

Peer-Review

Platt, Taylor E, and Kyle Lorenzo. 2026. "Association between Short-Term Sleep Duration and Cognitive Performance: Inhibitory Control, Problem Solving, Working Memory, and Idea Generation Rate." *Journal of High School Science* 10 (3): 85–104.
<https://doi.org/10.64336/001c.164834>.

I would reframe this study as an exploratory hypothesis generating study. Remove all causal claims and frame only as being associations. Since you yourself acknowledge bidirectionality, you cannot claim that sleep influences cognition.

Your strong contributions are that you measure all cognitive domains simultaneously, you (rightly) raise the reverse causality between sleep and cognition and you suggest that divergent thinking may behave differently from other domains (although see point 3).

Please address my specific concerns below.

1. Your method is inadequate for what you are measuring. Cognitive effects depend strongly on acute sleep (last 24–48h), not 2-week averages. You get participants to measure duration (recall bias). There is no timing relative to testing.

2. Even though you include distraction/device as covariates, this is probably not sufficient because online stroop timing validity is questionable, TOH can be solved using external aids, alternative uses test is scored by fluency - not by originality (therefore your result of less sleep associated with more divergent thinking is incorrect), your digit span is limited to 8; which creates ceiling effects in young cognitively competent adults (should have administered an adaptive digit span test - increasing in difficulty until failure).

3. Statistically, you have 4 cognition outcomes plus multiple covariates, yet there is no Bonferroni or FDR correction applied. Your results are highly vulnerable to Type I (false positive error).

Recalculate using Bonferroni.

Your regression models do not specify coefficient tables, confidence intervals or effect sizes beyond beta. For the effect sizes, there is no standardized effect interpretation or explanation of variance. Non linear effects are ignored; for example, the literature strongly supports inverted U-shaped relationship between sleep and cognition, however you have only used linear models. Fit $\text{cognition} = \beta_1(\text{sleep}) + \beta_2(\text{sleep}^2) + \epsilon$. If inverted-U emerges, your linear "negative effect" may be a truncated slice of a nonlinear curve. This can eliminate apparent reverse causality artifacts

You need to test non-linear models and perform a sensitivity analysis (with and without covariates).

E-value reasoning (how strong would an unmeasured confounder need to be?) or remove high-leverage points and re-run models. If results are unstable, your interpretation is likely spurious.

Perform a statistical power analysis. You have multiple covariates across only < 200 individuals.

This should have been performed before the start of the study.

4. Since your sample is primarily student based, you have selection bias, limited generalizability and possibly behavioral clustering. Please acknowledge in limitations.

5. Nomenclature is technically misused. Yours is not a 'sleep deprivation' study because you did not manipulate sleep deprivation.

6. You remove outliers (such as stroop) but do not justify with a sensitivity analysis or criteria.

Please do so.

7. Perform a residualization analysis. Regress sleep on covariates → get residual sleep. Regress cognition on covariates → get residual cognition. Examine residual–residual relationship. This reduces confounding from SES, age, etc. If the association disappears, your study is likely confounded, not associative (causal)

8. Perform a symmetry check by reversing the model. Report side-by-side: Model A: sleep → divergent thinking. Model B: divergent thinking → sleep. Then compare standardized β and partial R^2 (variance explained). If both are similar (between models), that would be indicative of reverse causation or bidirectionality. If one is stronger, then you can claim weak (not conclusive) evidence for uni-directionality.

Reviewer 1

R1.1. I would reframe this study as an exploratory hypothesis generating study. Remove all causal claims and frame only as being associations. Since you yourself acknowledge bidirectionality, you cannot claim that sleep influences cognition. Your strong contributions are that you measure all cognitive domains simultaneously, you (rightly) raise the reverse causality between sleep and cognition and you suggest that divergent thinking may behave differently from other domains (although see point 3).

Thank you for the suggestion. All causal claims have been revised to reflect association claims instead. The following language has been revised in the "Discussion" section, for example (page 14):

*The present findings in this study contribute to the debated findings in sleep and cognition literature; sleep duration patterns across two weeks **were only significantly associated** with divergent thinking, and no association was found between sleep duration and performance in the Stroop task, problem solving, or working memory.*

R1.2. Your method is inadequate for what you are measuring. Cognitive effects depend strongly on acute sleep (last 24–48h), not 2-week averages. You get participants to measure duration (recall bias). There is no timing relative to testing.

Thank you for your comment. We recognize that acute sleep (24h-48h) has been found to be associated with cognitive functioning. However, our primary goal of the study was to explore habitual sleep behaviors rather than one instance of sleep (Vezina-Im, 2023). Drawing from previous research looking at two-week timeframes (Bichon et al., 2025; Bradshaw et al., 2007), we wanted to focus on sleep patterns, especially because sleep duration can have high day-to-day variability among adults (Messman et al., 2022). The measure was adapted from the PSQI, which has been heavily used among adult samples. Although recall bias is always a limitation for self-report sleep, even subjective reports of sleep have been shown to be associated with cognitive functioning.

We have clarified this in the Introduction and Methodology and include a mention of sleep measures in the Limitations section.

R1.3. Even though you include distraction/device as covariates, this is probably not sufficient because online stroop timing validity is questionable, TOH can be solved using external aids, alternative uses test is scored by fluency - not by originality (therefore your result of less sleep associated with more divergent thinking is incorrect), your digit span is limited to 8; which creates ceiling effects in young cognitively competent adults (should have administered an adaptive digit span test - increasing in difficulty until failure).

Participants were instructed to complete the online tests independently and in a room without distraction as part of the procedure. Even so, we recognize the limitations in the online format of cognitive assessments and we have included these limitations in the limitations section.

The study was conducted via an online Qualtrics survey approximately 15-20 minutes long. Participants were instructed to complete the study independently and in a room with no distractions. After completing an online consent form, participants proceeded to the survey which was divided into sections: four digital cognitive tasks, demographic information (race/ethnicity, gender), distraction-level, and device type.

Terminology has been adjusted to specify the specific measure rather than broad constructs.

R1.4. Statistically, you have 4 cognition outcomes plus multiple covariates, yet there is no Bonferroni or FDR correction applied. Your results are highly vulnerable to Type I (false positive error). Recalculate using Bonferroni. Your regression models do not specify coefficient tables, confidence intervals or effect sizes beyond beta. For the effect sizes, there is no standardized effect interpretation or explanation of variance.

Thank you for your suggestion. The four cognition outcomes were tested using four separate models. The results were rerun with Bonferroni correction with consistent findings. Only a negative association between AUT fluency and sleep duration was significant.

Table 2. Results for sleep duration predicting cognitive task scores

Model	Unstandardized Beta	SE	95% CI	Standardized Beta	R ²
Stroop ~ Sleep Duration	0.18	.54	[-.88, 1.24]	0.03	.26
Tower of Hanoi ~ Sleep Duration	0.20	.26	[-.32, .71]	0.06	.25
Digit Span Task ~ Sleep Duration	0.02	.15	[-.28, .31]	-0.01	.28
AUT Fluency ~ Sleep Duration	-0.49*	.24	[-.95, -.02]	-0.18	.24

R1.5. Non linear effects are ignored; for example, the literature strongly supports inverted U-shaped relationship between sleep and cognition, however you have only used linear models. Fit cognition = beta1(sleep) + beta2(sleep^2) + epsilon. If inverted-U emerges, your linear “negative effect” may be a truncated slice of a nonlinear curve. This can eliminate apparent reverse causality artifacts You need to test non-linear models and perform a sensitivity analysis (with and without covariates). E-value reasoning (how strong would an unmeasured confounder need to be?) or remove high-leverage points and re-run models. If results are unstable, your interpretation is likely spurious.

Non-linear tests were conducted but yielded no significant results. We have included the results of non-linear regressions for clarity. Across all cognitive outcomes, quadratic terms for sleep duration were non-significant, both with and without covariates. Therefore, the current data did not find evidence for a curvilinear association between sleep duration and cognitive performance.

Cook’s distance values for all observations were below conventional thresholds ($D < 1$), suggesting no highly influential cases.

Table 4. Results for non-linear models for sleep duration predicting cognitive task scores

Model without covariates	Unstandardized Beta	SE	95% CI	Standardized Beta	R²
Stroop ~ Sleep Duration ²	0.26	.19	[-.13, .64]	.05	.00
Tower of Hanoi ~ Sleep Duration ²	-0.01	.09	[-.19, .17]	0.00	.01
Digit Span Task ~ Sleep Duration ²	-0.01	.14	[-.11, .11]	-0.01	.00
AUT Fluency ~ Sleep Duration ²	0.02	.08	[-.15, .18]	0.01	.01
Model with covariates	Unstandardized Beta	SE	95% CI	Standardized Beta	R²
Stroop ~ Sleep Duration ²	0.06	.20	[-.34, .46]	0.01	.08
Tower of Hanoi ~ Sleep Duration ²	-0.06	.10	[-.26, .13]	-0.03	.08
Digit Span Task ~ Sleep Duration ²	0.03	.16	[-.08, .14]	0.02	.12
AUT Fluency ~ Sleep Duration ²	0.05	.09	[-.12, .23]	0.02	.06

R1.6. Perform a statistical power analysis. You have multiple covariates across only < 200 individuals. This should have been performed before the start of the study.

Thank you for the comment and we apologize for the oversight. Statistical power was calculated but was omitted in the original submission. The power analysis for a general linear model set the Type I error rate at 0.05, a minimum desired power of .80, with a conservative effect size of 0.20 and 9 covariates, resulting in a required sample size of N = 72. Thus, our sample of 169 meets the criteria for statistical power. This information was added to the new “Analyses” section.

Power analysis was conducted for a general linear model, with minimum desired power set to .80, a Type I error rate of 0.05, nine covariates, and a small effect size of 0.20. This resulted in a minimum sample size of N = 72. To explore the hypothesis of whether sleep duration would be negatively associated with cognitive performance, four separate linear regression models were conducted. Sleep duration was modeled as the primary predictor variable, with race, age, gender, education level, region, device, and distraction included as covariates for all models. Scores on the Stroop task, Tower of Hanoi, digit span, and Alternative Uses Task fluency were included as separate outcomes.

R1.7. Since your sample is primarily student based, you have selection bias, limited generalizability and possibly behavioral clustering. Please acknowledge in limitations.

Thank you for the suggestion. We have acknowledged the limitations in the limitations section.

The participant population was also heavily skewed toward individuals aged 18–24, which limits the transfer of these findings to other age groups. The abundance of a younger population may be due to the process of participant outreach, which was conducted both on the St. Mary's College of Maryland campus and online; this selection bias allows for possible behavioral clustering.

R1.8. Nomenclature is technically misused. Yours is not a 'sleep deprivation' study because you did not manipulate sleep deprivation.

Thank you for the suggestion. We have replaced all uses of "sleep deprivation" and rephrased to be more accurate to the aims of the study (i.g., exploring sleep duration patterns).

*Experimental research suggests **short sleep duration** substantially impairs performance in reasoning (problem solving tasks), but does not affect short-term memory tasks (Ortega et al., 2009).*

R1.9. You remove outliers (such as stroop) but do not justify with a sensitivity analysis or criteria. Please do so.

Thank you for the suggestion. We have added more detail about both the general exclusion criteria and the outlier criteria.

Exclusion criteria were as follows: (1) participants failed to complete at least 90% of the survey or (2) provided illegible responses that could not be reliably interpreted by the researchers.

Outliers ($n = 3$) that claimed response times exceeding 1000 seconds were deemed too excessive to be true and excluded from analysis. These latencies were attributed to typographical errors such as misplaced decimals.

R1.10. Perform a residualization analysis. Regress sleep on covariates → get residual sleep. Regress cognition on covariates → get residual cognition. Examine residual–residual relationship. This reduces confounding from SES, age, etc. If the association disappears, your study is likely confounded, not associative (causal)

Thank you for the suggestion. We have run residualization analysis and found that a significant, negative association still exists between sleep duration and fluency when looking at the correlation ($r = -0.17$, $df = 167$, $p < .05$) and a simple regression ($\beta = -0.08$, $SE = .04$, $p < .05$). Thus, the relationship between fluency and sleep duration remains after reducing confounding variables.

R1.11. Perform a symmetry check by reversing the model. Report side-by-side: Model A: sleep → divergent thinking. Model B: divergent thinking → sleep. Then compare standardized β and partial R^2 (variance explained). If both are similar (between models), that

would be indicative of reverse causation or bidirectionality. If one is stronger, then you can claim weak (not conclusive) evidence for uni-directionality.

Thank you for the suggestion. We have run reverse models and found that AUT fluency scores were also significantly and negatively associated with sleep duration, though this was a weaker coefficient (-.07) compared to the model with sleep → AUT fluency (-.49). However, both R² values were similar (.23-.24), suggesting bidirectionality. We have added this information in the Results and Discussion section.

Table 3. Results for cognitive tasks predicting sleep duration

Model	Unstandardized Beta	SE	95% CI	Standardized Beta	R ²
Sleep Duration ~ Stroop	0.01	.01	[-.02, .34]	0.03	.20
Sleep Duration ~ Tower of Hanoi	0.02	.03	[-.04, .08]	0.07	.20
Sleep Duration ~ Digit Span Task	0.01	.05	[-.10, .11]	-0.01	.20
Sleep Duration ~ AUT Fluency	-0.07*	.03	[-.13, -.01]	-0.16	.23

Reviewer 2

R2.1 Please verify that all links to the references point to the correct source. Include enough recent references along with foundational ones. Ensure references directly support your claims. Avoid “padding.” Use credible sources (peer-reviewed journals, reputable books, official reports).

Thank you for the suggestion. All numbered citations have been re-done thoroughly and citation style has been corrected for outliers that were originally in MLA. All final citations are done in APA 7.

R2.2 All assumptions (including implicit assumptions) must be explicitly stated and clearly justified. The authors should explain why each assumption is reasonable and discuss its impact on the results and conclusions. Avoid overreaching conclusions that extend beyond what is supported by the data and analysis.

Thank you for the suggestion. All causal claims have been revised to reflect association claims instead.

R2.3 Verify that you have used past perfect tense and third person throughout the manuscript wherever applicable.

Thank you for your suggestion. There is no usage of present tense unless it is referencing current implications or the current state of the research field. For example: “Sleep is essential for optimal cognitive function” (Abstract). All references to past literature and the findings from this study are presented in past perfect tense.

Thank you for addressing most of my comments very comprehensively and completely. I am satisfied with all the revisions EXCEPT for the MOST important one:

My core problem still remains unaddressed, i.e you have conflated fluency with creativity. I realize that it may not be possible to go back to the drawing board and repeat the experiment. I am willing

to accept the following if you replace creativity with “idea generation rate” (fluency). Hence you can state something like “sleep is negatively associated with idea generation rate (fluency), NOT creativity or divergent thinking”.

You will need to change ALL instances in the manuscript where you continue to state “creativity” or “divergent thinking” (even if followed by the word fluency). Instead change all these to “idea generation rate” and explicitly state that “divergent thinking” and/or “creativity” were NOT measured.

This is a critical issue. Fluency reflects output quantity and is highly sensitive to disinhibition, response impulsivity, and reduced filtering—all of which may increase under sleep restriction. Therefore, the observed negative association between sleep and AUT performance may reflect increased response generation rather than increased creativity.

Reviewer 1:

Thank you for addressing most of my comments very comprehensively and completely. I am satisfied with all the revisions EXCEPT for the MOST important one:

My core problem still remains unaddressed, i.e you have conflated fluency with creativity. I realize that it may not be possible to go back to the drawing board and repeat the experiment. I am willing to accept the following if you replace creativity with “idea generation rate” (fluency). Hence you can state something like “sleep is negatively associated with idea generation rate (fluency), NOT creativity or divergent thinking”.

You will need to change ALL instances in the manuscript where you continue to state “creativity” or “divergent thinking” (even if followed by the word fluency). Instead change all these to “idea generation rate” and explicitly state that “divergent thinking” and/or “creativity” were NOT measured.

This is a critical issue. Fluency reflects output quantity and is highly sensitive to disinhibition, response impulsivity, and reduced filtering—all of which may increase under sleep restriction. Therefore, the observed negative association between sleep and AUT performance may reflect increased response generation rather than increased creativity.

Thank you for the suggestion. We have replaced all uses of “creativity” and “divergent thinking” and rephrased to be more accurate to the aims of the study (i.g., idea generation rate). We specifically state in the introduction that it is not to be confused with divergent thinking nor creativity.

Idea generation measures how many ideas are produced by an individual in a set time frame and is conceptually distinct from broader constructs such as creativity or divergent thinking (page 4).

The term “creativity” or such constructs is only retained in reference to the pre-established “dual pathway to creativity model.” Because this model’s underlying argument remains unchanged, we have kept it and changed “divergent thinking” to “idea generation rate,” which the model supports. We cannot alter the terminology of the model itself; however, we have clarified that this model refers to idea generation specifically.

The dual pathway model of creativity (42) offers a theoretical framework for the association between inhibitory control and idea generation: idea generation can emerge because of either flexibility (broad, associative thinking, which can create more solutions) or persistence in one idea (focused, goal-directed exploration, which can create deeper solutions, page 15).

Thank you for addressing my comments. My remaining concerns have to do with data and reporting inconsistencies, interpretation and over-reach. Please address these in the manuscript.

1. The abstract states “higher idea generation fluency ($\beta = -.65, p < .05$)”. But Table 2 reports Unstandardized $\beta = -0.49$ and Standardized $\beta = -0.18$. Neither value equals -0.65 .

2. Table 1 shows AUT fluency vs Stroop: $r = -0.160$. The discussion says “The positive association between increased performance in the Stroop task and fluency in the AUT...” I suggest modifying to “The positive association between increased performance (lower Stroop times) in the Stroop task and fluency in the AUT...”. This way, the negative association is clarified.

3. The interpretation of the bidirectionality analysis appears overstated. The manuscript argues that the association is stronger when sleep is modeled as the predictor because the unstandardized coefficient is larger in magnitude ($\beta = -0.49$ vs. $\beta = -0.07$). However, these coefficients are expressed in different units and are not directly comparable. The standardized coefficients (-0.18 vs. -0.16) and model R^2 values are very similar, suggesting a comparable association in both directions. Consequently, the analyses support the presence of an association between sleep duration and AUT fluency but do not provide evidence that one directional model is stronger than the other. This distinction should be clarified.

4. The reporting of model R^2 values in Table 2 may be misleading and would benefit from clarification. The manuscript reports R^2 values ranging from 0.24 to 0.28 for the regression models, despite sleep duration being a non-significant predictor in three of the four analyses. Readers may incorrectly interpret these R^2 values as indicating that sleep duration explains approximately 24–28% of the variance in cognitive performance. However, the reported R^2 values reflect the explanatory power of the entire model, including all covariates (e.g., age, gender, race, education, device type, distraction level), rather than the unique contribution of sleep duration itself.

Because the primary focus of the manuscript is the association between sleep duration and cognitive outcomes, it would be more informative to report the variance uniquely attributable to sleep.

Consider providing either (1) partial R^2 values for sleep duration or (2) incremental R^2 (ΔR^2) obtained by comparing models with and without sleep duration included. This would allow readers to assess the magnitude of the sleep effect independent of the covariates and better understand the practical significance of the reported associations.

5. The concluding statement that “highly creative individuals may be at risk of getting less sleep” is not supported by the present data. Creativity was not measured, only AUT fluency. Furthermore, the bidirectional analyses indicate that the direction of association remains unresolved. The conclusion should be revised to focus on the observed association between sleep duration and idea generation fluency rather than implying that creativity predisposes individuals to shorter sleep duration.

R1.1 Thank you for addressing my comments. My remaining concerns have to do with data and reporting inconsistencies, interpretation and over-reach. Please address these in the manuscript. The abstract states “higher idea generation fluency ($\beta = -.65, p < .05$)”. But Table 2 reports Unstandardized $\beta = -0.49$ and Standardized $\beta = -0.18$. Neither value equals -0.65 .

Thank you for your feedback. We aim to report the unstandardized beta here and this was a mistake, so the terminology has been changed as seen below:

A significant negative association was found between sleep duration and idea generation rate, such that participants who slept less showed higher idea generation fluency ($\beta = -.49, p < .05$).

R1.2 Table 1 shows AUT fluency vs Stroop: $r = -0.160$. The discussion says “The positive association between increased performance in the Stroop task and fluency in the AUT...” I suggest modifying to “The positive association between increased performance (lower Stroop times) in the Stroop task and fluency in the AUT...”. This way, the negative association is clarified.

Thank you for your feedback. This change has been applied. Here is the new version:

The positive association between increased performance (shorter Stroop times) in the Stroop task and fluency in the AUT aligns with prior evidence suggesting increased inhibitory control loosens associative constraints and fosters more flexible idea generation (42) ($r = .16$, Table 1).

R1.3 The interpretation of the bidirectionality analysis appears overstated. The manuscript argues that the association is stronger when sleep is modeled as the predictor because the unstandardized coefficient is larger in magnitude ($\beta = -0.49$ vs. $\beta = -0.07$). However, these coefficients are expressed in different units and are not directly comparable. The standardized coefficients (-0.18 vs. -0.16) and model R^2 values are very similar, suggesting a comparable association in both directions. Consequently, the analyses support the presence of an association between sleep duration and AUT fluency but do not provide evidence that one directional model is stronger than the other. This distinction should be clarified.

Thank you for your suggestion. The discussion section referring to this issue has been changed to reflect your concerns. The new version is as follows:

This suggests that participants with lower average sleep durations in the two weeks prior to taking the survey had higher idea generation rates. Follow-up analyses examining the reverse direction were also significant ($\beta = -.07$, $p < .05$, Table 3). The similarity in size of the standardized coefficients across both models ($\beta = -.18$, Table 2 versus $\beta = -.16$, Table 3) suggests the association between sleep duration and AUT fluency is comparable in both directions. The idea that increased idea generation rates may contribute to shorter sleep duration warrants further research.

R1.4 and R1.5 The reporting of model R^2 values in Table 2 may be misleading and would benefit from clarification. The manuscript reports R^2 values ranging from 0.24 to 0.28 for the regression models, despite sleep duration being a non-significant predictor in three of the four analyses. Readers may incorrectly interpret these R^2 values as indicating that sleep duration explains approximately 24–28% of the variance in cognitive performance. However, the reported R^2 values reflect the explanatory power of the entire model, including all covariates (e.g., age, gender, race, education, device type, distraction level), rather than the unique contribution of sleep duration itself.

Because the primary focus of the manuscript is the association between sleep duration and cognitive outcomes, it would be more informative to report the variance uniquely attributable to sleep. Consider providing either (1) partial R^2 values for sleep duration or (2) incremental R^2 (ΔR^2) obtained by comparing models with and without sleep duration included. This would allow readers to assess the magnitude of the sleep effect independent of the covariates and better understand the practical significance of the reported associations.

Thank you for your suggestion. Tables have been updated to include partial R^2 values. Partial R^2 for sleep duration on AUT fluency was 0.02, and the R^2 for sleep duration on all other cognitive variables was 0. The same is true for Table 3.

Table 2. Results for sleep duration predicting cognitive task scores

Model	Unstandardized Beta	SE	95% CI	Standardized Beta	R^2	Partial R^2
Stroop ~ Sleep Duration	0.18	.54	[-.88, 1.24]	0.03	.26	.00

Tower of Hanoi ~ Sleep Duration	0.20	.26	[-.32, .71]	0.06	.25	.00
Digit Span Task ~ Sleep Duration	0.02	.15	[-.28, .31]	-0.01	.28	.00
AUT Fluency ~ Sleep Duration	-0.49*	.24	[-.95, -.02]	-0.18	.24	.02

Table 3. Results for cognitive tasks predicting sleep duration

Model	Unstandardized Beta	SE	95% CI	Standardized Beta	R ²	Partial R ²
Sleep Duration ~ Stroop	0.01	.01	[-.02, .34]	0.03	.20	.00
Sleep Duration ~ Tower of Hanoi	0.02	.03	[-.04, .08]	0.07	.20	.00
Sleep Duration ~ Digit Span Task	0.01	.05	[-.10, .11]	-0.01	.20	.00
Sleep Duration ~ AUT Fluency	-0.07*	.03	[-.13, -.01]	-0.16	.23	.02

This sentence has been added before both tables:

Partial R² values indicated sleep duration explained no unique variance in Stroop performance, TOH, or DSMT scores (partial R² = .00), and accounted for a small proportion of variance in AUT fluency (partial R² = .02).

R1.6 The concluding statement that “highly creative individuals may be at risk of getting less sleep” is not supported by the present data. Creativity was not measured, only AUT fluency. Furthermore, the bidirectional analyses indicate that the direction of association remains unresolved. The conclusion should be revised to focus on the observed association between sleep duration and idea generation fluency rather than implying that creativity predisposes individuals to shorter sleep duration.

Thank you for your feedback. The conclusion has been changed to reflect these concerns. No mention of ‘creativity’ remains, and the bidirectionality of the significant results becomes a critical ending point.

Here is the new version as follows:

While inhibition, problem-solving, and working memory were not significantly related to sleep duration in this study, a negative association between sleep duration and idea generation rate was observed. These findings highlight the need for further research into the relationship between sleep duration and idea generation rate, and whether the association reflects bidirectionality or a directional effect.

Thank you for addressing my comments. Now that I see a partial R^2 of 0.02, your conclusions overstate the breadth of the findings. The data do not demonstrate differential associations across multiple cognitive domains; rather, sleep duration was unrelated to inhibitory control, problem solving, and working memory, and only a small association was observed with idea-generation fluency (partial $R^2 = .02$). Please revise the concluding interpretation to more directly reflect this pattern of results. Please also revise the abstract, and discussion as appropriate.

R1.1 Thank you for addressing my comments. Now that I see a partial R^2 of 0.02, your conclusions overstate the breadth of the findings. The data do not demonstrate differential associations across multiple cognitive domains; rather, sleep duration was unrelated to inhibitory control, problem solving, and working memory, and only a small association was observed with idea-generation fluency (partial $R^2 = .02$). Please revise the concluding interpretation to more directly reflect this pattern of results. Please also revise the abstract, and discussion as appropriate.

Thank you for your feedback. Every time the significance of the association between sleep duration and idea generation rate is suggested, the effect size is also mentioned to directly reflect the full picture of the results. The abstract has been changed to reflect this:

A small significant negative association was found between sleep duration and idea generation rate, such that participants who slept less showed higher idea generation fluency ($\beta = -.49$, $p < .05$, partial $R^2 = .02$).

The discussion has been changed to reflect this:

...there was a negative association between sleep duration and fluency in the AUT observed ($\beta = -.49$, $p < .05$, Table 2). This suggests that participants with lower average sleep durations in the two weeks prior to taking the survey had higher idea generation rates. Follow-up analyses examining the reverse direction were also significant ($\beta = -.07$, $p < .05$, Table 3). The similarity in size of the standardized coefficients across both models ($\beta = -.18$, Table 2 versus $\beta = -.16$, Table 3) suggests the association between sleep duration and AUT fluency is comparable in both directions. However, the unique variance in idea generation rate explained by sleep duration is small, suggesting that idea generation is not uniquely sensitive to sleep duration relative to other factors (partial $R^2 = .02$ in Table 2, 3). The partial R^2 value provides evidence of an observable but limited relationship between sleep duration and idea generation rate, rather than indicating sleep duration is a substantial determinant of idea generation rate. The small yet significant association warrants further research into the underlying reasons why this association emerged in idea generation rather than other cognitive domains.

Mention of the differential associations have been removed. The conclusion has been changed to reflect this:

Sleep duration was not significantly linked to inhibitory control, problem-solving, and working memory; a significant negative association between sleep duration and idea generation rate was observed, although the effect size was small.

Several statements continue to imply a degree of specificity or importance that is not supported by the reported effect size.

For example, the Abstract concludes:

“These findings highlight idea generation as particularly sensitive to sleep duration, whereas other domains of cognitive functioning remain intact.”

This statement remains too strong. The present study demonstrates only that one outcome reached statistical significance whereas three others did not. It does not demonstrate that idea generation is uniquely or particularly sensitive to sleep duration. Demonstrating domain specificity would require formal statistical comparisons between the observed regression coefficients (or interaction analyses), which were not performed. A more appropriate conclusion would simply state that, among the four cognitive measures examined, only idea-generation fluency showed a statistically significant association with sleep duration, although the unique variance explained was small. Similarly, several statements throughout the Discussion should be moderated.

For example:

“...the association emerged in idea generation rather than other cognitive domains.”

This wording implies that idea generation is genuinely different from the other cognitive domains. The present analyses do not establish this. Rather, they demonstrate that only one regression coefficient achieved statistical significance within this sample. Absence of statistical significance in the remaining outcomes should not be interpreted as evidence that those domains differ from idea generation.

Likewise, the statement:

“...idea generation is not uniquely sensitive to sleep duration relative to other factors...”

extends beyond what the reported statistics support. Partial R^2 quantifies the amount of unique variance explained by sleep duration, but it does not establish that other unmeasured variables are more important. The manuscript should instead state simply that sleep duration accounted for only a very small proportion of the observed variability in idea-generation fluency.

The following sentence should also be interpreted more cautiously:

“The small yet significant association warrants further research into the underlying reasons why this association emerged in idea generation rather than other cognitive domains.”

Although future research is certainly warranted, this wording again assumes that idea generation is genuinely distinct from the other domains. Given the very small effect size, the more scientifically appropriate interpretation is that this preliminary association should first be independently replicated before conclusions regarding cognitive-domain specificity are considered.

Finally, the Conclusion still contains language that somewhat overstates the findings:

“These findings highlight the need for further research into the relationship between sleep duration and idea generation rate...”

This statement would be strengthened by explicitly acknowledging that the observed relationship was weak. For example, it could state that the present findings provide preliminary evidence of a weak association that requires replication using larger samples and more objective measures of sleep before stronger conclusions can be drawn.

2. Because the primary outcome is derived from participants' Alternative Uses Task (AUT) responses, I encourage the authors to consider providing the de-identified AUT responses as Supplementary Material (or as an accompanying Excel/CSV file). Although not essential for publication, this would substantially enhance the transparency and long-term value of the dataset. Availability of the raw responses would enable future investigators to evaluate alternative scoring approaches, including originality, flexibility, elaboration, semantic-distance metrics, or other emerging computational methods that extend beyond the fluency measure analyzed in the present study. Such data sharing would facilitate methodological comparisons and allow future researchers to build upon the current work without requiring additional data collection.

R1.1 Several statements continue to imply a degree of specificity or importance that is not supported by the reported effect size. For example, the Abstract concludes: “These findings highlight idea generation as particularly sensitive to sleep duration, whereas other domains of cognitive functioning remain intact.” This statement remains too strong. The present study demonstrates only that one outcome reached statistical significance whereas three

others did not. It does not demonstrate that idea generation is uniquely or particularly sensitive to sleep duration. Demonstrating domain specificity would require formal statistical comparisons between the observed regression coefficients (or interaction analyses), which were not performed. A more appropriate conclusion would simply state that, among the four cognitive measures examined, only idea-generation fluency showed a statistically significant association with sleep duration, although the unique variance explained was small.

Thank you for your suggestion. I understand your qualms with the strength of the language and have changed it to reflect this. The final sentence of the abstract is now:

Idea-generation fluency was the only one of the four cognitive measures examined to show a statistically significant association with sleep duration, although the unique variance explained was small.

R1.2 Similarly, several statements throughout the Discussion should be moderated. For example: "...the association emerged in idea generation rather than other cognitive domains." This wording implies that idea generation is genuinely different from the other cognitive domains. The present analyses do not establish this. Rather, they demonstrate that only one regression coefficient achieved statistical significance within this sample.

Absence of statistical significance in the remaining outcomes should not be interpreted as evidence that those domains differ from idea generation.

Thank you for your suggestion. This sentence has been removed and replaced as per the instructions in R1.4.

R1.3 Likewise, the statement: "...idea generation is not uniquely sensitive to sleep duration relative to other factors..." extends beyond what the reported statistics support. Partial R^2 quantifies the amount of unique variance explained by sleep duration, but it does not establish that other unmeasured variables are more important. The manuscript should instead state simply that sleep duration accounted for only a very small proportion of the observed variability in idea-generation fluency.

Thank you for your suggestion. The sentence has been changed using this terminology:

However, sleep duration accounted for only a very small proportion of the observed variability in idea-generation fluency (partial $R^2 = .02$ in Table 2, 3).

R1.4 The following sentence should also be interpreted more cautiously: "The small yet significant association warrants further research into the underlying reasons why this association emerged in idea generation rather than other cognitive domains." Although future research is certainly warranted, this wording again assumes that idea generation is genuinely distinct from the other domains. Given the very small effect size, the more scientifically appropriate interpretation is that this preliminary association should first be independently replicated before conclusions regarding cognitive-domain specificity are considered.

Thank you for your suggestion. This sentence has been changed to reflect this:

The small yet statistically significant preliminary association warrants replication before conclusions are considered regarding why statistical significance emerged for idea generation specifically. Given the modest effect size, it remains unclear if this is a domain-specific association or a sample-specific variability.

R1.5 Finally, the Conclusion still contains language that somewhat overstates the findings: “These findings highlight the need for further research into the relationship between sleep duration and idea generation rate...” This statement would be strengthened by explicitly acknowledging that the observed relationship was weak. For example, it could state that the present findings provide preliminary evidence of a weak association that requires replication using larger samples and more objective measures of sleep before stronger conclusions can be drawn.

Thank you for your suggestion. The sentence has been changed to reflect your terminology. It now reads as:

The present findings offer preliminary evidence of a weak association, highlighting the need for replication using larger samples and more objective sleep measures before stronger conclusions can be drawn.

R1.6 Because the primary outcome is derived from participants’ Alternative Uses Task (AUT) responses, I encourage the authors to consider providing the de-identified AUT responses as Supplementary Material (or as an accompanying Excel/CSV file). Although not essential for publication, this would substantially enhance the transparency and long-term value of the dataset. Availability of the raw responses would enable future investigators to evaluate alternative scoring approaches, including originality, flexibility, elaboration, semantic-distance metrics, or other emerging computational methods that extend beyond the fluency measure analyzed in the present study. Such data sharing would facilitate methodological comparisons and allow future researchers to build upon the current work without requiring additional data collection.

Thank you for your suggestion. The excel sheet has been added as an attached document.

Thank you for addressing my comments. Accepted.