The emotional content of children's writing: a data-driven approach

COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 48 (2024) e13423

ISSN: 1551-6709 online
DOI: 10.1111/cogs.13423

The Emotional Content of Children's Writing: A Data-Driven Approach

Yuzhen Dong,[a] Yaling Hsiao,[b] Nicola Dawson,[a] Nilanjana Banerji,[c] Kate Nation[a]

[a]*Department of Experimental Psychology, University of Oxford*
[b]*School of Education, University of Birmingham*
[c]*Oxford University Press*

Received 1 February 2023; received in revised form 29 January 2024; accepted 16 February 2024

Abstract

Emotion is closely associated with language, but we know very little about how children express emotion in their own writing. We used a large-scale, cross-sectional, and data-driven approach to investigate emotional expression via writing in children of different ages, and whether it varies for boys and girls. We first used a lexicon-based bag-of-words approach to identify emotional content in a large corpus of stories (N>100,000) written by 7- to 13-year-old children. Generalized Additive Models were then used to model changes in sentiment across age and gender. Two other machine learning approaches (*BERT* and *TextBlob*) validated and extended these analyses, converging on the finding that positive sentiments in children's writing decrease with age. These findings echo reports from previous studies showing a decrease in mood and an increased use of negative emotion words with age. We also found that stories by girls contained more positive sentiments than stories by boys. Our study shows the utility of large-scale data-driven approaches to reveal the content and nature of children's writing. Future experimental work should build on these observations to understand the likely complex relationships between written language and emotion, and how these change over development.

Keywords: Children's writing; Emotion; Language production; Natural language processing; Data-driven approach; Sentiment analysis

Correspondence should be sent to Yuzhen Dong, Department of Experimental Psychology, University of Oxford, Anna Watts Building, Oxford OX2 6GG, UK. E-mail: yuzhen.dong@psy.ox.ac.uk

1. Introduction

As a form of expressive language, writing is a complex skill that draws on an array of cognitive factors from choosing words, constructing sentences, and building paragraphs, through to developing ideas that are appropriate for the discourse context and genre (Graham, 2006; Harris, Graham, Brindle, & Sandmel, 2009). For children, learning to write is difficult and takes time (e.g., Graham, 2018; McCutchen, 2006). In this paper, we focus on one important component of children's writing, namely, their use of language to express sentiments and emotional content. Our study builds from the hypothesis that emotional understanding is constructed through development, and that language plays a key role in this construction (Barrett, 2017; Hoemann, Xu, & Barrett, 2019; Lindquist, 2017). While this body of research has considered language in general, our focus is on written language—the language that children read or write. Analyses of the language content of children's books demonstrate that they are more linguistically complex than everyday conversational language, and that their emotional content is more nuanced (Dawson, Hsiao, Tan, Banerji, & Nation, 2021; Nation, Dawson, & Hsiao, 2022). If language is a key ingredient in the development of different aspects of emotion understanding (for review, see Shablack & Lindquist, 2019), it follows that experience with *written* language might be particularly associated with the construction of emotion through development. With this as a backdrop, we used natural language processing to analyze a large corpus of stories written by 7- to 13-year-old children. This data-driven cross-sectional approach allowed us to identify emotional content in the narrative writing of girls and boys at different ages.

1.1. Language and emotion development

According to the constructionist hypothesis of emotional development, language supports the acquisition of emotional knowledge and is considered a "key ingredient" in the development of different aspects of emotional understanding (Barrett, 2017; Hoemann et al., 2019; Lindquist, 2017; Shablack & Lindquist, 2019). Language is situated in social interaction from infancy onward and this is reflected in the association between children's experience with emotional language and their subsequent emotional understanding. For example, Dunn, Brown, and Beardsall (1991) coded "feeling-state" language (e.g., happy) from recordings of everyday mother−child interactions at 36 months of age. They noted that children who experienced more feeling-state language at 36 months were better able to understand others' emotions at 6 years of age. In line with the view that language is critical to the construction of emotion, there is also an association between language competence and emotion understanding in preschoolers (e.g., Widen & Russell, 2003, 2008) and older children (Beck, Kumschick, Eid, & Klann-Delius, 2012; S. Griffiths, Goh, & Norbury, 2020).

Emotion vocabulary grows through childhood and becomes more nuanced and multidimensional over time (e.g., Grosse, Streubel, Gunzenhauser, & Saalbach, 2021; Nook et al., 2020). Using a word generation task (e.g., "think about a child who is feeling sad and then write down as many words as you can to describe this feeling"), Doost, Moradi, Taghavi, Yule, and Dalgleish (1999) found that older children produced significantly more emotion words than

the younger children. This developmental period coincides with the onset of literacy and this might be important. Nation et al. (2022) argued that exposure to written language provides children with opportunities to experience language that is different from everyday usage, including the complex and nuanced language associated with emotions and mental states. They suggested that the language needed to drive emotional development through childhood might build directly from experience with written language. Consistent with this, Dawson et al. (2021) found that story books written for preschoolers (3−5 years old) contained more emotionally arousing words than child-directed speech. From mid-childhood onward, most new vocabulary is learned via reading, suggesting that the doubling of emotion word vocabulary seen every 2 years between 4 and 11 years of age (Baron-Cohen, Golan, Wheelwright, & Granader, 2010) is associated with reading experience. From this perspective, the extended development of connections between emotion and language (e.g., Grosse et al., 2021; Nook et al., 2020) reflects the time needed for these connections to accumulate from written language input. Reflecting on the importance of words in shaping the perception of emotion in faces, Doyle and Lindquist (2018, p.62) noted that the "scowls, frowns, grimaces and growls you see over time presumably develop into conceptual knowledge for what *anger* looks like, helping you to make meaning of new instances of facial actions as instances of anger." In the absence of an actual face, written language needs to describe the face so that the reader can infer its emotional state, as intended by the writer. Consequently, *scowl*, *frown*, *grimace*, and *growl* are specified in written language but might not be needed in a conversational exchange where the shared situation can provide the communicative intent, perhaps via facial expression or tone of voice. This partly explains why written language tends to be more linguistically complex than speech (Dawson et al., 2021; Nation et al., 2022), and indicates why written language (via reading or writing) might provide opportunities for children to experience emotional and mental state language, and in doing, drive their emotional development. Readers are required to navigate this complex language to understand social relationships and make inferences about the mental state of characters (Kim, Park, & Park, 2015; Mar, Oatley, Hirsh, dela Paz, & Peterson, 2006; Siller, Swanson, Serlin, & George, 2014; White, Hill, Happé, & Frith, 2009). Children's story books, therefore, provide opportunities to adopt the emotional perspectives of different characters and to understand empathy (e.g., Hogan, 2011; Kucirkova, 2019). Over time, this might build to the complex nexus between print exposure, literacy, fiction reading, mentalizing, emotion understanding, and theory of mind seen in older children and adults (e.g., Beck et al., 2012; Boerma, Mol, & Jolles, 2017; Kidd & Castano, 2013; Mar & Oatley, 2008; Schwering, Ghaffari-Nikou, Zhao, Niedenthal, & MacDonald, 2021).

1.2. Writing and emotional development

As children learn to read and experience text as a reader, they also begin to write. The act of writing provides opportunities for children to use emotional language in ways that are different from their everyday spoken interaction. In the absence of situational cues, children need to consider the emotional perspectives and mental states of characters and situations they wish to portray, and to nuance their language to communicate this intended meaning to the reader.

Potentially then, learning to write and writing experience might drive aspects of emotional development, and vice versa. While the relationship between *reading* and emotional development has been studied in the context of the theory of mind and the Event Indexing Model (e.g., Kidd & Castano, 2013; Mar & Oatley, 2008; Zwaan, Langston, & Graesser, 1995), the possible relationship between children's own *writing* and their emotional development has not been considered.

There are, however, reasons to consider production via writing as important, beyond the reading experience itself, and this suggestion is consistent with two very different lines of research. First, writing is a constructive act. To write requires active thinking by the author to communicate emotional nuance in a way that is understood by the reader as intended, and by doing so, writing may support emotion understanding. The Production-Distribution-Comprehension account of spoken language (MacDonald, 2013) sees establishing comprehension for the listener as central to the act of production by the speaker. Language experience flows from this production-comprehension loop, as sensitivity to the distributional properties of language builds over time and is drawn upon to support future processing encounters. This framework is attractive as it builds explicit links between language learning and language processing. Extending it to children's writing, the purposeful and productive act of selecting language to convey an intended message to the reader provides an opportunity for children to craft emotional language to capture shades of meaning (the scowls, frowns, and growls, for example). From this view, the act of writing might provide a vehicle for better understanding the emotional content of what children are writing about, and in turn, this will further tune and refine emotional understanding.

Second, a different body of work has noted that patterns of language use in expressive writing are associated with measures of well-being in adolescents, leading to the suggestion that linguistic analysis of young people's writing is a promising method for detecting mental health concerns (Cohen, Shroff, Nook, & Schleider, 2022; Nook, 2023; Nook, Hull, Nock, & Somerville, 2022; Shearer et al., 2021). On this view, a piece of writing can be seen as a reflection of a child's emotional understanding and internal mind, and it is, therefore, reasonable to ask whether developmental changes in the use of emotional language are evident in children's own writing. While this has not been examined in detail, previous studies have found that children express and experience their emotions differently through development (e.g., Bailen, Green, & Thompson, 2019; Thompson, 1991). Children's knowledge of emotion-specific vocabulary also increases across development, with the size of the emotional lexicon doubling every 2 years between 4 and 11 years old (Baron-Cohen et al., 2010). Using a novel emotion vocabulary assessment, Nook et al. (2020) provided further evidence that knowledge of emotion-specific vocabulary increases across childhood and while it plateaus at around 11 years of age, the abstractness of young people's description of emotion continues to increase through to age 18 years. Developmental change manifests not just as an increase in the amount of emotion vocabulary size, as older children from secondary schools produce significantly more negative emotion words than younger children from primary schools (Doost et al., 1999). This runs parallel with longitudinal evidence from self-reported mood measures showing that the frequency of negative emotion increases and positive emotion decreases

from childhood into late adolescence (Larson, Moneta, Richards, & Wilson, 2002; Weinstein, Mermelstein, Hankin, Hedeker, & Flay, 2007).

While these two lines of research give reason to propose that children's own writing is associated with their knowledge of emotion-specific vocabulary and emotional understanding, the current evidence base is severely limited, especially regarding large-scale writing samples. While future work should seek to establish the precise nature of any relationship between emotional language and emotion understanding through development, a critical first step is to capture the emotional content of children's writing, at scale and across age groups. We, therefore, sought to capture trends in the use of emotion-specific language across childhood, focusing on the special form of written language.

1.3. Gender differences in language and emotional development

There is evidence that gender differences in emotional development might be reflected in children's writing. In a small-scale study of 8- to 10-year-old's responses to a topic prompt, Kanaris (1999) found that girls focused more on description and elaboration and used more diverse verbs and adjectives, whereas boys were more egocentric, using more first-person singular pronouns. O'Kearney and Dadds (2004) asked 303 adolescents (12−18 years) to describe in writing how they felt after listening to vignettes designed to promote feelings of anger or fear. Analysis of the writing samples showed an increase in the use of complex emotion terms through adolescence overall, and girls used more inner-directed terms (e.g., sad, embarrassed, guilty, ashamed) than boys.

This difference in emotion word usage in writing is consistent with the claim that boys and girls may experience and express emotions differently. Chaplin and Aldao (2013) conducted a meta-analysis on gender differences in emotional expression from infancy to adolescence, using data from behavioral observations and questionnaires. Girls showed more positive and internalizing emotions, such as sympathy, while boys expressed more negative and externalizing emotions, such as anger. These observations have been related to girls having greater exposure to emotional talk from parents (Shablack & Lindquist, 2019). As children get older, experience with *written* language (i.e., the language content that children read or listen to in the context of shared reading) provides input that is gendered, as shown by the gender stereotypes that are pervasive in natural language, including books written for children (e.g., Charlesworth, Yang, Mann, Kurdi, & Banaji, 2021; Hsiao, Banerji, & Nation, 2021; Lewis, Cooper Borkenhagen, Converse, Lupyan, & Seidenberg, 2022). Relating language experience back to children's own written language production, the subtle and complex linguistic choices made by children through the process of writing may reflect or further refine gender differences in social-emotional development. For example, Hsiao et al. (2021) found that children's own stories reflected the gender biases seen in children's language experience, and this patterned differently for girls and boys.

It is important to note that the existing evidence base on gender differences in emotional expression in children's writing is small; it also comes from the analysis of texts that were elicited by explicit topic prompts or questions whereby adults request an emotional response based on a hypothetical situation. Children's responses may, therefore, be sensitive

to gender-related "display rules" and gender stereotypes (e.g., O'Kearney & Dadds, 2004). Topic prompts also tend to be based on a prototypical example, such as conflicts between close friends, and a broad emotion category, such as anger or fear. This might serve to constrain or direct children's emotion expression in writing. Our investigation was prompted by the need to better understand how emotion is expressed in writing by girls and boys in less primed contexts, using large datasets.

1.4. Large-scale language analysis

Data-driven approaches to analyzing language corpora can reveal patterns in how people use language, and from this, help address psychological questions (Jackson, Watts, List, Puryear, Drabble, & Lindquist, 2022, p. 807). Sentiment analysis seeks to identify attitudes and affect in natural language data (Wankhade, Rao, & Kulkarni, 2022). Relevant to our investigation, Hipson and Mohammad (2020) used a lexicon-based "bag of words" approach to sentiment analysis to study the emotional content of poems written by children in Grades 1 to 12 (6−18 years). This involved first obtaining sentiments of individual words from human judgment data, and then calculating a score for each poem, averaged across its words. They then analyzed sentiments along three emotion dimensions (valence, arousal, dominance) and four discrete emotion categories (anger, fear, sadness, joy). According to Warriner, Kuperman, and Brysbaert (2013), valence refers to the pleasantness of a word and the extent of its positivity or negativity. For example, "holiday" has a positive valence, whereas "prison" is negative. Arousal refers to the intensity of emotion provoked by a word with "calm" being low arousal and "active" high arousal. Dominance refers to the degree of control exerted by a word. This ranges from weak (e.g., "fatigue") to powerful (e.g., "grand"). To capture developmental change, Hipson and Mohammad (2020) modeled these different components across age. They found that poems written by older children showed more negative valence relative to those written by younger children, whereas arousal and dominance increased with age. Poems written by older children also showed increased intensities for all four discrete emotions. They also found that poems written by girls were more positive than poems written by boys. These findings show that sentiment analysis is a useful tool to explore children's writing development. However, poetry is a particular genre and other types of writing might pattern differently.

1.5. The present study

To capture the emotional content of children's writing at scale, we adapted Hipson and Mohammad's methods to track sentiments in a large sample of stories (N>100,000) written by 7- to 13-year-olds. This enabled us to explore two hypotheses. First, in keeping with the view that language is critical in the construction of emotion, we predicted that the children's stories would show decreasing positivity with age, in line with previous research on emotion development of increased frequency of negative emotions and the increased use of negative emotion words (Doost et al., 1999; Larson et al., 2002; Weinstein et al., 2007). To date, methods have probed emotions directly. In contrast, the stories in our corpus were not written in response to a prompt or instruction (see Methods). They, therefore, provide an opportunity to

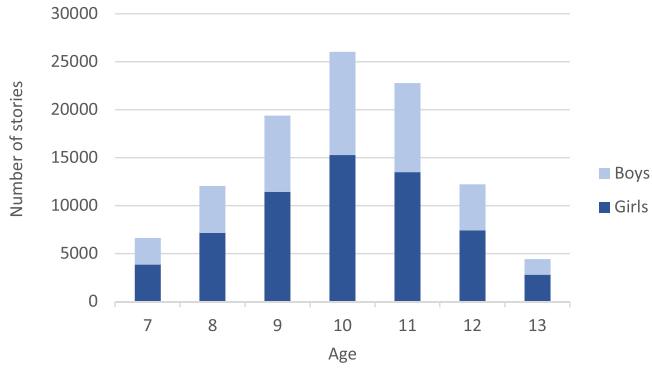


Fig. 1. Number of stories by age and gender.

track emotional content in more naturalistic circumstances. Our second hypothesis concerned gender differences. Following previous work in emotion development reviewed above, we predicted that stories by girls would contain more positive sentiments than those written by boys.

In summary, writing provides a vehicle for children to express emotion, and plausibly, the experience of communicating intent via writing might support the ongoing construction of emotion and language through childhood and adolescence. However, little is known about how and when children use language to communicate emotion in writing. As a first step to address this, our goal was to track sentiments in narrative writing across age and gender, using Hipson and Mohammad's (2020) methods. Note that this type of lexicon-based bag of words approach may not be good at managing the complexity of language in context. A sentence may mean different things and express different sentiments from its constituent words in isolation, more so if negation and other variations of language are considered. For example, after removing stop words such as "not," the sentence "I am happy" has the same emotional valence as "I am not happy"; clearly, this is not an accurate characterization of the sentiment being expressed across the two sentences. With this in mind, we validated our findings using two different sentiment analysis methods that are more robust to concerns about negation.

2. Methods

2.1. The Oxford children's language corpus

We analyzed 103,564 stories written by 7- to 13-year-old children and submitted to a UK children's writing competition in 2019. BBC Radio 2 500 Words was an annual national competition that invited 5- to 13-year-olds to submit a story on any topic; the only constraint was that it should be no longer than 500 words (we excluded entries from children aged 6 and under as the number of entries was relatively low). This resource provided a naturally occurring language sample not generated for assessments or any prompts or cues, allowing us to analyze children's own writing, free from constraints on time, instruction, or topic. Fig. 1

Table 1
Descriptive statistics of stories in different age groups

Age	Number of stories	Mean (SD) Number of emotion words per story
7	6605	118.14 (43.70)
8	12,064	127.47 (37.94)
9	19,372	134.77 (31.90)
10	26,021	138.11 (28.90)
11	22,771	140.20 (27.10)
12	12,234	141.79 (26.34)
13	4425	144.18 (25.73)
Total	103,492	136.12 (31.67)

shows the distribution of stories by age and gender. More girls (59.44%) contributed stories than boys (40.56%), and most stories came from 9- to 11-year-olds.

We first followed Hipson and Mohammad's (2020) procedures. Tokenization of the full corpus ($N_{\text{stories}} = 103{,}564$) resulted in 46,697,930 word tokens (218,914 unique). We excluded 23 stories (three empty files and 20 that contained only one word token; upon inspection, these contained a single letter or a random letter string). All other stories were retained for analysis ($N_{\text{stories}} = 103{,}541$; $M_{\text{words/story}} = 451.01$, $SD_{\text{words/story}} = 92.16$).

We removed stop words (e.g., articles, prepositions) using the in-built stop word list in R (R Core Team, 2022). Emotion associations for all remaining words were extracted from the National Research Council Valence, Arousal, and Dominance (NRC VAD) lexicon v1 (Mohammad, 2018b) and the NRC Emotion Intensity (NRC EI) lexicon v0.5 (Mohammad, 2018a).[1] The NRC VAD lexicon contains about 20,000 commonly used English words that have been scored on valence, arousal, and dominance, with each dimension ranging from 0 to 1. The NRC EI lexicon v0.5 contains around 6000 words from the NRC Emotion Lexicon (EmoLex) (Mohammad & Turney, 2010, 2013). These words are associated with anger, fear, sadness, or joy and each word is rated for the intensity of the associated emotion, ranging from 0 to 1.

At this stage, we excluded another 49 stories as they included fewer than five words from the NRC VAD. This ensured that each story contained enough words for averaging in subsequent analyses. Table 1 shows that the number of emotion words per story increased with age. For each story, we calculated its average valence, arousal, dominance, and emotion intensity scores, based on the words it contained.

2.2. Analyzing developmental trends and gender differences in sentiments

We used *Generalized Additive Models (GAMs)* to analyze the children's sentiments by age as these allowed us to model a smooth relationship between the predictor variable age and the outcome variables without imposing strict parameter values on the relationship (Hastie & Tibshirani, 1990; Wood, 2017). This flexibility for capturing nonlinear trends is very useful for exploratory analysis on a topic that is not well explored. Nonlinear trends in emotions over age were modeled using the mgcv package (Wood, 2017) in R, using Penalized Iterative Least Squares to penalize model fit as smoothing became more complex. This minimizes the

Generalized Cross Validation score (an index of model misfit that increases with respect to least squares and model complexity).

GAMs are additive models and do not allow for multiplicative AxB interaction terms. We investigated gender as a pseudo-interaction term, using the "by" argument: valence ~ s(Age, by = factor(Gender)). This tested for age effects at each level of gender.

2.3. Validating sentiment trends

The first validation approach used the in-built Python package *TextBlob* (Loria, 2018). This is another lexicon-based approach but takes into consideration negation, modifiers, and sentence context when determining the meaning and polarity of individual words. Polarity is the extent of positivity or negativity of the text, ranging from very negative (−1) to very positive (+1). We directly applied the *TextBlob* package without further finetuning.

Our second approach used a machine learning technique, namely, a submodel of *Bidirectional Encoder Representations from Transformers* (*BERT)* (Devlin, Chang, Lee, & Toutanova, 2019). *BERT* considers the context for each occurrence of a given word to generate word embedding representations for each word. For this validation approach, we directly applied the model *"nlptown/bert-base-multilingual-uncased-sentiment"* from Hugging Face, which was pretrained on product reviews and fine-tuned for sentiment analysis. The output of the model was one of five sentiment classes, from negative to positive. To get a more nuanced sentiment score, we used the softmax function to obtain the probability of all five sentiment classes by exponentiating the model's output logits and normalizing them so that they summed up to 1. From the resulting matrix, we assigned a weight to each sentiment class to obtain the resulting sentiment score for each story. The score was a scalar value that ranges from −1 (very negative) to 1 (very positive).

Analysis scripts are available on the Open Science Framework (https://osf.io/ywkrj/).

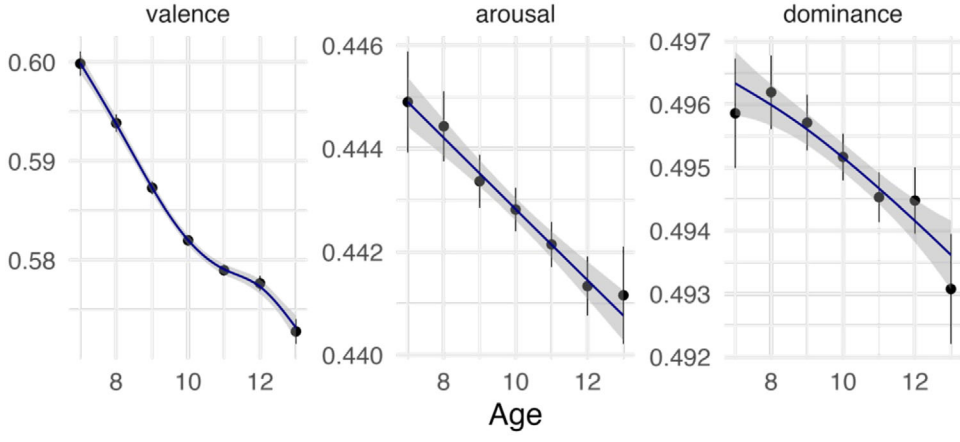
3. Results

3.1. Developmental trends in sentiments

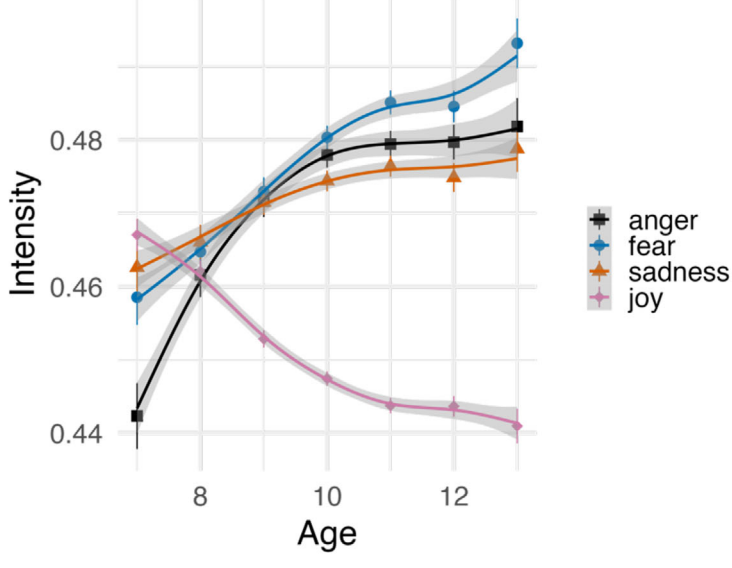
GAMs were run for each of the three dimensions (valence, arousal, dominance), and the four emotion categories (anger, fear, sadness, joy). Age was entered as a predictor in all models. We controlled for the increase in word count across ages (see Table 1). Fig. 2a shows that mean valence, arousal, and dominance scores decreased with age (all ps $< .001$). Fig. 2b shows the intensity of the three negative emotions (anger, fear, sadness) increased with age, whereas the intensity of the positive emotion (joy) decreased (all ps $< .001$). Thus, across emotional dimensions and categories, children's writing showed decreasing positivity with age.

3.2. Gender differences in sentiments

As plotted in Fig. 3, boys and girls showed a decline in valence scores with age ($p < .001$), but stories by girls were consistently higher in valence than those written by boys. The decline in arousal was mainly due to boys, with no significant change in arousal scores across age in



(a) Developmental trends of valence, arousal, and dominance

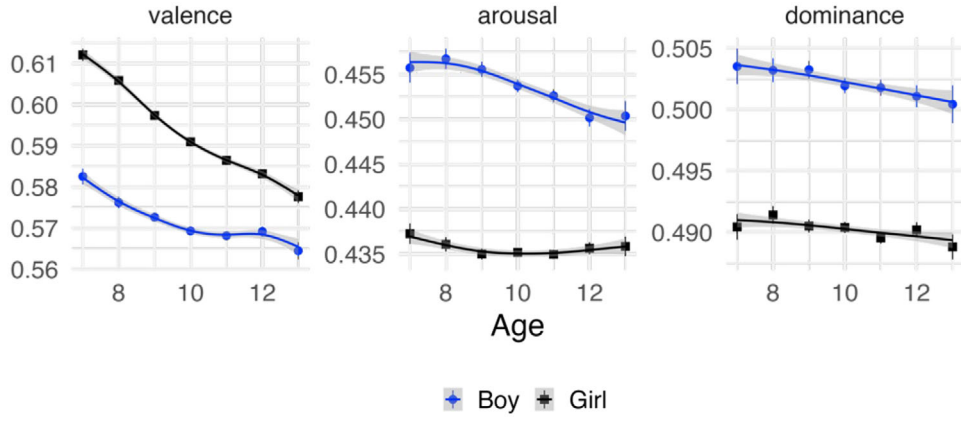


(b) Developmental trends of anger, fear, sadness, and joy

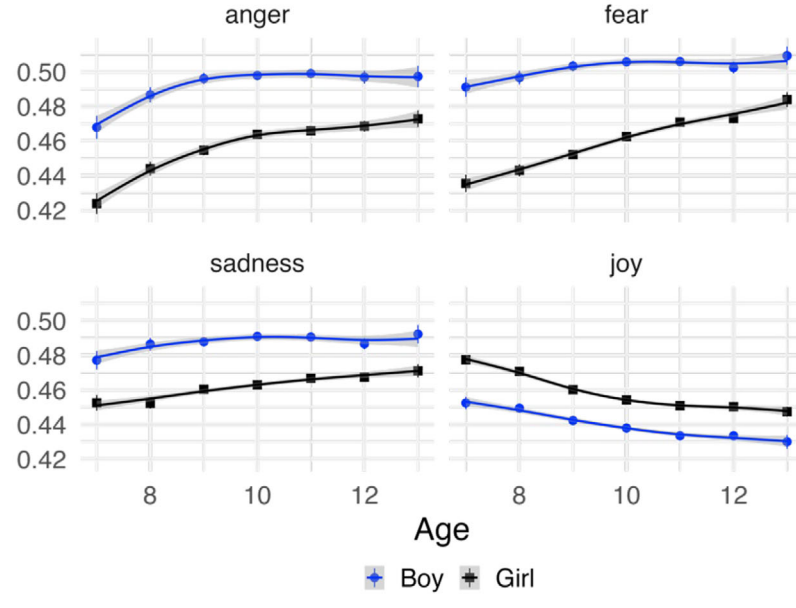
Fig. 2. Trends in (a) valence, arousal, dominance and (b) anger, fear, sadness, and joy intensity by age.

Note. The gray band represents 95% confidence intervals around the smooth fit.

the stories by girls ($F < 1$, $p = .36$). These effects were statistically significant ($p < .001$), with the exception of arousal scores for girls, as noted above. Stories by boys showed higher dominance scores, higher negative emotion intensity, and lower positive emotion intensity than those written by girls. The Supplementary Materials provide examples of usage by boys and girls (https://osf.io/ywkrj/).



(a) Gender difference in valence, arousal, and dominance



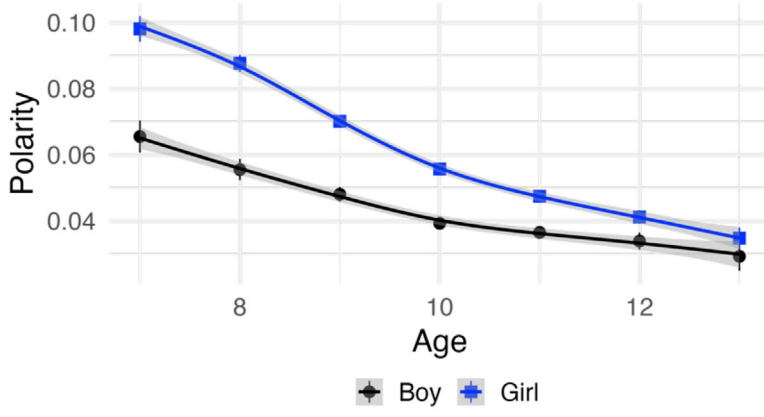
(b) Gender difference of anger, fear, sadness, and joy

Fig. 3. Trends in (a) valence, arousal, dominance and (b) anger, fear, sadness, and joy intensity by age and gender.

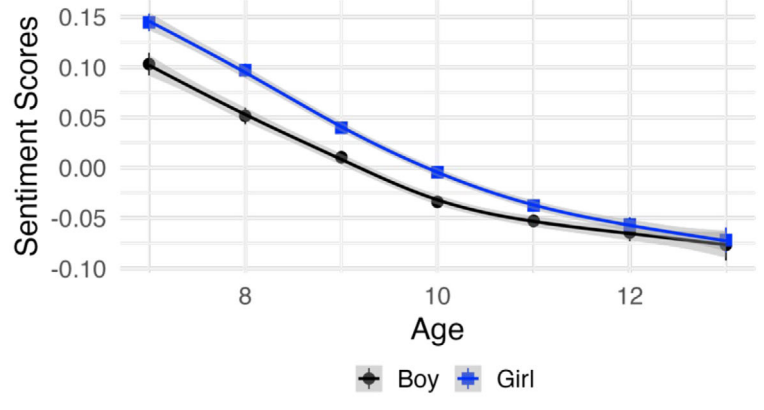
Note. The gray band represents 95% confidence intervals around the smooth fit.

3.3. Validating sentiment trends

Fig. 4a shows the trends in polarity scores of stories generated from the *TextBlob* model. As the age increased, the polarity of stories decreased. Fig. 4b shows that sentiment scores generated from the *BERT* model also declined across age. *GAMs* showed these trends to be



(a) Trends in polarity of stories (*TextBlob*) by age and gender



(b) Trends in sentiment scores of stories (*BERT*) by age and gender

Fig. 4. Trends in (a) polarity scores and (b) sentiment scores in children's story writing.

Note. The gray band represents 95% confidence intervals around the smooth fit.

statistically significant ($p < .001$). For both validation methods, stories by girls had more positive sentiments, especially for the younger age groups.

Finally, we computed correlations between the three measures of valence, polarity (returned by *TextBlob*), and sentiment (returned by *BERT*). As seen in Table 2, the three measures are positively correlated ($p < .001$).

4. Discussion

Although language is considered critical to the development of emotion, the contribution of written language has been largely ignored, and we know very little about how children

Table 2
Correlations between measures of valence, polarity, and sentiment ($N = 103{,}492$)

Measures	Valence	Polarity	Sentiment
Valence	–		
Polarity	0.56	–	
Sentiment	0.28	0.28	–

construct emotion in their own writing. As a first step to understanding this, our study used a data-driven approach to capture the sentiments in children's narrative writing, sampling from a large corpus of stories written by 7- to 13-year-old children and focusing on positivity over time as well as gender differences. We first adapted and extended Hipson and Mohammad's (2020) approach and then replicated our findings using more robust sentiment analysis methods. While the three measures differed in their scoring methods and abilities to handle negation, they converged to show that the sentiments expressed by children became more negative with age. Stories written by girls contained more positive sentiments than those written by boys, especially for the younger age groups.

4.1. Decreasing positive sentiments by age

The decrease in positivity with age mirrors the pattern seen in children's poems (Hipson & Mohammad, 2020). Although broadly similar, there were some differences in the age-related change in emotional content of stories versus poems, such as arousal, dominance, and joy intensity, which all decreased with age in children's stories, but increased with age in children's poems. These disparities might reflect inherent differences across the two genres; note also that stories were generally longer than the poems, and our age range was slightly younger too.

Why might children use more negative words as they get older? One explanation is that children experience more negative words with age. Using age-of-acquisition norms as a reference, Ponari, Norbury, and Vigliocco (2018) estimated that abstract words account for only 10% of vocabulary in 4-year-olds; this rises to over 40% for 12-year-olds. They also identified 8.5 years as a peak period for learning new abstract words, complementing the sharp increase in emotion knowledge seen at that time (Baron-Cohen et al., 2010). As child-directed speech contains many more positive than negative words (Ponari et al., 2018), the explosion of emotion vocabulary—and especially negative language—in mid-childhood is a likely consequence of accumulated experience with book language (Nation et al., 2022). As independent reading takes off, so too does the opportunity to learn new emotion words and use them in writing. The relative decrease in positivity with age also tracks patterns in language evolution. In English, for example, there are more negative words than positive words, and across different languages, Jackson, Lindquist, Drabble, Atkinson, and Watts (2023) found that negative valence correlates with faster cognate replacement over time. This results in a greater variety of negative emotion words being added to the lexicon and subsequently being available for children to use, and to shape ongoing emotional development,

in line with language being a key ingredient in the construction of emotion (Barrett, 2017; Hoemann et al., 2019; Lindquist, 2017).

Another interpretation of the increased negativity with age is that children may choose to write about darker and more complex topics as they grow older, with the increasing understanding that negative information draws more attention (Bahn, Vesker, García Alanis, Schwarzer, & Kauschke, 2017; Jackson et al., 2023; Pratto & John, 1991). These darker topics may invite the use of more negative words. Content analysis and topic modeling (e.g., Griffiths, Steyvers, & Tenenbaum, 2007) could be used to explore the nature of the topics children write about in future studies.

Alongside increased exposure to written language via reading, we speculate that learning to write and the productive act of writing provides powerful opportunities for children to construct their emotional worlds. Writing requires children to construct a language representation of the world they wish to convey into the mind of the reader. Language production shapes comprehension more generally (McDonald, 2013), and there is some evidence that writing drives the development of reading comprehension (Zagata et al., 2023). Emotion production may drive emotion comprehension, and vice versa. While children's stories provide a window to understand how they represent and describe the emotions of characters, it is important to note that the fictional world they create may not be representative of their own emotions. Although the decreasing positive sentiments in children's writing run parallel with their decreasing self-reported moods across development (e.g., Larson et al., 2002), our study is not able to establish a direct link between the two. It would be interesting to directly compare the linguistic content of children's writing when they are composing a story versus reflecting on their own feelings and experiences, especially given the interest in writing as a method to identify mental health concerns in adolescence (Cohen et al., 2022; Shearer et al., 2021). Experimental work is needed to establish whether there are direct relationships between children's experience with written language, their writing, and their own emotions.

4.2. Gender difference in sentiments in children's writing

In line with our prediction, stories by girls were more positive than those by boys, especially for the younger age group. This extends previous observations from elicited tasks and self-reports to narrative writing. Gender differences have been associated with differences in child-directed speech, with caregivers using more emotional language with girls (for review, Shablack & Lindquist, 2019). Adding to this, book language itself is also associated with systematic gender differences (Lewis & Lupyan, 2020). For example, Lewis et al. (2022) measured word-gender association in a children's book corpus and asked adults to rate the femininity or masculinity of words. They found that words rated as more feminine or more masculine were also more likely to be associated with characters of the respective gender. There is also evidence of androcentrism in children's books, and of this being reflected in children's own writing (Hsiao et al., 2021). Gender differences may be perpetuated in that girls are more likely to read books featuring girl characters and girl-related content, and vice versa for boys (Charlesworth et al., 2021; Lewis et al., 2022), and thus their own word choices

reflect the language they experience in stories they read. We suggest that children's writing may serve as a valuable resource for understanding emotional development in boys and girls (Chaplin & Aldao, 2013), and provide a foundation for research that seeks to identify potential young people's mental health concerns from analysis of their writing (Cohen et al., 2022; Shearer et al., 2021).

4.3. Limitations and future directions

While converging evidence demonstrated a decrease in positive sentiments in children's writing, some limitations need to be acknowledged. Using a naturally occurring dataset allowed us to examine real-world language use at scale, but we had no control over how the dataset was constructed, and metainformation is limited. For instance, the cross-sectional datasets obtained from a national competition might not fully represent the population. We note that the competition was open to all and was advertised extensively on local and national media and via schools. Entries were received from all parts of the UK, spanning rural and urban neighborhoods. Unfortunately, however, we were unable to access detailed demographic information for all entries. We note that using pretrained models without fine-tuning to children's stories may not fully capture the actual sentiment (Singh, Singh, & Singh, 2017), although Werlen, Imhof, Benites, and Bergamin (2019) argued that applying machine learning algorithms to different contexts was still effective. To validate the observed developmental trends in sentiments, future studies should employ longitudinal datasets and incorporate some hand-coded data for fine-tuning the language model. Note that we did not calculate effect size as it is not clear how this should be done or interpreted when using GAMs. Given the potential issue of overpowering given corpus size (Egbert et al., 2022), however, it is likely that our effect sizes are small and this is perhaps not surprising given the task was open-ended, without topic prompts or specific instructions. More generally, there is a pressing need for experimental work on the connections between children's writing and emotional development. Such work will enable causal inference, and help us understand the theoretical and practical significance of small effects; this is especially important if linguistic analysis is to be applied to mental health and clinical concerns (e.g., Cohen et al., 2022).

Many questions remain regarding the interplay between language and emotion and how this unfolds over development (e.g., Lindquist, 2021). We propose that experience with written language provides children with access to particular forms of emotional language that is then reflected in how children shape their own writing to communicate emotional content. Our study demonstrates the utility of large-scale data-driven approaches, but our findings need to be complemented by future work taking a more experimental approach. These will reveal whether and how experience with written language interacts with the acquisition of emotional knowledge and vice versa, and how this becomes reflected in children's own writing.

Acknowledgments

This research was funded by a British Academy Post-Doctoral Fellowship (PF2/180013) awarded to Yaling Hsiao, a grant from the Nuffield Foundation (EDO/43392) to Kate Nation,

and resources made available to Nilanjana Banerji by the Department of Children's Dictionaries and Children's Language Data at Oxford University Press. Data and code associated with this paper are available on the Open Science Framework website (https://osf.io/ywkrj/). The Oxford Children's Language Corpus is a growing database of writing for and by children developed and maintained by Oxford University Press for the purpose of children's language research.

Open Research Badges

This article has earned the Open Materials badge. Materials are available at https://osf.io/ywkrj/.

Note

1 Details of how the two lexicons were constructed can be found on this website: https://saifmohammad.com/WebPages/nrc-vad.html

References

Bahn, D., Vesker, M., García Alanis, J. C., Schwarzer, G., & Kauschke, C. (2017). Age-dependent positivity-bias in children's processing of emotion terms. *Frontiers in Psychology*, *8*, 1268. Retrieved from https://www.frontiersin.org/articles/10.3389/fpsyg.2017.01268

Bailen, N. H., Green, L. M., & Thompson, R. J. (2019). Understanding emotion in adolescents: A review of emotional frequency, intensity, instability, and clarity. *Emotion Review*, *11*(1), 63–73. https://doi.org/10.1177/1754073918768878

Baron-Cohen, S., Golan, O., Wheelwright, S., & Granader, Y. (2010). Emotion word comprehension from 4 to 16 years old: A developmental survey. *Frontiers in Evolutionary Neuroscience*, *2*, 109. Retrieved from https://www.frontiersin.org/articles/10.3389/fnevo.2010.00109

Barrett, L. F. (2017). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience*, *12*(1), 1–23. https://doi.org/10.1093/scan/nsw154

Beck, L., Kumschick, I. R., Eid, M., & Klann-Delius, G. (2012). Relationship between language competence and emotional competence in middle childhood. *Emotion*, *12*(3), 503–514. https://doi.org/10.1037/a0026320

Boerma, I. E., Mol, S. E., & Jolles, J. (2017). The role of home literacy environment, mentalizing, expressive verbal ability, and print exposure in third and fourth graders' reading comprehension. *Scientific Studies of Reading*, *21*(3), 179–193. https://doi.org/10.1080/10888438.2016.1277727

Chaplin, T. M., & Aldao, A. (2013). Gender differences in emotion expression in children: A meta-analytic review. *Psychological Bulletin*, *139*(4), 735–765. https://doi.org/10.1037/a0030737

Charlesworth, T. E. S., Yang, V., Mann, T. C., Kurdi, B., & Banaji, M. R. (2021). Gender stereotypes in natural language: Word embeddings show robust consistency across child and adult language corpora of more than 65 million words. *Psychological Science*, *32*(2), 218–240. https://doi.org/10.1177/0956797620963619

Cohen, K. A., Shroff, A., Nook, E. C., & Schleider, J. L. (2022). Linguistic distancing predicts response to a digital single-session intervention for adolescent depression. *Behaviour Research and Therapy*, *159*, 104220. https://doi.org/10.1016/j.brat.2022.104220

Dawson, N., Hsiao, Y., Tan, A. W. M., Banerji, N., & Nation, K. (2021). Features of lexical richness in children's books: Comparisons with child-directed speech. *Language Development Research*, *1*(1), Article 1. https://doi.org/10.34842/5we1-yk94

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding*. https://doi.org/10.48550/arXiv.1810.04805

Doost, H. T. N., Moradi, A. R., Taghavi, M. R., Yule, W., & Dalgleish, T. (1999). The development of a corpus of emotional words produced by children and adolescents. *Personality and Individual Differences*, *27*, 433–451. https://doi.org/10.1016/S0191-8869(98)00253-0

Doyle, C. M., & Lindquist, K. A. (2018). When a word is worth a thousand pictures: Language shapes perceptual memory for emotion. *Journal of Experimental Psychology: General*, *147*(1), 62–73. https://doi.org/10.1037/xge0000361

Dunn, J., Brown, J., & Beardsall, L. (1991). Family talk about feeling states and children's later understanding of others' emotions. *Developmental Psychology*, *27*(3), 448–455. https://doi.org/10.1037/0012-1649.27.3.448

Egbert, J., Biber, D., & Gray, B. (Eds.). (2022). Distribution considerations. In *Designing and evaluating language corpora: A practical framework for corpus representativeness* (pp. 122–155). Cambridge University Press. https://doi.org/10.1017/9781316584880.005

Graham, S. (2006). Writing. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (pp. 457–478). Lawrence Erlbaum Associates Publishers.

Graham, S. (2018). Introduction to conceptualizing writing. *Educational Psychologist*, *53*(4), 217–219. https://doi.org/10.1080/00461520.2018.1514303

Griffiths, S., Goh, S. K. Y., & Norbury, C. F. (2020). Early language competence, but not general cognitive ability, predicts children's recognition of emotion from facial and vocal cues. *PeerJ*, *8*, e9118. https://doi.org/10.7717/peerj.9118

Griffiths, T. L., Steyvers, M., & Tenenbaum, J. B. (2007). Topics in semantic representation. *Psychological Review*, *114*, 211–244. https://doi.org/10.1037/0033-295X.114.2.211

Grosse, G., Streubel, B., Gunzenhauser, C., & Saalbach, H. (2021). Let's talk about emotions: The development of children's emotion vocabulary from 4 to 11 years of age. *Affective Science*, *2*(2), 150–162. https://doi.org/10.1007/s42761-021-00040-2

Harris, K. R., Graham, S., Brindle, M., & Sandmel, K. (2009). Metacognition and children's writing. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 131–153). Routledge/Taylor & Francis Group.

Hastie, T., & Tibshirani, R. (1990). Exploring the nature of covariate effects in the proportional hazards model. *Biometrics*, *46*(4), 1005–1016. https://doi.org/10.2307/2532444

Hipson, W., & Mohammad, S. M. (2020). PoKi: A large dataset of poems by children. In *Proceedings of the 12th Language Resources and Evaluation Conference* (pp. 1578–1589). Retrieved from https://aclanthology.org/2020.lrec-1.196

Hoemann, K., Xu, F., & Barrett, L. F. (2019). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental Psychology*, *55*(9), 1830–1849. https://doi.org/10.1037/dev0000686

Hogan, P. C. (2011). *What literature teaches us about emotion*. Cambridge University Press. https://doi.org/10.1017/CBO9780511976773

Hsiao, Y., Banerji, N., & Nation, K. (2021). Boys write about boys: Androcentrism in children's reading experience and its emergence in children's own writing. *Child Development*, *92*(6), 2194–2204. https://doi.org/10.1111/cdev.13623

Jackson, J. C., Lindquist, K., Drabble, R., Atkinson, Q., & Watts, J. (2023). Valence-dependent mutation in lexical evolution. *Nature Human Behaviour*, *7*(2), Article 2. https://doi.org/10.1038/s41562-022-01483-8

Jackson, J. C., Watts, J., List, J.-M., Puryear, C., Drabble, R., & Lindquist, K. A. (2022). From text to thought: How analyzing language can advance psychological science. *Perspectives on Psychological Science*, *17*(3), 805–826. https://doi.org/10.1177/17456916211004899

Kanaris, A. (1999). Gendered journeys: Children's writing and the construction of gender. *Language and Education*, *13*(4), 254–268. https://doi.org/10.1080/09500789908666772

Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, *342*(6156), 377–380. https://doi.org/10.1126/science.1239918

Kim, Y.-S. G., Park, C., & Park, Y. (2015). Dimensions of discourse level oral language skills and their relation to reading comprehension and written composition: An exploratory study. *Reading and Writing*, *28*(5), 633–654. https://doi.org/10.1007/s11145-015-9542-7

Kucirkova, N. (2019). How could children's storybooks promote empathy? A conceptual framework based on developmental psychology and literary theory. *Frontiers in Psychology*, *10*, 121. https://doi.org/10.3389/fpsyg.2019.00121

Larson, R. W., Moneta, G., Richards, M. H., & Wilson, S. (2002). Continuity, stability, and change in daily emotional experience across adolescence. *Child Development*, *73*(4), 1151–1165. https://doi.org/10.1111/1467-8624.00464

Lewis, M., Cooper Borkenhagen, M., Converse, E., Lupyan, G., & Seidenberg, M. S. (2022). What might books be teaching young children about gender? *Psychological Science*, *33*(1), 33–47. https://doi.org/10.1177/09567976211024643

Lewis, M., & Lupyan, G. (2020). Gender stereotypes are reflected in the distributional structure of 25 languages. *Nature Human Behaviour*, *4*(10), 1021–1028. https://doi.org/10.1038/s41562-020-0918-6

Lindquist, K. A. (2017). The role of language in emotion: Existing evidence and future directions. *Current Opinion in Psychology*, *17*, 135–139. https://doi.org/10.1016/j.copsyc.2017.07.006

Lindquist, K. A. (2021). Language and emotion: Introduction to the special issue. *Affective Science*, *2*(2), 91–98. https://doi.org/10.1007/s42761-021-00049-7

Loria, S. (2018). *Textblob Documentation*. 73.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*, 226. https://doi.org/10.3389/fpsyg.2013.00226

Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, *3*(3), 173–192. https://doi.org/10.1111/j.1745-6924.2008.00073.x

Mar, R. A., Oatley, K., Hirsh, J., dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality*, *40*(5), 694–712. https://doi.org/10.1016/j.jrp.2005.08.002

McCutchen, D. (2006). Cognitive factors in the development of children's writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.) *Handbook of writing research* (pp. 115–130). The Guilford Press.

Mohammad, S. (2018a). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 174–184). https://doi.org/10.18653/v1/P18-1017

Mohammad, S. (2018b). Word affect intensities. In *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018)*. LREC 2018, Miyazaki, Japan. Retrieved from https://aclanthology.org/L18-1027

Mohammad, S. M., & Turney, P. D. (2010). Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text* (pp. 26–34). Retrieved from https://aclanthology.org/W10-0204

Mohammad, S. M., & Turney, P. D. (2013). *Crowdsourcing a word-emotion association lexicon*. https://doi.org/10.48550/arXiv.1308.6297

Nation, K., Dawson, N. J., & Hsiao, Y. (2022). Book language and its implications for children's language, literacy, and development. *Current Directions in Psychological Science*, *31*(4), 375–380. https://doi.org/10.1177/09637214221103264

Nook, E. C. (2023). The promise of affective language for identifying and intervening on psychopathology. *Affective Science*, *4*(3), 517–521. https://doi.org/10.1007/s42761-023-00199-w

Nook, E. C., Hull, T. D., Nock, M. K., & Somerville, L. H. (2022). Linguistic measures of psychological distance track symptom levels and treatment outcomes in a large set of psychotherapy transcripts. *Proceedings of the National Academy of Sciences*, *119*(13), e2114737119. https://doi.org/10.1073/pnas.2114737119

Nook, E. C., Stavish, C. M., Sasse, S. F., Lambert, H. K., Mair, P., McLaughlin, K. A., & Somerville, L. H. (2020). Charting the development of emotion comprehension and abstraction from childhood to adulthood using observer-rated and linguistic measures. *Emotion*, *20*(5), 773–792. https://doi.org/10.1037/emo0000609

O'Kearney, R., & Dadds, M. (2004). Developmental and gender differences in the language for emotions across the adolescent years. *Cognition and Emotion*, *18*, 913–938. https://doi.org/10.1080/02699930341000356

Ponari, M., Norbury, C. F., & Vigliocco, G. (2018). Acquisition of abstract concepts is influenced by emotional valence. *Developmental Science*, *21*(2), e12549. https://doi.org/10.1111/desc.12549

Pratto, F., & John, O. P. (1991). Automatic vigilance: The attention-grabbing power of negative social information. *Journal of Personality and Social Psychology*, *61*(3), 380–391. https://doi.org/10.1037//0022-3514.61.3.380

R Core Team. (2022). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Schwering, S. C., Ghaffari-Nikou, N. M., Zhao, F., Niedenthal, P. M., & MacDonald, M. C. (2021). Exploring the relationship between fiction reading and emotion recognition. *Affective Science*, *2*(2), 178–186. https://doi.org/10.1007/s42761-021-00034-0

Shablack, H., & Lindquist, K. A. (2019). The role of language in emotional development. In V. LoBue, K. Pérez-Edgar, & K. A. Buss (Eds.), *Handbook of emotional development* (pp. 451–478). Springer Nature Switzerland AG. https://doi.org/10.1007/978-3-030-17332-6_18

Shearer, N. J., Gillespie, A. N., Olds, T. S., Mensah, F. K., Edwards, B., Fernando, J. W., Wang, Y., Wake, M., & Lycett, K. (2021). Insights into adolescent well-being from computerised analysis of written language. *Acta Paediatrica*, *110*(6), 1880–1889. https://doi.org/10.1111/apa.15813

Siller, M., Swanson, M. R., Serlin, G., & George, A. (2014). Internal state language in the storybook narratives of children with and without autism spectrum disorder: Investigating relations to theory of mind abilities. *Research in Autism Spectrum Disorders*, *8*(5), 589–596. https://doi.org/10.1016/j.rasd.2014.02.002

Singh, J., Singh, G., & Singh, R. (2017). Optimization of sentiment analysis using machine learning classifiers. *Human-Centric Computing and Information Sciences*, *7*(1), 32. https://doi.org/10.1186/s13673-017-0116-3

Thompson, R. A. (1991). Emotional regulation and emotional development. *Educational Psychology Review*, *3*, 269–307. https://doi.org/10.1007/BF01319934

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(7), 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Warriner, A. B., Kuperman, V., & Brysbaert, M. (2013). Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behavior Research Methods*, *45*(4), 1191–1207. https://doi.org/10.3758/s13428-012-0314-x

Weinstein, S. M., Mermelstein, R. J., Hankin, B. L., Hedeker, D., & Flay, B. R. (2007). Longitudinal patterns of daily affect and global mood during adolescence. *Journal of Research on Adolescence*, *17*(3), 587–600. https://doi.org/10.1111/j.1532-7795.2007.00536.x

Werlen, E., Imhof, C., Benites, F., & Bergamin, P. B. (2019). *The reader's feeling and text-based emotions: The relationship between subjective self-reports, lexical ratings, and sentiment analysis*. SwissText 2019, Winterthur, 18–19 June 2019. https://doi.org/10.21256/zhaw-18859

White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the strange stories: Revealing mentalizing impairments in autism. *Child Development*, *80*(4), 1097–1117. https://doi.org/10.1111/j.1467-8624.2009.01319.x

Widen, S. C., & Russell, J. A. (2003). A closer look at preschoolers' freely produced labels for facial expressions. *Developmental Psychology*, *39*, 114–128. https://doi.org/10.1037/0012-1649.39.1.114

Widen, S. C., & Russell, J. A. (2008). Children acquire emotion categories gradually. *Cognitive Development*, *23*, 291–312. https://doi.org/10.1016/j.cogdev.2008.01.002

Wood, S. N. (2017). *Generalized additive models: An introduction with R* (2nd edition). CRC Press/Taylor & Francis Group.

Zagata, E., Kearns, D., Truckenmiller, A. J., & Zhao, Z. (2023). Using the features of written compositions to understand reading comprehension. *Reading Research Quarterly*, *58*(4), 624–654. https://doi.org/10.1002/rrq.503

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. *Psychological Science*, *6*(5), 292–297. https://doi.org/10.1111/j.1467-9280.1995.tb00513.x

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supplementary Materials
Supplementary Materials
Supplementary Materials