Impacts of Expressive Writing on Children's Anxiety and Mathematics Learning: Developmental and Gender Variability

Almaz Mesghina [a]

Lindsey Richland [b]

[a] *Department of Comparative Human Development, University of Chicago*

1126 East 59th St, Office 001, Chicago, IL, 60637, USA.

[b] *School of Education, University of California, Irvine*

3200 Education, University of California, Irvine. Irvine, CA 92636, USA.

Declarations of interest: none

Correspondence should be sent to Lindsey Richland, 3200 Education, University of California, Irvine. Irvine, CA 92636. E-mail: l.richland@uci.edu

Mesghina, A., & Richland, L. E. (2020). Impacts of expressive writing on children's anxiety and mathematics learning: Developmental and gender variability. *Contemporary Educational Psychology, 63.* https://doi.org/10.1016/j.cedpsych.2020.101926

Abstract

Expressive writing (EW), or writing about one's thoughts and feelings, has been posited to reduce the working memory (WM) load that pressure and anxiety can impose on test-takers. The mechanisms of EW are far from clear, however, and social and developmental questions about its role in shaping children's engagement with academic contexts remain. We report a study with 250 10-12-year-olds (M_{age} = 11.55; SD_{age} = 0.48; 127 females), exploring gender differences in these younger children's use of EW before a high demand mathematics lesson on ratio. In contrast to the literature with adults, children assigned to EW had *greater* anxiety relative to control. The highest-achieving higher-WM girls also were most impacted by EW, showing decreases in immediate learning and retention relative to their counterparts who did not write expressively. A mediation analysis indicated that EW reduced learning gains by increasing children's anxiety during the lesson, thereby suggesting that EW functioned quite differently from its use in older youth and adults. These data suggest that emotion regulation skills may be an under-considered mechanism that underpins benefits of EW in older adolescents and adults.

Keywords: expressive writing, anxiety, working memory, relational reasoning, learning, gender

Impacts of Expressive Writing on Anxiety and Children's Mathematics Learning: Developmental and Gender Variability

Elementary and middle school students regularly report experiencing high levels of anxiety in the classroom (Romano, 1997; Triplett & Barksdale, 2005), and external pressures such as stereotype threat can lead to reduced academic learning (Lyons et al., 2017). Further, high anxiety during testing is known to be detrimental to performance for some learners (see Beilock, 2008). Thus, feelings of worry may impede learning and test performance, with potentially serious implications for classroom achievement and engagement over the long term.

Expressive writing (EW) interventions, which prompt participants to free write about their inner emotional state prior to an emotional event, have been posited as a method for mitigating the harmful effects of worries on test performance, with successful interventions documented in adolescents and adults (e.g. Frattaroli et al., 2011; Ramirez & Beilock, 2011). With increasing attention to the role of stress and pressure in schools, more educators are seeking strategies to support students of all ages in their responses to worries, and it is reasonable to expect that some may turn to EW.

At the same time, no work has attempted to test the role of EW in reducing younger children's feelings of anxiety and worries in academic learning contexts. Because children's emotion regulation skills are still developing (Compas, 1987; Denham, 2007), we theorize that the benefits of EW may be less clear. This research builds upon and extends the literature on anxiety and performance by 1) investigating the role of active worries on learning (rather than on test performance), 2) testing whether an EW intervention can reduce anxiety and improve young students' learning, and 3) exploring the roles of working memory (WM) and gender in this relationship between worries and learning, both individual differences that have been shown to have relationships to mathematics, either through cognitive underpinnings (WM) or affective attitudes (gender).

The Role of WM in Learning while Anxious

Worries are the cognitive component of anxiety, and feeling anxious prior to a test can induce worries that compete for cognitive resources, such as working memory (WM), that are required for task performance (Eysenck & Calvo, 1992), including mathematical reasoning (Raghubar et al., 2010). Because WM is limited in capacity, any attention directed towards worries comes at a cost to other cognitive processes and negatively impacts test performance (Ashcraft & Kirk, 2001; Eysenck et al., 2007).

Indeed, WM has long been considered a moderator of the relationship between anxiety and performance (see Ashcraft & Kirk, 2001), particularly in children (Owens et al., 2008; Wang & Shah, 2014). Children and adults of high baseline WM capacity tend to be the most vulnerable to decrements in performance when feeling pressure to perform, with these results primarily holding for complex tasks that require WM resources (see Beilock, 2008). For example, in a special arithmetic task designed to pose a high WM load but that does not rely on prior knowledge, Beilock and DeCaro (2007) showed that in low pressure contexts, high-WM students outperformed low-WM students, but in anxiety-inducing situations, high-WM students' performance dropped considerably, whereas low-WM students' performance was unchanged. They find the mechanism to be the use of high versus low WM strategies, such that high WM individuals tended to use high WM strategies (e.g. using symbols and rule-based, step-by-step mathematical processing), but when under pressure, they switched to low WM strategies (e.g. guessing and estimations based on associations) which can be valid or invalid, but are less efficient (Beilock & DeCaro, 2007). In contrast, low WM individuals used the low WM strategies in all contexts, so their performance remained unchanged (Beilock & Carr, 2005, see Ramirez et al., 2016 for similar findings with children).

While it is well established that increased anxious ideation is related to decreases in test performance (e.g. Beilock et al., 2004; Hancock, 2001), less is known about how learning for content that taxes WM resources may be impacted by feelings of pressure, state anxiety and worries. In learning contexts, worry may load WM resources, reducing learners' ability to engage in higher order processing of more complex topics. Our work extends the literature by examining the role of anxiety on learning of a cognitively demanding lesson on ratio. The lesson compares two strategies to solve a ratio problem. Though a recommended teaching practice (Common Core Standards Initiative, 2010), instruction that compares strategies and requires relational reasoning is demanding of cognitive resources, including WM (Cho et al., 2007; Richland et al., 2017). Because the impact of anxiety on test performance has been shown to be greatest when tested content is demanding of WM (Beilock, 2008), this lesson provides an everyday, ecologically valid content area in which to examine the roles of student worries and an expressive writing intervention.

Expressive Writing Interventions

Expressive writing (EW) is typically used as a prompt that invites participants to write freely about their inner emotional state (see Pennebaker & Beall, 1986), and has been used before testing as a strategy for reducing worries. Compared to those who do not write or write about non-emotional content, adults who engage in EW displayed fewer worries and diminished anxiety (see Frattaroli, 2006; Smyth, 1998 for reviews), and EW prior to a high-anxiety test or adverse contexts resulted in better performance for adolescents and adults (Frattaroli et al., 2011; Ramirez & Beilock, 2011; Rozek et al., 2018). Writing about one's worries has been theorized to free up WM capacity (Klein & Boals, 2001; Yogo & Fujihara, 2008), and consequently, improve academic performance (Ramirez & Beilock, 2011) by reducing intrusive worries.

Developmental Constraints on the Benefits of Expressive Writing

The literature on EW and testing has focused exclusively on adults and adolescents, however, which has led to an assessment that EW may be helpful for students broadly (Ramirez & Beilock, 2011; Rozek et al., 2018), but there are reasons to believe the mechanisms may function differently in younger children. Because individuals are prompted to conjure up intense, emotional thoughts, EW necessitates some emotion regulation, or processes that change the “valence, intensity, or duration of an activated emotion” (Cole et al., 2004, p.320), so that the accessed emotions do not overwhelm the writer. If one cannot regulate the strong emotions they invoke during writing, then EW may serve to simply recall stressors and reactivate anxiety (Fivush et al., 2007).

Considering that children continue developing their emotion regulation skills well into adolescence (Compas, 1987; Denham, 2007; McRae et al., 2012), it is important to determine if emotional disclosure via writing will actually reduce the interference of worries into subsequent testing or learning for young students. For example, one study found that 9-13-year-olds who wrote most emotionally saw higher levels of anxiety and difficulties after EW (Fivush et al., 2007). A meta-analytic review of EW with adolescents provides further evidence that EW interventions may necessitate a developmental approach: compared to effect sizes for adults, the overall effect sizes of EW interventions on ten- to eighteen-year-olds’ socio-emotional and well-being outcomes were substantially smaller (Travagin et al., 2015). Importantly, this review found that EW interventions bared no significant effect on outcomes related to adolescents’ school performance (Travagin et al., 2015).

Gender and Social Development in Expressive Writing

A second developmental factor previously under-considered in the literature has been the role of gender socialization in understanding EW findings. Critiques of the EW literature argue that processes of gender socialization and gendered expectations of social roles influence the extent to which individuals may recognize, engage with, and consequently, benefit from emotions they disclose (Range & Jenkins, 2010; Stickney, 2010). Indeed, some work suggests gender may moderate the efficacy of EW interventions. In his meta-analysis of the early EW literature, Smyth (1998) found that the effect sizes for EW interventions increases as the proportion of males in the sample increased. Similarly, in a longitudinal design, Manier and Olivares (2005) manipulated EW in a group of college students and found that males demonstrated greater increases in subjective health and greater decreases in negative affect six weeks after the intervention. In explaining their gender differences, both studies argued that males showed greater gains than females because the EW context afforded a novel opportunity for males, who otherwise would have been less likely to disclose emotions as a means to recover from a stressful event (Manier & Olivares, 2005). Developmental work on children’s emotional processing lends further support to this argument. Specifically, because research on children’s personal narratives reveals that females write longer and more emotionally than males (e.g. Bohanek & Fivush, 2010; Cross & Madson, 1997; Fivush et al., 2007), and because females tend to ruminate on activated emotions longer than males (see Rood et al., 2009), young males and females may gain differently from EW.

Present Study

Therefore, the current study manipulated the presence of EW before learning about ratio in 10-12-year-old children. A low external pressure was also manipulated, though the focus in the current analysis is on the role of the EW manipulation, since the pressure manipulation did not systematically impact children's expressed worries. This study examined first whether writing about one's thoughts and feelings actually reduced these children's anxiety, as would be predicted by the adult literature, though might not be predicted by the developmental, emotion regulation literature. Second, we tested the hypothesis that children, like adults, would benefit from EW such that they would show improvements in learning for cognitively taxing content. Third, we examined whether WM capacity would moderate any effects of EW on learning, since WM moderates the relationship between anxiety and performance on cognitively taxing tasks (Beilock, 2008), particularly in elementary school children (Ramirez et al., 2013, 2016; Wang & Shah, 2014). We also explored the hypothesis that, because gender development is at a crucial turning point at this age of high social evaluation, gender would moderate the efficacy of EW.

Method

Participants

303 fifth- and sixth-grade students (151 males) from five schools in thirteen classrooms in the Chicago area were recruited for this study. Fifth and sixth grades were chosen because students at this age have been introduced to pre-requisite concepts necessary to solve ratio problems yet have not had a formal introduction to the concept (Common Core State Standards in Mathematics, 2010). 250 fifth- and sixth-graders (M_{age} = 11.55; SD_{age} = 0.48; 127 females) for whom we had complete data were included in this study. Students were excluded from analyses if they refused to participate (n = 1), did not speak English (n = 1), arrived late or were dismissed early and did not complete all measures (n = 21), or were absent during one of the study days (n = 29). Parents and guardians were informed of the study a few weeks before data collection and

were provided the opportunity to opt their child out. We also obtained children's written assent prior to data collection.

Design

Students were randomly assigned within-classroom to either the expressive writing (EW) condition (*n* = 137) or control (*n* = 113) prior to a video math lesson. Participants' gender ($\chi^2(1) = 0.83$, $p = .36$), racial/ethnic identity ($\chi^2(5) = 2.13$, $p = .83$), and age ($t(240) = 0.33$, $p = .74$) were evenly distributed between conditions. In order to induce anxieties about which students could write, we also randomly assigned participants to either receive an external pressure (*n* = 130) or no pressure (*n* = 120). Prior to writing, those assigned to receive pressure were told that the class would be rewarded with a pizza party only if they performed above 80% on the immediate post-test. Those assigned to no pressure did not know about the party, and were told that the researchers wanted to know how kids learn math. Though similar social-evaluative threats have been effective in inducing anxiety (Dickerson & Kemeny, 2004), students in our study felt equally anxious with or without pressure, suggesting our manipulation of pressure was no more anxiety-inducing than the learning context itself. Thus, we collapse across pressure conditions for the data analysis to investigate the effect of EW on anxiety and learning.

Writing Manipulation

Consistent with the typical EW paradigm (Pennebaker & Beall, 1986), students in the expressive condition were prompted to "write about all your deepest thoughts and feelings you may have about the math lesson and test you are about to receive." Students in the control were prompted to copy down an emotionally neutral paragraph from the introduction of the video lesson that was provided to all participants. All students wrote for five minutes and each was reminded that their writing would remain de-identified.

Intervention

The study spanned three visits to each of the participating classrooms. On Day 1, students received a pre-test on ratio. On Day 2, students received writing manipulations via laptops prior to watching a 30-minute video lesson on ratio, immediately followed by a post-test and a question gauging anxiety experienced during the lesson. About one week later, on Day 3, students completed a delayed post-test and a measure of baseline WM capacity. All classrooms were compensated for their time.

Ratio Lesson

All students watched the same 30-minute video lesson on ratio via laptops in their classrooms. The study design retained high ecological validity, as each student interacted with the video lesson and experimental manipulations individually on a laptop, while remaining seated and next to their peers, similarly to how they would engage with formal instruction in class (Begolli & Richland, 2016). The video lesson was modeled after mathematical discussion literature (see Kazemi & Hintz, 2014) as the teacher in the video used one example problem to compare between a correct strategy (least common multiple) and common misconception (subtraction) to solve ratio problems. To ensure active engagement, the lesson was broken into clips and the teacher in the lesson asked students to answer questions at the end of each clip, which they responded to on a paper-pencil worksheet. Instructions were presented on the screen and narrated through headphones.

Assessments

Learning was measured as the change in accuracy from pre-test to immediate post-test (immediate gains) and pre-test to delayed post-test (delayed gains or retention). Aside from counterbalanced ordering of items, students were administered the same 15-item assessment at all three timepoints. Multiple-choice and free-response items assessed students' procedural and conceptual understanding of ratio. See Begolli and Richland (2016) for test item construct analyses.

Working Memory

We assessed students' baseline WM capacity with the Shortened Symmetry Span Task (Foster et al., 2015) formatted for play on a 5th generation Apple iPad with a 9.7-inch active screen. Instructions were presented on the screen and narrated through headphones. Participants were required to hold target information (highlighted cells in a matrix) in memory while attending to distracting information (whether or not an unrelated matrix is symmetrical across its vertical axis). Participants completed four test trials which randomly varied in length from two to five cells to be memorized, for a total possible score of 14 cells correctly memorized in order.

Codes of EW Written Content

We used a developmentally appropriate coding scheme (see Fivush et al., 2007 for details) to analyze the written content of students' EW narratives. Specifically, we first calculated length of each narrative as the total number of propositions (independent subject-predicate clauses). Then, we categorized propositions into one of the mutually exclusive categories: problems, explanation, negative evaluations of others, positive evaluations of self, emotion, coping, and facts. We focused our analyses on problems (e.g. "Every time we have a test I think I will fail"), explanations (e.g. "[I am nervous] because I might not know the answer"), and negative evaluations (e.g. "Some of my classmates don't try hard enough"), as Fivush and colleagues (2007) have previously shown these factors to be related to subsequent well-being after EW. Negative evaluations were dropped from analyses due to the low frequency of occurrences (n = 2). Thus, we further expanded our analyses to include positive evaluations (e.g. "I am good at math") due to its relation to experienced anxiety (see below). Two trained researchers coded the narratives. To ensure interrater reliability, 20% of the narratives (n = 28) were coded by both researchers, who achieved excellent interrater reliability on all seven categories (Krippendorff's alphas ranging from .79 to 1.00; Krippendorff, 2009). Written content of each student's EW narrative was analyzed as the proportion of propositions containing each of the content categories above.

Experienced Anxiety

Feeling heightened pressure to perform is a common form of state anxiety (Baumeister, 1984). After the immediate post-test, students reported on their experienced anxiety during the lesson. Specifically, all students were asked to first "Please think about your experience during the lesson" then used a 5-point Likert scale (1: *Strongly Agree* to 5: *Strongly Disagree*) to report the extent to which they agreed with the statement "I felt a lot of pressure to do well on the test." We administered only one item due to time constraints during data collection, so we have validated these responses against learning and math anxiety data from all students, and written content data for those who expressively wrote. Using learning gains data, we found that our measure of anxiety was significantly and negatively related to both immediate ($r(248) = -.13$, $p = .04$) and delayed ($r(248) = -.16$, $p = .01$) learning gains. Our measure was also positively correlated with students' reported baseline mathematics anxiety, as measured by the extent to which they agreed with the statement "Math makes me nervous" using a 5-point Likert scale (1: *Not at all* to 5: *Very much*; $r(248) = .25$, $p < .001$). Furthermore, using written content data, our measure of anxiety was significantly and positively related to the proportion of their narratives that contained mentions of a problem ($r(136) =.23$, $p = .007$), and we find that our measure of anxiety was significantly negatively related to the proportion of their narratives that contained positive evaluations ($r(136) = -.20$, $p = .02$).

Results

Analytic Plan

We first provide an overview of the written content of students' EW narratives. Second, we test our research question regarding the role of EW on students' reported anxiety. Third, we show that reported anxiety is related to students' learning and retention, and subsequently examine how EW impacts students' learning gains. Lastly, we conclude by investigating the mediational role of anxiety on the relationship between EW and learning. For primary analyses of students' experienced anxiety and learning, we use a series of regressions to examine the impacts of condition (binary coded: 0 if control, 1 if EW), gender (binary coded: 0 if male, 1 if female), and WM (continuous). In Step 1, condition is first entered to explore main effects of the writing intervention. In Step 2, we add gender and WM to the model to explore how these factors may moderate any impacts of EW. We report standardized beta coefficients (β) for all regressions and unstandardized coefficients (b) in the mediation analysis.

Because students were nested within classrooms, there is always the possibility of nonindependence of students within the same class. We replicated our regression analyses using hierarchical linear modeling (HLM; Raudenbush & Bryk, 2002), a form of multilevel modeling that accounts for nested data (see Supplemental Materials for complete HLM results). The HLM results were entirely consistent with our findings from the original regression analyses. However, because our student-level and classroom-level sample sizes might not be large enough to properly estimate multilevel effects (Bryk & Raudenbush, 1992), and due to the insignificant between-class variance in students' reported anxiety (Hedges & Hedberg, 2007; Woltman et al., 2012), we believed that multilevel modeling might not be appropriate for this data (see Supplemental Materials for discussion). In lieu of multilevel modeling, we use robust clustered standard errors at the classroom level in all subsequent regression analyses to account for the possibility of nonindependence among students within the same class. We report non-standardized clustered standard errors (*SE*) for all regressions.

Written Content

In line with previous work, regressions revealed that girls assigned to expressively write (M = 71.15, SD = 24.93) wrote more words on average than boys (M = 62.85, SD = 23.90; β = 0.34, SE = 3.94, p = .05). Girls' EW narratives (M = 7.48, SD = 2.67) also contained more propositions than boys' (M = 6.88, SD = 2.47), though this difference did not reach significance (p = .28).

We regressed students' gender and WM capacity on each content category to test whether they predicted the written content of narratives for students in the EW condition (see Table 1 for descriptive statistics). There were no main effects or interactions of WM and gender on students' explanations. For positive evaluations, we found main effects of WM and gender, revealing that boys (β = -0.24, SE = 0.04, p = .02) and higher WM individuals (β = 0.41, SE = 0.04, p < .001) wrote more positive evaluations than their female and lower WM counterparts. We also found that WM capacity and gender interacted to predict positive evaluations (β = -0.51, SE = 0.07, p = .001). For boys, WM was positively related to their proportion of positive evaluations (t(138) = 5.40, p < .001). For girls, WM was unrelated to their positive evaluations (t(138) = -0.75, p = .47).

The interaction between WM capacity and gender also predicted students' problem statements (β = 0.44, SE = 0.13, p = .04). Split by gender, the patterns did not achieve significance, but they provide trending evidence to suggest that WM was positively related to girls' proportion of problem statements ($t(138) = 1.59$, $p = .14$), but negatively related to boys' problem statements ($t(138) = -1.77$, $p = .10$).

Experienced Anxiety

Anxiety by Condition

The analysis of written content revealed that the emotional valence of students' narratives was influenced by their WM capacity and gender. We next examined students' anxiety as a function of our condition manipulations. We found that the external prompt to raise feelings of pressure was not effective using a *t*-test comparison across added pressure versus no added pressure conditions ($t(248) = -0.78$, $p = .43$), so we collapsed across these conditions for all further analyses. This may have been due to all students feeling some pressure from the intervention. In Step 1, we first regressed condition on students' reported anxiety and found a main effect of condition, such that those who wrote expressively ($M = 3.00$, $SD = 1.34$) reported experiencing more anxiety than those in the control ($M = 2.50$, $SD = 1.42$; $\beta = 0.36$, $p = .03$). In Step 2, we added gender and WM to the model; no main effects or interactions emerged (for full regression table, see Table 2).

In sum, EW increased anxiety for all students, and this did not seem to change with WM capacity or gender. These results are in line with the developmental constraints on EW argument posited earlier, suggesting that the mechanism by which EW reduces anxiety in adolescents and adults functions differently in younger children.

Anxiety and Learning

Pearson correlations reveal that heightened anxiety during the lesson was related to lower learning gains both immediately after the lesson ($r(248) = -0.13$, $p = .04$) and after a week's delay ($r(248) = -0.16$, $p = .01$).[1] Congruent with the test performance literature, the more anxious students were, the less content they immediately learned and retained from the lesson. Because individual anxiety was related to lower gains, and because students reported greater anxiety after EW, we next explored the relation between condition and students' learning.

Learning

We first report on pre-test performance before analyzing students' immediate and delayed learning gains, respectively (see Table 3 for descriptive statistics for each outcome). As expected, average performance on the pre-test was low. Before the lesson, students scored an average of 18% ($SD = 20\%$). Pre-test performance did not differ between conditions ($t(248) = 0.13$, $p = .90$), nor by gender ($t(248) = -0.23$, $p = .82$). Students gained an average of 26% points ($SD = 29\%$) in total accuracy immediately after the lesson, and sustained average gains of 20% ($SD = 27\%$) from their pre-test after a week's delay. We first analyze students' immediate gains, followed by their delayed gains.

Immediate Learning

We first regressed condition on students' immediate learning gains and found no main effect of condition ($p = .44$; see Table 4 for full regression table). After adding gender and WM into the model in Step 2, we found a significant three-way interaction between condition, WM, and gender on students' immediate learning ($\beta = -0.73$, $p = .004$). We split the analysis by gender to further explore the three-way interaction (see Table 5). For girls, we found that the two-way interaction between condition and WM remained a significant predictor of their immediate learning gains ($\beta = -0.50$, $p = .007$). In the control condition, WM was positively related to girls' immediate learning gains ($t(127) = 5.20$, $p < .001$). However, WM was not related to learning for girls in the EW condition ($t(127) = -0.80$, $p = .44$; Figure 1). For boys, neither condition, WM, nor their interaction predicted their immediate learning gains.

Delayed Learning

We ran an identical series of regressions on students' delayed learning gains (see Table 4). Step 1 revealed no main effect of condition on delayed gains. Again, Step 2 revealed a three-way interaction between condition, WM, and gender ($\beta = -0.53$, $p = .01$). For girls, the two-way interaction between condition and WM remained a marginal predictor of their delayed learning gains ($\beta = -0.30$, $p = .06$; see Table 5). Again, WM positively predicted delayed learning gains for girls in the control condition ($t(127) = 3.39$, $p = .005$), but WM was no longer related to learning for those in the control ($t(127) = -0.09$, $p = .93$; Figure 1). For boys, neither condition, WM, nor their interaction predicted their delayed learning gains.

The results show that the impact of EW on learning is gendered, with the intervention only impacting girls', and not boys', immediate learning and retention. The patterns reveal that under normal learning conditions (without EW), WM predicted girls' learning, such that higher WM girls learned and retained more than lower WM girls. However, when asked to disclose their emotions prior to the lesson, WM no longer predicted girls' learning, and higher WM girls saw reduced learning gains and retention.

Mediation

Because students overall reported feeling more anxious in the EW condition, and because higher WM girls in the EW condition wrote the most negatively and saw significant, lasting decreases in learning relative their counterparts who did not expressively write, we next ask whether feelings of anxiety induced by writing may mediate the relationship between condition and learning from the lesson. We tested whether condition (IV) impacts learning (DV) through increased anxiety (mediator) using the Hayes (2013) PROCESS Model (model 4) with 10,000 bootstrapped samples.

As described earlier, condition predicted increased anxiety (b = 0.50, SE = 0.20, p = .03). Those in the EW condition reported feeling more anxious than those in the control. Additionally, increased anxiety during the lesson predicted both immediate and delayed learning gains ($b_{\text{immediate}}$ = - 0.03, $SE_{\text{immediate}}$ = 0.01, p =.02; b_{delayed} = - 0.03, SE_{delayed} = 0.01, p =.006), such that those who reported feeling more anxious after the lesson had lower immediate learning and retention. Furthermore, there was no direct effect of condition on either immediate or delayed learning ($b_{\text{immediate}}$ = .03, $SE_{\text{immediate}}$ = .03, p = .44; b_{delayed} = .02, SE_{delayed} = .02, p = .36). Rather, we found a significant indirect effect of reported anxiety on both immediate learning gains (b = -0.02, SE = .008, 95% confidence interval = [-.0333, -.0017]) and delayed learning gains ((b = -0.02, SE = .008, 95% confidence interval = [-.0352, -.0029]). Because the 95% confidence intervals do not contain zero, the indirect effects for immediate and delayed gains are significant at alpha = .05 (Hayes, 2013). Across all students, reported anxiety after the lesson mediated the effect of condition on both immediate and delayed learning gains – EW reduced immediate learning and retention through increasing anxiety.

Discussion

This study is the first to investigate how EW affects elementary school-aged children's learning. Similar to the literature on test performance, we found that increased anxiety during initial learning was negatively related to immediate learning and retention. However, we found evidence to suggest that EW may not operate similarly in children and adults: First, in contrast to the adult literature, writing about one's thoughts and feelings yielded increased anxiety in children, and this increase in anxiety mediated the relationship between EW and smaller immediate and sustained learning gains. Second, though our work aligns with literature that suggests WM moderates the relationship between EW and performance (e.g. Ramirez & Beilock, 2011), our findings suggest that these effects are gendered in EW contexts. While boys' learning was unchanged by condition, girls showed more sensitivity to our intervention. Specifically, higher WM girls assigned to EW disclosed more problems in their writing and saw greater decreases in learning and retention.

The mediation supports the inference that channeling their emotions via writing made students worry more and subsequently inhibited their learning. While adults are able to manage and appraise their stressful emotions when accessing them during writing, there is evidence to suggest that children at this age are unable to effectively process and create meaning of stressful experiences. A meta-analysis of personal narrative literature finds that children's autobiographical writings lack core elements that yield mature adult narratives (e.g. an understanding of temporality, a consistent theme in writing, causally and psychologically rich explanations; Habermas & Bluck, 2000). The authors argue that while children are able to recall content and access emotions specific to an event, they are still unable to organize and integrate isolated events into coherent, meaningful narratives until adolescence, during which individuals gain the requisite skills in emotional interpretation, problem solving, effortful coping, and cognitive control (Compas, 1987; Denham, 2007; Habermas & Bluck, 2000; McRae et al., 2012).

As such, for students who do write more emotionally, an EW intervention may serve to simply bring up worries without a way for children to effectively offload these concerns. Therefore, the very act of writing about emotions that reduces anxiety and frees up WM in adults

(Klein & Boals, 2001; Ramirez & Beilock, 2011) may in fact increase worry and consume WM in young children. In mathematical learning contexts, researchers predict that students who are higher in baseline WM capacity see greater learning gains than those who have lower WM capacity (Raghubar et al., 2010). We found this to be the case for girls in the control, but WM did not predict learning for those in the EW condition who reported feeling more anxious. Furthermore, our findings corroborate prior work suggesting that the performance of children and adults of higher, not lower, WM capacity is most impacted by increases in anxiety (Gimmig et al., 2006; Ramirez et al., 2013, 2016). Given that all students reported feeling more anxious in EW compared to control conditions, and given that higher WM girls assigned to the EW condition included the greatest proportion of personal problems and smallest proportion of positive statements in their narratives, it follows that higher WM girls' learning suffered most from EW.

The practical implications of our findings cannot be ignored. In line with Cohen's (1988) guidelines, our results showed moderate to large effect sizes. The magnitude of these effects is quite remarkable, given that these data were collected in students' actual, everyday mathematics classrooms, which introduce natural noisiness to our results. We attempted to reduce confounding variables and other sources of noise by randomizing students to condition within-classroom and using robust clustered standard errors to address the possibility of nonindependence within classrooms. However, unobserved student- and classroom-level factors (e.g. teacher quality, student engagement) may still impact students' learning from and anxiety towards the math lesson. Despite this, one brief, five-minute, independent writing intervention had a substantial impact on students' emotions, learning, and retention of rather difficult mathematics content.

Second, the patterns we observed in children are in contrast to the bulk of the literature characterizing improvements in adults' and adolescents' anxiety and depressive symptoms (see Frattaroli, 2006) and academic performance (e.g. Ramirez & Beilock, 2011; Rozek et al., 2018) after EW. Crucially, those finding positive impacts of EW on academic performance generally conclude that EW is a useful intervention to improve students' attitudes and performance, without specifying for whom EW may or may not work. Indeed, Kliewer et al. (2011) identified only nine studies examining the impacts of EW on children's mental health outcomes – none of which showed clear benefits of EW. Yet, in studies of EW in academic contexts, developmental factors remain largely under-considered. Our data show that EW interventions in academic contexts might not be suitable for children, and/or that researchers and practitioners must exercise caution when attempting to reduce children's anxieties and improve learning in the classroom through EW. With educators increasingly recognizing that children do regularly experience worries from factors such as performance pressure, there is a greater likelihood that EW may be tried as a strategy with younger children, so providing this caveat to the efficacy data is important.

Finally, there is a timeliness to these results. Students and teachers are grappling with an unfamiliar distance learning paradigm in the context of a global pandemic and its socioemotional and economic aftermath, which itself is a major source of stress, uncertainty, and anxiety. Now more than ever, it is important to understand how to assist students in downregulating their anxieties. Our ultimate goal for this work is to encourage researchers, parents, teachers, and policy makers to take pause when choosing strategies to assist students' emotion management. EW is a well-recognized and largely lauded emotion regulation strategy in a field that has few strategies to offer, particularly for younger students. Therefore, it is critical that any data on the efficacy of these strategies be made publicly available, particularly in this case where EW may in fact exacerbate anxiety and reduce learning for some students. It is possible that in other conditions, such as those in which EW were accompanied by training on how to regulate the emotions that emerged in the writing, or if a child co-constructs an EW narrative with an adult who can assist in making meaning of stressors (Sales & Fivush, 2005), such a paradigm could be quite effective for children of this age. But, as a tool for offloading worries prior to a learning or test moment, these data suggest that EW alone is unlikely to be successful and can be quite harmful.

Theoretically, these findings suggest that emotion regulation is a critical mechanism that likely underlies much of the benefits of expressive writing. Importantly, theoretical implications of these findings extend beyond the expressive writing intervention. In their model of Emotion Regulation in Achievement Situations, Harley and colleagues show that students experience many emotions in the classroom, and the strategies they choose to regulate their emotions have considerable consequences for their subsequent achievement and motivation (Harley et al., 2019). For example, a student who failed a math test may attempt to regulate their emotions effectively by reappraising the test as a learning opportunity, rather than a measure of one’s ability. Conversely, the student may ineffectively regulate their emotions by avoiding homework questions related to the test content they failed or by dropping the course altogether. Other work links emotion regulation to many social and emotional outcomes important for success in the classroom: students with more mature emotion regulation skills see improved peer relations, teacher relations, and fewer behavioral problems, all of which predict greater academic achievement (Graziano et al., 2007). Indeed, emotion regulation ability is considered an important predictor for school readiness (Blair, 2002; Blair & Raver, 2015) and ultimately, school success (Graziano et al., 2007; Gumora & Arsenio, 2002) and therefore necessitates continued attention.

Another explanation of our findings receiving less support from the literature is that anxious students could have been engaging in emotional regulation to mitigate such anxieties, but the process of emotion regulation itself takes up executive function resources, particularly WM (Hofmann et al., 2012; Schmeichel & Tang, 2015), and likely reduces the WM available for task-relevant processing, such as learning from the ratio lesson. Many efforts in the laboratory show that cognitive performance tends to decrease if individuals simultaneously attempt to manage negative emotions (e.g. Richards & Gross, 2000), and evidence from work on stereotype threat proposes that individuals experiencing high states of anxiety do engage in the highly taxing psychological work of managing their anxiety via suppression, resulting in worse mathematics test performance (Johns et al., 2008). Again, this literature samples primarily adult populations. In the current study, those who expressively wrote did report feeling more anxious, and certain higher WM girls who expressively wrote mentioned more problems than their peers, but we are unable to determine whether they differed in how they may have attempted to regulate their emotions as they attended to the lesson. Nevertheless, both explanations support the hypothesis that increased emotional writing is related to greater feelings of anxiety and hinders learning and retention of mathematical content.

Why are girls more sensitive to the writing-induced changes in anxiety and learning? Evidence from personal narrative literature suggests that, across the life course, girls are generally more verbose, detailed, and emotionally expressive in their writing (Bohanek & Fivush, 2010; Cross & Madson, 1997; Fivush & Buckner, 2003; Fivush et al., 2007). Consistent with these findings, girls in our sample wrote longer narratives than boys and higher WM girls in particular included a greater proportion of problem statements than all other students, which prior research suggests is inversely related to children's anxiety and overall health (Fivush et al., 2007). It follows that girls suffered most from EW.

Additionally, gender differences in habitual emotion regulation use may elucidate gender differences in EW gains. As Range and Jenkins (2010) argue, those socialized to believe they should not express their emotions may gain more from EW insofar as it provides a novel opportunity for them to cope with stressors. Girls are more likely to use or support emotion-focused strategies (e.g. disclosure, expression) as a means to regulate their emotions than boys (e.g. Patterson & McCubbin, 1987; Waters & Thompson, 2014). Indeed, studies often find that adult males do benefit more from EW interventions than females (Smyth, 1998; Manier & Olivares, 2005), though little work has yet to directly test gender differences in EW (Range & Jenkins, 2010; Stickney, 2010).

One alternative hypothesis for girls' sensitivity to the intervention may be that girls, particularly higher WM girls, were experiencing stereotype threat during the lesson. It is possible that writing about their deepest thoughts and feelings allowed girls to access underlying concerns about confirming negative gender stereotypes about math performance. Stereotype threat literature reliably demonstrates that while all individuals' performance (e.g. Spencer et al., 1999; Steele & Aronson, 1995) and engagement with the content (e.g. Lyons et al., 2017) suffers after their stereotyped identity is made salient, higher WM individuals are at most risk because stereotype threat operates by consuming WM through worry (Schmader & Johns, 2003). Although this study did not aim to make their gender salient, and although students did not make explicit statements about gender in their writing, it is possible that higher WM girls who wrote negatively had activated underlying concerns about how their performance may be evaluated relative to boys. There is some evidence that girls' underperformance under stereotype threat is related to the negative emotions they generate (Cadinu et al., 2005), but since we did not explicitly manipulate the salience of gender in the writing prompt we are unable to make such a conclusion.

Limitations

This study has a few possible limitations worth considering here. First, most EW interventions examining therapeutic benefits instruct participants to write for about 3 to 5 days and upwards to 30 minutes per session (Pennebaker, 1997). Research has shown that longer and more frequent writing sessions yield stronger effects (see Frattaroli, 2006; Smyth, 1998 for reviews). Though recent work finds positive effects of EW in shorter or less frequent sessions in reducing anxiety or improving performance (e.g. Frattaroli et al., 2011; Ramirez & Beilock, 2011; Rozek et al., 2018), which is what led to the current manipulation, the reviews do suggest that perhaps a longer and sustained writing intervention might afford more substantial effects on anxiety and learning than a one-time, five-minute task.

Second, every classroom study is limited by time to collect measures due to real classroom demands, but future iterations of this work would benefit from additional measures of anxiety, including state anxiety measured before the manipulation, and a more comprehensive measure at the end. We were pleased to find that our anxiety measure was validated against the performance data, mathematics anxiety, and written EW content, so we are confident that these data are valid, but more data would be helpful to better understand the participants' context.

Third, the intervention used an interactive videotaped lesson rather than having the students' own teacher teach the lesson in a manner that would impose a high WM demand, which would not have allowed for the same level of experimental control. Instead, the intervention was administered in everyday classrooms with all students present, and the videotaped lesson was of a real classroom lesson. This approach allowed for more ecological validity than a pull-out or lab-based paradigm, yet maintained full experimental control. Even so, it would be important in the future to move to train teachers in these interventions to assess wider validity of the work to explain how EW could function in everyday classroom usage.

Conclusion

In summary, expressive writing did not affect these late elementary students in the way predicted by studies of older youth and adults, which would have presumably meant improving performance by offloading worry ideation onto paper, and thereby making working memory resources available for the work at hand. Instead, these students – in particular the typically highest learning group: girls with higher working memory – decreased in learning and retention from a mathematics lesson that was designed to engage working memory resources. Theoretically, these developmental findings indicate that emotion regulation may be a crucial mechanism underlying the function of expressive writing that has not yet received adequate attention. The mechanism of offloading worries relies upon regulation skills to avoid perseveration about these worries, which may be fairly well automatized in many adolescents and adults, but less available as a resource earlier in development. The developmental implications of our findings cannot be ignored. As critiques of the expressive writing literature have previously argued, the roles of gender, age, and their intersection (Bohanek & Fivush, 2010; Fivush et al., 2007; Range & Jenkins, 2010; Stickney, 2010; Travagin et al., 2015) are key to understanding who gains from expressive writing. Though work on expressive writing has touted its ability to reduce anxiety, free up working memory, and improve test performance, prior work has focused only on adolescents and adults. Our data provide no evidence that expressive writing operates similarly in children and adults. In fact, data suggest expressive writing prior to learning may at best, provide a small, immediate, temporary learning boost, and at worst, put certain higher performing students at a great learning disadvantage. Given that mathematics concepts build on themselves, initial differences in learning may widen over time; therefore, emotion regulation interventions prior to children's learning demand continued consideration.

Acknowledgements

This work was supported by the Spencer Foundation, #201800012; the Institute of Education Sciences, U.S. Department of Education, through R305A170488 and R305B140048 to the University of Chicago; and the National Science Foundation, #32027447, to the University of California, Irvine. The opinions expressed are those of the authors and do not necessarily represent views of the funders. We thank the teachers and students who dedicated their class time to our research. We also thank Samuela Mouzaoir for her assistance with data analysis.

Notes

[1] Anxiety and learning are believed to be related in a nonlinear manner. However, when we first examined our data and graphed the bivariate associations between anxiety and learning, we did not find any evidence to suggest a possible nonlinear relationship between the two variables. Loess lines illustrated a linear relationship between self-reported anxiety and immediate and delayed learning gains. To test for nonlinearity in our models, we included a quadratic anxiety term as a predictor of learning. In line with evidence from the loess graphs, neither the linear nor the quadratic term predicted learning at either timepoint. (Immediate learning: $b_{\text{anxiety}} = 0.002$, $SE_{\text{anxiety}} = 0.06$, $p = .98$; $b_{\text{anxiety squared}} = -0.005$; $SE_{\text{anxiety squared}} = 0.01$; $p = .63$. Delayed learning: $b_{\text{anxiety}} = 0.007$, $SE_{\text{anxiety}} = 0.06$, $p = .91$; $b_{\text{anxiety squared}} = -0.007$; $SE_{\text{anxiety squared}} = 0.01$; $p = .51$). The linear predictor seemed to best fit the observed patterns and is used in all regression analyses.

References

Ashcraft, M. H. & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology*, *130*(2), 224-237. https://doi.org/10.1037/0096-3445.130.2.224

Baumeister, R. F. (1984). Choking under pressure: Self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 46*, 610-620. https://doi.org/10.1037//0022-3514.46.3.610.

Begolli, K.N. & Richland, L.E. (2016). Teaching mathematics by comparison: Analog visibility as a double-edged sword. *Educational Psychology*, *108*(2), 194-213. https://doi.org/10.1037/edu0000056.

Beilock, S. L. (2008). Math performance in stressful situations. *Current Directions in Psychological Science*, *17*(5), 339–343. https://doi.org/10.1111/j.1467-8721.2008.00602.x.

Beilock, S. L. & Carr, T. H. (2001). On the fragility of skilled performance: What governs choking under pressure? *Journal of Experimental Psychology: General, 130*(4), 701-725. https://doi.org/10.1037/0096-3445.130.4.701.

Beilock, S. L. & Carr, T. H. (2005). When high-powered people fail: Working memory and “choking under pressure.” *Psychological Science, 16*(2), 101-105. https://doi.org/10.1111/j.0956-7976.2005.00789.x.

Beilock, S. L. & DeCaro, M. S. (2007). From poor performance to success under stress: Working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(6), 983-998. https://doi.org/10.1037/0278-7393.33.6.983.

Beilock, S. L., Kulp, C. A., Holt, L. E., & Carr, T. H. (2004). More on the fragility of performance: Choking under pressure in mathematical problem solving. *Journal of Experimental Psychology, 133*(4), 584-600. https://doi.org/10.1037/0096-3445.133.4.584

Blair, C. (2002). School readiness: Integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist, 57*(2), 111-127. https://doi.org/10.1037//0003-066x.57.2.111.

Blair, C. & Raver, C. C. (2015). School readiness and self-regulation: A developmental psychobiological approach. *Annual Review of Psychology*, *66*, 711–731. https://doi.org/10.1146/annurev-psych-010814-015221.

Bohanek, J. G. & Fivush, R. (2010). Personal narratives, well-being, and gender in adolescence. *Cognitive Development, 25*(4), 368-379. https://doi.org/10.1016/j.cogdev.2010.08.003.

Bryk, S. W. & Raudenbush, A. S. (1992). *Hierarchical linear models: Applications and data analysis methods.* Sage.

Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why do women underperform under stereotype threat? Evidence for the role of negative thinking. *Psychological Science*, *16*(7), 572-578. https://doi.org/10.1111/j.0956-7976.2005.01577.x.

Cho, S., Holyoak, J. K., & Cannon, T. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory and Cognition*, *35*, 1445-1455. https://doi.org/10.3758/BF03193614.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.

Cole, P. M., Martin, S. E., & Dennis, T.A. (2004). Emotion regulation as a scientific construct: Methodological challenges and directions for child development research. *Child Development, 75*(2), 317-333. https://doi.org/10.1111/j.1467-8624.2004.00673.x.

Common Core State Standards Initiative. (2010). Common core state standards for mathematics. National Governors Association Center for Best Practices and the Council of Chief State School Office.

Compas, B. E. (1987) Coping with stress during childhood and adolescence. *Psychological Bulletin, 101*(3), 393-403. https://doi.org/10.1037/0033–2909.101.3.393.

Cross, S. E. & Madson, L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin, 122*(1), 5-37. https://doi.org/10.1037/0033-2909.122.1.5.

Denham, S. A. (2007). Dealing with feelings: How children negotiate the worlds of emotions and social relationships. *Cognition, Brain, & Behavior*, *11*(1), 1-48.

Dickerson, S. S. & Kemeny, M. E. (2004). Acute stressors and cortisol Responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin*, *130*(3), 355-391. https://doi.org/10.1037/0033-2909.130.3.355.

Eysenck, M. W. & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition and Emotion, 6*(6), 409-434. http://dx.doi.org/10.1080/02699939208409696.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*(2), 336-353. https://dx.doi.org/10.1037/1528-3542.7.2.336.

Fivush, R., & Buckner, J. (2003). Constructing gender and identity through autobiographical narratives. In R. Fivush & C. Haden (Eds.), *Autobiographical memory and the construction of a narrative self: Developmental and cultural perspectives* (pp. 149–167). Erlbaum.

Fivush, R., Marin, K., Crawford, M., Reynolds, M. & Brewin, C. R. (2007) Children’s narratives and well-being. *Cognition and Emotion*, *21*(7), 1414-1434. https://doi.org/10.1080/02699930601109531.

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory and Cognition*, *43*, 226-236. https://doi.org/10.3758/s13421-014-0461-7.

Frattaroli, J. (2006). Experimental disclosure and its moderators: A meta-analysis. *Psychological Bulletin*, *132*(6), 823-865. https://doi.org/10.1037/0033-2909.132.6.823.

Frattaroli, J., Thomas, S., & Lyubomirsky, S. (2011). Opening up in the classroom: Effects of expressive writing on graduate school entrance exam performance. *Emotion*, *11*(3), 691-696. https://doi.org/10.1037/a0022946.

Gimmig, D., Huguet, P., Caverni, J.P., & Cury, F. (2006). Choking under pressure and working memory capacity: When performance pressure reduces fluid intelligence (Gf). *Psychonomic Bulletin & Review*, *13*, 1005–1010. https://doi.org/10.3758/bf03213916.

Graziano, P. A., Reavis, E. D., Keane, S. P., & Calkins, S. D. (2007). The role of emotion regulation and children’s early academic success. *Journal of School Psychology*, *45*(1), 3-19. https://doi.org/10.1016/j.jsp.2006.09.002.

Gumora, G., & Arsenio, W. F. (2002). Emotionality, emotion regulation, and school performance in middle school children. *Journal of School Psychology*, *40*(5), 395-413. https://doi.org/10.1016/S0022-4405(02)00108-5.

Habermas, T., & Bluck, S. (2000). Getting a life: The development of the life story in adolescence. *Psychological Bulletin*, *126*, 748–769. https://doi.org/10.1037/0033-2909.126.5.748.

Hancock, D. R. (2001). Effects of test anxiety and evaluative threat on students’ achievement and motivation. *The Journal of Educational Research*, *94*(5), 284-290. https://doi.org/10.1080/00220670109598764.

Harley, J. M., Pekrun, R., Taxer, J. L., & Gross, J. J. (2019). Emotion regulation in achievement situations: An integrated model. *Educational Psychologist*, *54*(2), 106-126. https://doi.org/10.1080/00461520.2019.1587297.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press.

Hedges, L. V. & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis. 29,* 60–87. https://doi.org/10.3102/0162373707299706

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences*, *16*, 174–180. https://doi.org/10.1016/j.tics.2012.01.006.

Johns, M., Inzlicht, M., & Schmader, T. (2008). Stereotype threat and executive resources depletion: Examining the influence of emotion regulation. *Journal of Experimental Psychology: General, 137*(4), 691-705. https://doi.org/10.1037/e506102012-359.

Kazemi, E. & Hintz, A. (2014). *Intentional Talk: How to Structure and Lead Productive Mathematical Discussions.* Stenhouse Publishers.

Klein, K. & Boals, A. (2001). Expressive writing can increase working memory capacity. *Journal of Experimental Psychology*, *130*(3), 520-533. https://doi.org/10.1037/0096-3445.130.3.520.

Kliewer, W., Lepore, S. J., Farrell, A. D., Allison, K. W., Meyer, A. L., Sullivan, T. N., & Greene, A. Y. (2011). A school-based expressive writing intervention for at-risk urban adolescents' aggressive behavior and emotional lability. *Journal of Clinical Child &*

Adolescent Psychology, *40*(5), 693-705. https://doi.org/10.1080/15374416.2011.597092.

Krippendorff, K. (2009). Testing the reliability of content analysis data: What is involved and why. In K. Krippendorff & M. A. Bock (Eds.), *The content analysis reader* (pp. 350–357). Sage.

Lyons, E. M., Simms, N., Begolli, K. N., & Richland, L. E. (2017). Stereotype threat effects on learning from a cognitively demanding mathematics lesson. *Cognitive Science*, *42*(2), 678-690. https://doi.org/10.1111/cogs.12558.

Manier, D. & Olivares, A. (2005). Who benefits from expressive writing? Moderator variables affecting outcomes of emotional disclosure interventions. *Counseling and Clinical Psychology Journal, 2*(1), 15-28.

McRae, K., Gross, J. J., Weber, J., Robertson, E. R., Sokol-Hessner, P., Ray, R. D., … Ochsner, K. N. (2012). The development of emotion regulation: An fMRI study of cognitive reappraisal in children, adolescents, and young adults. *Social Cognitive and Affective Neuroscience, 7*(1), 11-22. https://doi.org/10.1093/scan/nsr093.

Owens, M., Stevenson, J., Norgate, R., & Hadwin, J. A. (2008). Processing efficiency theory in children: Working memory as a mediator between trait anxiety and academic performance. *Anxiety, Stress, and Coping*, *21*(4), 417-430. https://doi.org/10.1080/10615800701847823.

Patterson, J. M. & McCubbin, H. I. (1987). Adolescent coping style and behaviors: Conceptualization and measurement. *Journal of Adolescence*, *10*(2), 163-186. https://doi.org/10.1016/s0140-1971(87)80086-6.

Pennebaker, J. W (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science*, *8*(3), 162-166. https://doi.org/10.1111/j.1467-

9280.1997.tb00403.x.

Pennebaker, J. W. & Beall, S. K. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. *Journal of Abnormal Psychology*, *95*(3), 274-281. https://doi.org/10.1037/0021-843x.95.3.274.

Raghubar, K. P., Barnes, M. A., & Hecht, S. A. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*(2), 110-122. https://doi.org/10.1016/j.lindif.2009.10.005.

Ramirez, G. & Beilock, S.L. (2011). Writing about testing worries boosts exam performance in the classroom. *Science*, *331*(6014), 211-3. https://doi.org/10.1126/science.1199427.

Ramirez, G., Chang, H., Maloney, E. A., Levine, S. C., & Beilock, S. L. (2016). On the relationship between math anxiety and math achievement in early elementary school: The role of problem solving strategies. *Journal of Experimental Child Psychology, 141*(2016), 83-100. https://doi.org/10.1016/j.jecp.2015.07.014.

Ramirez, G., Gunderson, E. A., Levine, S. C., & Beilock, S. L. (2013). Math anxiety, working memory, and math achievement in early elementary school. *Journal of Cognition and Development*, *14*(2), 187-202. https://doi.org/10.1080/15248372.2012.664593.

Range, L. M. & Jenkins, S. R. (2010). Who benefits from Pennebaker's expressive writing paradigm? Research recommendations from three gender theories. *Sex Roles*, *63*, 149-164. https://doi.org/10.1007/s11199-010-9749-7.

Raudenbush, S. W.,& Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Sage.

Richards, J. M., & Gross, J. J. (2000). Emotion regulation and memory: The cognitive costs of keeping one's cool. *Journal of Personality and Social Psychology*, *79*, 410–424. https://doi.org/10.1037/0022-3514.79.3.410.

Richland, L. E., Begolli, K. N., Simms, N., Frausel, R. R., & Lyons, E. A. (2017). Supporting mathematical discussions: The roles of comparison and cognitive load. *Educational Psychology Review*, *29*(1), 41-53. https://doi.org/10.1007/s10648-016-9382-2.

Romano, J. L. (1997). Stress and coping: A qualitative study of fourth and fifth graders. *Elementary School Guidance & Counseling*, *31*(4), 273-282.

Rood, L., Roelofs, J., Bogels, S. M., Nolen-Hoeksema, S., & Schouten, E. (2009). The influence of emotion-focused rumination and distraction on depressive symptoms in non-clinical youth: A meta-analytic review. *Clinical Psychology Review, 29*(7), 607-616. https://doi.org/10.1016/j.cpr.2009.07.001.

Rozek, C. S., Ramirez, G., Fine, R. D., & Beilock, S. L. (2018). Reducing socioeconomic disparities in the STEM pipeline through student emotion regulation. *Proceedings of the National Academy of Sciences, 116*(5), 1553-1558. https://doi.org/10.1073/pnas.1808589116.

Sales, J. M. & Fivush, R. (2005). Social and emotional functions of mother-child reminiscing about stressful events. *Social Cognition, 23*(1), 70-90. https://doi.org/10.1521/soco.23.1.70.59196.

Sattizahn, J. R., Moser, J. S., & Beilock, S. L (2016). A closer look at who “chokes under pressure”. *Journal of Applied Research in Memory and Cognition, 5*(4), 470-477. https://doi.org/10.1016/j.jarmac.2016.11.004.

Schmader, T. & Johns, M. (2003). Converging evidence that stereotype threat reduces working memory capacity. *Journal of Personality and Social Psychology, 85*(3), 440-452. https://doi.org/10.1037/0022-3514.85.3.440.

Schmeichel, B. J. & Tang, D. (2015). Individual differences in executive functioning and their relationship to emotional processes and responses. *Current Directions in Psychological Science, 24*(2), 93-98. https://doi.org/10.1177/0963721414555178.

Smyth, J. M. (1998). Written emotional expression: Effect sizes, outcome types, and moderating variables. *Journal of Consulting and Clinical Psychology, 66*(1), 174-184. https://doi.org/10.1037/0022-006x.66.1.174.

Spencer, S., Steele, C., & Quinn, D (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 351*, 4-28. https://doi.org/10.1006/jesp.1998.1373.

Steele, C. & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology, 695*, 369-389. https://doi.org/10.1037/0022-3514.69.5.797.

Stickney, L. (2010). Who benefits from Pennebaker's expressive writing? More research recommendations: A Commentary on Range and Jenkins. *Sex Roles, 63*(3), 165-172. https://doi.org/10.1007/s11199-010-9771-9.

Travagin, G., Margola, D., & Revenson, T. A. (2015). How effective are expressive writing interventions for adolescents? A meta-analytic review. *Clinical Psychology Review, 36*, 42-55. https://doi.org/10.1016/j.cpr.2015.01.003.

Triplett, C. F. & Barksdale, M. A. (2005). Third through sixth graders' perceptions of high-stakes testing. *Journal of Literacy Research, 37*(2), 237-260. https://doi.org/10.1207/s15548430jlr3702_5.

Wang, Z. & Shah, P. (2014). The effect of pressure on high- and low-working-memory students: An elaboration of the choking under pressure hypothesis. *British Journal of Educational Psychology*, *8*, 226-238. https://doi.org/10.1111/bjep.12027.

Waters, S. F. & Thompson, R. A. (2014). Children's perceptions of the effectiveness of strategies for regulating anger and sadness. *International Journal of Behavioral Development, 38*(2), 174-181. https://doi.org/10.1177/0165025413515410.

Woltman, H., Feldstain, A., MacKay, J. C., & Rocchi, M. (2012). An introduction to hierarchical linear modeling. *Tutorials in Quantitative Methods for Psychology*, *8*(1), 52–69.

Yogo, M. & Fujihara, S. (2008). Working memory capacity can be improved by expressive writing: A randomized experiment in a Japanese sample. *British Journal of Health Psychology, 13*(1), 77-80. https://doi.org/10.1348/135910707x252440.

Table 1

Average proportion of propositions containing each written content category. Mean (SD).

		Problem	Explanation	Positive Evaluation
Boys (n = 72)				
	Low WM (*n* = 33)	0.16 (0.19)	0.16 (0.16)	0.10 (0.16)
	High WM (*n* = 39)	0.13 (0.18)	0.14 (0.16)	0.18 (0.18)
Girls (n = 66)				
	Low WM (*n* = 45)	0.15 (0.19)	0.11 (0.12)	0.11 (0.16)
	High WM (*n* = 21)	0.27 (0.17)	0.12 (0.12)	0.06 (0.11)
Overall (n = 137)		0.17 (0.19)	0.13 (0.15)	0.12 (0.17)

Note. WM is analyzed as a continuous variable in regression analyses, but for descriptive purposes is displayed here as a median split.

Table 2

Full regression table for students' average reported anxiety experienced during the lesson.

	Model 1					Model 2				
	β	*b*	SE	*t*	Sig.	β	*b*	SE	*t*	Sig.
Condition	0.36	0.50	0.20	2.48	.03	0.25	0.68	0.61	1.11	.29
WM						0.05	0.23	0.63	0.37	.72
Gender						-0.19	-0.30	0.48	-0.62	.55
WM*Gender						0.01	0.06	0.91	0.07	.95
Condition*WM						-0.14	-0.69	0.89	-0.65	.53
Condition*Gender						0.32	-0.46	0.73	-0.52	.61
Condition*WM*Gender						0.37	1.90	1.33	1.23	.24

Note. β and *b* indicate standardized and unstandardized regression coefficients, respectively. Standard errors are clustered at the level of classroom. The degrees of freedom for Models 1 and 2 are (1, 248) and (7, 242), respectively.

Table 3

Students' average pre-test scores and immediate and delayed learning gains. Mean (SD).

	Control (*n* = 113)	EW (*n* = 137)	Overall (*N* = 250)
Pretest Accuracy			
Boys (n = 123)			0.19 (0.21)
Low WM (*n* = 62)	0.14 (0.20)	0.17 (0.22)	
High WM (*n* = 61)	0.27 (0.24)	0.19 (0.19)	
Girls (n =127)			0.18 (0.20)
Low WM (*n* = 73)	0.16 (0.17)	0.14 (0.14)	
High WM (*n* = 54)	0.20 (0.22)	0.28 (0.27)	
Overall (N=250)	0.19 (0.21)	0.18 (0.20)	0.18 (0.20)
Immediate Gains			
Boys (n = 123)			0.21 (0.25)
Low WM (*n* = 62)	0.17 (0.26)	0.16 (0.22)	
High WM (*n* = 61)	0.18 (0.21)	0.28 (0.27)	
Girls (n =127)			0.32 (0.32)
Low WM (*n* = 73)	0.16 (0.23)	0.36 (0.36)	
High WM (*n* = 54)	0.43 (0.29)	0.24 (0.29)	
Overall (N = 250)	0.25 (0.28)	0.27 (0.30)	0.26 (0.29)
Delayed Gains			
Boys (n=123)			0.15 (0.24)
Low WM (*n* = 62)	0.12 (0.21)	0.14 (0.24)	
High WM (*n* = 61)	0.08 (0.21)	0.22 (0.26)	
Girls (n =127)			0.25 (0.30)
Low WM (*n* = 73)	0.15 (0.20)	0.26 (0.33)	
High WM (*n* = 54)	0.36 (0.32)	0.18 (0.23)	
Overall (N=250)	0.19 (0.27)	0.21 (0.28)	0.20 (0.27)

Note. WM was analyzed as a continuous variable in regression analyses, but for descriptive purposes is displayed here as a median split.

Table 4

Full regression table for students' immediate (top) and delayed (bottom) learning gains.

	Model 1					Model 2				
Immediate Gains	*β*	*b*	SE	*t*	Sig.	*β*	*b*	SE	*t*	Sig.
Condition	0.09	0.03	0.03	0.80	.44	0.15	-0.06	0.07	-0.79	.45
WM						-0.06	-0.06	0.12	-0.51	.62
Gender						0.42	-0.13	0.12	-1.07	.31
WM*Gender						0.50	0.53	0.17	3.12	.009
Condition*WM						0.20	0.21	0.13	1.64	.13
Condition*Gender						-0.10	0.34	0.12	2.91	.01
Condition*WM*Gender						-0.73	-0.78	0.22	-3.49	.004
Delayed Gains	*β*	*B*	SE	*t*	Sig.	*β*		SE	*t*	Sig.
Condition	0.07	0.02	0.02	0.95	.36	0.29	-0.02	0.05	-0.40	.70
WM						-0.10	-0.10	0.10	-1.05	.32
Gender						0.58	-0.04	0.07	-0.54	.60
WM*Gender						0.41	0.41	0.11	3.61	.004
Condition*WM						0.21	0.21	0.13	1.61	.13
Condition*Gender						-0.36	0.15	0.06	2.35	.04
Condition*WM*Gender						-0.53	-0.53	0.18	-2.97	.01

Note. *β* and *b* indicate standardized and unstandardized regression coefficients, respectively. Standard errors are clustered at the level of classroom. The degrees of freedom for Models 1 and 2 are (1, 248) and (7, 242), respectively.

Table 5

Regression analyses of boys' and girls' immediate (top) and delayed (bottom) learning gains

	Boys (n = 123)					Girls (n = 127)				
Immediate Gains	*β*	*b*	SE	*t*	Sig.	*β*	*b*	SE	*t*	Sig.
Condition	0.20	-0.06	0.07	-0.79	.44	0.11	0.29	0.09	3.14	.008
WM	-0.06	-0.06	0.12	-0.51	.62	0.41	0.47	0.09	5.20	<.001
Condition*WM	0.22	0.21	0.13	1.65	.13	-0.50	-0.57	0.18	-3.23	.007
Delayed Gains	*β*	*b*	SE	*t*	Sig.	*β*	*b*	SE	*t*	Sig.
Condition	0.35	-0.02	0.05	-0.40	.70	-0.04	0.13	0.05	2.52	.03
WM	-0.11	-0.10	0.10	-1.05	.32	0.29	0.31	0.09	3.39	.005
Condition*WM	0.23	0.21	0.13	1.62	.13	-0.30	-0.32	0.15	-2.07	.06

Note. *β* and *b* indicate standardized and unstandardized regression coefficients, respectively. Standard errors are clustered at the level of classroom. The model degrees of freedom for boys and girls are (3, 119) and (3, 123), respectively.

Figure 1

Illustrative pattern of results: WM significantly predicted girls' immediate learning (A) and retention (B) in the control condition, but WM was unrelated to learning in the EW condition. Gains are percentage point increases in test score.