COGNITION AND EMOTION
2024, VOL. 38, NO. 1, 171–179
https://doi.org/10.1080/02699931.2023.2258589

Sadness facilitates “deeper” reading comprehension: a behavioural and eye tracking study

Caitlin Mills [a], Rosy Southwell[b] and Sidney K. D’Mello[b,c]

[a]Department of Educational Psychology, University of Minnesota, Minneapolis, MN, USA; [b]Institute of Cognitive Science, University of Colorado Boulder, Boulder, CO, USA; [c]Department of Computer Science, University of Colorado Boulder, Boulder, CO, USA

ABSTRACT

Reading is one of the most common everyday activities, yet research elucidating how affective influence reading processes and outcomes is sparse with inconsistent results. To investigate this question, we randomly assigned participants ($N = 136$) to happiness (positive affect), sadness (negative affect), and neutral video-induction conditions prior to engaging in self-paced reading of a long, complex science text. Participants completed assessments targeting multiple levels of comprehension (e.g. recognising factual information, integrating different textual components, and open-ended responses of concepts from memory) after reading and after a week-long delay. Results indicated that the Sadness (vs. Happiness) condition had higher comprehension scores, with the largest effects emerging for assessments targeting deeper levels comprehension immediately after reading. Eye-tracking analyses revealed that such benefits may be partly driven by sustained attentional focus over the 20-minute reading session. We discuss results with respect to theories on affect, cognition, and text comprehension.

ARTICLE HISTORY
Received 22 September 2022
Revised 26 May 2023
Accepted 27 June 2023

KEYWORDS
reading comprehension; eye-tracking; negative affect; text processing

How do emotions impact the way we read and understand texts? Research thus far offers a murky picture with some studies suggesting positive affect facilitates text comprehension (Bohn-Gettler & Rapp, 2011; Megalakaki et al., 2019; Scrimin & Mason, 2015), while others highlighting the benefits of negative affect (Mills et al., 2019; Trevors et al., 2021). These disparate findings are perhaps not surprising given that the broader literature suggests dissociable benefits of both positive and negative affect across various stages of information processing. For example, positive affect enhances working memory (Storbeck & Maswood, 2016; Yang et al., 2013) and promotes expanded attentional focus with a bias toward more global processing (Basso et al., 1996; Derryberry & Tucker, 1994; Fredrickson & Branigan, 2005). In contrast, negative affect elucidates more analytical, local processing (Gasper & Clore, 2002) and leads to more vivid and detailed memories (Beukeboom & Semin, 2006; Kensinger, 2007).

Whereas experimental paradigms can target these individual cognitive processes, text comprehension is a complex activity which entails integration information across multiple levels of processing (Kintsch, 1988; Zwaan & Radvansky, 1998). First, at the *surface* code level, readers must decode symbols, identify words, and parse sentences (e.g. recognising the letter “e” and the word “bear” in the text below). Next, they must create literal semantic representations –called the *textbase* – to represent facts from the text (e.g. Goldilocks slept “after” eating). Lastly, readers must construct an overall mental model (i.e. *situation model*) that integrates information across multiple textual components (*text-based* inferences; “their” refers to the bears in the example text below), a and from prior-knowledge (*knowledge-*

CONTACT Caitlin Mills cmills@umn.edu University of Minnesota, Minneapolis, MN, 55455 USA.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/02699931.2023.2258589.

based inferences; e.g. encounters with bears can be dangerous).

> Goldilocks entered the bears' house while they were away.
>
> She ate their porridge and fell asleep.
>
> Later, Goldilocks breathed a sigh of relief.

In recognition of such complexities, Bohn-Gettler's (2019) Process-Emotion-Task (PET) framework examines how affect might influence different levels of text processing. Their key hypotheses are derived from the affect infusion model (Forgas 1995; 2002), which suggests that affect can shape cognition, and the extent to which affect is "infused" will depend on various factors such as complexity, effort, and novelty. For example, affect will have more impact when processing is more difficult and generative compared to easy and familiar (e.g. direct retrieval of memories). Accordingly, the PET framework posits that effects of emotion are more likely to emerge when comprehension assessments require more constructive processing at the situation model level compared to the textbase level. In line with this, Mills et al. (2019) found that participants who were experimentally induced with sadness (negative affect) before reading outperformed those who were induced with happiness (positive affect) on assessments that required them to connect different parts of the text (i.e. situation model), whereas no differences emerged for assessments targeting factual memory (i.e. text base level). However, this study stopped short of assessing deeper levels of comprehension, for example, assessments that required readers to construct explanations of the text from memory.

There is also an important question about whether positive vs. negative affect is more beneficial to text comprehension. Once again drawing from the affect infusion model, the PET framework predicts that positive and negative affect might differentially benefit comprehension through different processing styles. Specifically, positive affect is thought to facilitate *global*, *elaborative*, and *assimilative* processing, where textual information is integrated within existing knowledge structures (i.e. making connections between the text and experiences). This processing style is considered to be more efficient, as incoming information is incorporated into existing beliefs and schemas, and details can either be ignored or revised to maintain a consistent schema (Bless et al., 2006; Fiedler & Bless, 2000). For example, a new type of flower could be assimilated into the category of "pretty things." Negative affect, in contrast, is hypothesised to facilitate more *local*, *analytical,* and *accommodative* processing; rather than fitting the text to prior knowledge, existing internal representations are adapted to incorporate the new information. This mode of processing is thought to stem from an evolutionary need to more carefully evaluate perceived threats, thus narrowing attentional focus to focus on details (Forgas, 1995). Such adaptive mechanisms also purportedly update mental models in order to learn from, avoid, or appropriately react to potentially harmful or negative stimuli. For example, a person who thought bees were cute but gets stung by one might now recategorize them (i.e. accommodation or conceptual change) as being both "cute things" as well as a "harmful thing."

Of the studies have that have induced affective states prior to reading, at least two studies reported a comprehension benefit for readers induced with negative affect (Mills et al., 2019; Trevors et al., 2021), whereas two other studies found benefits after a positive affect induction (Bohn-Gettler & Rapp, 2011; Scrimin & Mason, 2015). Notably, and directly related to the first hypothesis above, the comprehension assessments of these studies widely varied (e.g. true/false, fact-based multiple choice, retrospectively coded think-alouds), and none focused on systematic variation of assessments targeting different levels of comprehension. Further, extant research has focused on relatively short texts (under 1000 words) so it is unclear whether any effects of emotions will persist across long, connected texts and on comprehension assessed after memory consolidation has occurred (i.e. after a delay).

Accordingly, we designed a study to examine the effects of experimentally induced positive vs. negative affect on comprehension of a long, connected, real-world, expository (informational) text using assessments of comprehension at multiple levels. Negative and positive affect were manipulated using videos that targeti sadness and happiness, respectively, in line with previous literature (Bohn-Gettler & Rapp, 2011, Mills et al, 2019). Our main hypothesis was that there would be a benefit of negative affect because reading a long, complex text might be facilitated by the ability to focus on details and sustain attention over time, which are benefitted by more local processing engendered by negative affect (Sinclair & Mark, 1995; Von Hecker & Meiser, 2005). This hypothesis aligns with, but expands upon, ideas from the PET framework; negative affect may provide a particular benefit by inspiring readers to maintain systematic processing throughout the whole text, allowing them to construct a more complete mental model of the whole text. If this is the case, we would expect any benefits of affect to be observed on assessments targeting deeper levels of comprehension as predicted by the PET framework.

Finally, we also explored differences in reading times and eye-gaze patterns to provide a complementary account of how affective states shape the reading process over time, in line with our predictions above. Many of the accounts about how affect influences comprehension processes imply a crucial time element, including how the reading process unfolds over time (i.e. sustained attention over the course of the text). We thus aim to provide a more comprehensive account of how the reading process unfolds under the influence of affect while reading a long, connected text.

Methods

Additional details on methods and analyses can be found in the supplemental materials. Data and code are available upon request.

Participants

Simulation studies suggest that approximately $N =$ 100 is adequate for medium effect sizes in multilevel modeling (Maas & Hox, 2005). Data were therefore collected from 136 participants from a Northeastern University 1 ($N = 85$) and Midwestern University ($N = 51$). Participants had a mean age of 20.7 (range 18–44), 69% female, 30% male, 1.4% Other gender; 1.4% African American, 8.6% Asian, 79% Caucasian, 5.0% Hispanic, and 5.8% Other ethnicity.

Materials

Affective state induction

We used a validated 2.5-minute videos to induce Negative ("The Champ"), Positive ("Whose Line is it Anyway"), and Neutral affect ("The Lover" and "L'Amant.") before reading (Gilman et al., 2017; Gross & Levenson, 1995). The Negative and Positive emotion inductions were specific to *sadness* and *happiness*, respectfully, which are the most commonly induced emotions in the reading comprehension studies cited above, with a number of studies using the same videos.

Text

Participants read a 6500-word excerpt from a book on surface tension in liquids (Boys, 1959) split across 57 screens of text (pages). The text demonstrates various properties of water and surface tension by describing a series of experiments, and had a Flesh-Kincaid grade score 11.8, on par with an average high-school reading level.

Eye-tracking

Both sites used a Tobii TX 300 for eye tracking, which uses infrared eye tracking and samples at 120 Hz. Stimuli were displayed on a 23″ monitor integrated with the eye-tracker and had a resolution of 1920 × 1080. Participants were seated such that eyes were approximately 65 cm from the screen, and participants were free to move however they would like. We computed four gaze features per page: (1) *number of fixations* per page; (2) *mean fixation duration* in seconds; (3) *mean saccade amplitude* in pixels; and (4) *regression fixation proportion*, i.e. fixations on earlier words than the previous fixation – see Figure 1. These features are widely used and of theoretical relevance (Rayner et al., 2003); for instance, regressions indicate corrective processing (Rayner et al., 2006), saccades increase in amplitude during skim reading (Strukelj & Niehorster, 2018), and fixation durations are longer while reading difficult passages (Rayner et al., 2006). We also examined *reading time per page* in seconds as a measure of processing depth (Mills et al., 2017).

Eye gaze data was unavailable for 5 of the participants. Of 7467 potential pages of eye-tracking data (131 × 57), 6147 (82%) were deemed valid (excluding pages with no eyetracker data [e.g. due to a failure to detect a pupil], pages with fewer than three fixations, or that were read for less than two seconds). The raw gaze data was fixation filtered to extracted gaze fixations (when gaze remains on given location) and saccades (when gaze moves to a new location) using the open-source OGAMA software (Voßkühler et al., 2008) using a 31-pixel radius, minimum fixation duration of 100 ms, and minimum distance between fixations of 57 pixels. An examination of the distributions of the eye gaze features

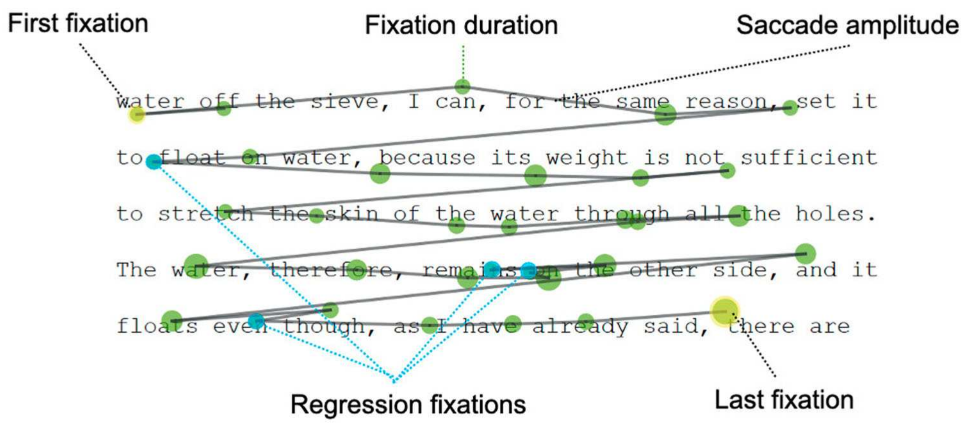

Figure 1. Gaze on a page of text showing fixations (circles), saccades (lines) and regressions (BLUE fixations and saccades) overlaid on the text. Fixation circles have radius proportional to each fixation duration.

indicated some extreme outliers for all five variables so we winsorized the top 1% of the data. Refer to Table S1-S3 for descriptive statistics and inter-variable correlations.

Comprehension assessments

The text was divided into 15 concepts, each spanning multiple pages, and with a focal theme. Comprehension was assessed immediately after reading, using two multiple-choice *factual* and one *text-integration* question for each of 15 textual concepts (scored as correct [1] or incorrect [0]). An additional 15 *open-ended generation* prompts (one per concept) allowing freeform textual responses designed to elicit elaboration and generalisation were interspersed with the multiple-choice assessments. These items were scored by two trained raters (blind to condition) on a continuum between 0 and 1. Six researchers independently wrote archetypal answers to each question. Criteria for scoring included similarity between each response and the archetypal answers, and inclusion of particular keywords deemed critical by the researchers. Two raters marked a subset of the responses, with a high convergence between the scorers ($r = 0.891$). Given the strong convergence between scorers, a single researcher scored the remainder of responses.

Additional multiple-choice questions (two factual and one text-integration per concept) were presented after a 1-week delay with items (immediate vs. delay) counterbalanced across participants. See Table 1 for examples. We averaged scores on learning assessments to the participant level. Five participants were missing behavioural data (e.g. posttests), whereas 17 failed to complete the delayed posttest assessments, independent of condition ($\chi^2 = 1.47$, $p = .48$).

Procedure

Participants provided consent and then were calibrated on the eye-tracker. They were randomly assigned to watch one of the emotion induction videos. Self-reported valence (very unpleasant to very pleasant) and arousal (sleepy to active) on 9-point scales were collected before and after the video. Participants then read the text (self-paced) and could advance back and forth using the arrow keys. Participants in the Negative video condition were also shown the positive emotion video at then end of the study to assuage any residual negative affect. All participants were debriefed and given credit/compensation for participating. Finally, participants completed the comprehension assessments *without* access to the text after reading and a week later.

Results

We used linear mixed effects regression models for the analyses in R (version 4.3.0; R Core Team, 2021), using the lme4 package (version 1.1-33; Bates et al., 2007) with Satterthwaite's degrees of freedom method from lmerTest. Condition was included as a three-level categorical fixed effect with the Negative condition as the reference group as we were primarily concerned with the positive vs. negative contrast

Table 1. Example text and corresponding assessments for a single concept. For brevity, text is abridged here with ellipses, to show relevant text for the example assessment items.

Text	*… I have melted a quantity of paraffin in a dish and dipped this gauze into the melted paraffin so as to coat the wire all over with it, but I have shaken it well while hot to knock the paraffin out of the holes. […] I have laid a small piece of paper in the sieve, and am pouring the water on to the paper, which breaks the fall. I have now poured in about half a tumbler of water, and I might put in more, I take away the paper but not a drop runs through. […] If now I shake the water off the sieve, I can, for the same reason, set it to float on water, because its weight is not sufficient to stretch the skin of the water through all the holes. The water, therefore, remains on the other side, and it floats …*
Factual	Does the sieve the author used in an experiment float on water? a) yes, the weight of the sieve is not sufficient to stretch the skin of the water through the holes b) no, the water passes through the holes c) no, the sieve is too heavy to float d) none of the above *correct response: a)*
Text-integration	Which of the following is the most similar to how water behaves if you poured it in a wax-coated thimble covered with holes? a) like a colander for draining b) carrying a pail of water with a leak c) like a regular cup of water d) none of the above *correct response: c)*
Open-ended generative	**Prompt:** How does "going to sea in a sieve" seem possible based on the authors demonstration? **Response:** *"If the holes of the sieve are small enough such that the elastic skin of water could form across each of them, and if the sea was fairly calm, and of course if the sieve itself was large enough to hold you (i.e. displace enough water such that the weight of the water is at least equal to the combined weight of you and the sieve)."*

based on prior work. However, we report all three pairwise contrasts using estimated marginal means. The random effect structure varied across analyses as detailed below.[1] We used two-tailed tests with a $p < .05$ cutoff for significance.

Manipulation checks

Regressing *pre-manipulation* valence and arousal scores on Condition with cohort (i.e. university testing site) as a random intercept (formula: $pre \sim Condition + 1 \mid cohort$) indicated no significant differences ($ps > .28$). Next, we regressed *post-manipulation* scores after accounting for pre-manipulation levels (formula: $post \sim Condition + arousal\_pre + valence\_pre + 1 \mid cohort$) and found that the Negative and Positive conditions did not report different levels of arousal ($\beta = .15$, $p = .21$), which were both significantly higher than the Neutral condition, more so for the Positive ($\beta = .38$, $p < .01$) than Negative ($\beta = .24$, $p < .05$) conditions. The Positive condition had quantitatively higher valence scores than Neutral ($\beta = .25$, $p = .08$), both of which were higher than Negative ($\beta = 1.11$ for Positive; 0.86 for Neutral; $ps < .0001$). We also compared post- vs. pre- change in scores using paired-samples t-tests, finding a significant increase in valence for Positive ($p < .01$), decrease for Negative ($p < .01$), and no change for Neutral ($p = .61$). For arousal, there was an increase for Positive ($p < .01$), but not the other two condition ($ps > .12$). Thus, the manipulation was successful at eliciting the desired emotions (see contour plots in S2), more so for the Negative vs. Neutral distinction than Positive vs. Neutral, which is consistent with the literature (Schaefer et al., 2010).

There were also no significant differences across condition for age ($F\ (2) = 1.03$, $p = .359$), gender (χ^2 (4) = 5.5, Fisher's $p = .19$), and ethnicity (χ^2 (8) = 7.2, Fisher's $p = .56$), suggesting successful random assignment (see Table S1 for demographic summary). Please refer to Tables S1-S2 for descriptive statistics related to random assignment and manipulation checks. See also Figure S1 histograms of behavioural scores and Table S3 for correlations among these.

Comprehension outcomes

We regressed each of the five comprehension measures (open-ended generation [for immediate assessment only]; factual and text-integration memory [for immediate and delayed assessments]) on Condition. We included cohort as a random intercept; more complex random effect structures resulted in convergence errors – formula: Comprehension measure ∼ condition + (1 | cohort).

In general, the Negative group outperformed the Positive group ($\beta_{avg} = .22$) with a significant difference ($\beta = .42$, $p = .02$) for the open-ended generation assessment only. The Negative group also had quantitatively higher (albeit not significantly) scores than the Neutral group ($\beta_{avg} = .22$), which similarly outperformed, but not significantly, the Positive group ($\beta_{avg} = -.20$). The above pattern replicated when restricted to the 119 participants who also completed a delayed assessment (i.e. Negative was significantly higher than Positive [$\beta = -.49$; $p = .012$] for open-ended generation items; with no other significant differences). Tables S5-S9 provides complete results and model formulae.

Reading time and Eye-gaze

We analyzed reading time and the four eye gaze measures at the page level to ascertain overall condition differences as well as changes over time (pages). Specifically, we regressed each variable on the Condition × Page Number [1 to 57] interaction term with random slopes and intercepts for page number and participant nested within cohort; more complex random effect structures yielded convergence errors. Because reading times were moderately correlated with the eye gaze measures (|rs| from .096 to .536), we added reading time as an additional covariate for the gaze models[2], such that the final models were constructed with the formula: DV ~ Condition * Page number + Reading time + (Page number | Cohort: Participant) + (1 | Cohort)

There was no significant main effect of condition for any of the models, but there were significant Condition × Page Number interactions for reading time, saccade amplitude, and proportion of regressive fixations. Simple slopes analyses indicated that reading times significantly decreased across pages for the Positive ($\beta = -.15$, $p < .001$) and Neutral ($\beta = -.09$, $p = .01$) conditions whereas they were stable for the Negative condition ($\beta = -.02$, $p = .58$). Conversely, saccade amplitudes increased across pages for Positive and Neutral conditions ($\beta s = .07$, $ps < .01$) but were stable for the Negative condition ($\beta = -.04$, $p = .13$). Lastly, there were more regressive fixations for the Neutral condition ($\beta = .08$, $p < .01$) as time progressed, whereas they remained consistent for the other conditions (βs of $-.02$ for Negative and .021 for Positive, $ps > .49$). The Negative reading time and saccade amplitude slopes were significantly ($ps < .01$) different than the corresponding Positive slopes, whereas the Negative and Neutral conditions differed with respect to the saccade amplitude and regression fixation slopes ($ps < .01$); there were no differences between any of the Positive and Neutral slopes.[3]

Discussion

Our study aimed to address several gaps in the literature on the influence of affect and reading comprehension. Novel aspects of our work include a long, connected, real-world, expository text, assessments targeting multiple levels of comprehension and across time, and an analysis of reading time and eye-gaze. Our results suggest that sadness (Negative affect induction) selectively benefitted comprehension as assessed by open-ended generation items in comparison to happiness (Positive affect induction). There was also a non-significant trend in favour of negative affect for text-integration items, both after reading and a week later compared to positive affect, replicating a similar finding from Mills et al. (2019). Comparisons to the neutral condition revealed trends of a small facilitative effect for negative affect and a detrimental effect for positive affect. Lastly, differences among conditions increased as a function of comprehension depth: factual ($|\beta_{avg}| = .07$) < text-integration memory ($|\beta_{avg}| = .18$) < open-ended generative ($|\beta_{avg}| = .28$), confirming a key hypothesis of the PET framework regarding when affect influences text comprehension (Bohn-Gettler, 2019).

These effect sizes are both practically and theoretically important given this relatively nascent area of research, and they can be useful for future studies incorporating multiple comprehension assessments. As Kraft (2020) points out, traditional effect size "benchmarks" (i.e. those proposed by Cohen in 1969) may be somewhat outdated or misleading, particularly in the context of educationally-relevant studies where "small" effects are the norm and can have meaningful long-term impacts. In particular, the significant difference between the negative (sadness) and positive (happiness) conditions for open-ended generation items corresponded to an effect size of 0.42, which is more than double the 0.17 average effect (computed from 1200 effects) across 495 reading studies (Kraft, 2020).

Our findings also extend to more general theories on the influence of affect on information processing; negative affect (sadness) had stronger effects for building mental representations perhaps by facilitating attentional focus over longer periods of time (i.e. a 20-minute-long minute reading session). This interpretation is supported by previous theoretical and empirical work suggesting that negative emotions tend to help narrow focus and remember more details (Fredrickson, 2004; Gasper & Clore, 2002; Kensinger, 2007; Storbeck & Clore, 2005). Evolutionary-based arguments, for example, argue that negative emotions help narrow attention to potentially harmful situations to enable a person to act accordingly or learn for future situations. Our work is in line with this account, yet extends it to a longer, non-threatening environment; though our emotion induction itself was not threatening, it may have nevertheless primed participants to more deeply attend to the text.

We also examined how eye-gaze behaviours were influenced by affect during reading. Previous studies have had mixed results, with one study indicating that participants induced with positive affect spent longer time reading and had longer regressive fixations compared to negative affect (Scrimin & Mason, 2015). However, this particular study used a paradigm where four out of six comprehension items were assessed *both* before and after reading, which likely altered reading strategies and presumably eye gaze patterns, i.e. the pretesting effect (Richland, et al., 2009). In contrast, there is also support for the general idea that positive affect may induce readers to alter their reading strategies to be less systematic – and perhaps less detail-oriented – as they progress through the text, including our results. For example, one study found that positive affect may lead to faster reading times and shorter fixations (compared with a neutral mood induction; Ballenghein et al., 2019), whereas readers had longer first-pass fixation times while reading negative textual content (once again, compared to neutral; Arfé et al., 2022). Our results are consistent with these effects, namely that the positive condition sped up their reading, whereas the negative condition demonstrated more consistent reading behaviours across time.

Lastly, because it is unclear how long the mood induction lasted, there are at least two possibilities to explain our results. On the one hand, the mood induction may have lasted throughout the reading. On the other hand, even if the mood induction wore off after a few minutes, the results could be explained by a cascading effect, whereby readers in the negative (sadness) condition may have read more attentively from the beginning, which helped them maintain a stronger mental model to sustain attention and aid comprehension across time (Kopp, Mills, & D'Mello, 2015).

Like all studies, ours has limitations. One limitation with our study is that we only included the open-ended items in the immediate assessment because these questions could not be re-asked in a different manner a week later without being susceptible to testing effects. Another is that due to the complexity of collecting eye gaze data, our sample size was underpowered to detect some of the smaller effects. We also opted to use a remote eye tracker and did not constrain movements to enable a more naturalistic reading experience. However, this precluded computation of some gaze features (e.g. first-past fixations) that require high precision eye tracking. We also conceptualised affect along the valence dimension to be consistent with the PET framework, which precluded an analysis from a discrete emotion perspective (Lerner & Keltner, 2000). Future work may consider including other negative emotions, while also clarifying the possible role of arousal since it increased from pre–post for the positive affect (happiness) condition, compared to the other too, though levels of arousal after viewing the videos were consistent across the positive and negative conditions.

Future work should consider manipulating multiple negative emotions (e.g. anger, fear, sadness) as well as arousal and emotional intensity (Megalakaki et al., 2019). It might also be prudent to manipulate the reading goal (i.e. reading for extrinsic reward vs. reading for pleasure). Such work may help elucidate the potential interaction between the reader and the text (Bohn-Gettler, 2019), particularly if certain texts evoke either accommodative or assimilative processing themselves (e.g. in the case of unfamiliar vs. familiar content; Bohn Gettler, 2019). Finally, manipulating the emotional valence of the text may also be important, given that the text itself can alter participants' subjective feelings of valence (Arfé et al., 2022; Lerner & Keltner, 2000; Smith et al., 2022).

In summary, our results lend support to the idea that negative affect (specifically sadness) can be beneficial in the context of reading for understanding and learning by sustaining attention over time.

Author notes

All materials and code are available at the following link: https://osf.io/mz4dp/?view_only=93196740d61e4b0f8625ac523585b1a7. Data is available upon re quest and approval from IRB. A previous publication (Southwell et al., in press, *User Modeling and User-Adapted Interaction*) used the same dataset as part of a larger machine learning project focused on eye

-gaze, but it does not overlap with this paper in any way (i.e. affect is not mentioned in that study).

Notes

1. We began with maximal random effects structures, but simplified them to address convergence issues. For the manipulations checks and learning measures models, we began with (Condition | Cohort). Our initial models for the reading time and eye-gaze measures included the following structure: (Page number | Participant: Cohort) + (Condition * Page number | Cohort).
2. Regarding the results of the models with and without reading time as a covariate, we found that the results were the same with one minor exception. Specifically, whereas the difference between the positive and neutral slopes for the proportion of regressive fixations yielded a $p = .07$ with reading time in the model, this difference changed to $p = .058$ without reading time included.
3. Two of the models, reading time and fixation duration, did not converge with (1 | Cohort) included. We thus re-ran the models without Cohort as a random intercept, and the results did not change.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by the National Science Foundation (NSF) (DRL 1235958, DRL 1920510). Any opinions, findings and conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of NSF.

ORCID

Caitlin Mills http://orcid.org/0000-0003-4498-0496

References

Arfé, B., Delatorre, P., & Mason, L. (2022). Effects of negative emotional valence on readers' text processing and memory for text: An eye-tracking study. *Reading and Writing*, 1–26.

Ballenghein, U., Megalakaki, O., & Baccino, T. (2019). Cognitive engagement in emotional text Reading: Concurrent recordings of eye movements and head motion. *Cognition and Emotion*, *33*(7), 1448–1460. https://doi.org/10.1080/02699931.2019.1574718

Basso, M. R., Schefft, B. K., Ris, M. D., & Dember, W. N. (1996). Mood and global-local visual processing. *Journal of the International Neuropsychological Society*, *2*(3), 249–255. https://doi.org/10.1017/S1355617700001193

Bates, D., Maechler, M., & Dai, B. (2007). The lme4 package. *R Package Version*, *2*(1), 74.

Beukeboom, C. J., & Semin, G. R. (2006). How mood turns on language. *Journal of Experimental Social Psychology*, *42*(5), 553–566. https://doi.org/10.1016/j.jesp.2005.09.005

Bless, H., Fiedler, K., & Forgas, J. P. (2006). Mood and the regulation of information processing and behavior. *Affect in Social Thinking and Behavior*, 6584.

Bohn-Gettler, C. M. (2019). Getting a grip: The PET framework for studying how reader emotions influence comprehension. *Discourse Processes*, *56*(5-6), 386–401.

Bohn-Gettler, C. M., & Rapp, D. N. (2011). Depending on my mood: Mood-driven influences on text comprehension. *Journal of Educational Psychology*, *103*(3), 562. https://doi.org/10.1037/a0023458

Boys, C. V. (1959). *Soap-bubbles and the Forces which Mould Them*. Garden City, NY: Doubleday Anchor Books.

Derryberry, D., & Tucker, D. M. (1994). *Motivating the focus of attention.*

Fiedler, K., & Bless, H. (2000). *The formation of beliefs at the interface of affective and cognitive processes.*

Forgas, J. P. (1995). Mood and judgment: the affect infusion model (AIM). *Psychological Bulletin*, *117*(1).

Forgas, J. P. (2002). Feeling and doing: Affective influences on interpersonal behavior. *Psychological Inquiry*, *13*(1), 1–28.

Fredrickson, B. L. (2004). The broaden-and-build theory of positive emotions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*(1449), 1367–1377. https://doi.org/10.1098/rstb.2004.1512

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition & Emotion*, *19*(3), 313–332. https://doi.org/10.1080/02699930441000238

Gasper, K., & Clore, G. L. (2002). Attending to the big picture: Mood and global versus local processing of visual information. *Psychological Science*, *13*(1), 34–40. https://doi.org/10.1111/1467-9280.00406

Gilman, T. L., Shaheen, R., Nylocks, K. M., Halachoff, D., Chapman, J., Flynn, J. J., & Coifman, K. G. (2017). A film set for the elicitation of emotion in research: A comprehensive catalog derived from four decades of investigation. *Behavior Research Methods*, 49, 2061–2082.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition and Emotion*, *9*(1), 87–108.

Kensinger, E. A. (2007). Negative emotion enhances memory accuracy: Behavioral and neuroimaging evidence. *Current Directions in Psychological Science*, *16*(4), 213–218. https://doi.org/10.1111/j.1467-8721.2007.00506.x

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a construction-integration model. *Psychological Review*, *95*(2), 163.

Kopp, K., Mello, S., & Mills, C. (2015). Influencing the occurrence of mind wandering while reading. *Consciousness and Cognition*, *34*, 52–62.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & Emotion*, *14*(4), 473–493. https://doi.org/10.1080/026999300402763

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86. https://doi.org/10.1027/1614-2241.1.3.86

Megalakaki, O., Ballenghein, U., & Baccino, T. (2019). Effects of valence and emotional intensity on the comprehension and memorization of texts. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00179

Mills, C., Graesser, A., Risko, E. F., & D'Mello, S. K. (2017). Cognitive coupling during reading. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/xge0000309

Mills, C., Wu, J., & D'Mello, S. (2019). Being sad is not always bad: The influence of affect on expository text comprehension. *Discourse Processes*, *56*(2), 99–116. https://doi.org/10.1080/0163853X.2017.1381059

Rayner, K., Chace, K. H., & Slattery, T. J. (2006). Eye movements as reflections of comprehension processes in reading. *Scientific Studies of Reading*, *10*(3), 241–255.

Rayner, K., Pollatsek, A., & Reichle, E. D. (2003). Eye movements in reading: Models and data. *Behavioral and Brain Sciences*, *26*(4), 507–518. https://doi.org/10.1017/S0140525X03520106

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning?. *Journal of Experimental Psychology: Applied*, *15*(3), 243.

Schaefer, A., Nils, F., Sanchez, X., & Philippot, P. (2010). Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers. *Cognition & Emotion*, *24*(7), 1153–1172. https://doi.org/10.1080/02699930903274322

Scrimin, S., & Mason, L. (2015). Does mood influence text processing and comprehension? Evidence from an eye-movement study. *British Journal of Educational Psychology*, *85*(3), 387–406. https://doi.org/10.1111/bjep.12080

Sinclair, R. C., & Mark, M. M. (1995). The effects of mood state on judgemental accuracy: Processing strategy as a mechanism. *Cognition & Emotion*, *9*(5), 417–438. https://doi.org/10.1080/02699939508408974

Smith, S. L., Ward, R. T., Allen, L. K., Wormwood, J. B., & Mills, C. (2022). Mind your words: Affective experience during reading mediates the effect of textual valence on comprehension. *Applied Cognitive Psychology*, *36*(5), 1131–1141. https://doi.org/10.1002/acp.3983

Storbeck, J., & Clore, G. L. (2005). With sadness comes accuracy; With happiness, false memory mood and the false memory effect. *Psychological Science*, *16*(10), 785–791. https://doi.org/10.1111/j.1467-9280.2005.01615.x

Storbeck, J., & Maswood, R. (2016). Happiness increases verbal and spatial working memory capacity where sadness does not: Emotion, working memory and executive control. *Cognition and Emotion*, *30*(5), 925–938. https://doi.org/10.1080/02699931.2015.1034091

Strukelj, A., & Niehorster, D. C. (2018). One page of text: Eye movements during regular and thorough reading, skimming, and spell checking. *Journal of Eye Movement Research*, *11*(1).

Trevors, G., Bohn-Gettler, C., & Kendeou, P. (2021). The effects of experimentally induced emotions on revising common vaccine misconceptions. *Quarterly Journal of Experimental Psychology*, *74*(11), 1966–1980. https://doi.org/10.1177/17470218211017840

Von Hecker, U., & Meiser, T. (2005). Defocused attention in depressed mood: Evidence from source monitoring. *Emotion*, *5*(4), 456. https://doi.org/10.1037/1528-3542.5.4.456

Voßkühler, A., Nordmeier, V., Kuchinke, L., & Jacobs, A. M. (2008). OGAMA (Open Gaze and Mouse Analyzer): Open-source software designed to analyze eye and mouse movements in slideshow study designs. *Behavior Research Methods*, *40*(4), 1150–1162. https://doi.org/10.3758/BRM.40.4.1150

Yang, H., Yang, S., & Isen, A. M. (2013). Positive affect improves working memory: Implications for controlled cognitive processing. *Cognition & Emotion*, *27*(3), 474–482. https://doi.org/10.1080/02699931.2012.713325

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*(2), 162.