SI

Supplemental Information on Methods

Timeline of Study Activities

See Figure S1 for a graphical display of a timeline of study activities, including writing exercises paired with semester final exams and the end of the year survey.

Additional Information on Randomization and Student Writing Exercise Materials

Participants were randomly assigned to receive one of four different writing exercises, including three intervention exercises and one control exercise. Students were randomized to condition with a randomized block design. In each year of data collection, students were randomized to condition within blocking factors of gender and economic background in order to ensure equal numbers of male and female and higher and lower income students across conditions. Regardless of the writing exercise, students were given ten minutes to complete the writing exercise before their semester final exams.

Expressive writing intervention. When assigned to the expressive writing condition, students were asked to write freely about the exam they were about to take. The prompt was as follows:

> *Please take the next 10 minutes to write as openly as possible about your thoughts and feelings about the exam you are about to take. In your writing, really let yourself go and explore your emotions and thoughts as you are getting ready to start the exam. You might write about your current thoughts or write about how you have felt during other similar situations at school or in other situations in your life. Please try to be as open as possible as you write about your thoughts at this time.*

Reappraisal intervention. When assigned to the reappraisal condition, students were asked to read a short passage about reappraising the physiological arousal associated with stress and anxiety. Following the passage, they were asked to answer two free response questions to help internalize the message of the passage. A different passage was used for the fall and spring semester final exams in order to avoid repetitiveness.

Fall Semester Reappraisal Instructions, Passage, and Questions:

> *As you might expect, taking a test can be a stressful experience. Before starting your exam, we are going to go through a procedure designed to help you perform at your best.*
>
> *First, please read over some information taken from a scientific journal article. The information is about how our body's response to stress helps improve performance. While you read, we would like you to think about how these bodily reactions can help your test performance today.*
>
> *After the reading, you will be asked to answer two questions about what you just read.*
>
> *This reading is based on Nock's 2011 study in the Journal of Psychophysiology*
>
> *Sometimes in important situations, people notice that they have a faster heartbeat, sweaty palms, shortness of breath, butterflies in their stomach, and lots of energy running through their body. People usually think that this means that they are nervous, anxious, or worried. However, these feelings happen for all kinds of reasons, and it does not mean that we need to feel worried or nervous. For example, we feel this same way when we are*

excited about a surprise, when we are getting ready for a fun sports competition, or when we fall in love. So, feeling a faster heartbeat, for example, doesn't mean you will perform badly. Having these feelings could actually help you!

This is because when people care about something, such as doing well on a test, our body's nervous system tells the body to release energy and deliver more oxygen to the brain. This helps you to stay alert and pay attention to the important thing that is going on in your life. Therefore, experiencing a faster heartbeat, heavy breathing, or sweaty palms could actually be a good thing. It is your body's way of pumping you full of energy and attention! But it all depends on whether you choose to use this energy.

As you are getting ready to take this test, just keep in mind that the feelings and symptoms you may be experiencing are normal. It is just your body's way of getting you prepared to tackle and deal with something important.

Remember that the increase in energy you are experiencing is helping you, so take advantage of all that extra energy and attention!

Question 1: How do people sometimes feel in important situations?

Question 2: How can the way a person feels in important situations help them do well in those situations?

Spring Semester Reappraisal Instructions, Passage, and Questions:

As you might expect, taking a test can be a stressful experience. Before starting your exam, we are going to go through a procedure designed to help you perform at your best.

First, please read over some information taken from a scientific journal article. The information is about how our body's response to stress helps improve performance. While you read, we would like you to think about how these bodily reactions can help your test performance today.

After the reading, you will be asked to answer two questions about what you just read.

This reading is based on Craske and Barlow's 2012 Client Workbook for Anxiety and Panic

Anxiety or feeling nervous is a normal physical reaction that helps you deal with things that make you stressed. It is not harmful. In fact, if we did not have these reactions we could not survive. If feeling nervous is helpful, why can it feel so bad?

Feeling nervous because of tests is the result of how we think about our body's responses. When the fight-flight system activates, our brain tries to find if there is something dangerous around us. However, in modern society there is often no obvious physical danger or threat. When no danger can be found, then our brains can start to invent reasons such as, "There must be something wrong with me." Nothing could be further from the truth.

During times when you feel nervous or anxious, remember that your body's responses are good. A faster heartbeat, sweating, and heavy breathing all help bring oxygen where it is needed. For example, this means that your body is releasing energy and more oxygen to the brain, which helps you stay alert and pay attention to whatever is happening currently. That means that a faster heartbeat, heavy breathing, or sweaty palms could actually be a useful thing. It is your body's way of filling you with energy and improving your attention!

So, if you find yourself feeling nervous or anxious in while taking a test, think about how your body's responses can actually energize and help you.

Question 1: How do people sometimes react when they feel nervous? Why does this happen?

Question 2: How can the way a person feels in stressful situations help them do well in those situations?

Combined expressive writing and reappraisal intervention. Participants assigned to the combined condition saw the prompts from both the expressive writing and reappraisal interventions (identical prompts as described above). However, because students still only had ten minutes to complete the writing exercise, they were given less space for their expressive writing response, making it possible to complete the combined intervention within the allotted time.

Active control condition. In the active control condition, participants read a passage that instructed students to ignore their stress while taking the exam. When assigned to the active control condition, students were asked to read a short passage about ignoring stress. Following the passage, they were asked to answer two free response questions to help internalize the message of the passage. A different passage was used for the fall and spring semester final exams in order to avoid repetitiveness.

Fall Semester Control Instructions, Passage, and Questions:

> *As you might expect, taking a test can be a stressful experience. Before starting your exam, we are going to go through a procedure designed to help you perform at your best.*
>
> *First, please read over some information taken from a scientific journal article. The information is about the benefits of ignoring stress when you feel it. While you read, we would like you to think about how ignoring your stress can help your test performance today.*
>
> *After the reading, you will be asked to answer two questions about what you just read.*
>
> *This reading is based on Craske and Barlow's 2001 Clinical Handbook of Psychological Disorders*

Worry or anxiety is a common way to react to stress that can be decreased by not paying attention to what makes us worried. Ignoring things that make your worry is not harmful. In fact, if we were not able to do this, we could not function in a modern society where there are many possible sources of harm.

Worrying about if we will do well on a test hurts our ability to do well on the test because we spend time thinking about all the bad things that could happen. To decrease these bad thoughts, people should try to pay attention to something other than their stress.

Once again, if you find yourself feeling stressed during today's exam, focus on ignoring the stress you feel.

Question 1: How do people sometimes feel in stressful situations? How can this affect their test performance?

Question 2: How can changing what a person pays attention to help during stressful situations? What advice does the article give about what to do if a person feels stressed during the test?

Spring Semester Control Instructions, Passage, and Questions:

As you might expect, taking a test can be a stressful experience. Before starting your exam, we are going to go through a procedure designed to help you perform at your best.

First, please read over some information taken from a scientific journal article. The information is about the benefits of ignoring stress when you feel it. While you read, we

would like you to think about how ignoring your stress can help your test performance today.

After the reading, you will be asked to answer two questions about what you just read.

This reading is based on S.L. Smith's 2008 article in the Journal of Experimental Psychology

In important situations, people tend to feel worried which activates of the nervous system. The nervous system affects blood flow, breathing, and your hormones. Heart rate increases as our body tries to move blood to muscles that will help us extract ourselves from stressful situations. Breathing becomes faster and deeper because of our body's need for more oxygen.

This research shows that focusing on something safe in one's environment can help to decrease the negative impact of stress on our bodies. In other words, when people do not notice the cause of their worries, their body's responses were much healthier. Thus, sometimes it is good to be "blissfully unaware" and to try to ignore your stressors.

During your exam today, remember that if you do not pay attention to your stress, you will be less likely to experience negative physiological symptoms.

Question 1: How do people sometimes feel in important situations? How does the nervous system react?

Question 2: How does ignoring stress affect a person's physical responses? How can this help on a test?

Additional Information on Measures

Student economic background. Our student economic background measure, which we used to divide students into higher- and lower-income groups, was based on whether they qualified for free or reduced lunch status at their school. In total, 285/1175 students were in the lower-income group (233 received free lunch, and 52 received reduced lunch). Because the reduced lunch group was so small, we could not conduct subgroup analyses within only this group. Instead, we combined students who received either free or reduced lunch into one lower-income group of students for our analyses.

Biology exams. All students received the same final exam each semester regardless of their classroom or teacher, which allowed for straightforward comparisons across classes. Participants received an average exam score of 75.75% (SD = 15.75).

Academic records. We received information for student free or reduced lunch status, student race, and student gender from their official school records.

Post-Intervention Survey. Students were given a survey at the end of the school year to assess their attitudes about test-taking; in particular, the survey measured whether they viewed test-related anxiety as potentially helpful instead of uniformly harmful (39). This scale was based on a subscale in a measure of test anxiety developed for students in this age range (39). In total, 73% of students completed the survey, and we tested whether the students who did not complete

the survey differed from the students who did complete the survey. Importantly, attrition analyses revealed that the survey was not missing at different rates across experimental condition, χ^2 (3, N = 1175) = 3.37, p = .34. However, lower-income students (61.40%) were significantly less likely to complete the survey than higher-income students (77.10%), χ^2 (1, N = 1175) = 26.29, p < .001. Because of this differential rate of survey completion by student economic background, it is possible that analyses on student reappraisal of test anxiety are less generalizable than analyses on student achievement measures, which we have for all participating students.

Supplemental Information on Primary Analyses, Secondary Analyses, and Robustness Checks

Randomization Checks

To assess whether students were successfully randomized to condition, we tested whether there was similarity in student demographics across the four conditions. We found no significant differences between conditions in terms of the proportion of lower- and higher-income students, χ^2 (3, N = 1175) = .61, p = .90, male/female students, χ^2 (3, N = 1175) = 2.18, p = .54, or students from White/Asian and African American/Latino racial groups, χ^2 (3, N = 1153) = .31, p = .96. Additionally, we found that prior achievement did not significantly differ across experimental conditions, F(3,1172) = .50, p = .68.

Primary Analyses: Tests for Differences in Effectiveness Between Intervention Groups

In the primary analyses reported in the paper (see Table S1), there were two orthogonal contrasts in each model that tested differences between intervention groups: one tested whether the two individual intervention conditions differed from the combined intervention condition and the other compared the expressive writing condition to reappraisal condition. Specific results

from those tests are included here. There were no significant effects of either of those contrasts on average exam score, $F(1,1174) = .38$, $p = .54$; $F(1,1174) = .38$, $p = .54$, respectively, or course passing rate, $\chi^2 (1, N = 1175) = .30$, $p = .59$; $\chi^2 (1, N = 1175) = 2.67$, $p = .10$, respectively. On reappraisal of test anxiety, there was no significant difference when comparing the expressive writing and reappraisal conditions to the combined condition, $F(1,860) = .04$, $p = .83$. There was a difference between the expressive writing condition and reappraisal condition on reappraisal of test anxiety, $F(1,860) = 4.62$, $p = .03$. Students in the reappraisal condition showed higher levels of reappraisal of test anxiety than students in the expressive writing condition.

Primary Analyses: Effects of Prior Achievement in Main Models

As would be expected, the covariate in the models presented in the main paper (i.e., prior year achievement) was a significant predictor of exam performance, $F(1,1174) = 1223.83$, $p < .001$, course passing rate, $\chi^2 (1, N = 1175) = 66.41$, $p < .001$, and reappraisal of test anxiety, $F(1,860) = 16.96$, $p < .001$.

Primary Analyses: Hierarchical Regression to Test Additional Interactions Between Contrasts that Compare Treatment Groups and Students' Economic Background

In the models used in the primary analyses, students' economic background was only interacted with the all interventions as compared to the control group contrast and not with the two other centered contrasts, which tested for differences between intervention groups (one comparing expressive writing and reappraisal to the combined condition and the other comparing expressive writing to reappraisal). With respect to our hypotheses, we did not expect that any intervention group differences would differ by student income. That is, even though, for example, we expected that the combined intervention group might produce stronger benefits than either the expressive writing or reappraisal groups, we did not expect any additional benefits of

the combined group over those two other intervention groups to differ for higher- and lower-income students. Beyond not having hypotheses involving those interactions, empirical evidence also did not show that any additional significant variance would be explained by including student income interactions with the two centered contrasts that tested for differences between intervention groups. Using hierarchical regression, we found no significant improvement in predicting any of the three outcomes when adding these two interaction terms to the models (exam score: r^2 change < .001, F change (2,1166) = .66, p = .52; course passing rate: χ^2 (2, N = 1175) = .81, p = .67; reappraisal of test anxiety: r^2 change = .001, F change (2,852) = .54, p = .58).

Secondary Analyses: Tests for Differential Intervention Effects Across Semesters

Although our primary analyses focused on the effects on students' average exam score, we ran secondary analyses to assess whether those effects differed across semesters (i.e., the semester 1 exam score and semester 2 exam score). We ran a mixed ANCOVA to assess this possibility with semester of the exam as the within-subjects factor and then the between subject factors of student income, the interventions vs. control contrast, the interaction between these two variables, and also prior achievement as a control variable. There was a significant main effect of the intervention, F(1,1170) = 8.02, p < .01, and intervention by student income interaction, F(1,1170) = 8.92, p < .01, indicating that students performed better on their exams when given an intervention and that this effect was specific to lower-income students. However, there was neither a significant intervention by time two-way interaction, F(1,1170) = .14, p = .71, nor a significant intervention by student income by time three-way interaction, F(1,1170) = .42, p = .52, indicating that the benefits of the interventions on students' exam scores did not differ across the two semesters.

Secondary Analyses: Comparisons of Individual Intervention Groups to the Control Group

Our a priori hypothesis was that all three interventions tested would be beneficial as compared to the control group. Thus, our primary analyses compared students who were assigned to receive any of the three interventions to students who were assigned to receive the control writing exercise. These primary analyses allow for the most reliable results because students are not broken into smaller groups, especially for our subgroup of interest, lower-income students. However, in secondary analyses reported here, we used dummy coded variables to assess the effect of each individual intervention as compared to the control group (Table S2).

These regression models regressed each outcome on the following set of base predictors: three dummy codes (one for each intervention vs. control contrast), student income (dummy coded as well with 0 coded as lower-income since that would allow the main models to show intervention effects for this targeted reference group), the interaction terms between each intervention/control contrast and student income, and a control measure of students' prior achievement. When there were significant interactions, we explored whether there were significant effects of each intervention group as compared to the control group for students from higher- and lower-income backgrounds.

Effects on exam performance. The base predictors were used to predict students' average semester final exam performance (Figure S2, Panel A). There was a significant effect of student income, $F(1,1174) = 85.22$, $p < .001$, such that students from lower-income backgrounds performed worse on the exams than students from higher-income backgrounds. There were significant positive main effects of the expressive writing intervention, $F(1,1174) = 10.64$, $p = .001$, reappraisal intervention, $F(1,1174) = 4.47$, $p = .04$, and the combined expressive writing and reappraisal intervention, $F(1,1174) = 7.20$, $p = .01$. These effects should be interpreted as the

effect of each of those interventions, as compared to the control group, for students from lower-income backgrounds because lower-income students were the reference group (i.e., coded as 0). There were also significant interactions between student income and the expressive writing contrast, $F(1,1174) = 8.63$, $p < .01$, reappraisal contrast, $F(1,1174) = 3.76$, $p = .05$, and the combined expressive writing and reappraisal contrast, $F(1,1174) = 5.27$, $p = .02$, indicating that the intervention group effects significantly differed between students from higher and lower-income backgrounds. There were no significant effects of any of the interventions for higher-income students. Finally, students' prior achievement was significantly positively associated with mean exam performance, $F(1,1174) = 1219.03$, $p < .001$.

Effects on course passing rate. In addition to exam performance, the base predictors were used to predict students' course passing rate in a logistic regression, which was utilized because of the binary nature of the data (i.e., students either passed both semesters or failed at least one semester). Results are graphed in Figure S2, Panel B. There was a significant effect of student income, χ^2 (1, N = 1175) = 18.74, $p < .001$, such that a higher proportion of students from higher-income backgrounds passed the course than students from lower-income backgrounds. There were significant positive effects of the expressive writing intervention, χ^2 (1, N = 1175) = 11.17, $p = .001$, and reappraisal intervention, χ^2 (1, N = 1175) = 3.92, $p = .048$, but not the combined expressive writing and reappraisal intervention, χ^2 (1, N = 1175) = 1.07, $p = .30$. These effects should be interpreted as the effect of each of those interventions, as compared to the control group, for students from lower-income backgrounds because lower-income students were the reference group (i.e., coded as 0). There were also significant interactions between student income and the expressive writing contrast, χ^2 (1, N = 1175) = 4.28, $p = .04$, but not with the reappraisal contrast, χ^2 (1, N = 1175) = 3.69, $p = .06$, or the combined expressive writing and

reappraisal contrast, χ^2 (1, N = 1175) = .64, p = .42, indicating that the expressive writing intervention effect significantly differed between higher- and lower-income students. There were no significant effects of the expressive writing intervention for students from higher-income backgrounds. Finally, students with higher levels of prior achievement were more likely to pass the course, χ^2 (1, N = 1175) = 66.57, p < .001.

Effects on students' reappraisal of test anxiety. Students completed a survey at the end of the school year to assess their attitudes about test anxiety (Figure S2, Panel C). There was a significant effect of student income, F(1,860) = 5.50, p = .02, such that students from lower-income backgrounds had lower levels of reappraisal of test anxiety than students from higher-income backgrounds. There was not a significant main effect of the expressive writing contrast, F(1,860) = 2.74, p = .10; however, there were significant effects for the reappraisal contrast, F(1,860) = 6.79, p = .01, and the combined expressive writing and reappraisal contrast, F(1,860) = 7.18, p = .01. These effects should be interpreted as the effect of each of those interventions, as compared to the control group, for students from lower-income backgrounds because students from lower-income backgrounds were the reference group (i.e., coded as 0). There was only a significant interaction between student income and the combined expressive writing and reappraisal contrast, F(1,860) = 5.12, p = .02, indicating that the effect of the combined expressive writing and reappraisal intervention differed between students from higher and lower-income backgrounds. No interaction between student income and the reappraisal or expressive writing contrast indicates that the reappraisal and expressive writing intervention effects did not significantly differ between students from higher and lower-income backgrounds. There was no significant effect of the combined expressive writing and reappraisal intervention for students

from higher-income backgrounds. Finally, there was a significant positive association between students' prior achievement and reappraisal of test anxiety, $F(1,860) = 16.51$, $p < .001$.

Secondary Analyses: Effect Sizes

Effect sizes and 95% confidence intervals for intervention effects among lower-income students for all primary and secondary intervention and control group comparisons are shown in Table S3.

Secondary Analyses: Do Intervention Effects Differ by Race or Class?

Although 84% of lower-income students in this study were underrepresented minority (URM) students (African American, Latino, multiracial, American Indian, or Pacific Islander), only 55% of URM students were from lower-income backgrounds. Therefore, in this school, being part of an underrepresented minority group was not highly overlapping with students' economic background, as it is in many schools. That is, even though most lower-income students were also members of underrepresented minority groups, only about half of underrepresented minority students came from lower-income backgrounds in this school.

In terms of the benefits of the interventions, we saw that they are specific to lower-income students. We conducted supplemental analyses to tease apart the effects of race and economic background, at least for URM students. Because there were not enough lower-income White and Asian students to test as a separate group, we created three groups of students: White and Asian students (across all economic backgrounds), URM students from higher-income backgrounds, and URM students from lower-income backgrounds. We tested regression models with the following independent variables as predictors of all outcomes: prior performance, a dummy coded contrast for White and Asian students, a dummy coded contrast for higher-income URM students, an all interventions vs. control group contrast (0 for control group, 1 for students

assigned to any of the three intervention groups), and both two-way interactions between the intervention contrast and the race/class dummy codes (i.e., the White/Asian student dummy code and the higher-income URM dummy code). The aforementioned dummy codes mean that the reference group is lower-income URM students. This also means that the main effect of the intervention contrast in this model represented the effect of the interventions for lower-income URM students. Results showed that intervention effects depended on students' economic background but not their race. This is because benefits of the interventions were significant for the lower-income (URM) group of students but not for higher-income students, regardless of race.

For exam scores, there were significant main effects of the White/Asian contrast, $F(1,1174) = 98.95$, $p < .001$, the higher-income URM contrast, $F(1,1174) = 53.70$, $p < .001$, the intervention contrast, $F(1,1174) = 10.63$, $p = .001$, the interaction between the White/Asian and intervention contrast, $F(1,1174) = 7.37$, $p = .01$, and the interaction between the higher-income URM contrast and the intervention contrast, $F(1,1174) = 10.81$, $p = .001$. These results showed that the intervention had significant benefits for lower-income URM students, $F(1,238) = 5.68$, $p = .02$, but not White/Asian students, $F(1,738) = .002$, $p = .97$, or higher-income URM students, $F(1,196) = 1.70$, $p = .19$. Finally, students' prior achievement was significantly positively associated with mean exam performance, $F(1,1174) = 1058.26$, $p < .001$.

For course passing rate, there were significant main effects of the White/Asian contrast, $\chi^2 (1, N = 1175) = 10.78$, $p = .001$, the higher-income URM contrast, $\chi^2 (1, N = 1175) = 9.35$, $p < .01$, the intervention contrast, $\chi^2 (1, N = 1175) = 8.42$, $p < .01$, no significant effect of the interaction between the White/Asian and intervention contrast, $\chi^2 (1, N = 1175) = 1.02$, $p = .31$, and a significant interaction between the higher-income URM contrast and the intervention

contrast, χ^2 (1, N = 1175) = 4.76, p = .03. These results showed that the intervention had significant benefits for lower-income URM students, χ^2 (1, N = 239) = 8.79, p < .01, but not White/Asian students, χ^2 (1, N = 739) = .03, p = .86, or higher-income URM students, χ^2 (1, N = 197) = 1.07, p = .30. Finally, students' prior achievement was significantly positively associated with course passing rate, χ^2 (1, N = 1175) = 49.71, p < .001.

For reappraisal of test anxiety, there was a significant main effect of the White/Asian contrast, F(1,860) = 5.06, p = .03, no significant effect of the higher-income URM contrast, F(1,860) = 1.87, p = .17, a significant effect of the intervention contrast, F(1, 860) = 6.26, p = .01, no significant effect of the interaction between the White/Asian and intervention contrast, F(1,860) = 3.35, p = .07, and no significant effect of the interaction between the higher-income URM contrast and the intervention contrast, F(1,860) = 3.10, p = .08. Subgroup analyses showed that the intervention had significant benefits for lower-income URM students, F(1,144) = 5.33, p = .02, but not White/Asian students, F(1,579) = .80, p = .37, or higher-income URM students, F(1,135) < .001, p = .99. Finally, students' prior achievement was significantly positively associated with reappraisal of test anxiety, F(1, 860) = 12.97, p < .001.

Robustness Checks: Hierarchical Linear Modeling

Hierarchical linear modeling (HLM) allowed us to account for students being nested within classrooms, as an additional robustness check on these results. We preferred single level models for our primary analyses for a variety of reasons. First, multilevel models are most helpful when a study involves only a sample of level-2 units (i.e., classrooms in this case), but our study includes all 9th grade biology classrooms in the school. Second, when an empty model was estimated to assess the variance explained at the classroom level (as compared to at the student level), we found that very little of the variance was explained at the classroom level

(~2% on our study outcomes). This suggests that a multilevel model accounting for classrooms would not have much of an effect on the results of our primary analyses because most of the variance is between students as compared to between classrooms. Finally, as a robustness check, we reran our primary analyses with HLM with students nested within classrooms, and the results were consistent with what we found in our single level models that were reported in the main manuscript, which is supports the aforementioned point about most of the differences on student outcomes being at the student level vs. at the classroom level.

Robustness Checks: Cohort Effects

Tests for differences across cohorts. As a robustness check, we tested additional models on our three main dependent variables that also included interactions between year of data collection (cohort) and our key base predictors (the intervention groups vs. control group contrast, students' economic background, and the intervention groups x students' economic background interaction) and found no significant cohort two- or three-way interactions, suggesting that the effects of the interventions did not differ by year of data collection (cohort).

Combined expressive writing and reappraisal intervention effects within a single cohort. Since the combined expressive writing and reappraisal intervention condition was only fielded in one of the two years of data collection, it is possible that the effects of this condition could be confounded with the year of data collection. To address this concern, we re-ran the dummy coded analysis that compares each intervention group to the control group just within the one year of data collection in which the combined intervention was one of the conditions. We found that, even just within this one year, there was a significant effect of the combined intervention group as compared to the control group on exam performance (main effect of the combined expressive writing and reappraisal intervention: $p = .02$, student income by combined

expressive writing and reappraisal intervention interaction $p = .047$). This suggests that the combined condition effect was not because of a comparison between students from different school years.

Robustness Checks: Standardizing Exam Scores

In order to test whether the effects of our primary analyses held when exam scores were standardized across semesters, we re-ran the primary analysis on average exam scores on a newly calculated dependent variable: the average of the z-scores of each semester exam. Results of this model were consistent with what was found on raw exam scores.

Table S1. Unstandardized coefficients and SEs for primary analyses.

VARIABLES	Average Exam Scores	Course Passing Rate	Reappraisal of Test Anxiety
Intervention Groups (vs. Control)	0.53**	0.06	0.06**
	(0.18)	(0.10)	(0.02)
Students' Economic Background	-4.18***	-0.85***	-0.03
	(0.35)	(0.16)	(0.04)
Intervention Groups x Students' Economic Background	0.53**	0.19*	0.05*
	(0.18)	(0.09)	(0.02)
Expressive Writing and Reappraisal (vs. Combined)	-0.18	0.08	-0.01
	(0.29)	(0.15)	(0.03)
Expressive Writing (vs. Reappraisal)	0.22	0.30	-0.08*
	(0.36)	(0.18)	(0.04)
Prior Achievement	0.69***	0.11***	0.01***
	(0.02)	(0.01)	(0.00)
Constant	-99.53***	-23.53***	0.43
	(4.89)	(3.14)	(0.54)
Observations	1175	1175	861

Standard errors in parentheses
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$
Each Column represents an analysis on a specific student outcome.

Table S2. Unstandardized coefficients and SEs for secondary dummy code analyses.

VARIABLES	Average Exam Scores	Course Passing Rate	Reappraisal of Test Anxiety
Expressive Writing (vs. Control)	4.79**	1.46**	0.29
	(1.47)	(0.44)	(0.18)
Reappraisal (vs. Control)	3.08*	0.79*	0.46**
	(1.46)	(0.40)	(0.17)
Combined Expressive Writing and Reappraisal (vs. Control)	5.30**	0.59	0.57**
	(1.97)	(0.57)	(0.21)
Students' Economic Background (0=lower-income)	11.55***	2.85***	0.34*
	(1.25)	(0.66)	(0.14)
Expressive Writing x Economic Background	-4.94**	-1.82*	-0.32
	(1.68)	(0.88)	(0.20)
Reappraisal x Economic Background	-3.25†	-1.58††	-0.31
	(1.67)	(0.82)	(0.20)
Combined Condition x Economic Background	-5.18*	-0.89	-0.55*
	(2.26)	(1.11)	(0.24)
Prior Achievement	0.69***	0.11***	0.01***
	(0.02)	(0.01)	(0.00)
Constant	-106.88***	-25.28***	0.11
	(4.82)	(3.15)	(0.53)
Observations	1175	1175	861

Standard errors in parentheses

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, † $p = 0.05$, †† $p = 0.06$

Each Column represents an analysis on a specific student outcome. Independent variables are dummy codes with the control group as the reference group for the intervention contrasts and lower-income students as the reference group for the economic background variable. This means that the main effects of the intervention contrasts can be interpreted as the effects of the interventions for lower-income students. The interaction terms can be interpreted as the difference in the intervention effect between lower- and higher-income students.

Table S3. Effect Sizes and 95% Confidence Intervals for Raw Mean Differences for Intervention Contrasts for Lower-Income Students.

Metric of Effect Size	Outcome	Contrast	Effect Size	95% CI LL	95% CI UL
Cohen's *d*	Exam Score	All Interventions vs. Control	0.35	0.09	0.61
		Expressive Writing vs. Control	0.33	0.03	0.64
		Reappraisal vs. Control	0.27	-0.04	0.57
		Combined vs. Control	0.55	0.13	0.96
Odds Ratio	Course Passing Rate	All Interventions vs. Control	2.87	1.62	5.10
		Expressive Writing vs. Control	3.85	1.80	8.22
		Reappraisal vs. Control	2.30	1.17	4.53
		Combined vs. Control	2.74	1.01	7.42
Cohen's *d*	Reappraisal of Test Anxiety	All Interventions vs. Control	0.46	0.12	0.79
		Expressive Writing vs. Control	0.31	-0.09	0.71
		Reappraisal vs. Control	0.49	0.09	0.89
		Combined vs. Control	0.68	0.18	1.16

Note: Cohen's *d*s are reported for exam score and reappraisal of test anxiety outcomes, and odds ratios are reported for the binary course passing rate outcome. This table shows effect sizes and 95% confidence intervals for raw mean differences between experimental and control groups for lower-income students only since the benefits of the intervention were generally specific to this group of students. "All interventions" means the entire group of students assigned to the three intervention groups (expressive writing, reappraisal, and combined expressive writing and reappraisal). "Combined" means the combined expressive writing and reappraisal group.

Figure S1. Timeline of Study Activities. Study activities are displayed temporally across the school year. Events above the line indicate typical school events, and events below the line indicate timing of the student writing exercises and survey. Students completed writing exercises for 10 minutes directly before they received their final exams in both the fall and spring semesters.

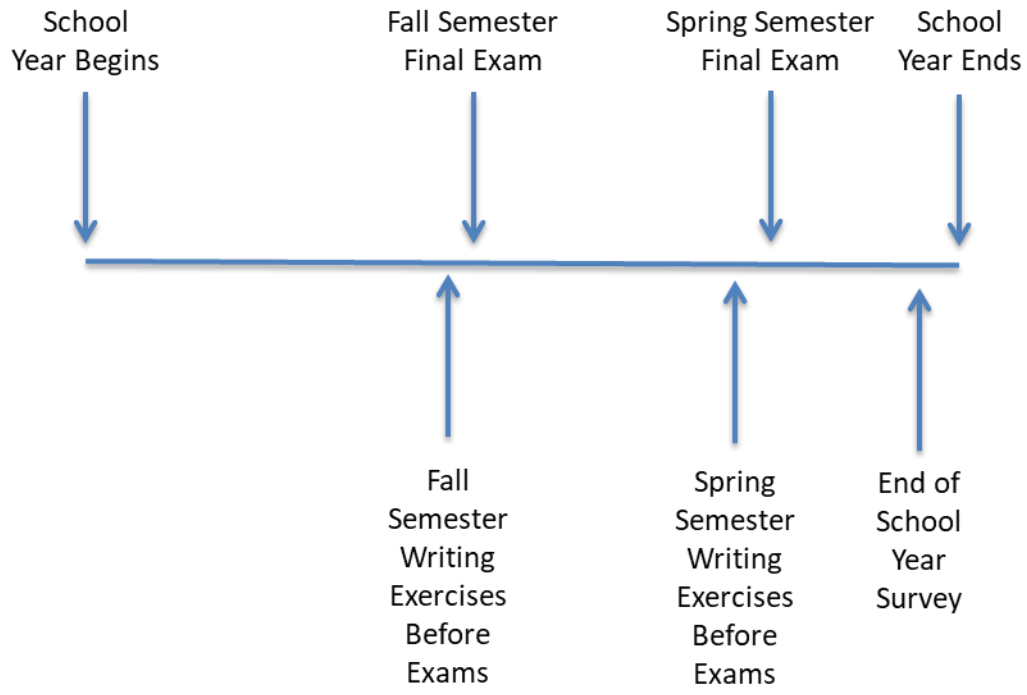
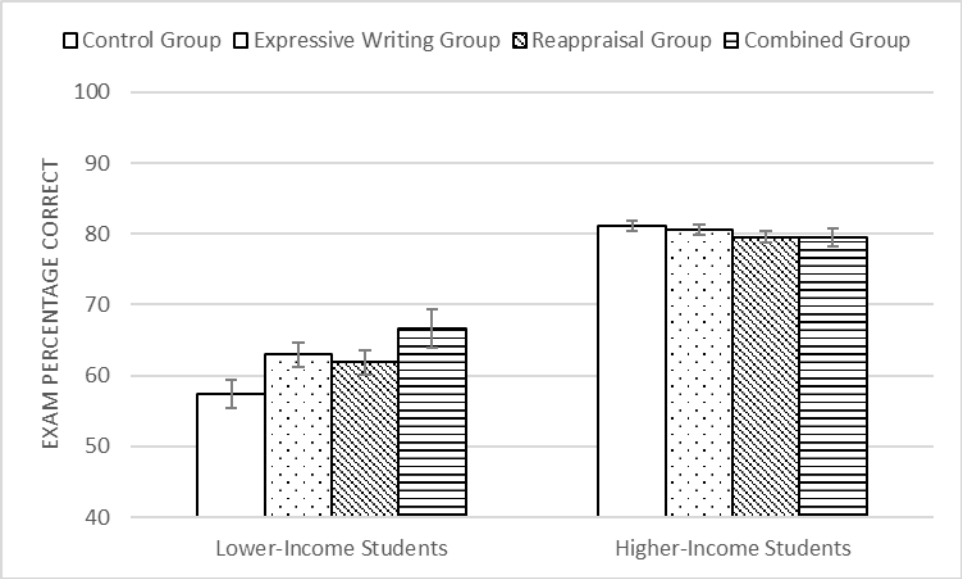
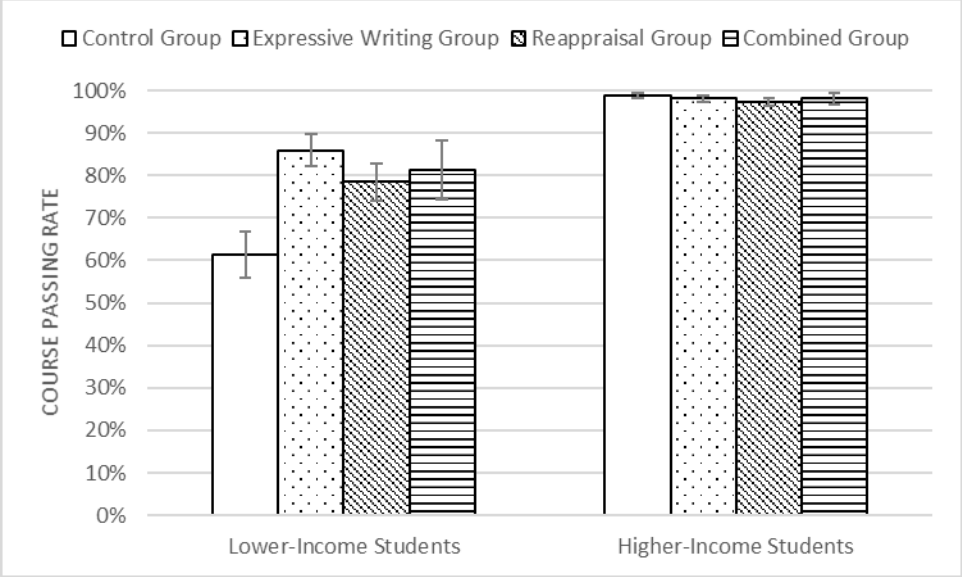


Figure S2. Secondary Analyses Comparing Individual Intervention Groups to the Control Group. Control group students are compared to each of the intervention groups here individually on exam performance (A), course passing rate (B), and reappraisal of test anxiety (C). Students are defined as lower- or higher-income based on free or reduced lunch status. Error bars represent +/- 1 SE of the mean.

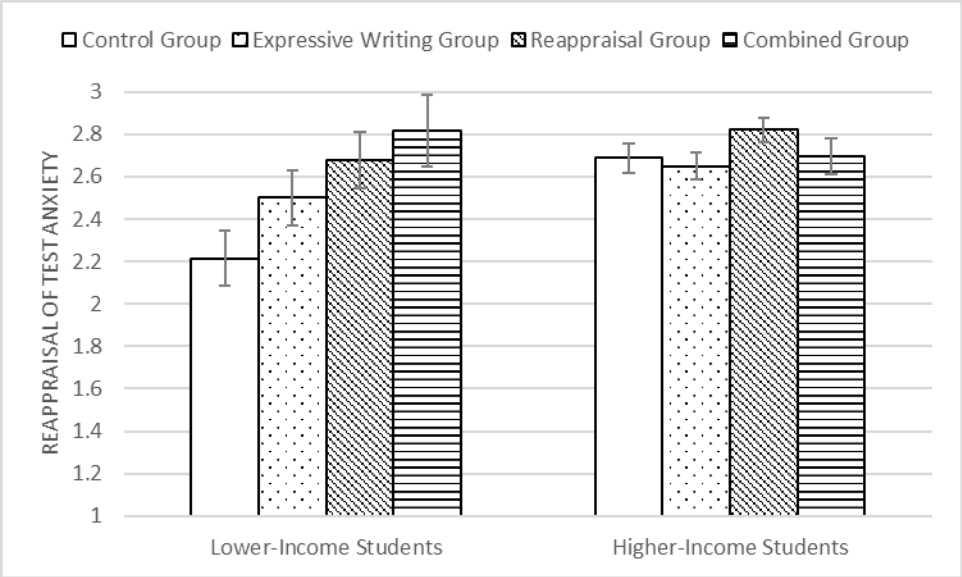
A)



B)



C)


Control Group
Expressive Writing Group
Reappraisal Group
Combined Group
REAPPRAISAL OF TEST ANXIETY
3
2.8
2.6
2.4
2.2
2
1.8
1.6
1.4
1.2
1
Lower-Income Students
Higher-Income Students