporting heterogeneity? As explained above, individuals with depressive symptoms could report severe sleep difficulties, either because they are indeed deprived of good sleep as a result of depression or simply because they are apt to rate everything more negatively. Reporting heterogeneity could indeed stem from the fact that the meaning of the two end-points of our VAS, ‘no difficulty’, and ‘extreme difficulty’, could be interpreted very differently depending on whether a person suffers from depression or not. We now focus our analysis on reporting heterogeneity and try to answer this question.

Reporting heterogeneity in sleep difficulties

As we are most concerned about whether individuals with depressive symptoms could report sleep difficulty on a different scale from others, we start our analysis about the presence of reporting heterogeneity by exploring the bivariate relationship between vignette evaluations (which measure reporting scale) and depressive symptoms. [Fig 4](#) shows the distribution of the evaluations of the five anchoring vignettes used in our study by depressive symptoms severity. If the vignette equivalence assumption holds, any difference in vignette evaluations by individuals with different depressive symptoms would suggest the presence of reporting heterogeneity in sleep difficulty assessment. The five plots of [Fig 4](#) show that most vignette evaluations are relatively homogeneous, irrespective of the depressive symptoms college students are suffering from. This holds true except perhaps among those with severe depressive symptoms, for which noticeable differences can be seen in the evaluations of vignettes 2 and 4. While purely descriptive, this bivariate analysis already shed light on the possible presence of reporting heterogeneity in self-assessed sleep difficulty among college students with and without depressive symptoms.

We then resort to multivariate analysis and regress the five vignette evaluations on our set of control variables, including depressive symptoms. Under the vignette equivalence assumption, if the coefficient of a variable turns out to be statistically significant, then it would mean that the vignettes are not perceived equally by respondents who have the characteristic associated with that particular coefficient. In other words, there would be some reporting heterogeneity associated with that specific characteristic x (See [Eq 6](#)). [Table 3](#) reports the results of the regressions of our five vignettes on our set of control variables. It is reassuring that very few

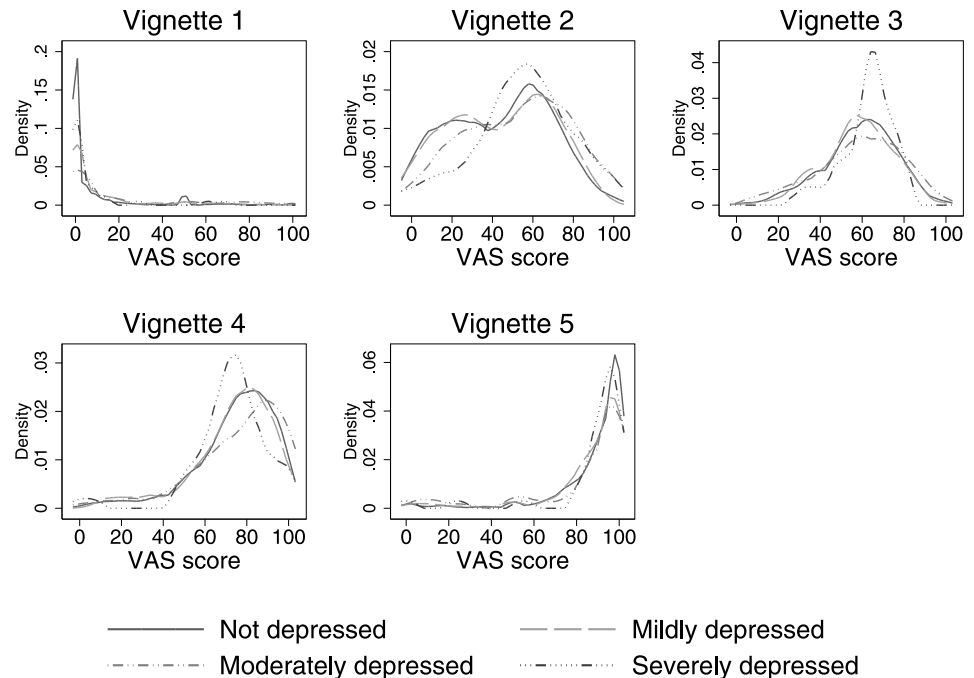


Fig 4. Distribution of vignette evaluations by depressive symptoms. Note: 0 means ‘no difficulty’ and 100 means ‘extreme difficulty’ with sleep.

<https://doi.org/10.1371/journal.pone.0246370.g004>

coefficients are significantly different from 0, suggesting that individuals evaluate vignettes in a rather similar way. One exception clearly stands out, however. Individuals suffering from ‘moderate’ and ‘severe’ depressive symptoms appear to evaluate vignette 2 more negatively, with their evaluations being on average 8.3 ($p < 0.01$) and 10.6 ($p < 0.05$) points closer to ‘extreme difficulty’, respectively, compared to those with ‘no or minimal’ depressive

Table 3. Linear regressions of vignette evaluations on our set of control variables.

	(1) Vignette 1	(2) Vignette 2	(3) Vignette 3	(4) Vignette 4	(5) Vignette 5
Female	2.41** (0.97)	0.48 (1.23)	-1.46* (0.85)	-0.97 (0.96)	-1.67* (0.99)
Age	0.20 (0.31)	0.33 (0.39)	-0.25 (0.28)	-0.26 (0.31)	-0.26 (0.31)
Mild depressive symptoms	0.87 (1.29)	-0.15 (1.67)	-0.91 (1.11)	-0.87 (1.31)	-1.51 (1.38)
Moderate depressive symptoms	5.96** (2.78)	8.26*** (2.62)	0.15 (2.17)	2.09 (2.19)	-3.34 (2.56)
Severe depressive symptoms	-4.58 (3.17)	10.6** (4.37)	3.76* (2.25)	-1.32 (3.19)	1.92 (3.78)
Constant	3.48 (7.77)	38.0*** (9.95)	66.7*** (7.33)	85.6*** (7.88)	95.0*** (8.01)

Note: Robust standard errors are in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). We also control for parent’s income, education level, nationality, number of siblings, relationship status, relationship with parents, number of close friends and school performance. These coefficients are not reported in the table but are available upon request.

<https://doi.org/10.1371/journal.pone.0246370.t003>

symptoms. This suggests the presence of reporting heterogeneity in self-reported sleep difficulties in these two subgroups of individuals. College students with moderate depressive symptoms also tend to evaluate vignette 1 higher than college students without any depressive symptoms (6 points, $p < 0.05$).

Corresponding results when using PHQ-9 score instead of depressive symptoms as measure of mental health are presented in [S2 Table](#). We show in the Appendix that PHQ-9 score has a statistically significant association with the evaluations of vignette 2 (0.55 points, $p < 0.01$) but has no significant statistical correlation (at conventional level) with the evaluations of vignettes 1, 3, 4 and 5. These results are very much in line with results presented in [Table 3](#).

To further determine the existence of reporting heterogeneity, we add vignette evaluations directly in the sleep difficulties regression as control variables. Indeed, assuming that individuals use the same reporting scale to evaluate their sleep difficulties as the scale they use to evaluate the vignettes, reporting heterogeneity can be directly controlled for by adding vignette evaluations as regressors in the econometric specification. [Table 4](#) reports the results of these

Table 4. Linear regressions of sleep difficulties on our set of control variables, including the vignette evaluations in columns 2 and 4.

	(1) Sleep difficulty	(2) Sleep difficulty	(3) Sleep difficulty	(4) Sleep difficulty
Female	1.24 (1.20)	1.08 (1.19)	0.21 (1.09)	0.053 (1.08)
Age	0.20 (0.39)	0.18 (0.39)	0.14 (0.36)	0.12 (0.36)
Mild depressive symptoms	19.6*** (1.69)	19.6*** (1.70)	14.1*** (1.57)	14.1*** (1.57)
Moderate depressive symptoms	32.4*** (2.40)	30.9*** (2.40)	24.7*** (2.27)	23.2*** (2.21)
Severe depressive symptoms	37.4*** (3.69)	36.8*** (3.64)	28.5*** (3.40)	27.9*** (3.28)
Vignette 1		0.098*** (0.035)		0.10*** (0.032)
Vignette 2		0.091*** (0.026)		0.088*** (0.024)
Vignette 3		0.032 (0.040)		0.047 (0.035)
Vignette 4		0.075* (0.044)		0.072* (0.039)
Vignette 5		-0.0022 (0.044)		0.00078 (0.041)
Take more than 20 mins to fall asleep			17.7*** (1.17)	17.7*** (1.15)
Sleep less than 7 hours			6.76*** (1.42)	6.79*** (1.41)
Low sleep efficiency			9.07*** (1.48)	9.15*** (1.48)
Constant	17.5* (9.93)	5.38 (10.4)	11.9 (9.18)	-1.21 (9.68)

Note: Robust standard errors are in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). We also control for parent's income, education level, nationality, number of siblings, relationship status, relationship with parents, number of close friends and school performance. These coefficients are not reported in the table but are available upon request.

<https://doi.org/10.1371/journal.pone.0246370.t004>

Table 5. Linear regressions of sleep difficulties on our set of control variables, accounting for reporting heterogeneity.

	(1) "True" effect b_0	(2) Reporting Heterogeneity γ	(3) "True" effect b_0	(4) Reporting Heterogeneity γ
Female	1.48 (1.25)	-0.24 (0.56)	0.42 (1.14)	-0.22 (0.56)
Age	0.25 (0.41)	-0.049 (0.18)	0.19 (0.38)	-0.056 (0.18)
Mild depressive symptoms	20.1*** (1.78)	-0.51 (0.72)	14.5*** (1.66)	-0.44 (0.75)
Moderate depressive symptoms	29.8*** (2.60)	2.62** (1.32)	22.0*** (2.43)	2.70** (1.35)
Severe depressive symptoms	35.3*** (3.85)	2.07 (2.20)	26.4*** (3.56)	2.12 (2.21)
Take more than 20 mins to fall asleep			18.2*** (1.21)	-0.52 (0.56)
Sleep less than 7 hours			6.46*** (1.49)	0.29 (0.71)
Low sleep efficiency			9.05*** (1.61)	0.015 (0.76)
Constant	17.5* (9.88)		11.9 (9.13)	

Note: Cluster robust standard errors at the respondent level reported in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$). We also control for parent's income, education level, nationality, number of siblings, relationship status, relationship with parents, number of close friends, and school performance. These coefficients are not reported in the table but are available upon request. Columns 1 and 3 show the effects of the control variables on sleep difficulties, net of reporting heterogeneity, while columns 2 and 4 show the effects of the control variables on reporting heterogeneity.

<https://doi.org/10.1371/journal.pone.0246370.t005>

regressions. Column 1 is the result of regressing sleep difficulties on our set of control variables (same regression as column 1 of Table 2) and column 2 represents the results of the same regression including this time vignettes evaluations as control variables. Although vignettes 1's and 2's evaluations are themselves highly significant (0.10 points with $p < 0.01$ and 0.09 points with $p < 0.01$, respectively), the inclusion of vignettes as control variables does not affect the associations between depressive symptoms and sleep difficulties. The conclusion is identical when we add vignette evaluations (column 4) to the model that also controls for sleep patterns and efficiency (column 3). Our results are robust to using PHQ-9 score as a continuous measure instead of descriptive depression categories, as shown in S2 Table. These results indicate that the associations between sleep difficulties and depressive symptoms remain large and statistically significant even after controlling for reporting heterogeneity.

Because there nonetheless appears to be some reporting heterogeneity in self-assessed sleep difficulties, as evidenced by the statistically significant effect of some of our vignettes, it would be interesting to understand what factors drive this reporting heterogeneity. Since we observe both self-assessed sleep difficulties and vignette evaluations, we can run a double-index model that allows us to estimate the associations between our control variables and both sleep difficulties (free from any reporting heterogeneity) and reporting heterogeneity itself, hence enabling us to determine the factors explaining reporting heterogeneity.

Results from this model are summarized in Table 5. Odd-numbered columns represent the "true" associations between our control variables, net of any reporting heterogeneity (vector of coefficients b_0), and sleep difficulties, and even-numbered columns show the associations

between our set of regressors and reporting heterogeneity (vector of coefficients γ). From column 1, it is clear that the associations between being mildly, moderately, and severely depressed and sleep difficulties remain very strong (with increases of 20.1, 29.8, and 35.3 points, respectively, compared to those with no depressive symptoms, all with $p < 0.01$) and are very close to the ones obtained in column 2 of Table 4. The same is true when we include sleep patterns and efficiency in the model (column 3). Columns 2 and 4 show that individuals suffering from mild and severe depressive symptoms do not appear to evaluate the VAS scale differently from those with no depressive symptoms. Respondents with moderate depressive symptoms, however, do perceive the VAS scale differently, with their sleep difficulty evaluations being 2.6–2.7 points higher than what it “really should be”. That being said, this effect is very small, less than 10% in magnitude, and the associations between sleep difficulties and depressive symptoms remain very strong even after controlling for reporting heterogeneity. Again, these results are robust to using PHQ-9 score as a continuous measure instead of descriptive depression categories (S2 Table).

Discussion and conclusion

Many studies have analyzed the association between sleep difficulties and depression [2, 67, 68]. All found that individuals who suffer from depression have more severe sleep difficulties. However, very few studies have investigated whether that association could be explained, at least partially, by reporting heterogeneity [44]. Based on a sample of college students (18–25 years of age) asked to complete a survey online, we confirm that sleep difficulties and depressive symptoms are strongly related. Individuals with mild, moderate, and severe depressive symptoms rated their sleep difficulties to be 20, 32, and 37 points higher, on a 0–100 Visual Analogue Scale, compared with those who have no or minimal depressive symptoms. These associations hold even after accounting for differences in sleep patterns and efficiency.

We then ask whether this association finds its explanation, at least partially, in reporting heterogeneity. Our analysis does indicate the presence of reporting heterogeneity in self-assessed sleep difficulties. We show that college students who suffer from moderate depressive symptoms tend to be pessimistic about their sleep difficulties, rating them 2.6–2.7 points higher than if they were to use the same reporting scale as college students with no or minimal symptoms (reference group). Given the large magnitude of the associations between sleep difficulties and depressive symptoms, the difference in reporting patterns between college students with no or minimal symptoms and those with moderate depressive symptoms appear to be minor (less than 10% in magnitude). Interestingly, we do not find corresponding differences in reporting patterns between individuals with no or minimal depressive symptoms and those with severe depressive symptoms. Individuals with severe depressive symptoms appear to rate their sleep difficulties about 2 points higher than if they were to use the same reporting scale as those with no or minimal symptoms, but this effect is not statistically significant. A possible explanation for this absence of statistically significant reporting heterogeneity among individuals with severe depressive symptoms could stem from the fact that we have very few individuals in our sample that fall in that category (1.8% of our sample).

Our study investigates the existence of reporting heterogeneity in self-assessed sleep difficulties among college students with and without depressive symptoms. While focusing on this subpopulation is highly relevant because of the important changes in sleep patterns and dramatic increases in depression onset many college students are facing [91, 92]—with consequences for depression and bipolar disorder during adulthood [24]—our results only speak to the college student population and it is therefore difficult to generalize our findings to the overall population. This constitutes an important limitation and we hope to encourage further

researchers to explore whether our results hold in other settings and in the general population. Moreover, our results indicate minor differences in reporting behaviors between individuals with and without symptoms when it comes to evaluating their sleep difficulties. Future research could however explore the presence of reporting heterogeneity across these groups in other domains such as self-reported quality of life and well-being, for which corresponding reporting patterns are likely to be different.

Another limitation of our study is that our vignette adjustment to control for reporting heterogeneity rely on two assumptions, i.e., response consistency and vignette equivalence, that cannot be directly tested in our application. That being said, these two assumptions are very commonly used in the vignette literature [45, 47, 55, 62, 63] and we present in S3 and S4 Appendices. some evidence that suggests that they could hold in our setting as well. In addition, although widely used and validated, PHQ-9 is the sole instrument used in this study to determine the presence of depressive symptoms of our respondents and we therefore cannot rule out possible misclassification errors. Finally, we can not rule out the possibility that college students with depressive symptoms were less likely to participate in our study and complete our survey, which can lead to selection bias.

Overall, the results of our study suggest that comparisons of self-assessed sleep difficulties between college students are meaningful, even between those with and without depressive symptoms, and that reporting heterogeneity plays a minor role in explaining the relationship between sleep difficulties and depressive symptoms.

Supporting information

S1 Data.

(ZIP)

S1 Appendix. Vignettes used in this study.

(PDF)

S2 Appendix. The PHQ-9 questionnaire.

(PDF)

S3 Appendix. Response consistency.

(PDF)

S4 Appendix. Vignette equivalence.

(PDF)

S5 Appendix. Test of linear specification and fractional probit regression.

(PDF)

S1 Table. Descriptive statistics of sleep patterns.

(PDF)

S2 Table. PHQ-9 as a continuous measure.

(PDF)

Author Contributions

Conceptualization: Zhiyong Huang, Fabrice Kämpfen.

Data curation: Zhiyong Huang, Fabrice Kämpfen.

Formal analysis: Zhiyong Huang, Fabrice Kämpfen.

Funding acquisition: Zhiyong Huang, Fabrice Kämpfen.

Investigation: Zhiyong Huang, Fabrice Kämpfen.

Methodology: Zhiyong Huang, Fabrice Kämpfen.

Project administration: Zhiyong Huang, Fabrice Kämpfen.

References

1. Hyppä MT, Kronholm E. Quality of sleep and chronic illnesses. *Journal of clinical epidemiology*. 1989; 42(7):633–638. [https://doi.org/10.1016/0895-4356\(89\)90006-1](https://doi.org/10.1016/0895-4356(89)90006-1)
2. Augner C. Associations of subjective sleep quality with depression score, anxiety, physical symptoms and sleep onset latency in students. *Central European journal of public health*. 2011; 19(2):115. <https://doi.org/10.21101/cejph.a3647>
3. Lombardi DA, Folkard S, Willetts JL, Smith GS. Daily sleep, weekly working hours, and risk of work-related injury: US National Health Interview Survey (2004–2008). *Chronobiology international*. 2010; 27(5):1013–1030. <https://doi.org/10.3109/07420528.2010.489466>
4. Schwartz S, Anderson WM, Cole SR, Comoni-Huntley J, Hays JC, Blazer D. Insomnia and heart disease: a review of epidemiologic studies. *Journal of psychosomatic research*. 1999; 47(4):313–333. [https://doi.org/10.1016/S0022-3999\(99\)00029-X](https://doi.org/10.1016/S0022-3999(99)00029-X)
5. LeBlanc M, Beaulieu-Bonneau S, Mérette C, Savard J, Ivers H, Morin CM. Psychological and health-related quality of life factors associated with insomnia in a population-based sample. *Journal of psychosomatic research*. 2007; 63(2):157–166. <https://doi.org/10.1016/j.jpsychores.2007.03.004>
6. Niet de GJG, Tiemens BGB, Lendemeijer HB, Hutschemaekers GJMG. Perceived sleep quality of psychiatric patients. *Journal of psychiatric and mental health nursing*. 2008; 15(6):465–470. <https://doi.org/10.1111/j.1365-2850.2008.01250.x>
7. Algul A, Ates MA, Semiz UB, Basoglu C, Ebrinc S, Gecici O, et al. Evaluation of general psychopathology, subjective sleep quality, and health-related quality of life in patients with obesity. *The International Journal of Psychiatry in Medicine*. 2009; 39(3):297–312. <https://doi.org/10.2190/PM.39.3.f> PMID: 19967901
8. Mayers AG, Grabau EAS, Campbell C, Baldwin DS. Subjective sleep, depression and anxiety: interrelationships in a non-clinical sample. *Human Psychopharmacology: Clinical and Experimental*. 2009; 24(6):495–501. <https://doi.org/10.1002/hup.1041>
9. Hayashino Y, Yamazaki S, Takegami M, Nakayama T, Sokejima S, Fukuhara S. Association between number of comorbid conditions, depression, and sleep quality using the Pittsburgh Sleep Quality Index: results from a population-based survey. *Sleep medicine*. 2010; 11(4):366–371. <https://doi.org/10.1016/j.sleep.2009.05.021>
10. Schulenberg JE, Sameroff AJ, Cicchetti D. The transition to adulthood as a critical juncture in the course of psychopathology and mental health. *Development and psychopathology*. 2004; 16(4):799–806. <https://doi.org/10.1017/S0954579404040015>
11. Furr SR, Westefeld JS, McConnell GN, Jenkins JM. Suicide and depression among college students: A decade later. *Professional Psychology: Research and Practice*. 2001; 32(1):97. <https://doi.org/10.1037/0735-7028.32.1.97>
12. Garlow SJ, Rosenberg J, Moore JD, Haas AP, Koestner B, Hendin H, et al. Depression, desperation, and suicidal ideation in college students: results from the American Foundation for Suicide Prevention College Screening Project at Emory University. *Depression and anxiety*. 2008; 25(6):482–488. <https://doi.org/10.1002/da.20321> PMID: 17559087
13. Stein MB, Fuetsch M, Müller N, Höfler M, Lieb R, Wittchen HU. Social anxiety disorder and the risk of depression: a prospective community study of adolescents and young adults. *Archives of general psychiatry*. 2001; 58(3):251–256. <https://doi.org/10.1001/archpsyc.58.3.251>
14. Roenneberg T, Kuehnle T, Pramstaller PP, Ricken J, Havel M, Guth A, et al. A marker for the end of adolescence. *Current biology*. 2004; 14(24):R1038–R1039. <https://doi.org/10.1016/j.cub.2004.11.039> PMID: 15620633
15. Roenneberg T, Kuehnle T, Juda M, Kantermann T, Allebrandt K, Gordijn M, et al. Epidemiology of the human circadian clock. *Sleep medicine reviews*. 2007; 11(6):429–438. <https://doi.org/10.1016/j.smrv.2007.07.005> PMID: 17936039
16. Wong ML, Lau EYY, Wan JHY, Cheung SF, Hui CH, Mok DSY. The interplay between sleep and mood in predicting academic functioning, physical health and psychological health: a longitudinal study. *Journal of psychosomatic research*. 2013; 74(4):271–277. <https://doi.org/10.1016/j.jpsychores.2012.08.014>

17. Glozier N, Martiniuk A, Patton G, Ivers R, Li Q, Hickie I, et al. Short sleep duration in prevalent and persistent psychological distress in young adults: the DRIVE study. *Sleep*. 2010; 33(9):1139–1145. <https://doi.org/10.1093/sleep/33.9.1139> PMID: 20857859
18. Asarnow LD, McGlinchey E, Harvey AG. The effects of bedtime and sleep duration on academic and emotional outcomes in a nationally representative sample of adolescents. *Journal of Adolescent Health*. 2014; 54(3):350–356. <https://doi.org/10.1016/j.jadohealth.2013.09.004>
19. Baglioni C, Nanovska S, Regen W, Spiegelhalter K, Feige B, Nissen C, et al. Sleep and mental disorders: A meta-analysis of polysomnographic research. *Psychological bulletin*. 2016; 142(9):969. <https://doi.org/10.1037/bul0000053> PMID: 27416139
20. Lund HG, Reider BD, Whiting AB, Prichard JR. Sleep patterns and predictors of disturbed sleep in a large population of college students. *Journal of adolescent health*. 2010; 46(2):124–132. <https://doi.org/10.1016/j.jadohealth.2009.06.016>
21. Feige B, Baglioni C, Spiegelhalter K, Hirscher V, Nissen C, Riemann D. The microstructure of sleep in primary insomnia: an overview and extension. *International Journal of Psychophysiology*. 2013; 89(2):171–180. <https://doi.org/10.1016/j.ijpsycho.2013.04.002>
22. Becker SP, Dvorsky MR, Holdaway AS, Luebke AM. Sleep problems and suicidal behaviors in college students. *Journal of psychiatric research*. 2018; 99:122–128. <https://doi.org/10.1016/j.jpsychires.2018.01.009>
23. Mojtabai R, Olfson M, Han B. National trends in the prevalence and treatment of depression in adolescents and young adults. *Pediatrics*. 2016; p. e20161878. <https://doi.org/10.1542/peds.2016-1878>
24. Sabaté E. Depression in young people and the elderly. Priority medicines for Europe and the world, public health approach to innovation. 2004; 6.
25. Kessler RC, Amminger GP, Aguilar-Gaxiola S, Alonso J, Lee S, Ustun TB. Age of onset of mental disorders: a review of recent literature. *Current opinion in psychiatry*. 2007; 20(4):359. <https://doi.org/10.1097/YCO.0b013e32816ebc8c>
26. Kaplan KA, Hirshman J, Hernandez B, Stefanick ML, Hoffman AR, Redline S, et al. When a gold standard isn't so golden: Lack of prediction of subjective sleep quality from sleep polysomnography. *Biological psychology*. 2017; 123:37–46. <https://doi.org/10.1016/j.biopsycho.2016.11.010> PMID: 27889439
27. Suzuki H, Kaneita Y, Osaki Y, Minowa M, Kanda H, Suzuki K, et al. Clarification of the factor structure of the 12-item General Health Questionnaire among Japanese adolescents and associated sleep status. *Psychiatry research*. 2011; 188(1):138–146. <https://doi.org/10.1016/j.psychres.2010.10.025> PMID: 21145596
28. Johns MW, Gay TJ, Goodyear MD, Masterton JP. Sleep habits of healthy young adults: use of a sleep questionnaire. *British journal of preventive & social medicine*. 1971; 25(4):236.
29. Bobes J, González MP, Vallejo J, Sáiz J, Gibert J, Ayuso JL, et al. Oviedo Sleep Questionnaire (OSQ): A new semistructured interview for sleep disorders. *European Neuropsychopharmacology*. 1998; 8: S162. [https://doi.org/10.1016/S0924-977X\(98\)80199-5](https://doi.org/10.1016/S0924-977X(98)80199-5)
30. Webb WB, Bonnet M, Blume G. A post-sleep inventory. *Perceptual and Motor Skills*. 1976; 43(3):987–993. <https://doi.org/10.2466/pms.1976.43.3.987>
31. Åkerstedt T, Hume KEN, Minors D, Waterhouse JIM. The subjective meaning of good sleep, an intraindividual approach using the Karolinska Sleep Diary. *Perceptual and motor skills*. 1994; 79(1):287–296. <https://doi.org/10.2466/pms.1994.79.1.287>
32. Sonnegg A, Faul JD, Ofstedal MB, Langa KM, Phillips JW, Weir DR. Cohort profile: the health and retirement study (HRS). *International journal of epidemiology*. 2014; 43(2):576–585. <https://doi.org/10.1093/ije/dyu067>
33. Börsch-Supan A, Brandt M, Hunkler C, Kneip T, Korbmacher J, Malter F, et al. Data resource profile: the Survey of Health, Ageing and Retirement in Europe (SHARE). *International journal of epidemiology*. 2013; 42(4):992–1001. <https://doi.org/10.1093/ije/dyt088> PMID: 23778574
34. Kowal P, Chatterji S, Naidoo N, Biritwum R, Fan W, Lopez Ridaura R, et al. Data resource profile: the World Health Organization Study on global AGEing and adult health (SAGE). *International journal of epidemiology*. 2012; 41(6):1639–1649. <https://doi.org/10.1093/ije/dys210> PMID: 23283715
35. Zhao Y, Hu Y, Smith JP, Strauss J, Yang G. Cohort profile: the China health and retirement longitudinal study (CHARLS). *International journal of epidemiology*. 2012; 43(1):61–68. <https://doi.org/10.1093/ije/dys203>
36. Buysse DJ, Hall ML, Strollo PJ, Kamarck TW, Owens J, Lee L, et al. Relationships between the Pittsburgh Sleep Quality Index (PSQI), Epworth Sleepiness Scale (ESS), and clinical/polysomnographic measures in a community sample. *Journal of Clinical Sleep Medicine*. 2008; <https://doi.org/10.5664/jcsm.27351>

37. Jackson CL, Patel SR, Jackson WB, Lutsey PL, Redline S. Agreement between self-reported and objectively measured sleep duration among white, black, Hispanic, and Chinese adults in the United States: Multi-ethnic study of atherosclerosis. *Sleep*. 2018; <https://doi.org/10.1093/sleep/zsy057> PMID: [29701831](https://pubmed.ncbi.nlm.nih.gov/29701831/)
38. Lemola S, Ledermann T, Friedman EM. Variability of sleep duration is related to subjective sleep quality and subjective well-being: an actigraphy study. *PloS one*. 2013; 8(8):e71292. <https://doi.org/10.1371/journal.pone.0071292>
39. Baker FC, Maloney S, Driver HS. A comparison of subjective estimates of sleep with objective polysomnographic data in healthy men and women. *Journal of psychosomatic research*. 1999; 47(4):335–341. [https://doi.org/10.1016/S0022-3999\(99\)00017-3](https://doi.org/10.1016/S0022-3999(99)00017-3)
40. Jean-Louis G, Kripke DF, Ancoli-Israel S. Sleep and quality of well-being. *Sleep*. 2000; 23(8):1115. <https://doi.org/10.1093/sleep/23.8.1k>
41. Manzar MD, BaHammam AS, Hameed UA, Spence DW, Pandi-Perumal SR, Moscovitch A, et al. Dimensionality of the Pittsburgh Sleep Quality Index: a systematic review. *Health and quality of life outcomes*. 2018; 16(1):89. <https://doi.org/10.1186/s12955-018-0915-x> PMID: [29743066](https://pubmed.ncbi.nlm.nih.gov/29743066/)
42. Bliwise DL, Friedman L, Yesavage JA. Depression as a confounding variable in the estimation of habitual sleep time. *Journal of clinical psychology*. 1993; 49(4):471–477. [https://doi.org/10.1002/1097-4679\(199307\)49:4%3C471::AID-JCLP2270490403%3E3.0.CO;2-7](https://doi.org/10.1002/1097-4679(199307)49:4%3C471::AID-JCLP2270490403%3E3.0.CO;2-7)
43. Tarrasch R, Laudon M, Zisapel N. Cross-cultural validation of the Leeds sleep evaluation questionnaire (LSEQ) in insomnia patients. *Human Psychopharmacology: Clinical and Experimental*. 2003; 18(8):603–610. <https://doi.org/10.1002/hup.534>
44. Tareque MI, Ikeda N, Koshio A, Hasegawa T. Is adjustment for reporting heterogeneity necessary in sleep disorders? Results from the Japanese World Health Survey. *BMC psychiatry*. 2016; 16:25. <https://doi.org/10.1186/s12888-016-0733-9>
45. King G, Murray CJL, Salomon JA, Tandon A. Enhancing the Validity and {Cross-Cultural} Comparability of Measurement in Survey Research. *Am Polit Sci Rev*. 2004; 98(01):191–207. <https://doi.org/10.1017/S000305540400108X>
46. Kapteyn A, Smith JP, Van Soest A. Are Americans Really Less Happy with Their Incomes? Review of Income and Wealth. 2013; 59(1):44–65. <https://doi.org/10.1111/j.1475-4991.2012.00532.x>
47. Bago d'Uva T, D'Uva TB, Van Doorslaer E, Lindeboom M, O'Donnell O. Does reporting heterogeneity bias the measurement of health disparities? *Health Econ*. 2008; 17(3):351–375. <https://doi.org/10.1002/hec.1269>
48. Bago d'Uva T, Lindeboom M, O'Donnell O, Van Doorslaer E. Education-related inequity in healthcare with heterogeneous reporting of health. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*. 2011; 174(3):639–664. <https://doi.org/10.1111/j.1467-985X.2011.00706.x>
49. Salomon JA, Tandon A, Murray CJL. Comparability of self rated health: cross sectional multi-country survey using anchoring vignettes. *BMJ*. 2004; 328(7434):258. <https://doi.org/10.1136/bmj.37963.691632.44>
50. Jürges H. True health vs response styles: exploring cross-country differences in self-reported health. *Health Econ*. 2006; 16(2):163–178.
51. Hanandita W, Tampubolon G. Does reporting behaviour bias the measurement of social inequalities in self-rated health in Indonesia? An anchoring vignette analysis. *Quality of Life Research*. 2016; 25(5):1137–1149. <https://doi.org/10.1007/s11136-015-1152-y>
52. Molina T. Reporting Heterogeneity and Health Disparities Across Gender and Education Levels: Evidence From Four Countries. *Demography*. 2016; 53(2):295–323. <https://doi.org/10.1007/s13524-016-0456-z>
53. Knott RJ, Black N, Hollingsworth B, Lorgelly PK. Response-Scale Heterogeneity in the EQ-5D. *Health economics*. 2017; 26(3):387–394. <https://doi.org/10.1002/hec.3313>
54. Knott RJ, Lorgelly PK, Black N, Hollingsworth B. Differential item functioning in quality of life measurement: An analysis using anchoring vignettes. *Social Science & Medicine*. 2017; 190:247–255. <https://doi.org/10.1016/j.socscimed.2017.08.033>
55. Rice N, Robone S, Smith PC. Vignettes and health systems responsiveness in cross-country comparative analyses. *J R Stat Soc Ser A Stat Soc*. 2012; 175(2):337–369. <https://doi.org/10.1111/j.1467-985X.2011.01021.x>
56. Malhotra C, Do YK. Socio-economic disparities in health system responsiveness in India. *Health Policy Plan*. 2013; 28(2):197–205. <https://doi.org/10.1093/heapol/czs051>
57. Angelini V, Viola A, Danilo C, Luca C, Omar P. Do Danes and Italians Rate Life Satisfaction in the Same Way? Using Vignettes to Correct for {Individual-Specific} Scale Biases. *Oxf Bull Econ Stat*. 2013; 76(5):643–666. <https://doi.org/10.1111/obes.12039>

58. Crane M, Rissel C, Greaves S, Gebel K. Correcting bias in self-rated quality of life: an application of anchoring vignettes and ordinal regression models to better understand {QoL} differences across commuting modes. *Qual Life Res.* 2016; 25(2):257–266. <https://doi.org/10.1007/s11136-015-1090-8>
59. Ravallion M, Himelein K, Beegle K. Can Subjective Questions on Economic Welfare Be Trusted? *Econ Dev Cult Change.* 2016; 64(4):697–726.
60. Huang Z, Kämpfen F. Assessing (and Addressing) Reporting Heterogeneity in Visual Analogue Scales (VAS) with an Application to Gender Difference in Quality of Life. University of Pennsylvania Population Center Working Paper (PSC/PARC), 2019-26. 2019;.
61. Hopkins DJ, King G. Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability. *Public Opin Q.* 2010; 74(2):201–222. <https://doi.org/10.1093/poq/nfq011>
62. Kristensen N, Johansson E. New evidence on cross-country differences in job satisfaction using anchoring vignettes. *Labour Econ.* 2008; 15(1):96–117. <https://doi.org/10.1016/j.labeco.2006.11.001>
63. Kapteyn A, Smith JP, van Soest A. Vignettes and {Self-Reports} of Work Disability in the United States and the Netherlands. *Am Econ Rev.* 2007; 97(1):461–473. <https://doi.org/10.1257/000282807780323596>
64. Wang J. Rural-to-urban Migration and Rising Evaluation Standards for Subjective Social Status in Contemporary China. *Soc Indic Res.* 2016;.
65. van Soest A, Delaney L, Harmon C, Kapteyn A, Smith JP. Validating the use of anchoring vignettes for the correction of response scale differences in subjective questions. *J R Stat Soc Ser A Stat Soc.* 2011; 174(3):575–595. <https://doi.org/10.1111/j.1467-985X.2011.00694.x>
66. Studer R. Does it Matter How Happiness is Measured? Evidence From a Randomized Controlled Experiment. Working Paper Series, Working Paper No 49. 2011;.
67. Harvey AG, Stinson K, Whitaker KL, Moskovitz D, Virk H. The subjective meaning of sleep quality: a comparison of individuals with and without insomnia. *Sleep.* 2008; 31(3):383–393. <https://doi.org/10.1093/sleep/31.3.383>
68. Milette K, Razykov I, Pope J, Hudson M, Motivala SJ, Baron M, et al. Clinical correlates of sleep problems in systemic sclerosis: the prominent role of pain. *Rheumatology.* 2010; 50(5):921–925. <https://doi.org/10.1093/rheumatology/keq411> PMID: 21169344
69. Arnett JJ. Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist.* 2000; <https://doi.org/10.1037/0003-066X.55.5.469> PMID: 10842426
70. Sawyer SM, Azzopardi PS, Wickremarathne D, Patton GC. The age of adolescence. *The Lancet Child and Adolescent Health.* 2018; [https://doi.org/10.1016/S2352-4642\(18\)30022-1](https://doi.org/10.1016/S2352-4642(18)30022-1)
71. World Health Organization. World Health Survey 2002: survey manual. World Health Organization, Geneva. 2002;.
72. Jürges H, Winter J. Are anchoring vignettes ratings sensitive to vignette age and sex? *Health Economics.* 2013; 22(1):1–13.
73. Grol-Prokopczyk H, Freese J, Hauser RM. Using Anchoring Vignettes to Assess Group Differences in General Self-Rated Health. *Journal of Health and Social Behavior.* 2011; 52(2):246–261. <https://doi.org/10.1177/0022146510396713>
74. Au N, Lorgelly PK. Anchoring vignettes for health comparisons: an analysis of response consistency. *Quality of Life Research.* 2014; 23(6):1721–1731. <https://doi.org/10.1007/s11136-013-0615-2>
75. Shrivastava D, Jung S, Saadat M, Sirohi R, Crewson K. How to interpret the results of a sleep study. *Journal of community hospital internal medicine perspectives.* 2014; 4(5):24983. <https://doi.org/10.3402/jchimp.v4.24983>
76. Spitzer RL, Williams JBW, Kroenke K. Test review: Patient Health Questionnaire–9 (PHQ-9). *Rehabilitation Counseling Bulletin.* 2014; 57(4):246–248. <https://doi.org/10.1177/0034355213515305>
77. Kohler IV, Payne CF, Bandawe C, Kohler HP. The demography of mental health among mature adults in a low-income, high-HIV-prevalence context. *Demography.* 2017; 54(4):1529–1558. <https://doi.org/10.1007/s13524-017-0596-9>
78. Arrieta J, Aguerrebere M, Raviola G, Flores H, Elliott P, Espinosa A, et al. Validity and utility of the Patient Health Questionnaire (PHQ)-2 and PHQ-9 for screening and diagnosis of depression in rural Chiapas, Mexico: A cross-sectional study. *Journal of clinical psychology.* 2017; 73(9):1076–1090. <https://doi.org/10.1002/jclp.22390> PMID: 28195649
79. Adewuya AO, Ola BA, Afolabi OO. Validity of the patient health questionnaire (PHQ-9) as a screening tool for depression amongst Nigerian university students. *Journal of affective disorders.* 2006; 96(1-2):89–93. <https://doi.org/10.1016/j.jad.2006.05.021>

80. Huang Z, Kohler IV, Kämpfen F. A single-item Visual Analogue Scale (VAS) measure for assessing depression among college students. *Community mental health journal*. 2020; 56(2):355–367. <https://doi.org/10.1007/s10597-019-00469-7>
81. Erbe D, Eichert HC, Rietz C, Ebert D. Interformat reliability of the patient health questionnaire: validation of the computerized version of the PHQ-9. *Internet Interventions*. 2016; 5:1–4. <https://doi.org/10.1016/j.invent.2016.06.006>
82. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *Journal of general internal medicine*. 2001; 16. <https://doi.org/10.1046/j.1525-1497.2001.016009606.x>
83. Nunnally JC. *Psychometric theory* 3E. Tata McGraw-Hill Education; 1994.
84. World Health Organization. Adolescent mental health; 2019. <https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health>.
85. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. *Psychiatry research*. 1989; 28(2):193–213. [https://doi.org/10.1016/0165-1781\(89\)90047-4](https://doi.org/10.1016/0165-1781(89)90047-4)
86. Talero-Gutiérrez C, Moreno CB, González-Reyes R, Palacios L, de la Peña F, Muñoz-Delgado J. Sleep evaluation scales and questionnaires: a review. *Actas Esp Psiquiatr*. 2008; 36(1):50–59.
87. Monk TH, Buysse DJ, Kennedy KS, Potts JM, DeGrazia JM, Miewald JM. Measuring sleep habits without using a diary: the sleep timing questionnaire. *Sleep*. 2003; 26(2):208–212. <https://doi.org/10.1093/sleep/26.2.208>
88. Weiss TM, Atanasov S, Calhoun KH. The association of tongue scalloping with obstructive sleep apnea and related sleep pathology. *Otolaryngology—Head and Neck Surgery*. 2005; 133(6):966–971. <https://doi.org/10.1016/j.otohns.2005.07.018>
89. Ohayon M, Wickwire EM, Hirshkowitz M, Albert SM, Avidan A, Daly FJ, et al. National Sleep Foundation's sleep quality recommendations: first report. *Sleep Health*. 2017; 3(1):6–19. <https://doi.org/10.1016/j.sleh.2016.11.006> PMID: 28346153
90. Øyane NMF, Bjorvatn B. Sleep disturbances in adolescents and young adults with autism and Asperger syndrome. *Autism*. 2005; 9(1):83–94. <https://doi.org/10.1177/1362361305049031>
91. Stone LB, Hankin BL, Gibb BE, Abela JR. Co-rumination predicts the onset of depressive disorders during adolescence. *Journal of abnormal psychology*. 2011; 120(3):752. <https://doi.org/10.1037/a0023384>
92. Avenevoli S, Knight E, Kessler RC, Merikangas KR. Epidemiology of depression in children and adolescents. *Handbook of depression in children and adolescents*. 2008;.