(How) Do We Teach Emotions?*

Anjali Adukia Matthew Bonci Paula Dastres
Emileigh Harrison Jake Nicoll Teodora Szasz

May 22, 2026
Active Working Draft

Abstract

Emotional intelligence is a key component of human capital, shaped in part by the educational materials children consume. In this study, we apply tools from machine learning to examine the modeling of emotions in public school textbooks and award-winning children's literature. Our analysis reveals a stark and robust mismatch: while textual content exposes children to a diverse emotional landscape, images depict almost exclusively happiness and calm, and "negatively" valenced emotions rarely appear. This pattern persists across time, genre categories, and demographic subgroups. We present empirical evidence that smiling faces attract purchasers, possibly at the expense of emotional diversity in images.

Keywords: Culture, Emotions, Curriculum Analysis, Education Policy

JEL Category: I20, I21, I28, J15, Z13

*Affiliation: University of Chicago, University of Wisconsin-Madison, University of Texas at Austin. For helpful feedback, we thank Alex Eble, James Evans, Yuan Chang Leong, Lesley Turner, Isabeau Dasho, and seminar participants at Notre Dame, Zurich AI + Economics, Hong Kong AI Research in the Social Sciences and Humanities, Association for Education Finance and Policy, the Society for Research on Educational Effectiveness, and The University of Chicago. For technical and research assistance, we thank Ping-Jung Liu. For financial support, we thank the Becker Friedman Institute at UChicago. The research reported here was also supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A200478 to the University of Chicago. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education. Researcher(s)' own analyses calculated (or derived) based in part on data from Market Track, LLC dba Numerator and marketing databases provided through the Numerator Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Numerator data are those of the researcher(s) and do not reflect the views of Numerator. Numerator is not responsible for and had no role in analyzing or preparing the results reported herein.

Introduction

The centrality of emotion to human experience has long been recognized across cultures, traditions, and disciplines (Izard, 1977; Lutz and White, 1986; Elster, 1998; Marcus, 2000; Gross, 2015; Denham, 2019). Beyond their psychological importance, emotional skills constitute a key component of human capital. Individuals with greater emotional intelligence exhibit stronger stress management, prosocial behavior, academic performance, and workplace functioning (Mayer and Geher, 1996; Parker et al., 2004; Austin, Saklofske and Mastoras, 2010; Brackett et al., 2012). These skills are likely to be especially valuable in modern labor markets; as work has become more team-based and service-oriented, employers increasingly reward workers who can communicate effectively, coordinate with others, and manage interpersonal frictions (Hochschild, 1983; Deming, 2017; Bonhomme, 2021; Hansen et al., 2021; Hoffman and Tadelis, 2021). Consequently, the return to social-emotional skills has risen even as the return to some cognitive skills has stagnated (Castex and Kogan Dechter, 2014; Edin et al., 2022).

Schools play a crucial role in teaching emotional skills through direct instruction, classroom interactions, teacher expectations, and curricular materials. Programs that explicitly teach emotional intelligence can improve student behavior and academic outcomes (Brackett et al., 2012). Children's media, while not explicitly designed for instruction, is another important site of learning. Exposure to emotionally rich storybooks and educational television programs have been shown to improve children's emotional vocabulary, perspective-taking skills, empathy, emotion recognition, and coping skills (Kumschick et al., 2014; Rasmussen et al., 2016; Foulds, 2023). Because these emotional skills are relevant for later educational and labor market success, the way children's materials present emotions is critical.

In this study, we examine how emotions are presented in children's educational materials across formal and informal settings: (i) three decades of public school textbooks from California and Texas and (ii) a century of award-winning children's books commonly found in homes, libraries, and schools (Smith, 2013; Cockcroft, 2018; Adukia et al., 2023). Using these two sets of influential children's media, we analyze how emotions are presented across modalities, settings, characters, and over time. We also incorporate data on both consumer and public library purchasing behavior to assess how preferences for particular emotional portrayals influence book acquisition decisions, which may in turn affect the supply of representation of emotions in children's media.

To extract data on emotions from books, we apply artificial intelligence methods from natural language processing (applied to text) and computer vision (applied to images).

Specifically, we employ large multimodal models (LMMs) to classify eight primary emotions across both text and images: anger, calm, confusion, disgust, fear, happiness, sadness, and surprise.[1] To assess classification robustness, we compare outputs from our primary models with alternative approaches, including a lexicon-based emotion classification method applied to text, a RoBERTa model fine-tuned on an emotion-specific text dataset, and a pre-trained convolutional neural network (CNN) used for feature-based emotion classification in images. We examine heterogeneity by context (e.g. corpora, state, grade, subject) and character identities (e.g. putative race, skin color, gender, and age).

Our analysis reveals a systematic divergence between textual and visual emotional content: written text exposes children to a broad emotional range including happiness, sadness, anger, and fear. By contrast, images almost exclusively depict happiness and calm while rarely representing emotions traditionally categorized as "negative." These patterns persist across time periods, contexts, and demographic groups both in text and in images. Moreover, we find a striking misalignment between the emotions portrayed in the text and images on the same page. Regardless of which emotions characters display in the text, images disproportionately depict characters as calm or happy. This demonstrates that the discrepancy between emotional representation in text and images does not simply reflect choices by illustrators to highlight certain positive components of their narratives; rather, the disconnect reflects a genuine incongruence between the emotions that are written about and the emotions that are shown. Almost uniformly, these books suggest that regardless of the emotions characters experience in the text, the appropriate outward response is to display a calm or happy appearance.

We next examine whether consumer demand helps explain the visual overrepresentation of calm and happiness in children's books. By merging information about the emotions in children's books with purchaser decisions – both of households and public libraries – we find that, controlling for book attributes, purchasers are less likely to choose children's books whose covers depict emotions traditionally considered "negative." This suggests that the lack of a range of emotions in images might be driven by consumer demand and reflects the societal value placed on displays of "positive" emotions. As a result, these emotions capture attention and make books look more appealing (Bloch, 1995; Underwood and Klein, 2002; Pieters and Wedel, 2004). Additionally, if consumers have limited bandwidth to accu-

[1]This set of emotions was derived from the six basic emotions identified by Ekman, Sorenson and Friesen (1969) with two added emotions: calm and confusion, particularly relevant to children's media contexts. Given that we are predicting these emotions in the text and images of children's media, we also allow for the model to predict "unsure" or "no emotion" to allow for cases that are not obvious or cases where a character is detected but their emotion is not visible.

rately assess the quality of emotional representation in books and the messages they convey, then authors and illustrators could be incentivized to oversupply happy and calm images in equilibrium even relative to the underlying preference for these emotions (Ellison, 2006; Cusumano, Fabbri and Pieroth, 2024).

Our findings contribute to multiple strands of literature. First, we expand prior research on emotional representation in children's books (Nikolajeva, 2014; Ku, 2019; Kardum, Dadic and Horvat, 2021), which has largely relied on manual content analysis and focused narrowly on text. Second, we add to the literature on cultural economics, which has recently rapidly expanded through the use of novel data sources to study naming conventions (Bazzi, Fiszbein and Gebresilasse, 2020), folklore (Michalopoulos and Xue, 2021), movies (Michalopoulos and Rauh, 2024), style (Yanagizawa-Drott and Voth, 2023), and now emotions. Third, we inform literature on non-cognitive skill development (Groves, 2005; Heckman, Stixrud and Urzua, 2006; Duckworth et al., 2007; Lundberg, 2017; Kautz et al., 2014), highlighting how curricular materials can shape early emotional intelligence. Fourth, we illustrate how books can effectively advertise themselves to consumers with limited attention to a book's content (Dröge and Darmon, 1987; Eliaz and Spiegler, 2011), in this case, through the emphasis of certain emotions. Finally, we contribute to computational social science by developing a multimodal approach to measuring emotional representation in children's media. Building on work that uses natural language processing and computer vision to study social representation, stereotypes, and cultural content (Ash et al., 2021; Adukia et al., 2023; Ludwig and Mullainathan, 2024; Adukia and Harrison, 2025), we examine the emotional content in the text and images of children's media: public school textbooks from the last three decades and influential books from a century of children's literature.

The paper proceeds as follows: Section I discusses background related to the study and teaching of emotions. Section II details our data sources. Section III explains the processes used to transform the images and text into data, along with the tools employed for analysis. The main results are presented in Section IV. In Section IV.D, we further probe the apparent contrast between emotions in text and images by considering both modalities on the same page simultaneously. In Section V.A, we demonstrate robustness to alternative approaches to measurement, while Section V.B discusses our model validation strategy, which we hope offers guidance for others using LMMs for measurement in economics and in other disciplines. In Section VI, we explore how consumer demand may contribute to the patterns we uncover. Finally, we offer a discussion in Section VII.

I Background

Emotions constitute a fundamental dimension of human experience. In a smartphone-based survey that prompted participants to report their emotions multiple times per day, Trampe, Quoidbach and Taquet (2015) found that individuals commonly reported experiencing a range of emotions — including joy, anxiety, love, contentment, sadness, disgust, and anger — all within a single day. This importance of emotion has long been recognized across cultures and traditions, making it a natural subject of sustained inquiry across a wide range of disciplines. In psychology, foundational work has examined the nature, classification, and development of emotions (Izard, 1977; Plutchik and Kellerman, 1980; Averill, 1983), while more recent research has emphasized emotion regulation, emotional competence, and the role of early socialization in shaping how children learn to recognize, manage, and express emotions (Gross, 2015; Eisenberg, Cumberland and Spinrad, 1998; Denham, 2019). Sociology has highlighted how social roles, status relations, institutions, and labor markets shape emotional expression, including norms regarding which emotions are appropriate to display in different contexts (Kemper, 1978; Hochschild, 1983; Thoits, 1989). Anthropological research has emphasized that emotional experiences and displays reflect culturally shared meanings and expectations, including norms governing appropriate emotional responses across social settings (Lutz and White, 1986). Political science has examined the role of emotions in political judgment, participation, and collective behavior (Marcus, 2000). Philosophy has long debated the nature, moral significance, and rationality of emotions (Solomon, 1976; Rorty et al., 1980), while economics has examined how emotions influence preferences, decision-making and social interactions (Elster, 1998; Loewenstein, 2000).

Given the importance of emotion to human experience, emotional skills are an important component of human capital. A large literature shows that non-cognitive skills shape educational attainment, employment, and earnings (Heckman and Rubinstein, 2001; Groves, 2005; Heckman, Stixrud and Urzua, 2006). Emotional intelligence—the ability to recognize, understand, and regulate emotions in oneself and others—is one such skill. Individuals with greater emotional intelligence exhibit stronger stress management, prosocial behavior, academic performance, and workplace functioning (Mayer and Geher, 1996; Parker et al., 2004; Austin, Saklofske and Mastoras, 2010; Brackett et al., 2012). The economic value of emotional skills is likely to be especially high in modern labor markets. As work has become more team-based and service-oriented, employers increasingly reward workers who can communicate effectively, coordinate with others, and manage interpersonal frictions (Hochschild, 1983; Deming, 2017; Bonhomme, 2021). Furthermore, manipulating emotional displays to exhibit happiness and calmness is an essential feature of what customer service workers provide (Hochschild, 1983). Emotional intelligence is central to these tasks: regulating your own emotions and recognizing emotions in others can reduce frictions in teams and contribute to managerial effectiveness, which in turn improves earnings (Hansen et al., 2021; Hoffman and Tadelis, 2021). Consequently, the value of these skills is growing even as the return to cognitive skills stagnates (Castex and Kogan Dechter, 2014; Edin et al., 2022).

Schools are a central institution through which children acquire these skills. In particular, these institutions foster emotional intelligence both through direct instruction (Brackett et al., 2012), classroom interactions, teacher expectations, and education materials. Children's books are another important site of rich emotional language (Green and Sun, 2025). According to the constructionist hypothesis in developmental psychology, emotions are abstract concepts learned by children through an accumulation of encounters with emotions (Hoemann, Xu and Barrett, 2019). From this perspective, the images in children's books can provide a model for the rich accompanying emotion language that can help children calibrate both their emotion recognition skills and their understanding of how to express emotions during precisely the age range that these skills are developing. Indeed, children rapidly learn to recognize emotions between the ages of two and eight, although the skill is still malleable after the age of eight (Widen and Russell, 2008; Ruffman et al., 2023). This calibration could improve emotion recognition skills, which have been shown to translate to improved social relationships (Brackett et al., 2012). Consistent with this hypothesis, exposure to emotionally rich storybooks and educational television programs improves children's emotional vocabulary, perspective-taking skills, empathy, emotion recognition, and coping skills (Kumschick et al., 2014; Rasmussen et al., 2016; Foulds, 2023).

Additionally, research demonstrates that visual depictions of emotions and emotional cues enhance children's emotion identification and response capabilities relative to text-based presentations alone (Arizpe and Styles, 2015; Nikolajeva, 2017). This ability to identify emotions has important social consequences: children who struggle to identify emotions like sadness tend to exhibit lower prosocial behavior (Dickerson and Quas, 2021). Because children often lack firsthand experience with many emotions and may not yet have fully developed the capacity to understand others' feelings (Doherty, 2008), both formal and informal educational materials play a crucial role in this developmental process, particularly when they represent a broad and realistic distribution of emotional experiences.

Because emotional norms are relevant for later educational and labor market success, the way educational materials represent emotions is potentially important. Ex ante, however, it is not obvious which emotions these materials will emphasize. On the one hand, educational materials may aim to build emotional intelligence by exposing children to the richness and complexity of real-life emotional experience. Consistent with the constructionist hypothesis, this could involve stories and images that portray a broad range of emotions and model how characters recognize, express, and regulate them (Hoemann, Xu and Barrett, 2019). On the other hand, educational materials may also reflect prevailing social and labor market norms about which emotions are appropriate to display. Many social situations and occupations discourage the expression of so-called "negative" emotions and instead reward calmness, compliance, or a façade of positivity (Hochschild, 1983; Deming, 2017). To the extent that schools socialize children into future social and labor market roles, these norms may be reflected in materials that downplay or omit emotions such as sadness, anger, or fear (Bowles, Gintis et al., 1976; Entorf and Dohmen, 2025).

This ambiguity creates an empirical question: do children's educational materials expose children to a broad emotional repertoire, or do they reproduce a narrower set of socially and economically rewarded emotional displays? Moreover, because children encounter these materials through both text and images, it is important to examine not only which emotions are described, but also which emotions are visually modeled. If images do not align with the emotions described in the accompanying text, materials may provide weaker or inconsistent cues for developing emotion recognition skills (Brady, Ogren and Johnson, 2024). Despite the importance of children's materials in the formation of emotional skills, there is limited systematic evidence on how emotions are represented across both modalities.

We address this gap by examining the emotional content of both textbooks and children's literature, comparing patterns across formal and informal educational materials. We also incorporate data on consumer and public library purchasing behavior to assess whether demand for particular emotional portrayals influences book acquisition decisions and, in turn, the emotional representations children are likely to encounter.

II Data Sources

We analyze two categories of children's media representing both formal and informal educational contexts: state-adopted public school textbooks and award-winning children's books. These materials constitute two prominent sources of content to which children are regularly exposed during critical developmental periods.

For the public school curricular materials, we analyze state-adopted textbooks from Texas and California in the subjects of Science, Social Studies, and Reading/Language Arts for 3rd and 5th grades. The selection of these states is justified by three key factors. First, California and Texas are the two largest public school systems in the United States, enrolling approximately 5.92 million and 5.53 million students, respectively, in fall 2023. Second, California and Texas both use state-level review processes in which appointed reviewers evaluate publisher materials for standards alignment and may prompt revisions, whereas most states leave textbook selection primarily to local districts (Doan and Kaufman, 2024). Third, these states occupy opposing positions on the political spectrum with Texas advisory panels being more likely to reflect Republican perspectives and California panels being more likely to reflect Democratic perspectives. Differences across states could reflect the influence of ideology on the representation of emotions and identities in educational materials (Biasi et al., 2025).

For children's literature, we focus on books recognized for being particularly distinguished or exemplary. Specifically, we include titles that have won awards featured by the Association for Library Service to Children. These awards honor books for their "exceptional literary" or "artistic" value, or for centering on the experiences of people with underrepresented identities. We analyze two sets of children's books: those that have won "Mainstream" awards, which emphasize literary quality and influence, and those recognized with "Diversity" awards, which specifically celebrate books featuring underrepresented identities. The corpus encompasses both fiction and nonfiction works, representing realistic and imaginary settings, thereby capturing the full spectrum of narratives available to young readers in curated, high-quality literature.

III Methods

Previous research has established robust computational methods for emotion detection in textual data (Bollen, Mao and Pepe, 2011; Thelwall, Buckley and Paltoglou, 2011; Hasan, Rundensteiner and Agu, 2014; Demszky et al., 2020). However, emotions are communicated not only through language, but also through non-verbal cues such as gestures and body language. Facial expressions in particular have been extensively documented as reliable indicators of emotional states (Sullivan and Masters, 1988; Lucas, Diener and Larsen, 2003; Öhman, 2007, 2009), so only measuring emotional representation in text may miss a central channel through which children encounter and learn about emotions. We therefore develop a multimodal framework that systematically quantifies emotional content in both the textual and visual components of children's media.

We focus our analysis on eight emotion categories: anger, calm, confusion, disgust, fear, happiness, sadness, and surprise. These classifications incorporate the six basic emotions originally delineated by Ekman, Sorenson and Friesen (1969) while extending the framework to include calm and confusion. To accommodate ambiguous cases, we include two additional classifications: "unsure" for instances where multiple emotions appear equally plausible or emotional expression is ambiguous and "no emotion" for affectively neutral content, where a character's emotion is not obvious or visible. While the concept of universal basic emotions remains contested in the literature (Russell, 2003), empirical evidence supports the existence of discrete emotional categories (Cowen and Keltner, 2020). However, the precise set of discrete emotions is not agreed upon (Ekman, 2016), and our primary set aligns with taxonomies accounting for displays of emotions across different contexts (Cowen and Keltner, 2017), including in faces (Cowen and Keltner, 2020).

III.A Text Analysis

We developed three complementary approaches to quantify emotional content in the text of children's printed content. Our primary measure uses an LMM for character-level emotion detection. We supplement this with a lexicon-based frequency analysis and a fine-tuned transformer model. This multi-method strategy allows us to test the sensitivity of our findings to different measurement strategies. Figure 1a provides an overview of the data extraction process applied to the text on book pages.

LMM. We use GPT-4o-mini to construct page-level measures of character emotions in text, motivated by recent evidence that LMMs can support scalable social-science annotation tasks, including emotion classification (Ziems et al., 2024). For each page, the model identifies characters, assigns each character one of eight emotion labels—anger, calm, confusion, disgust, fear, happiness, sadness, or surprise—and records "unsure" or "no emotion" when appropriate. We also extract character attributes, including age, gender, race, and whether the character is human, allowing us to analyze heterogeneity in emotional representation across demographic groups. More details are found in Appendix C.B.

Lexicon-Based Analysis. We also implement a lexicon-based approach which involves counting the frequency of predetermined emotion-related words, offering a straightforward and interpretable baseline. To do so, we construct comprehensive keyword dictionaries[2]. For each sentence, we enumerate emotion-specific keyword occurrences and aggregate counts to our analytical units of interest (e.g., complete corpus, subject matter categories, demographic subgroups). This frequency-based approach provides a transparent, replicable baseline measurement independent of neural network architectures.

Transformer Model Prediction. As a second text-based robustness check, we apply a RoBERTa-base transformer model fine-tuned on the GoEmotions dataset created by Demszky et al. (2020).[3] The model returns a probability score for each of the 28 GoEmotions

[2]We construct these dictionaries using three sources: WordNet semantic expansion, ChatGPT-assisted generation, and manual curation; more details are found in Appendix C.B. Examples of keywords related to each emotion can be found in Appendix Table B.1. The full lexicon can be viewed at `https://github.com/miielab/emotions_paper`

[3]The model can be found at `https://huggingface.co/SamLowe/roberta-base-go_emotions`.

labels, which include 27 emotion labels and a neutral category. In our implementation, we assign each sentence to the label with the highest predicted probability and drop sentences classified as "neutral," since our goal is to characterize the distribution of emotional content.[4] We then map the remaining fine-grained GoEmotions labels to the corresponding aggregate emotion categories used in our analysis.[5] This supervised transformer-based approach provides an additional benchmark for our primary LMM-based text classification method.

III.B Image Analysis

Social scientists have increasingly utilized images as primary sources of information (Ash et al., 2021; Adukia et al., 2023; Yanagizawa-Drott and Voth, 2023; Adukia and Harrison, 2025). Building on this emerging methodological literature, we developed two complementary computational pipelines to systematically extract structured data on emotional expressions and character attributes from scanned pages of children's books. We use an LMM and, separately, a two-step face detection model and feature classification model. Figure 1b provides an overview of the data extraction process applied to the images on book pages.

LMM Classification. We use Gemini 2.5 Pro Preview to construct character-level measures of emotions and attributes from book images. We process scanned pages in two stages: first, we classify each page as having only text or containing an image; only pages with images proceed to the second stage. For each illustrated page, the model provides information for the illustration as a whole such as a description of the setting as well as information on each depicted character including bounding-box coordinates, race, gender, age, skin color, human/non-human status, and emotion labels. Mirroring our text analysis, our primary emotion taxonomy includes anger, calm, confusion, disgust, fear, happiness, sadness, and surprise, with "unsure" and "no emotion" allowed for ambiguous or neutral cases. More detailed information is provided in Appendix C.B.

Convolutional Neural Network-Based Face Detection and Feature Classification. As a second approach to measuring emotions and character attributes in images, we use a computer-vision pipeline that builds on the face-detection and feature-classification models developed in Szasz et al. (2022). We first apply FDAI, a face-detection model trained on illustrated faces from children's books, to identify faces in book images. We then use an

[4]Dropping neutral sentences parallels our lexicon-based approach, in which we only examine sentences containing emotion-related words.

[5]Because GoEmotions does not include a "calm" label, this approach maps the model outputs to the remaining aggregate emotion categories. We view this limitation as useful for assessing whether our conclusions are sensitive to the set of emotions available to the text classifier. Our mapping is adapted from the correspondence categories provided at https://github.com/google-research/google-research/blob/master/goemotions/data/ekman_mapping.json.

extended version of our feature-classification model, originally developed for Szasz et al. (2022), to classify race, gender, age, and emotion for each detected face. We use these predicted labels as an alternative to our LMM-based classifications to ensure that our results are robust to different measurement choices.

IV Results

In this section, we present our findings on the emotional content expressed through the text and images of influential children's materials.

IV.A Emotions Overall and Over Time

We first examine the overall patterns of emotions in text and in images, respectively. Figure 2 shows the distributions of detected emotions in text and images across all books in our sample. In the text, the most common emotions are happiness and sadness, which account for approximately 29% and 20% of characters' emotions, respectively (Figure 2a). With respect to the distribution of the other emotions (anger, calm, confusion, disgust, fear, and surprise), we observe a meaningful number of depictions of each in text (ranging between 4% and 12%), reflecting a balance between "positively" and "negatively" valenced emotions. In sharp contrast, there is a high degree of homogeneity in the emotions displayed in images. In images, calm and happiness together account for approximately 90% of the depicted emotions by characters (Figure 2b).[6] Notably, negatively valenced emotions, which feature prominently in these books based on the text analysis, are largely absent from images. This pattern underscores a stark incongruence between the emotional narratives in the text and the limited emotional range portrayed in the images. Figure 3 shows that these findings are stable over time.

IV.B Emotions in Different Contexts

To better understand how emotions are represented across different types of books, we examine heterogeneity across a variety of different subcollections of our data. Overall, we find similar patterns across each context: images remain concentrated in calm and happiness, whereas a wider variety of emotions is displayed in text.

In Figure 4 we see that the distribution of emotions in the text of children's literature and textbooks are broadly similar, with textbooks showing a slightly stronger overrepresentation of happiness compared to children's literature. As may be expected, children's

[6] Note that these proportions reflect the proportion of all characters that have an emotion. We drop characters that are classified as having "No Emotion" or an "Unsure" Emotion. The proportion of pages that fit these categories are approximately 30% for characters in the text and 27% for characters in the images. Distributions including these classifications can be found in Figure A.1. Dropping these characters does not change our results in any substantive way and improves the interpretability of our analyses.

literature displays a somewhat richer and more diverse range of emotions compared to textbooks, though both collections cover a variety of emotions in text.

Within the subcollections of these two sources of children's text, the patterns remain qualitatively similar. In particular, the Mainstream and Diversity subcollections of children's literature show comparable proportions of each emotion, the largest difference between the two is that the Diversity subcollection exhibits a slightly greater proportion of sadness (approximately 5 percentage points higher). With respect to textbooks, we observe similar patterns across those from California and Texas.

We also note some differences by grade and subject in textbooks. For example, happiness is represented more frequently in the text of third-grade textbooks relative to fifth-grade textbooks (10 percentage point difference). Conversely, emotions such as anger, fear, and sadness appear slightly more frequently in fifth-grade textbooks.

We also observe differences by subject. Science textbooks contain less anger and sadness compared to reading and social studies textbooks. This may be consistent with the instructional goals of each subject: science textbooks may not require as rich an emotional narrative to convey concrete concepts, whereas reading and social studies naturally incorporate a broader range of emotions. Even acknowledging small differences across subjects, we still observe that the representation of emotions in text remains broadly consistent across all subcollections.

Figure 5 displays the same heterogeneity analysis for images. We again find qualitatively similar patterns across subcollections, with calm and happiness consistently constituting the vast majority of emotional depictions. These two emotions are even more overrepresented in textbooks than in children's books. As in the textual analysis, this is consistent with a somewhat richer display of distinct emotions in children's literature compared to textbooks. This overrepresentation of happiness and calm is especially strong in science and social studies books. Reading textbooks still overrepresent happiness and calm, but the distribution of emotions in these books' images more closely resembles the distribution in children's literature than that in textbooks from the other subjects. This could reflect that these textbooks include age-appropriate stories, possibly designed to resemble the award-winning books in our sample. We also observe that third-grade textbooks display more happiness and less calm than fifth-grade textbooks. The discrepancies nearly balance each other, so the total percentage in these two categories is similar across the two grade levels.

IV.C Emotions Displayed by Different Identities

Next, we examine how displays of emotion vary for characters of different identities. As described in Section III, we obtain the predicted race and gender of characters, whether they are children or adults, and whether they are human or non-human. For both text and images, these attributes are predictions made by the respective models described in Section III.

Figure 6 displays this heterogeneity analysis in text. Descriptively, we see that females are slightly more likely than males to show sadness in text, whereas males are slightly more likely to show anger. Additionally, adults are more likely to express anger or calm, whereas children are more likely to express fear or confusion.

Figure 7 presents the analogous heterogeneity analysis for images. We see that male faces are more likely to show calm than female faces. This is consistent with gender stereotypes that females are less calm than males. However, female faces are more likely, to show happiness than male faces by a similar amount. A similar pattern holds for the other identities; groups that are more frequently shown as calm are less frequently shown as happy and vice versa. All groups display calm or happy facial expressions in almost all instances, and no group shows any other emotion with comparable frequency.

IV.D Concurrent Text and Image Analysis

To further investigate our text and image data, we directly compare the emotions expressed through each modality on the same page. Figure 8 shows the co-occurrence of emotions for text and images on the same page, normalized by the text columns. In other words, each cell shows the proportion of a specific emotion label in images, given a particular emotion label in the text of the same page. When comparing the presence of emotions in images and text on the same page, calm is the emotion most likely to align: when calm is mentioned in the text, 70% of emotions in images on the corresponding page are calm. Following calm, when a character described in the text is happy, 27% of emotions in images show happy. In contrast, emotions such as anger, fear, sadness, and surprise when mentioned in the text, are visually represented with shares of only 13%, 9%, 10%, and 8% respectively. When there is an emotion other than calm in the text, images on that page still depict a share of calm between 58% and 67%. This indicates that, not only are kids not exposed to a variety of emotional displays in images, but there is frequently a mismatch between what characters say, do, and feel in the text compared to what they convey in images. Moreover, this mismatch exhibits a consistent direction: negative emotions described in the text are accompanied by positive expressions in the visual cues. This incongruence may complicate social emotional learning and possibly reinforce an implicit message that suppressing negative emotions is desirable.

It could be the case that, despite broad differences between emotions in text and images, the images do faithfully convey the emotions of at least one of the characters in the text. To probe this possibility, we recompute the co-occurrence matrix for a sample of pages in which we identify a unique character of the same predicted race and gender in both the text and image of a particular page. Figure A.3 illustrates that the distributions of emotions in text and images in this matched sample are similar to the overall distributions displayed in Figure 2, indicating that the matched sample is similar to the full sample of text and images.

Figure A.4 further illustrates the misalignment of emotions in the text and images corresponding to these matched characters. The broad pattern of misalignment persists. We note that there is similar weight on the diagonal compared to Figure 8, reflecting the fact that emotions displayed in text and images frequently fail to align even when we restrict to characters matched uniquely on race and gender. Regardless of the emotions a character displays in the text, they are most likely to be depicted as calm in the accompanying image.

We further probe the overall text-image discordance by computing a page-level measure of the distance between the set of emotions in text and the set of emotions in images. In particular, we combine the page-level indicators for the presence of different emotions into a single eight-dimensional vector separately for text and images. Each entry in these vectors corresponds to an emotion and takes a value of one if and only if the emotion is present on the page. After normalizing, we compute the total variation distance (TVD) between the page-level distributions of text and images as $\delta(p, q) = \frac{1}{2}\|p - q\|_1$.

Figure A.5 displays a histogram of the total variation distances for the distributions of emotions in text and images on individual pages. In about 10% of cases, we see that the TVD is zero, indicating that the emotions in the text and images align perfectly. However, for nearly half of the pages, the TVD is one, indicating that there is no overlap between emotions featured in text and images on the same page.

Overall, these results demonstrate that the discrepancy between the distributions of emotions in text and images is not driven merely by illustrations that selectively emphasize specific positively valenced emotions in the text; rather, the illustrations convey happiness and calm even when these emotions are absent in the text. These patterns speak directly to the mechanisms discussed in Section I. Given the role of visual cues in children's emotion recognition, this mismatch may narrow the range of emotional displays children see modeled and may reinforce norms about which emotions are appropriate to express.

V Robustness and Validation

V.A Alternative Approaches to Measurement

To assess robustness, we compare our primary LMM measures with alternative approaches for both text and images. These include a lexicon-based text measure, a fine-tuned transformer text classifier, and a CNN-based image classifier, described in Sections III.A and III.B. These alternative approaches differ from our primary LMM measures in their unit of analysis. The CNN-based image classifier predicts emotions for each detected face, whereas the LMM analyzes full images and the characters they contain. Similarly, the alternative text measures count emotion-related words or classify emotions at the sentence level, while the primary text LMM produces character-level emotion labels. Therefore, we do not expect the results of these analyses to align perfectly with our main results. Nonetheless, the findings demonstrate the robustness of our main findings to alternative ways of classifying emotion. Moreover, the use of different models with varying complexity and transparency provides a source of validation for the findings we obtain leveraging generative AI.

Figure A.6a shows the results of our lexicon-based approach. We see similar levels of anger, disgust, fear, and sadness in comparison to our main approach. In contrast, it appears that our main model identifies around 10% more happiness and around 10% less calm in our text than the lexicon method. This may be an artifact of the lexicon approach's sensitivity to keyword choice. Indeed, calm and happiness may be lexically difficult to differentiate. For example, consider the word "contentment" (which we use as one of the words related to the emotion of "calm"). Depending on the context of the surrounding text, this word could either indicate a state of calm or of happiness. We thus expect there to be some fluidity between these two categories with the lexicon model.

Figure A.6b shows the results of the GoEmotions model. Here we note qualitatively similar patterns to Figure 2a in all but surprise, to which the GoEmotions model attributes 20% more of its classifications. As noted earlier, this may be due to the structure of our inputs. The GoEmotions model could, for example, classify sentences associated with discovery and learning as expressing surprise. (To see this, consider a history textbook chapter on the Age of Discovery.) This expression would not be counted by the LMM because there is no specific character who is exhibiting surprise. Nonetheless, the results from this method are arguably more comparable to those of the LMM than the Lexicon-based results since the GoEmotions model and LMM both use surrounding context to classify emotions.

Figure A.6c shows the results from our alternative image model. This model predicts less calm and more happiness than the LMM for images, but the sum of emotions in these two categories is consistent across models. This further reflects the overlap between these two categories that we have already discussed. Overall, we find a qualitatively similar distribution to the one obtained using the LMM, which lends further confidence to our main results.

In summary, across a variety of different models ranging in complexity and interpretability, we obtain similar distributions of emotions in text and images to the ones we find using generative AI. This provides one way to validate the results from these more complex models. Additionally, while our main textual results only examine emotions in the text displayed by characters in order to be more comparable to our image results, our robustness approaches allow for different levels of analysis. We observe that our findings are qualitatively consistent across levels of analysis. In particular, regardless of the models used, we find that happiness and calm are consistently overrepresented in images relative to text.

V.B Manual Validation

In this section, we describe the validation procedures for our primary LMM-based emotion measures. Traditional metrics such as F1 scores or correlations between labels produced by models and humans exhibit several shortcomings. First, these measures do not admit an interpretable threshold above which the models are deemed acceptable. Second, some tasks will give rise to more disagreement between human annotators, and this uncertainty in the ground truth should be incorporated in an assessment of model performance. Third, even in settings where multiple labelers label the same example, traditional comparisons of the correlations between model and human responses versus those between human and human responses are limited in that comparable correlations may be neither necessary nor sufficient for assessing a model's effectiveness for producing measurements used in downstream applications. Moreover, to the extent that human disagreement is the result of differing but legitimate interpretations of a given example, it may be admissible for a model to agree with *some* humans, even if it does not agree with all or even the modal annotator.

To address these concerns, we proceed in the spirit of Calderon, Reichart and Dror (2025) by examining how our models compare with human labelers in terms of their ability to agree with other human labelers. Our procedure directly tests the comparability of the model with each human annotator, allowing us to determine whether the model is statistically interchangeable with that annotator. We obtain a "pass rate," indicating the fraction of human annotators with whom the model is statistically interchangeable. The procedure is outlined in Figure A.7.[7]

[7]We also report traditional macro F1 scores in Table B.3.

More concretely, suppose we have a set of examples denoted $\mathcal{X}$ and human annotators $\mathcal{H}$. For a given example $x \in \mathcal{X}$, let $\mathcal{H}(x)$ denote the set of human annotators who labeled example x. Similarly, for $h \in \mathcal{H}$, let $\mathcal{X}(h)$ denote the set of examples labeled by annotator h. For annotator $h \in \mathcal{H}(x)$, we can also define $\mathcal{H}_{-h}(x)$ to be the set of annotators who labeled example x, excluding annotator h. For $h \in \mathcal{H}(x)$, let $h(x)$ denote the label that human annotator h assigns to example x. As in Calderon, Reichart and Dror (2025), let $f(x)$ denote the model's label assigned to example x. Then, for a given annotator $h_1 \in \mathcal{H}$ who labeled example $x \in \mathcal{X}$, we define the leave-one-out agreement score with annotator $h_2 \in \mathcal{H}$ excluded as

$$\mathrm{ACC}_{h_1}(x, h_2) = \frac{1}{\left|\mathcal{H}_{-h_2}(x)\right|} \sum_{h' \in \mathcal{H}_{-h_2}(x)} \mathbf{1}\{h_1(x) = h'(x)\}.$$

With this definition, we construct a measure of how the model performs compared to human h on example x in terms of relative agreement with all other human annotators who labeled x:

$$d(x, h) = \mathrm{ACC}(f, x, h) - \mathrm{ACC}(h, x, h).$$

Then, we compute

$$\mu(h) = \frac{1}{|\mathcal{X}(h)|} \sum_{x \in \mathcal{X}(h)} d(x, h).$$

We then perform a one-sided t-test of the null $H_0 : \mu(h) \geq 0$. A low (negative) value of $\mu(h)$ provides evidence that the model agrees less with other annotators than annotator h on the set of examples labeled by h.

We can perform the same test for every human annotator. Of course, there exists dependence across the test statistics. We apply the Benjamini-Yekutieli procedure to obtain adjusted p-values that control the false discovery rate under arbitrary dependence among the test statistics. For a fixed level α, we determine the fraction W of adjusted p-values that are above α. This fraction represents the "pass rate," which intuitively is the fraction of human annotators with whom we cannot reject that the model is interchangeable.

We note that our procedure deviates from the alternative annotator test proposed by Calderon, Reichart and Dror (2025) in two important ways. First, they define an example-level win rate in place of our relative agreement. That is, they define

$$\tilde{d}(x, h) = \mathbf{1}\{\mathrm{ACC}(f, x, h) - \mathrm{ACC}(h, x, h) \geq 0\}.$$

This binarization discards information in the sense that the magnitude of the discrepancy between the model's agreement and h's agreement with other humans is not accounted for. The second departure we make from Calderon, Reichart and Dror (2025) is that we test for whether the model exhibits significantly lower agreement with other annotators rather than significantly higher agreement. Due to the fact that having equal agreement with humans is arguably the theoretical limit on how well a model can perform in a classification task, testing whether the model agrees with other human annotators significantly more than humans agree with each other sets an implausibly high bar. As a result, these authors introduce a cost parameter ε that biases the test in favor of the model. Given our definitions, the analogue of their implementation would be to test $H_0 : d(x, h) \leq -\varepsilon$. The spirit of this cost parameter is to tolerate lower model performance since the model offers a cheaper source of measurement than human annotators. In practice, however, the choice of ε is somewhat arbitrary. We find it more natural and interpretable to test whether the model is statistically underperforming relative to human annotators and avoid the introduction of this cost parameter.

We display the results of this validation exercise in Table B.4. For each exercise, we show the adjusted p-values and pass rates at different significance levels. For several exercises, the LMM exceeds the 0.5 pass rate proposed by Calderon, Reichart and Dror (2025) as the threshold above which the model can justifiably replace human annotators. We comment that the model generally performs well at measuring demographic characteristics. In many settings (and especially in text), demographic characteristics may simply be unavailable, so we are only assessing model performance on a selected set of examples. Nonetheless, in cases where demographic information can be ascertained, the model performs well.

At the same time, while the model performs comparably to some annotators in terms of emotion labeling, the 0.5 threshold is not always exceeded. The model's emotion labels in text are not statistically distinguishable from two out of six annotators. It performs better in image labeling, meeting or exceeding the 0.5 pass rate at each significance level. To the extent that humans can hold distinct but reasonable interpretations of an emotion, we view these pass rates favorably. However, after binning the emotions into two valence categories ($\{\text{happy, calm}\}$ and $\{\text{any other emotion}\}$, as well as $\{\text{no emotion}, \text{unsure}\}$), we see that the pass rate in image labeling declines. While the models may perform satisfactorily overall, these results motivate an analysis to verify that the types of mistakes the models make are not driving our primary results.

To do this, we use the validation sample to estimate the simulated distribution of human emotion labels we predict would be obtained if humans had labeled the entire dataset.

We let $\mathcal{C}$ denote the set of emotion classifications. We observe that

$$\mathbb{P}(h(x) = c) = \sum_{c' \in \mathcal{C}} \mathbb{P}(h(x) = c | f(x) = c') \mathbb{P}(f(x) = c').$$

We obtain $\mathbb{P}(f(x) = c')$ from the model labels on the entire dataset, and we estimate $\mathbb{P}(h(x) = c | f(x) = c')$ using the validation sample. We obtain consistent estimates since the validation sample was obtained by stratified random sampling from the full sample (with LMM class labels as strata to ensure representation of each class). The results are shown in Figure A.8 with 95% confidence intervals obtained via stratified bootstrap resampling from the validation data. We observe that people ascribe calm to more examples in text, and, compared to the model, they label more emotions in faces as happy rather than calm. Nonetheless, we recover the same pattern from the predicted human labels that text displays both positively and negatively valenced emotions while images show happy and calm almost uniformly. This is further confirmed in Figure A.9, which shows the distribution of emotions predicted by humans and the model after binning into Happy/Calm and Non-Happy/Calm.

In Figure A.10, we also show an additional robustness check for the LMM classifications in the text where we subset to emotion labels that the model has high confidence in (≥ 0.8 on a scale of 0 to 1).[8] This is a different subset of characters (likely characters where the emotion is clear). When we examine the overall representation of emotions in this subset, we see the largest changes in the proportion of happiness (10% increase) and confusion (9% decrease), but our main takeaways remain the same.

Overall, the results of the validation exercise lend confidence in the quality of the measurements produced by our LMMs. The model is never statistically worse than all annotators, and it exceeds the prescribed 0.5 threshold in several instances. To the extent that human disagreement reflects legitimate alternative interpretations, this means that the models are not out of line with the labels humans might propose. The model is not identical to human annotators, but we nevertheless recover the same qualitative conclusions after simulating the distribution of emotion labels that we predict humans would produce if they were to label the entire dataset.

[8] We note that the self-reported confidence score has no probabilistic interpretation, but model performance increases with this score. In particular, at this threshold, we obtain a pass rate of 5/6. We retain 282 validation examples after filtering on the confidence score, so this is not driven by a decrease in power of the test. We also see that the macro F1 score and LLM-human agreement relative to human-human agreement both increase with the confidence threshold.

VI Consumer Demand

Having documented a systematic oversupply of happy and calm emotions in images relative to the text in two influential sources of children's media, we next ask whether these patterns are consistent with economic forces shaping the content of children's books. A simple mechanism is that purchasers may prefer books whose visible features promise a pleasant emotional experience. In the context of literature, narratives associated with "negative" emotions may be perceived as more demanding or less enjoyable, prompting readers under stress or seeking relaxation to favor covers promising pleasant experiences (Becker, 1996; Elster, 1998; Hogan, 2011). If purchasers are more likely to select books whose visible features convey positive emotions, then authors, illustrators, and publishers may be incentivized to emphasize positive expressions, especially in images and on covers. This logic is consistent with evidence that consumer demand shapes media content (Gentzkow, Glaeser and Goldin, 2006; Gentzkow and Shapiro, 2010; Sacerdote, Sehgal and Cook, 2020) and that salient product attributes influence choice when consumers have limited attention (Ellison, 2006; Bordalo, Gennaioli and Shleifer, 2013).

Book covers provide a natural setting to study this mechanism. Covers are highly visible at the point of purchase and can function both as attention-grabbing product attributes and as signals of a book's tone and content (Bloch, 1995; Pieters and Wedel, 2004; Underwood and Klein, 2002; Eliaz and Spiegler, 2011). This is especially relevant for online purchases included in our sample, where consumers often cannot inspect interior pages before buying. If consumers favor happy or calm images, this could create market incentives to supply books with these positive emotions on the cover, even when the text contains a broader range of emotions.

This discussion yields a simple empirical prediction: conditional on book characteristics, titles whose covers depict only happy or calm characters should be selected more often than titles whose covers include a character displaying another emotion. We test this prediction using two measures of selection in the children's literature sample: household purchases from the Numerator data and holdings in the Seattle Public Library system.

VI.A Empirical Analysis of Consumer Demand

We limit our analysis to the books in our children's literature collection because their consumption is more a measure of consumer demand than the consumption of textbooks, which are chosen at the school level.[9]

[9] Of course, parents can select into different schools, but the content in textbooks is only one dimension of this multi-faceted choice.

We study consumer demand using two separate data sources. First, we estimate a conditional logit model using micro-level book purchase data from the Numerator OmniPanel. These data contain information on purchases by individuals and speak directly to consumers' revealed preferences. One challenge with using these data is that we cannot measure consumers' consideration sets. Although it is plausible that online purchasers – who comprise the majority of our sample – could select any book in the collection that we study, we still address this limitation by estimating a similar model using data on book stocks from the Seattle Public Library system. We expect that a public library would consider purchasing a much broader range of books than a typical individual consumer. Therefore, the assumption that every award-winning book in our sample is also in the library's consideration set becomes more plausible. Moreover, although public libraries may have different objective functions than individual consumers, we expect that they would generally stock more copies of books that readers prefer, so these stocks should indirectly reflect readers' choices.

In order to specifically study the seeming overrepresentation of happy and calm in images, we now classify emotions into valence categories of happy/calm or non-happy/calm, with the latter category indicating that some emotion besides happy or calm is displayed.[10] We estimate a conditional logit model using the Numerator data to explicitly study consumers' revealed preferences, and we subsequently use the public library data to regress the number of copies of a book held by the library against the same set of features. Our models include an indicator for whether there are exclusively happy/calm characters on the cover, an indicator for whether there is any character on the cover, and a set of controls including indicator variables for the award(s) the book won, the book's copyright year, whether the book is a picture book, whether it contains no pictured characters, the number of pages, and the average number of words per page.[11] We do not expect consumers to accurately perceive the proportions of emotions displayed in the text or images *within* the books. However, we still run two alternative specifications in which we control for the proportions of happy/calm emotions in both images and text by (1) directly including these proportions as controls and (2) by adding indicators for quintiles of these proportions within their respective distributions. These additional controls allow for consumers who attend more closely to the emotional content to respond to the emotional content of the books, and they can potentially proxy for unobserved confounders.

Table 1 displays the results from the Numerator data. We observe that the emotional representation on the cover plays a substantive role in shaping consumers' purchase decisions.

[10] "No emotion" and "unsure" labels are not included in this category.

[11] These last two features serve as a proxy for the age of the target audience of the book.

In particular, across all specifications, we see that consumers place a significant premium on books depicting happy and calm characters on the cover compared to books depicting other emotions. Moreover, consumers significantly favor books with no characters on the cover to books with a character displaying an emotion which is neither happy nor calm. This suggests an explicit aversion to books that have these other emotions on the cover. In turn, this aversion could disincentivize book creators from illustrating non-happy/calm emotions, which would produce the patterns we uncover.

In Table 2, we display an analogous set of results using the Seattle Public Library system data. Overall, the results are qualitatively similar. As before, we find higher stocks of books with exclusively happy or calm cover characters or no character at all than books with non-happy/calm characters on the cover. This suggests that an inability to properly specify consumers' consideration sets does not drive our results in Table 1. Instead, these results jointly speak to consistent purchasing patterns of readers favoring books with depictions of happiness and calm over books with other emotional depictions.

Additionally, in order to connect our findings to purchasers' decisions about happy and calm emotions in general and not just on covers, we find that the emotions displayed on the cover do provide an informative signal about the emotions in the content of the book. We estimate another conditional logit model to study the probability that a cover contains only happy/calm characters as a function of the proportion of happy/calm emotions in text and images in the content of the book. We also perform an analogous analysis to study the probability a cover contains only non-happy/calm characters. The results are displayed in Table 3. We find that having higher proportions of happy/calm emotions in text and images predicts having exclusively happy/calm characters on the cover. Likewise, lower proportions of happy/calm emotions in text and images predict having exclusively non-happy/calm characters on the cover.[12] These results indicate that covers do provide consumers with an informative signal about the emotional content within the book (e.g. they can indeed judge a book by its cover). Therefore, consumers' preference for books with more happy/calm characters pictured on the cover is also consistent with wanting these emotions displayed in the content of the book.

In light of these results, the patterns we uncover in Section IV are consistent with a supply-side response to readers' observed purchasing patterns regarding certain emotions. Moreover, since images – especially those on covers – are particularly salient compared to

[12]The reason the latter result is not mechanical is because these two categories are not mutually exclusive. Covers could have a mix of happy/calm and non-happy/calm emotions, or they could display no characters at all.

text, authors would face a stronger disincentive to depict non-happy/calm emotions in the images compared to text. This is consistent with our finding that emotions are relatively diverse in text while images predominantly show characters who are either happy or calm. We also note that this market equilibrium is not necessarily innocuous; if consumers are boundedly rational but are attracted to depictions of positive emotions, then the resulting equilibrium may lead to an oversupply of happy/calm emotions relative to what is best for young readers.

VII Conclusion

Our analysis yields important insights into how emotions are taught to children through both formal and informal educational materials. One of the most striking results is the disproportionate emphasis on calm and happiness in images, which together account for about 90% of visually depicted emotions. While these emotions are crucial for fostering a positive and serene learning environment, the underrepresentation of other emotions—such as anger, sadness, or confusion—limits students' exposure to the full spectrum of emotional experiences. These images are particularly influential in shaping young readers' understanding of emotions, as they provide direct and intuitive cues that text alone cannot. Therefore, whether inadvertently or intentionally, this imbalance may signal to students that some emotions are more acceptable or valued than others, potentially affecting the way they recognize, understand, and navigate complex emotional landscapes. Moreover, the near absence of visual models of a broader range of emotions prevents these books from effectively teaching emotion recognition, which is an important component of emotional intelligence.

Our findings also reveal that this imbalance between depictions of "positive" and "negative" emotions in images relative to the text persists across time, context, and character identity. Despite growing awareness of the importance of emotional intelligence in education, the representation of emotions in educational materials appears stagnant, with no clear trend toward broader inclusion of underrepresented emotions. Moreover, happiness and calm are overrepresented in all subcollections and by characters of different identities. However, there are some notable sites of heterogeneity. For example, female characters are more often depicted as happy, while male characters are more often shown as calm, suggesting a subtle reinforcement of gendered emotional stereotypes.

While text offers a more diverse range of emotional representation compared to images, it often displays emotions that conflict with those depicted in images on the same page. Even within the same character on the same page, no matter which emotions the character displays in the text, they are disproportionately likely to be shown as either calm or happy. This incongruence could have important consequences for how children learn to interpret and respond to emotions. In the books they are most likely to be exposed to, they are implicitly shown that only calm and happy reactions are appropriate. When it comes to public school textbooks in particular, these patterns suggest that curriculum designers and publishers may not yet fully recognize the potential of educational materials to foster emotional literacy. Alternatively, these materials may intentionally emphasize the importance of positivity regardless of the situation.

In the case of children's literature, we uncover evidence that the disproportionate representation of happy and calm emotions is consistent with a supply-side response to purchasers' revealed preferences. Both individual households and public libraries (whose inventories can be expected to reflect the interests of the public) are significantly less likely to purchase books that feature characters with "negative" emotions on the cover. As a result, authors and illustrators include images of happy and calm characters in an effort to attract consumers. These patterns may reflect the societal value placed on "positive" emotions. Importantly, we note that even if purchasers did want a greater diversity of emotions in their kids' books, they may be imperfectly attentive or unable to carefully peruse the emotional content of a book before purchase. In this case, smiling faces may capture their limited attention and carry appeal, and this could result in an inefficient overrepresentation of happy and calm emotions (Ellison, 2006; Gabaix and Laibson, 2006; Cusumano, Fabbri and Pieroth, 2024). We do not disentangle underlying preferences from the effects of imperfect attention, but simply illustrate that consumers' revealed preferences for happy and calm emotions on book covers could drive the patterns we uncover.

More broadly, our findings underscore the role of children's materials in the formation of emotional skills and norms. Emotional intelligence develops through repeated exposure to emotional language, visual cues, and models of expression. We show that textbooks and children's books provide a broad emotional vocabulary in text while visually reinforcing a narrower repertoire centered on happiness and calm. This pattern is consistent with social and market incentives that reward positive emotional displays, but it may also limit the range of emotional expressions children see modeled in influential educational materials. From an educational perspective, the results point to the value of more intentional emotional diversity in both text and images, so that children encounter not only examples of emotional regulation, but also visual models of the broader range of emotions that shape everyday life.

References

Adukia, Anjali, Alex Eble, Emileigh Harrison, Hakizumwami Birali Runesha, and Teodora Szasz. 2023. "What we teach about race and gender: Representation in images and text of children's books." The Quarterly Journal of Economics, 138(4): 2225–2285.

Adukia, Anjali, and Emileigh Harrison. 2025. "Separation of church and state curricula? Examining public and religious private school textbooks." National Bureau of Economic Research.

Arizpe, Evelyn, and Morag Styles. 2015. Children reading picturebooks: Interpreting visual texts. Routledge.

Ash, Elliott, Ruben Durante, Maria Grebenschikova, and Carlo Schwarz. 2021. "Visual representation and stereotypes in news media."

Austin, Elizabeth J, Donald H Saklofske, and Sarah M Mastoras. 2010. "Emotional intelligence, coping and exam-related stress in Canadian undergraduate students." Australian Journal of Psychology, 62(1): 42–50.

Averill, James R. 1983. "Studies on anger and aggression: Implications for theories of emotion." American Psychologist, 38(11): 1145.

Bazzi, Samuel, Martin Fiszbein, and Mesay Gebresilasse. 2020. "Frontier culture: The roots and persistence of "rugged individualism" in the United States." Econometrica, 88(6): 2329–2368.

Becker, Gary S. 1996. Accounting for Tastes. Harvard University Press: Cambridge.

Biasi, Barbara, Minseon Park, John D Singleton, and Seth D Zimmerman. 2025. "Identity and Ideology in the School Boardroom." National Bureau of Economic Research.

Bloch, Peter H. 1995. "Seeking the Ideal Form: Product Design and Consumer Response." Journal of Marketing, 59(3): 16–29.

Bollen, Johan, Huina Mao, and Alberto Pepe. 2011. "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena." Vol. 5, 450–453.

Bonhomme, Stéphane. 2021. "Teams: Heterogeneity, sorting, and complementarity." arXiv preprint arXiv:2102.01802.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer. 2013. "Salience and consumer choice." Journal of Political Economy, 121(5): 803–843.

Bowles, Samuel, Herbert Gintis, et al. 1976. Schooling in capitalist America. Vol. 57, New York: Basic Books.

Brackett, Marc A, Susan E Rivers, Maria R Reyes, and Peter Salovey. 2012. "Enhancing academic performance and social and emotional competence with the RULER feeling words curriculum." Learning and Individual Differences, 22(2): 218–224.

Brady, Shannon M, Marissa Ogren, and Scott P Johnson. 2024. "Effects of conflicting emotional cues on toddlers' emotion perception." British Journal of Developmental Psychology, 42(3): 376–391.

Calderon, Nitay, Roi Reichart, and Rotem Dror. 2025. "The alternative annotator test for llm-as-a-judge: How to statistically justify replacing human annotators with llms." 16051–16081.

Castex, Gonzalo, and Evgenia Kogan Dechter. 2014. "The changing roles of education and ability in wage determination." Journal of Labor Economics, 32(4): 685–710.

Cockcroft, Marlaina. 2018. "Caldecott and Newbery Medal wins bring instant boost to book sales." School Library Journal, 64(2).

Cowen, Alan S, and Dacher Keltner. 2017. "Self-report captures 27 distinct categories of emotion bridged by continuous gradients." Proceedings of the national academy of sciences, 114(38): E7900–E7909.

Cowen, Alan S, and Dacher Keltner. 2020. "What the face displays: Mapping 28 emotions conveyed by naturalistic expression." American Psychologist, 75(3): 349.

Cusumano, Carlo M, Francesco Fabbri, and Ferdinand Pieroth. 2024. "Competing to commit: Markets with rational inattention." American Economic Review, 114(1): 285–306.

Deming, David J. 2017. "The growing importance of social skills in the labor market." The quarterly journal of economics, 132(4): 1593–1640.

Demszky, Dorottya, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade, and Sujith Ravi. 2020. "GoEmotions: A dataset of fine-grained emotions." arXiv preprint arXiv:2005.00547.

Denham, Susanne A. 2019. "Emotional competence during childhood and adolescence." In Handbook of emotional development. 493–541. Springer.

Dickerson, Kelli L, and Jodi A Quas. 2021. "Emotional awareness, empathy, and generosity in high-risk youths." Journal of Experimental Child Psychology, 208: 105151.

Doan, Sy, and Julia Kaufman. 2024. "What Role Do States Play in Selecting K–12 Textbooks?" State Education Standard, 24(1).

Doherty, Martin. 2008. Theory of mind: How children understand others' thoughts and feelings. psychology press.

Dröge, Cornelia, and Rene Y Darmon. 1987. "Associative positioning strategies through comparative advertising: Attribute versus overall similarity approaches." Journal of Marketing Research, 24(4): 377–388.

Duckworth, Angela L, Christopher Peterson, Michael D Matthews, and Dennis R Kelly. 2007. "Grit: perseverance and passion for long-term goals." Journal of personality and social psychology, 92(6): 1087.

Edin, Per-Anders, Peter Fredriksson, Martin Nybom, and Björn Öckert. 2022. "The rising return to noncognitive skill." American Economic Journal: Applied Economics, 14(2): 78–100.

Eisenberg, Nancy, Amanda Cumberland, and Tracy L Spinrad. 1998. "Parental socialization of emotion." Psychological inquiry, 9(4): 241–273.

Ekman, Paul. 2016. "What scientists who study emotion agree about." Perspectives on psychological science, 11(1): 31–34.

Ekman, Paul, E Richard Sorenson, and Wallace V Friesen. 1969. "Pan-cultural elements in facial displays of emotion." Science, 164(3875): 86–88.

Eliaz, Kfir, and Ran Spiegler. 2011. "Consideration sets and competitive marketing." The Review of Economic Studies, 78(1): 235–262.

Ellison, Glenn. 2006. "Bounded rationality in industrial organization." Econometric Society Monographs, 42: 142.

Elster, Jon. 1998. "Emotions and economic theory." Journal of economic literature, 36(1): 47–74.

Entorf, Antonia K, and Thomas J Dohmen. 2025. "The effect of compulsory education on non-cognitive skills: Evidence from low-and middle-income countries." Economics of Education Review, 107: 102654.

Foulds, Kim. 2023. "Co-viewing mass media to support children and parents' emotional ABCs: An evaluation of Ahlan Simsim." Early Childhood Education Journal, 51(8): 1479–1488.

Gabaix, Xavier, and David Laibson. 2006. "Shrouded attributes, consumer myopia, and information suppression in competitive markets." The Quarterly Journal of Economics, 121(2): 505–540.

Gentzkow, Matthew, and Jesse M Shapiro. 2010. "What drives media slant? Evidence from US daily newspapers." Econometrica, 78(1): 35–71.

Gentzkow, Matthew, Edward L Glaeser, and Claudia Goldin. 2006. "The rise of the fourth estate. How newspapers became informative and why it mattered." In Corruption and reform: Lessons from America's economic history. 187–230. University of Chicago Press.

Green, Clarence, and He Sun. 2025. "Picturebooks increase the frequency and diversity of emotion vocabulary in children's language environments: Modeling potential benefits to emotional literacy, with pedagogical resources." Early Education and Development, 36(3): 568–586.

Gross, James J. 2015. "Emotion regulation: Current status and future prospects." Psychological inquiry, 26(1): 1–26.

Groves, Melissa Osborne. 2005. "How important is your personality? Labor market returns to personality for women in the US and UK." Journal of Economic Psychology, 26(6): 827–841.

Hansen, Stephen, Tejas Ramdas, Raffaella Sadun, and Joe Fuller. 2021. "The demand for executive skills." National Bureau of Economic Research.

Hasan, Maryam, Elke Rundensteiner, and Emmanuel Agu. 2014. "Emotex: Detecting emotions in twitter messages."

Heckman, James J, and Yona Rubinstein. 2001. "The importance of noncognitive skills: Lessons from the GED testing program." American economic review, 91(2): 145–149.

Heckman, James J, Jora Stixrud, and Sergio Urzua. 2006. "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior." Journal of Labor Economics, 24(3): 411–482.

Hochschild, Arlie Russell. 1983. "The Managed Heart: Commercialization of Human Feeling."

Hoemann, Katie, Fei Xu, and Lisa Feldman Barrett. 2019. "Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis." Developmental psychology, 55(9): 1830.

Hoffman, Mitchell, and Steven Tadelis. 2021. "People management skills, employee attrition, and manager rewards: An empirical analysis." Journal of political economy, 129(1): 243–285.

Hogan, Patrick Colm. 2011. What Literature Teaches Us About Emotion. Cambridge University Press.

Izard, Carroll E. 1977. "Differential emotions theory." In Human Emotions. 43–66. Springer.

Kardum, Rona Bušljeta, Katarina Dadic, and Martina Horvat. 2021. "Education for the cultivation of emotions through textbooks: the example of Croatian high school history textbooks." Educational process: international journal, 10(2): 28–41.

Kautz, Tim, James J Heckman, Ron Diris, Bas Ter Weel, and Lex Borghans. 2014. "Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success." , (Working Paper 20749).

Kemper, Theodore D. 1978. "Toward a sociology of emotions: Some problems and some solutions." The American Sociologist, 30–41.

Kumschick, Irina R, Luna Beck, Michael Eid, Georg Witte, Gisela Klann-Delius, Isabella Heuser, Rüdiger Steinlein, and Winfried Menninghaus. 2014. "Reading and feeling: The effects of a literature-based intervention designed to increase emotional competence in second and third graders." Frontiers in Psychology, 5: 1448.

Ku, Yunjung. 2019. "Collocations of English Emotion Expressions in High School Textbooks." , 20(4): 63–73.

Loewenstein, George. 2000. "Emotions in economic theory and economic behavior." American economic review, 90(2): 426–432.

Lucas, Richard E, Ed Diener, and Randy J Larsen. 2003. "Measuring positive emotions." In Positive Psychological Assessment: A Handbook of Models and Measures. , ed. Shane J Lopez and Charles Richard Snyder. American Psychological Association.

Ludwig, Jens, and Sendhil Mullainathan. 2024. "Machine learning as a tool for hypothesis generation." The Quarterly Journal of Economics, 139(2): 751–827.

Lundberg, Shelly. 2017. "Non-cognitive skills as human capital." Education, Skills, and Technical Change: Implications for Future US GDP Growth, 77.

Lutz, Catherine, and Geoffrey M White. 1986. "The anthropology of emotions." Annual Review of Anthropology, 405–436.

Marcus, George E. 2000. "Emotions in politics." Annual Review of Political Science, 3(1): 221–250.

Mayer, John D, and Glenn Geher. 1996. "Emotional intelligence and the identification of emotion." Intelligence, 22(2): 89–113.

Michalopoulos, Stelios, and Christopher Rauh. 2024. "Movies." , (Working Paper 32220). Accessed: 2024-08-06.

Michalopoulos, Stelios, and Melanie Meng Xue. 2021. "Folklore." The Quarterly Journal of Economics, 136(4): 1993–2046.

Nikolajeva, Maria. 2014. "33 Emotion ekphrasis: representation of emotions in children's picturebooks." Visual communication, 711.

Nikolajeva, Maria. 2017. "Emotions in picturebooks." In The Routledge companion to picturebooks. 110–118. Routledge.

Öhman, Arne. 2007. "Has evolution primed humans to "beware the beast"?" Proceedings of the National Academy of Sciences, 104(42): 16396–16397.

Öhman, Arne. 2009. "Of snakes and faces: An evolutionary perspective on the psychology of fear." Scandinavian journal of psychology, 50(6): 543–552.

Parker, James DA, Laura J Summerfeldt, Marjorie J Hogan, and Sarah A Majeski. 2004. "Emotional intelligence and academic success: Examining the transition from high school to university." Personality and individual differences, 36(1): 163–172.

Pieters, Rik, and Michel Wedel. 2004. "Attention capture and transfer in advertising: Brand, pictorial, and text-size effects." Journal of Marketing, 68(2): 36–50.

Plutchik, Robert, and Henry Kellerman. 1980. Emotion Theory, Research, and Experience. Academic press.

Rasmussen, Eric E, Autumn Shafer, Malinda J Colwell, Shawna White, Narissra Punyanunt-Carter, Rebecca L Densley, and Holly Wright. 2016. "Relation between active mediation, exposure to Daniel Tiger's Neighborhood, and US preschoolers' social and emotional development." Journal of Children and Media, 10(4): 443–461.

Rorty, Amélie Oksenberg, et al. 1980. Explaining Emotions. Vol. 22, University of California Press Berkeley.

Ruffman, Ted, Qiuyi Kong, Hui Mei Lim, Kangning Du, and Emilia Tiainen. 2023. "Recognition of facial emotions across the lifespan: 8-year-olds resemble older adults." British Journal of Developmental Psychology, 41(2): 128–139.

Russell, James A. 2003. "Core affect and the psychological construction of emotion." Psychological review, 110(1): 145.

Sacerdote, Bruce, Ranjan Sehgal, and Molly Cook. 2020. "Why is all COVID-19 news bad news?" National Bureau of Economic Research.

Smith, Vicky. 2013. "The 'Caldecott effect'." Children and Libraries: The Journal of the Association for Library Service to Children, 1(1): 9–13.

Solomon, Robert C. 1976. The Passions: The Myth and Nature of Human Emotion. Anchor.

Sullivan, Denis G, and Roger D Masters. 1988. "“Happy warriors”': Leaders' facial displays, viewers' emotions, and political support." American Journal of Political Science, 345–368.

Szasz, Teodora, Emileigh Harrison, Ping-Jung Liu, Ping-Chang Lin, Hakizumwami Birali Runesha, and Anjali Adukia. 2022. "Measuring representation of race, gender, and age in children's books: Face detection and feature classification in illustrated images." Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision.

Thelwall, Mike, Kevan Buckley, and Georgios Paltoglou. 2011. "Sentiment in Twitter events." Journal of the American Society for Information Science and Technology, 62(2): 406–418.

Thoits, Peggy A. 1989. "The sociology of emotions." Annual review of sociology, 15(1): 317–342.

Trampe, Debra, Jordi Quoidbach, and Maxime Taquet. 2015. "Emotions in everyday life." PloS One, 10(12): e0145450.

Underwood, Robert L, and Nicole M Klein. 2002. "Packaging as brand communication: Effects of product pictures on consumer responses to the package and brand." Journal of Marketing Theory and Practice, 10(4): 58–68.

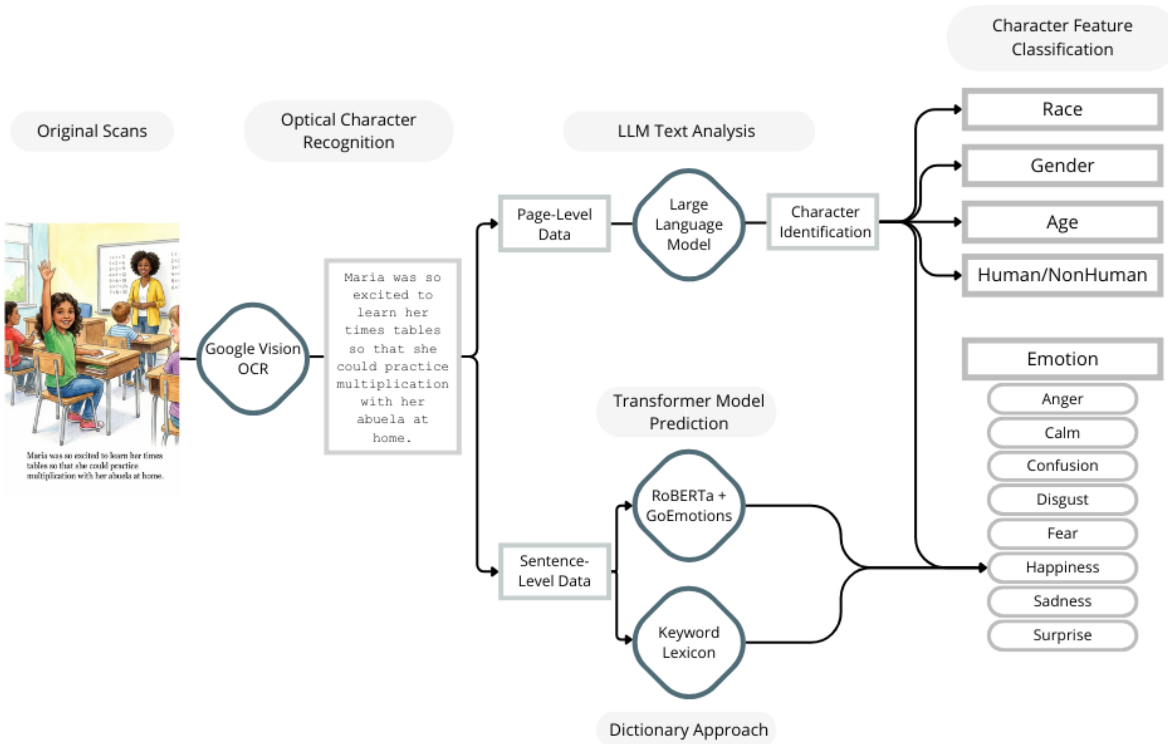
Widen, Sherri C, and James A Russell. 2008. "Children acquire emotion categories gradually." Cognitive development, 23(2): 291–312.

Yanagizawa-Drott, David H., and Hans-Joachim Voth. 2023. "Image(s)." Working Paper, University of Zurich.

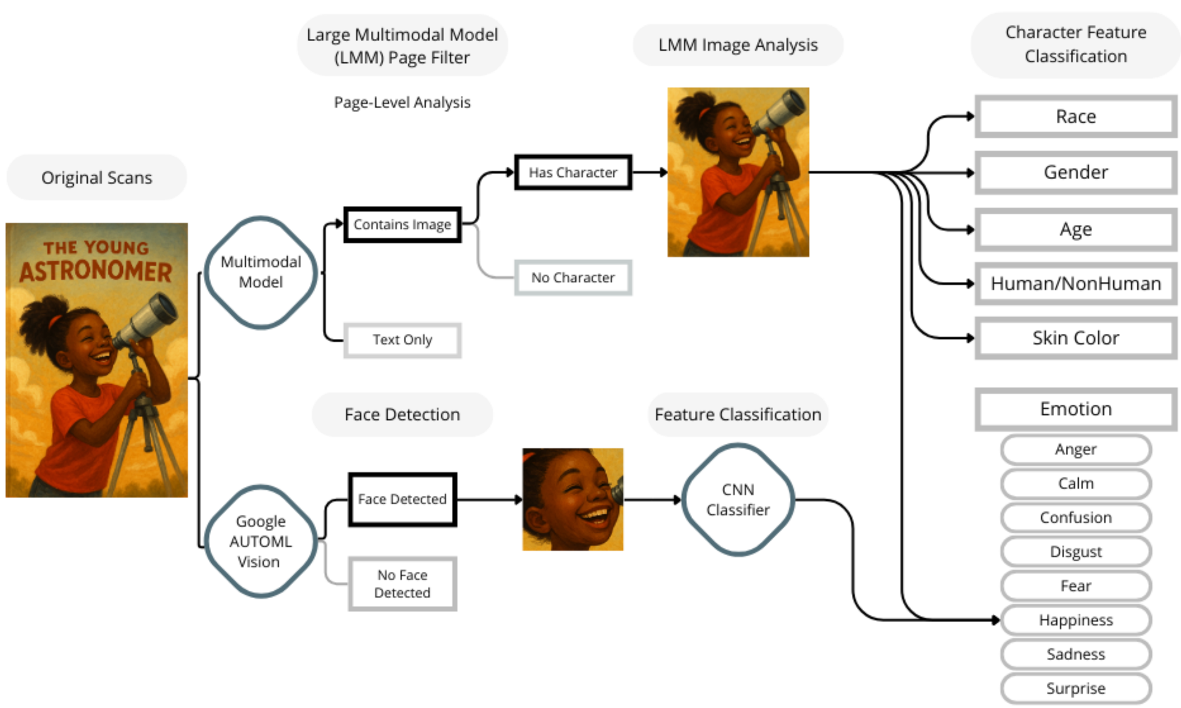
Ziems, Caleb, William Held, Omar Shaikh, Jiaao Chen, Zhehao Zhang, and Diyi Yang. 2024. "Can large language models transform computational social science?" Computational Linguistics, 50(1): 237–291.

FIGURE 1
Data Extraction Pipelines

(a) Text Analysis

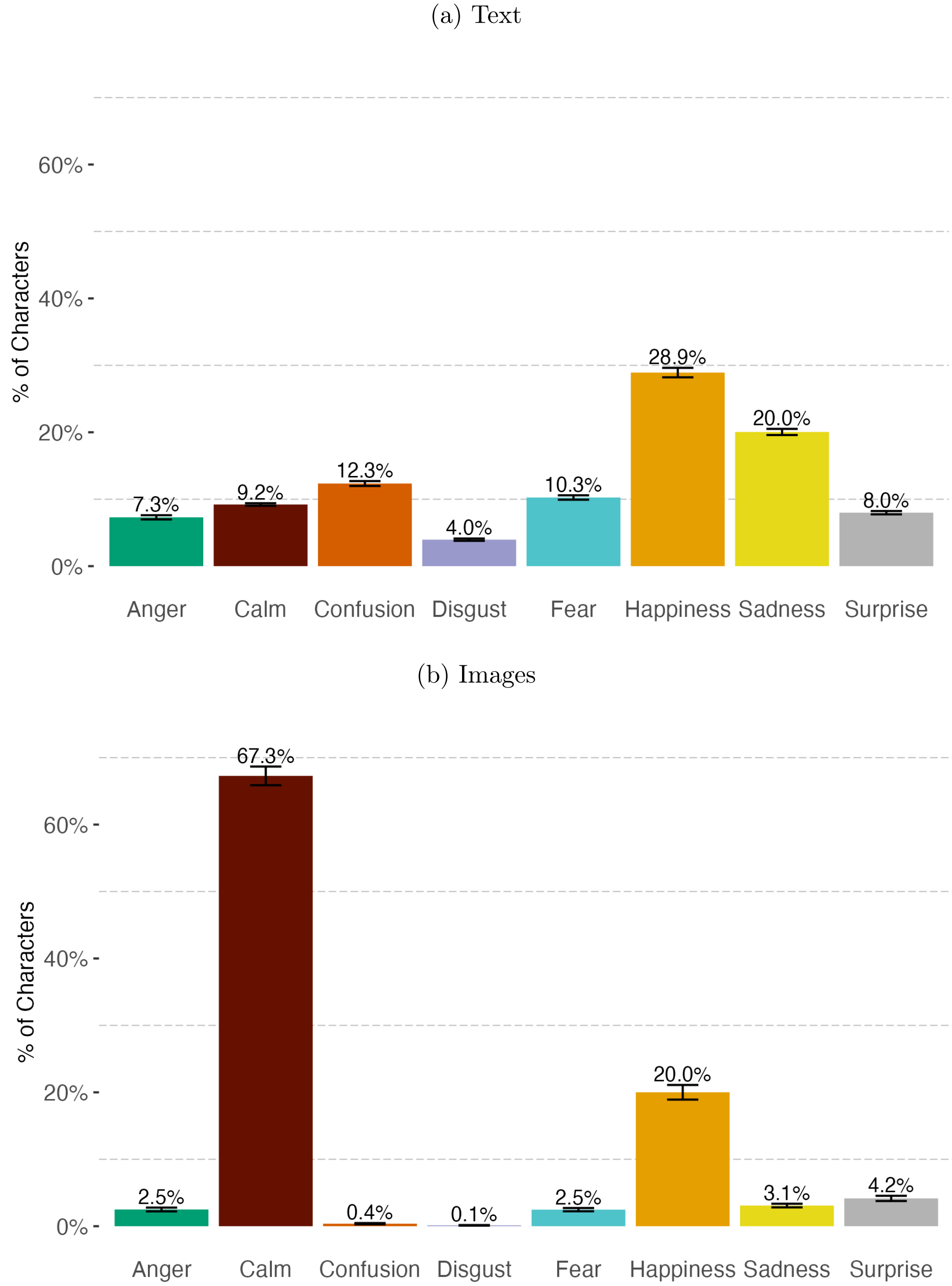


(b) Image Analysis



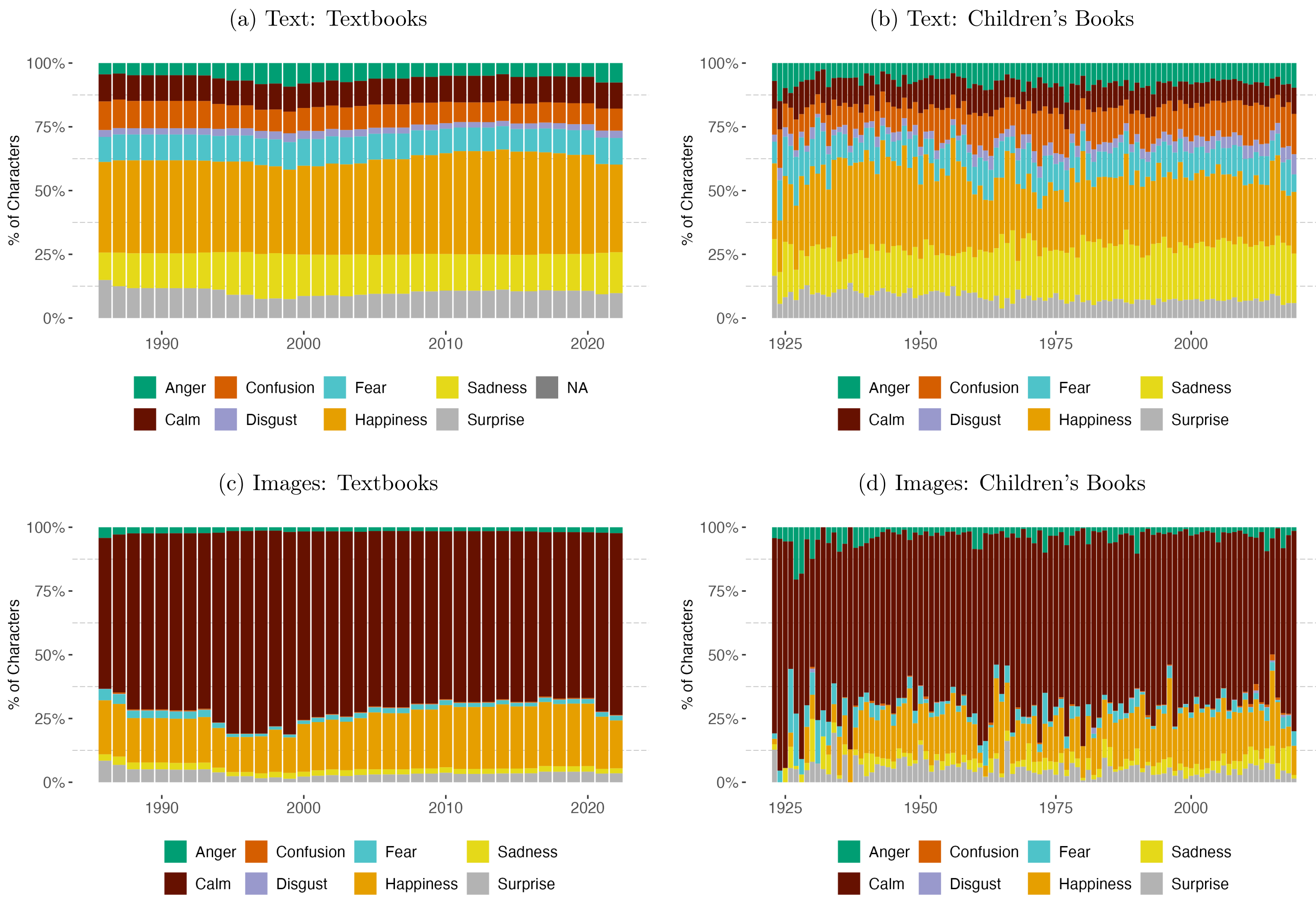
Notes: This figure illustrates the steps in our text analysis (top) and image analysis (bottom) pipelines beginning with the raw scan of each page in our sample of books.

FIGURE 2
Overall Representation of Emotions



Notes: This figure shows the aggregate displays of emotions in text and images. The top panel shows the percentage of emotions conveyed by characters as a share of all emotions detected in the text of the books (literature and textbooks) using a large language model. The bottom panel shows the percentage of emotions conveyed by faces as a share of all emotions detected in the images of books (literature and textbooks) using a large multimodal model. Error bars depict 95% confidence intervals constructed from standard errors that are clustered at the book level.

FIGURE 3
Representation of Emotions Over Time



Notes: This figure shows the percentage of emotions detected over time. Panel (a) shows the share of all emotions detected in the sample of textbooks, and panel (b) shows the shares in literature. Emotions are detected using OpenAI. Emotions for each textbook are counted for each year the textbook is in use, and emotions for each children's book are counted for the book's copyright year. Panels (c) and (d) show the analogous results for emotions detected in faces in both corpora. Children's books are associated with their respective copyright years, whereas textbooks are associated with each year they are in use. Note that textbooks are in use for multiple years, whereas the children's books only appear one time in the time series panel.

FIGURE 4

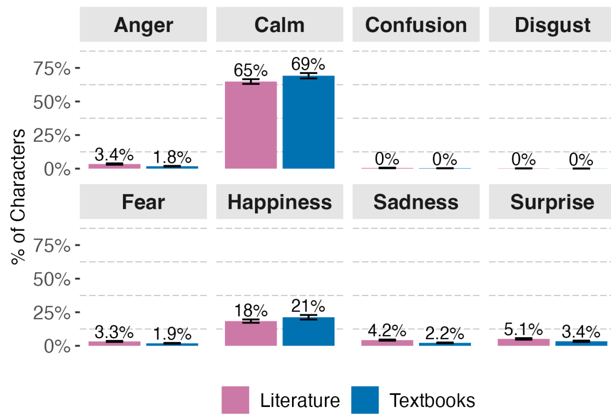
Representation of Emotions in the Text of Different Subcollections

Notes: This figure shows the percentage of emotions conveyed by characters as a share of all emotions detected by OpenAI separately for different subcollections. Panel (a) compares textbooks and children's books, panel (b) compares California and Texas textbooks, panel (c) compares science, reading, and social studies textbooks, panel (d) compares grade three and grade five textbooks, and panel (e) compares the Mainstream and Diversity collections of children's books. Error bars depict 95% confidence intervals constructed from standard errors that are clustered at the book level.

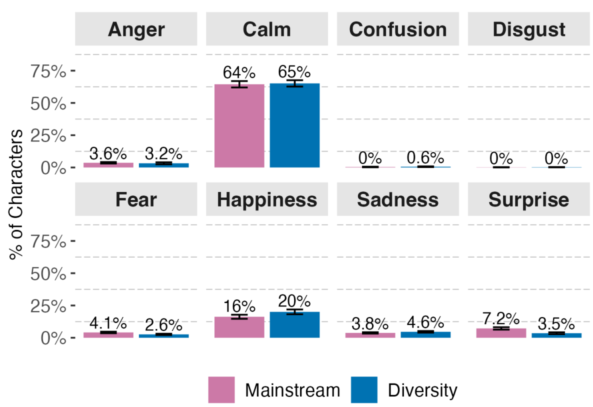
FIGURE 5

Representation of Emotions in the Images of Different Subcollections

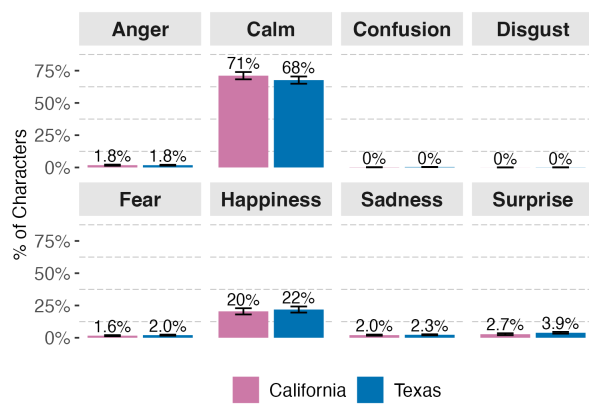
(a) Children's Literature vs. Textbooks



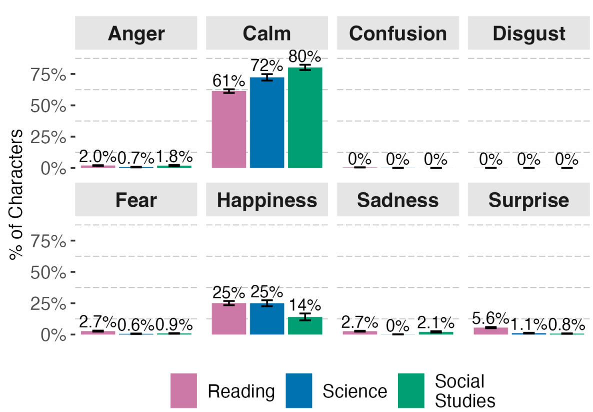
(b) Mainstream vs. Diversity



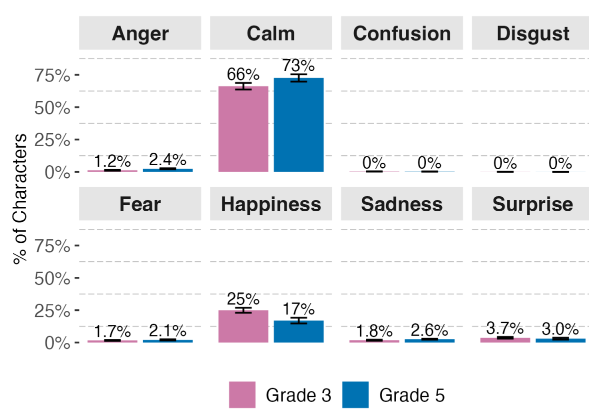
(c) California vs. Texas Textbooks



(d) By Textbook Subject



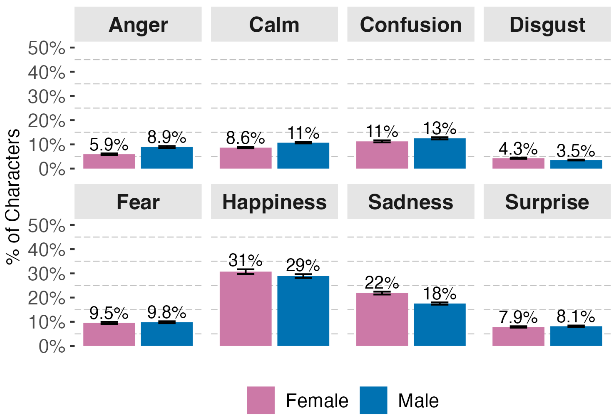
(e) By Textbook Grade



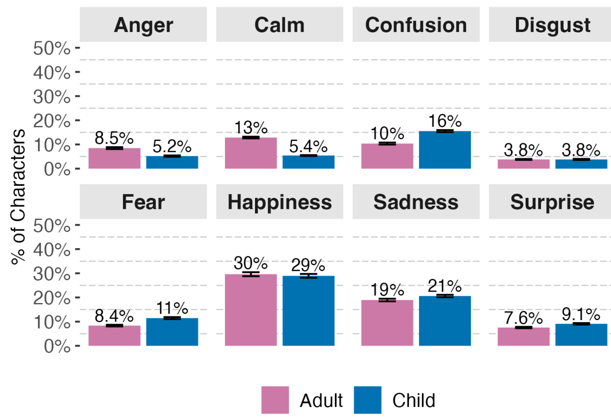
Notes: This figure shows the percentage of emotions conveyed by detected faces as a share of all emotions detected by an LMM separately for different subcollections. Panel (a) compares textbooks and children's books, panel (b) compares California and Texas textbooks, panel (c) compares science, reading, and social studies textbooks, panel (d) compares grade three and grade five textbooks, and panel (e) compares the Mainstream and Diversity collections of children's books. Error bars depict 95% confidence intervals constructed from standard errors that are clustered at the book level.

FIGURE 6
Displays of Emotions in the Text for Characters with Different Identities

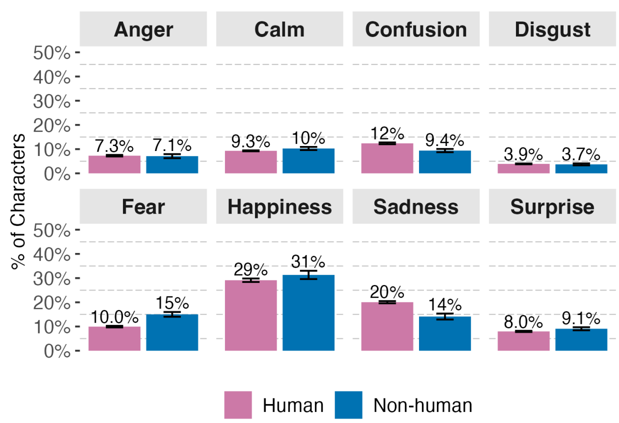
(a) Gender



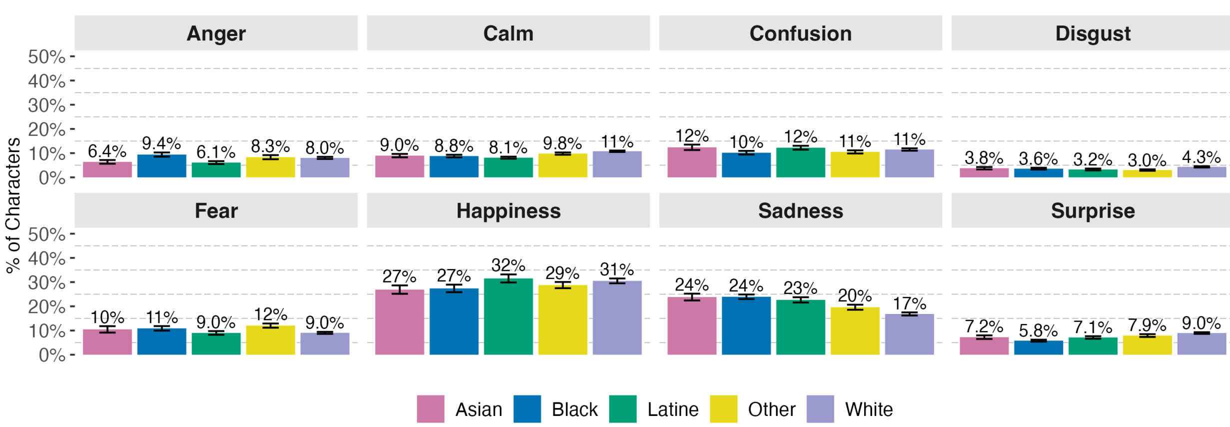
(b) Age



(c) Human vs. Non-Human

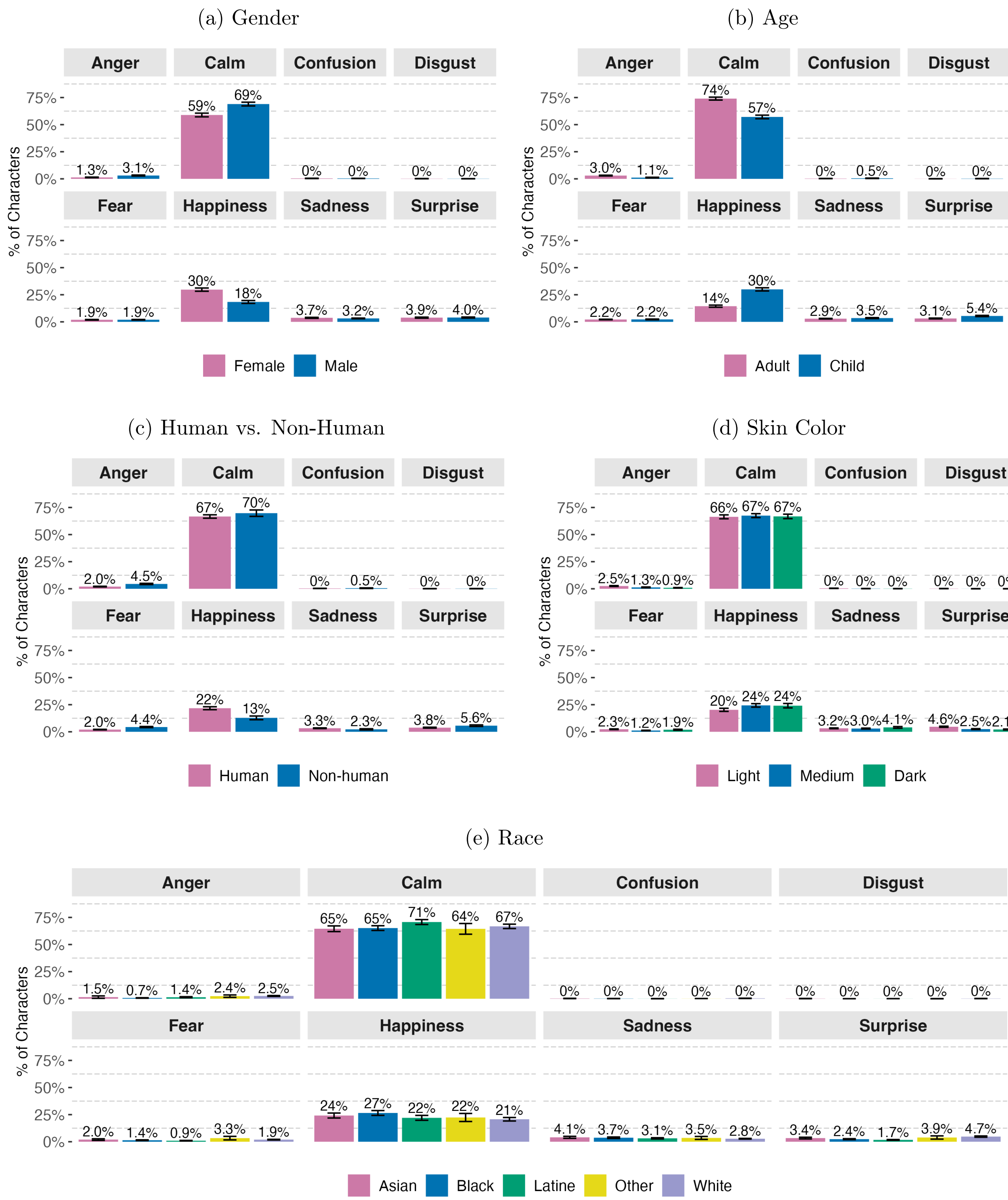


(d) Race



Notes: This figure shows the percentage of emotions conveyed in the text by characters of different identities. Panel (a) compares female and male characters, panel (b) compares human characters of different races, panel (c) compares characters of different ages, and panel (d) compares human and non-human characters. Error bars depict 95% confidence intervals constructed from standard errors that are clustered at the book level.

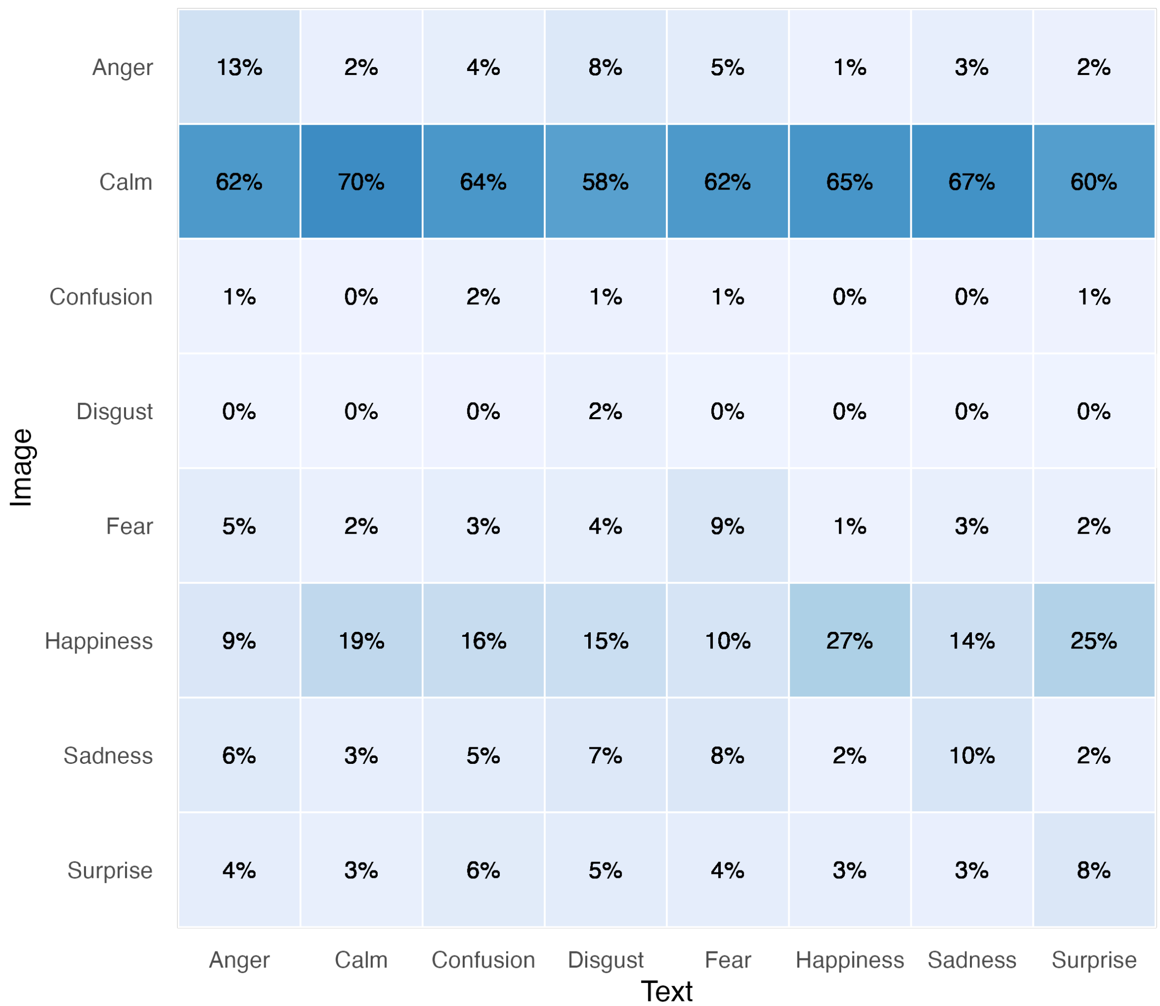
FIGURE 7
Displays of Emotions in the Faces of Characters with Different Identities



Notes: This figure shows the percentage of emotions conveyed in the detected faces of characters of different identities. Panel (a) compares female and male characters, panel (b) compares human characters of different races, panel (c) compares characters of different ages, panel (d) compares human and non-human characters, and panel (e) compares characters with different skin color. Error bars depict 95% confidence intervals constructed from standard errors that are clustered at the book level.

FIGURE 8

Normalized Co-occurrence Matrix



Note: This figure shows the co-occurrence of emotions for text and images on the same page, normalized by the text columns. In other words, each cell shows the proportion of times a specific image label occurred, given a particular text label. For example, the top left cell shows that when anger is found in text, an image on the same page will depict anger 22% of the time.

TABLE 1
Book Purchases

	Dependent variable: Whether Purchased		
	(1)	(2)	(3)
Happy or Calm Only Cover	0.160***	0.554***	0.667***
	(0.025)	(0.031)	(0.031)
No Cover Character	1.186***	0.490***	0.542***
	(0.024)	(0.030)	(0.031)
Book Controls		X	X
Linear Emotion Controls		X	
Quintile Emotion Controls			X
Observations	28,521	28,521	28,521
AIC	351,716.6	296,521.2	295,216.7
BIC	351,745.1	298,218.3	296,999.3
RMSE	0.05	0.05	0.05

Notes: This table presents results from a logistic regression of the binary purchase outcome on cover emotion measures. The sample is restricted to literature books and characters that express a defined emotion. "Happy or Calm Only Cover" equals 1 when the cover's characters display only calmness or happiness; "No Cover Character" equals 1 when no characters are depicted on the cover; the omitted group is books that contain at least one character who portrays an emotion other than happy or calm. Book controls include a variety of book characteristics including indicator variables for the award(s) the book won, the book's copyright year, the number of pages, average words per page, whether the book is a picture book (characters with emotions only appear in images), and whether it contains no images (characters with emotions appear only in text). Observations refer to the number of times we observe a unique book being purchased by a household across all households. We cluster standard errors by household. In this table, ***, **, and * represent significance at the 0.01, 0.05, and 0.1 levels, respectively.

Researcher(s)' own analyses calculated (or derived) based in part on data from Market Track, LLC dba Numerator and marketing databases provided through the Numerator Datasets at the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Numerator data are those of the researcher(s) and do not reflect the views of Numerator. Numerator is not responsible for and had no role in analyzing or preparing the results reported herein. LLMs and AI tools are never applied to the Numerator data.

TABLE 2
Library Inventory

	Dependent variable:		
	Number of Books in Branch Inventory		
	(1)	(2)	(3)
Happy or Calm Only Cover	0.1334***	0.1174***	0.0760***
	(0.0191)	(0.0280)	(0.0270)
No Cover Character	0.1213***	0.1839***	0.1366***
	(0.0206)	(0.0420)	(0.0372)
Book Controls		X	X
Linear Emotion Controls		X	
Quintile Emotion Controls			X
Observations	6,944	6,944	6,944
Squared Correlation	0.00313	0.38340	0.38359
Pseudo R^2	0.00132	0.20337	0.20350
BIC	16,514.9	14,196.2	14,247.1

Notes: This table presents results from a linear regression of the library checkout outcome on cover and content emotion measures, controlling for the distribution of emotions in both text and images. The sample is restricted to literature books and characters that express a defined emotion. "Happy or Calm Only Cover" equals 1 when the cover's characters display only calmness or happiness; "No Cover Character" equals 1 when no characters are depicted on the cover; the omitted group is books that contain at least one character who portrays an emotion other than happy or calm. Book controls include a variety of book characteristics including indicator variables for the award(s) the book won, the copyright year of the book, the number of pages, average words per page, whether the book is a picture book (characters with emotions only appear in images), and whether it contains no images (characters with emotions appear only in text). We also add fixed effects for the specific library and cluster standard errors by specific library. In this table, ***, **, and * represent significance at the 0.01, 0.05, and 0.1 levels, respectively.

TABLE 3

Relationship Between the Book Cover and Content

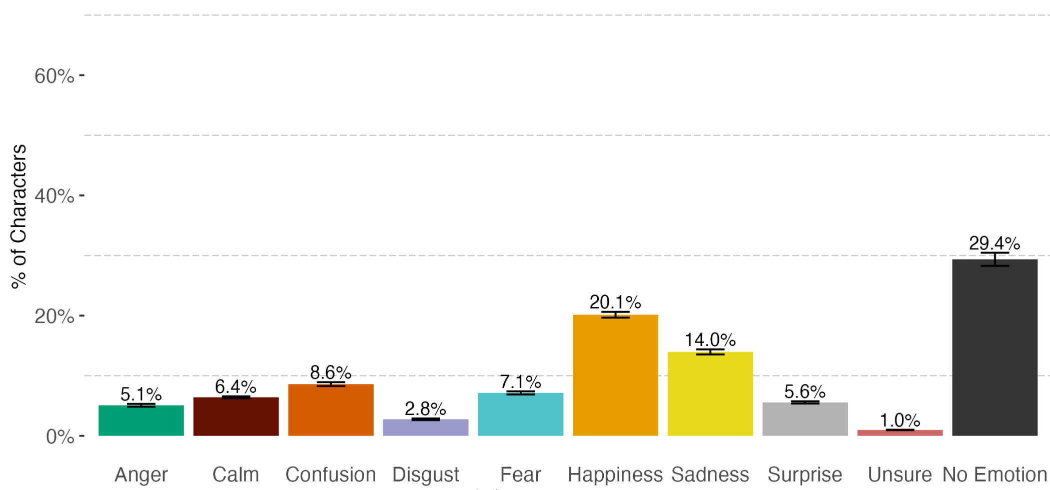
	Book Cover:		
Proportion Happy or Calm in Book Content:	Happy or Calm Cover (1 = yes) (1)	Mixed Cover (1 = yes) (2)	Not Happy or Calm Cover (1 = yes) (3)
Images	3.420***	-3.370***	-5.520***
	(0.471)	(0.867)	(0.982)
Text	2.202***	1.403	-3.510*
	(0.659)	(1.250)	(1.803)
Book Controls	X	X	X
Observations	1,113	1,113	1,113
AIC	1,310.5	557.3	461.4
BIC	1,917.3	1,164.1	1,068.2
Log Likelihood	-534.246	-157.658	-109.700
RMSE	0.40	0.20	0.17

Notes: This table presents results from a logistic regression where the outcome in column one indicates whether a book cover has only happy or calm characters, the outcome in column two indicates whether a book cover has a mix of happy or calm characters and characters that are neither happy nor calm, and the outcome in column three indicates whether a book cover has exclusively characters that are neither happy nor calm. The two variables shown are the proportion of happy or calm characters in images and text in the entire book. The book controls include the number of pages, average words per page, whether there are no characters in the text or images with an emotion, and indicator variables for the award(s) the book won and the copyright year. In this table, ***, **, and * represent significance at the 0.01, 0.05, and 0.1 levels, respectively.

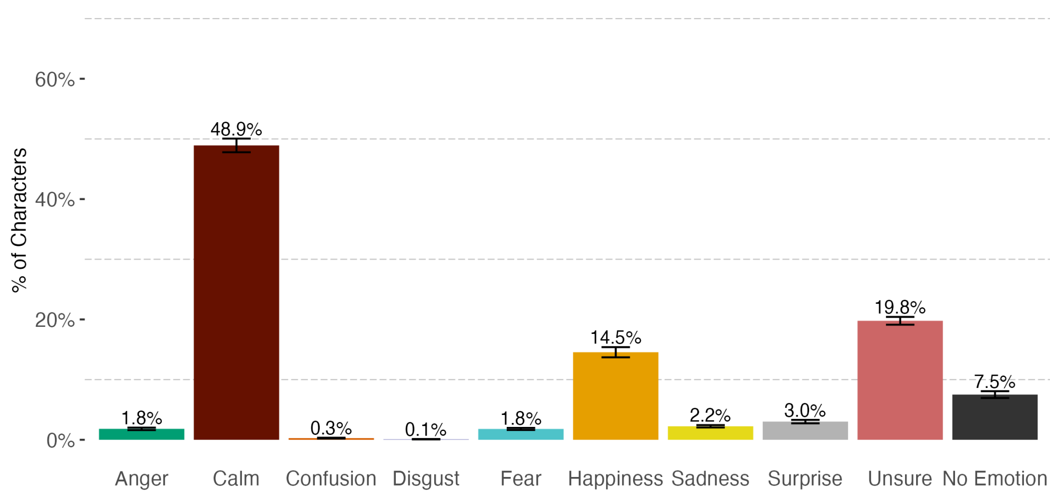
A Appendix Figures

FIGURE A.1
Overall Representation of Emotions (Including No Emotion and Unsure)

(a) Text



(b) Images



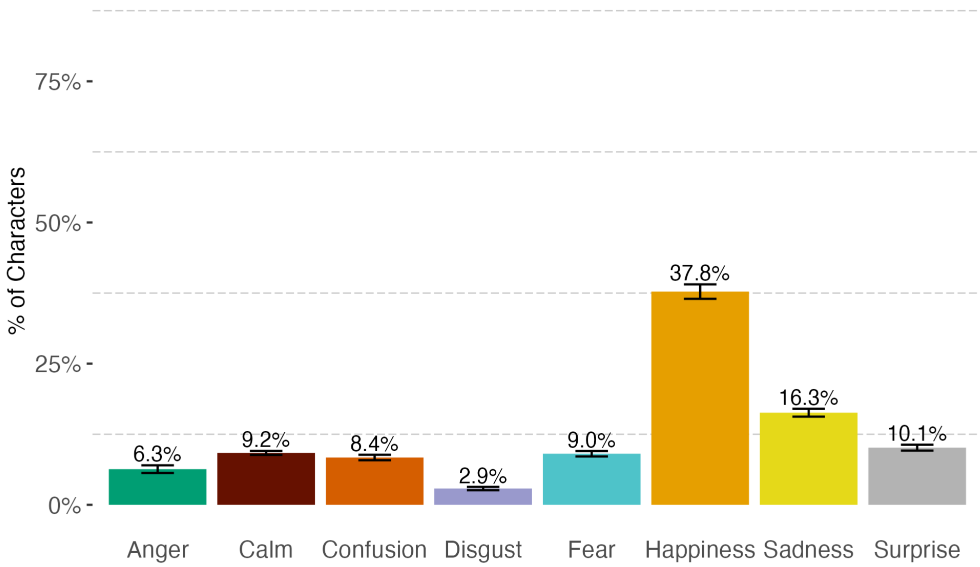
Notes: This figure replicates Figure 2 with additional categories for "No Emotion" and "Unsure." The top panel shows the percentage of emotions conveyed by characters as a share of all emotions detected in the sample of books (literature and textbooks). Emotions are detected using an LMM. The bottom panel shows the percentage of emotions conveyed by faces as a share of all emotions detected in the sample of books (literature and textbooks). Emotions are detected using an LMM.

FIGURE A.2
Representation of Emotions Over Time (including No Emotion and Unsure)

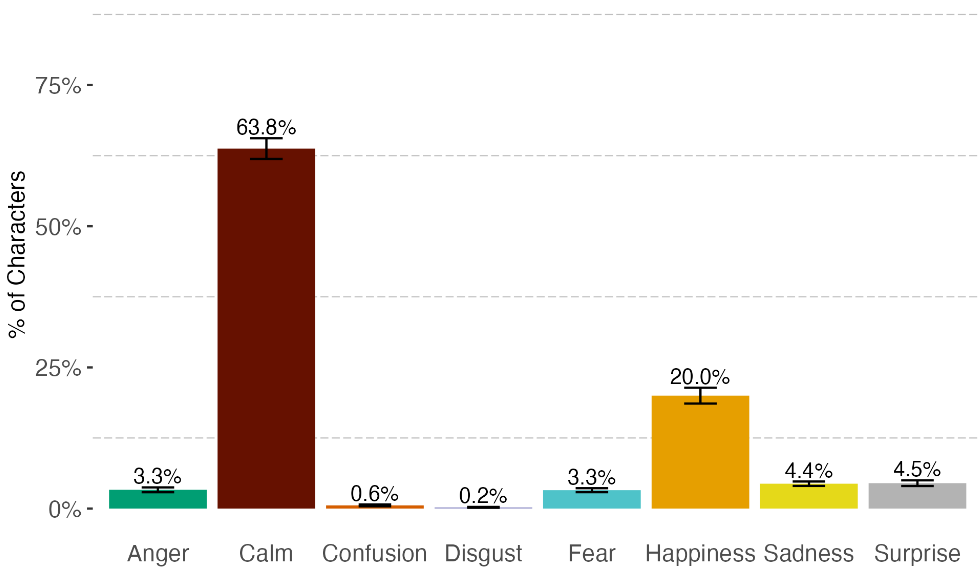
Notes: This figure replicates Figure 3 with additional categories for "No Emotion" and "Unsure." Panel (a) shows the share of all emotions detected in the sample of textbooks, and panel (b) shows the shares in literature. Emotions are detected using OpenAI. Emotions for each textbook are counted for each year the textbook is in use, and emotions for each children's book are counted for the book's copyright year. Panels (c) and (d) show the analogous results for emotions detected in faces in both corpora. Children's books are associated with their respective copyright years, whereas textbooks are associated with each year they are in use. Note that textbooks are in use for multiple years, whereas the children's books only appear one time in the time series panel.

FIGURE A.3
Overall Representation of Emotions (Matched Sample)

(a) Text

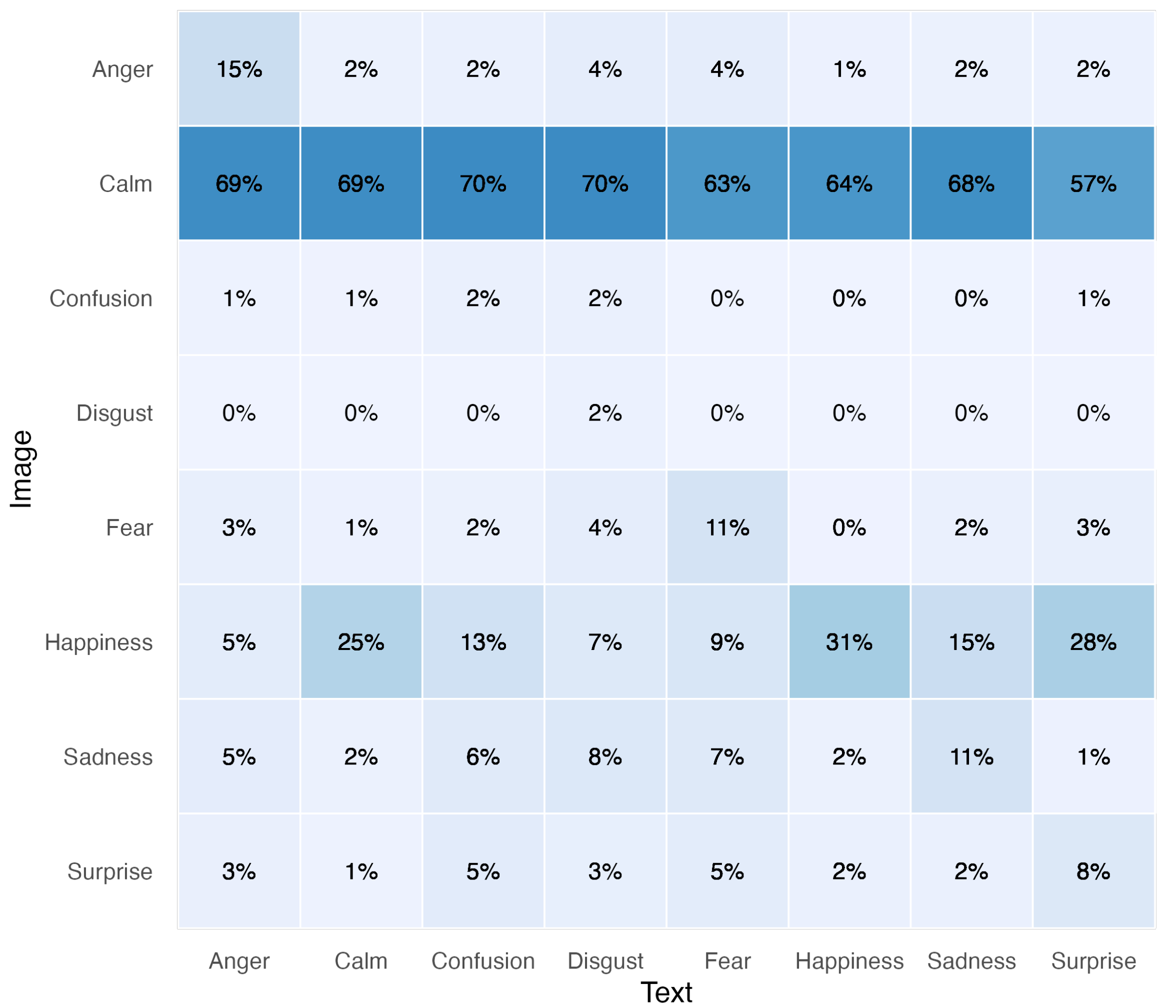


(b) Images



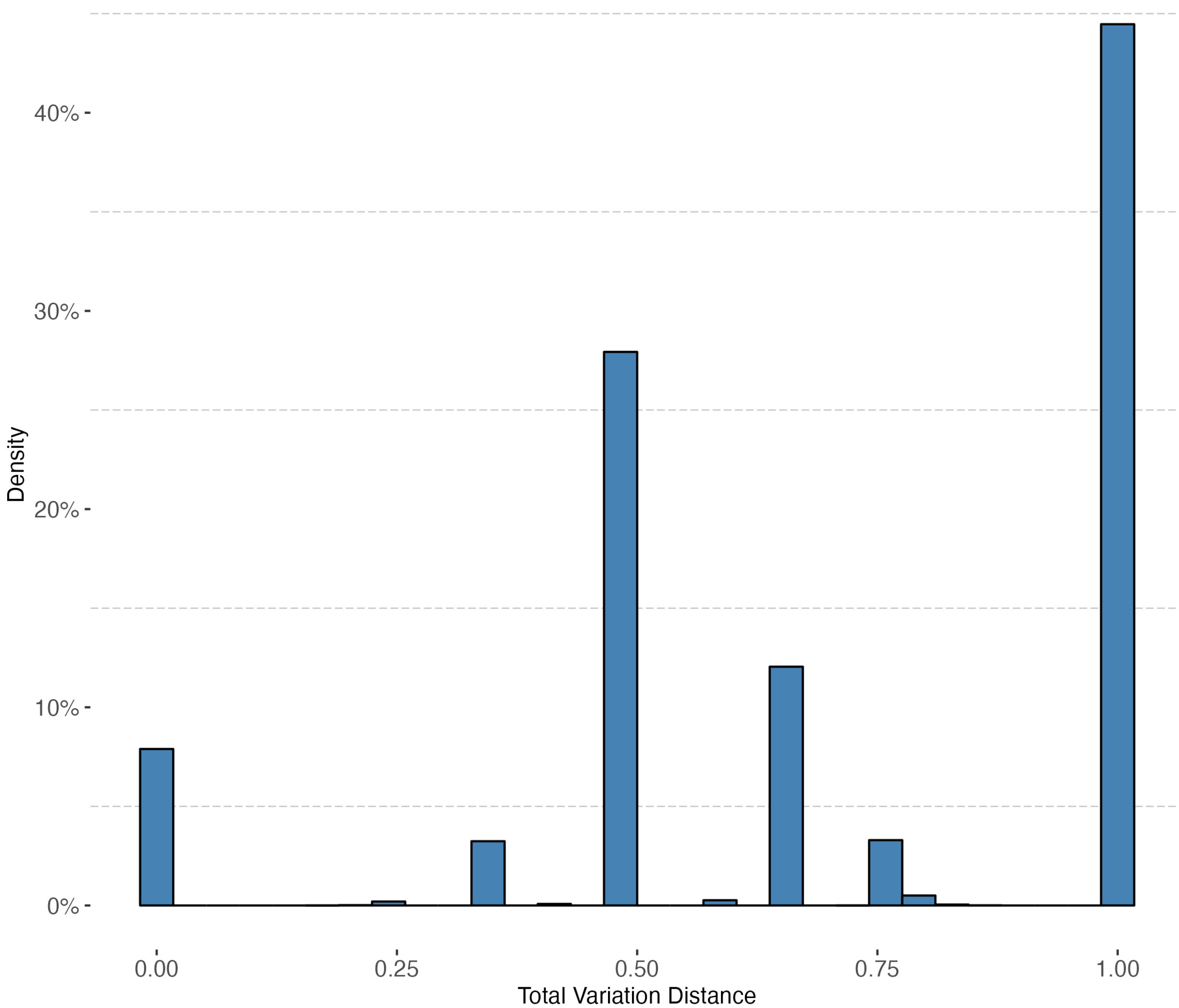
Notes: This figure shows the aggregate displays of emotions in text and images for the sample of matched pages. The top panel shows the percentage of emotions conveyed by characters as a share of all emotions detected in the sample of books (literature and textbooks); emotions are detected using an LMM. The bottom panel shows the percentage of emotions conveyed by faces as a share of all emotions detected in the sample of books (literature and textbooks); emotions are detected using an LMM.

FIGURE A.4
Normalized Co-occurrence Matrix for Characters Matched in Text and Images



Note: This figure shows the co-occurrence of emotions for text and images on the same page, where unique human characters are matched on their race and gender. Co-occurrence is normalized by the text columns. In other words, each cell shows the proportion of times a specific image label occurred, given a particular text label. For example, the top left cell shows that for pages where there is one Black male character in the text and the image, when the character is shown as angry in text,the matched character will be depicted as angry in an image only 15% of the time.

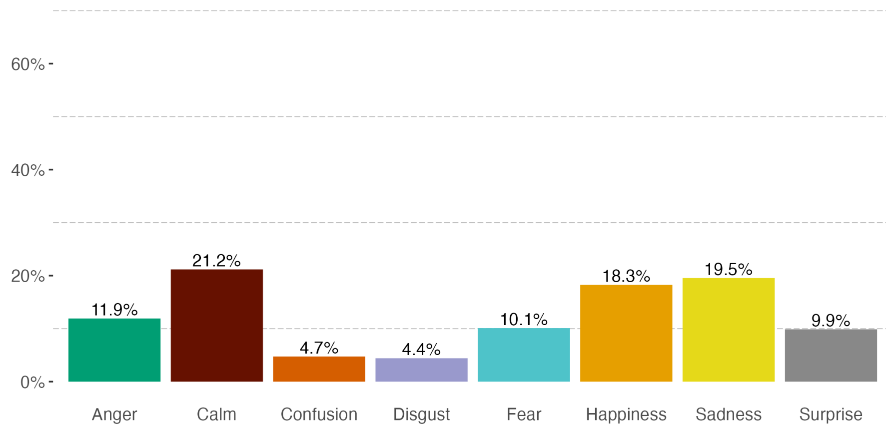
FIGURE A.5
Total Variation Distance Between Text and Images



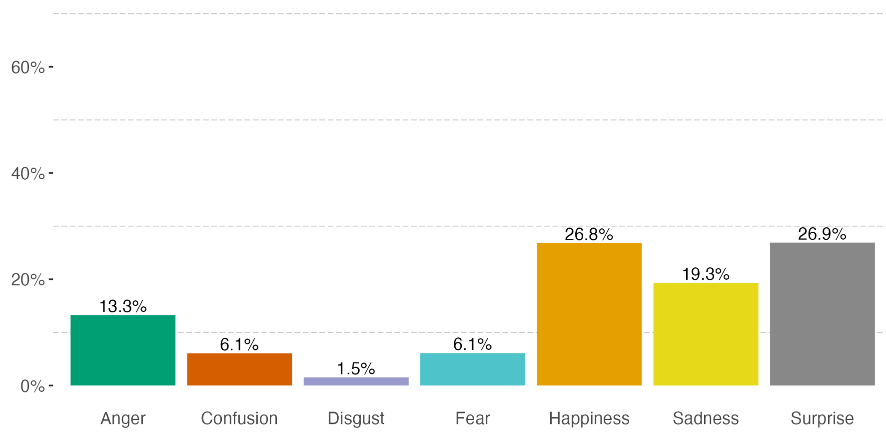
Note: This figures shows the total variation distance (tvd) between the emotional content of images and text on the same page. We calculate tvd by taking the l^1 distance between the empirical distributions of our 8 emotions in text and in images on a given page and multiply by 1/2.

FIGURE A.6
Robustness: Overall Representation of Emotions

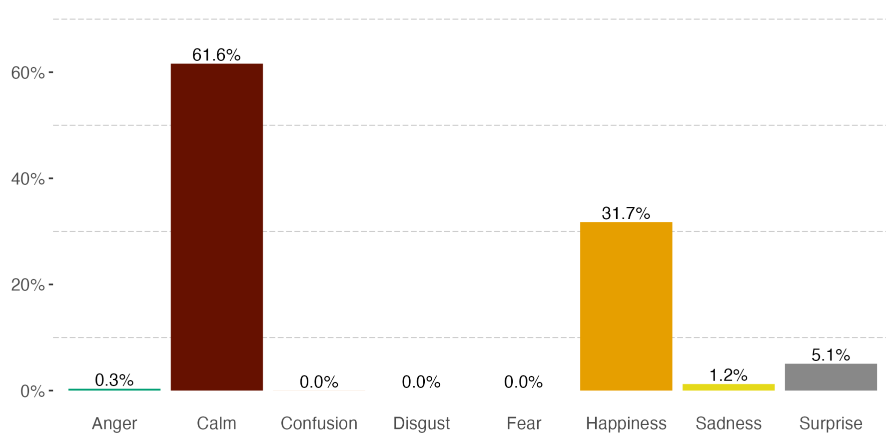
(a) Text: Lexicon Approach



(b) Text: GoEmotions AI Model

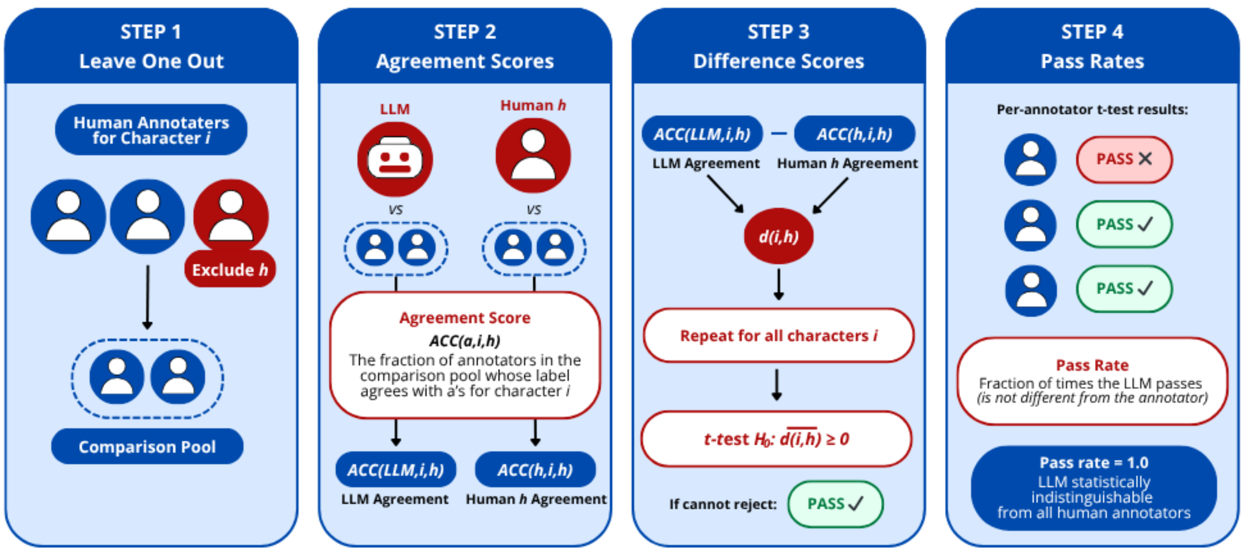


(c) Images: Feature Classification



Notes: This figure shows the aggregate displays of emotions in text and images using alternative approaches. Panel (a) shows the percentage of emotion-specific words as a share of all emotion words from a pre-specified lexicon. Panel (b) shows the percentage of emotion-specific sentences as a share of all emotion sentences as classified by the GoEmotions model. Panel (c) shows the percentage of detected faces.

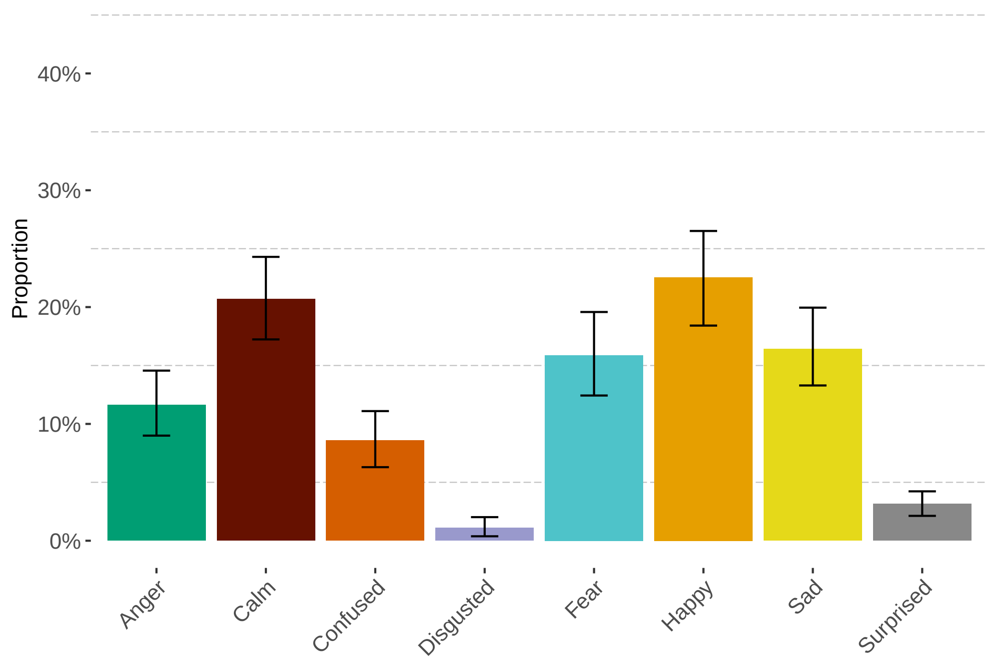
FIGURE A.7
Alternative Annotator Test Procedure



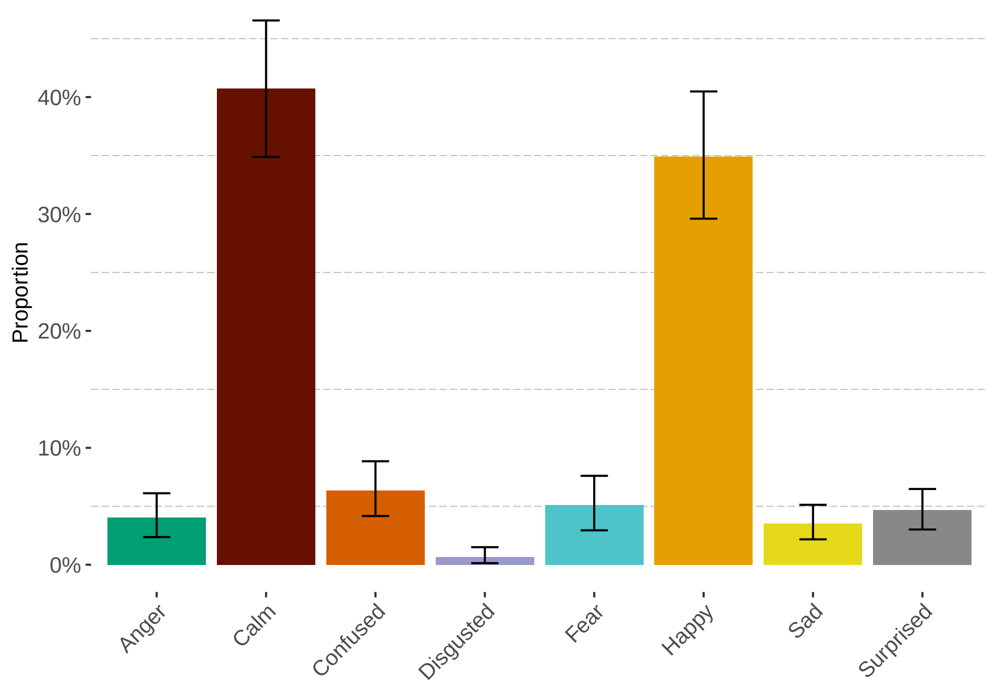
Note: This figures shows the procedure used to validate the AI output. The process involves assessing the model's reliability relative to human annotators in terms of its relative capacity to agree with other raters on the same set of examples.

FIGURE A.8
Simulated Distribution of Human Labels

(a) Text



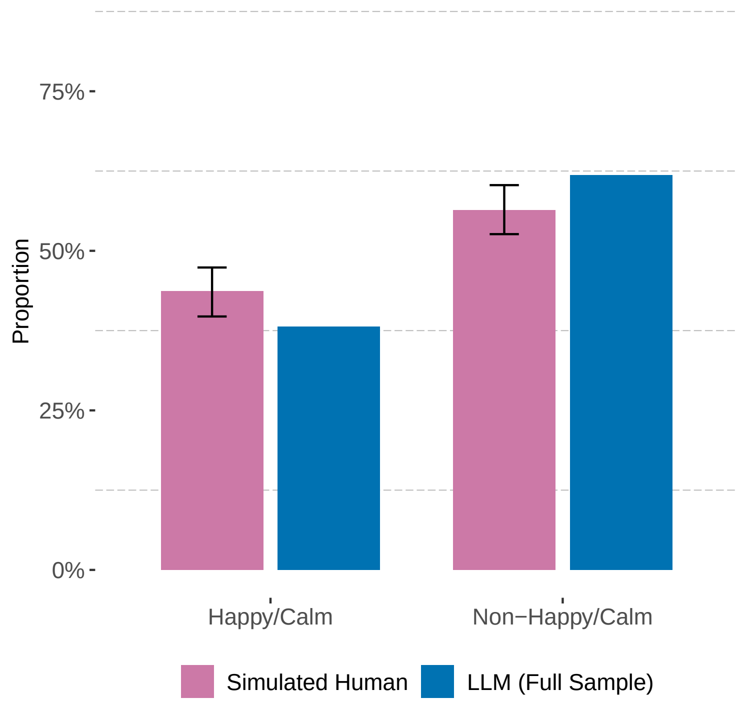
(b) Images



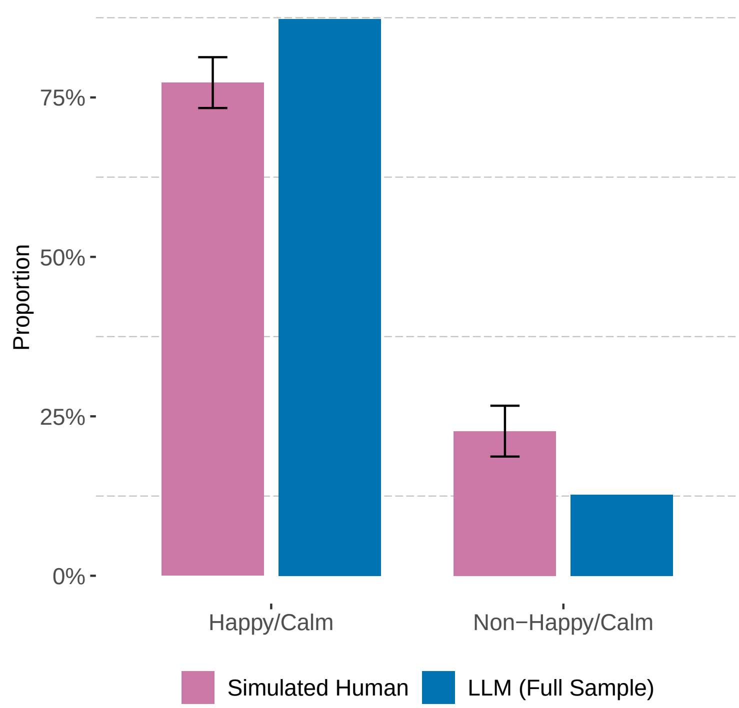
Notes: This figure shows the distribution of emotions predicted to be obtained by human labelers on the full sample. The simulation procedure is described in Section V.B. The top panel shows the distribution in text, and the bottom shows the distribution in images.

FIGURE A.9
Simulated Human Labels versus Models on Happy/Calm and Non-Happy/Calm

(a) Text

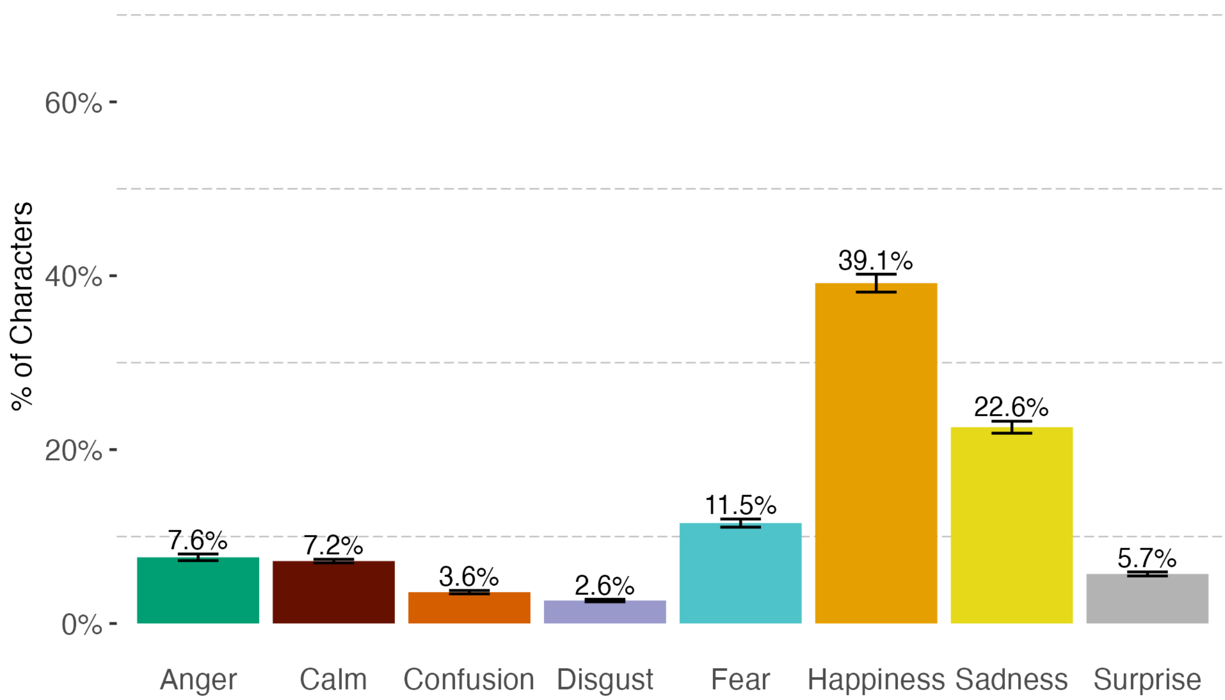


(b) Images



Notes: This figure compares the human and model predictions on the Happy/Calm and Non-Happy/Calm split. The top panel shows the distribution in text, and the bottom shows the distribution in images.

FIGURE A.10
Overall Representation of High Confidence Emotions In Text



Note: This figure shows the aggregate displays of emotions in text when we subset to emotion labels that the model has high confidence in (≥ 0.8). Error bars depict 95% confidence intervals constructed from standard errors that are clustered at the book level.

B Appendix Tables

TABLE B.1
Sample of Mapping from Keyword Lexicon to Emotions in Main Analysis

anger	**calm**	**confusion**	**disgust**	**fear**	**happiness**	**sadness**	**surprise**
acerb	abide	absentminde	abhor	affright	affection	abandon	aah
acerbic	amiable	absurd	abhorrence	afraid	amuse	abandonment	aflutter
acerbity	aplomb	absurdity	abhorrent	aghast	amusing	ache	ah
acrid	asleep	addle	abhorringly	alarm	banter	afflict	aha
acrimonious	assuage	baffle	abominable	alarum	beam	affliction	ahha
...	...	...	...	...	...	...	...

Note: This table presents a sample of our lexicon dictionary. Each of the eight emotion categories is associated with a mutually exclusive set of keywords. The full lexicon can be viewed here

TABLE B.2
Mapping from GoEmotions to Emotions in Main Analysis

anger	**confusion**	**disgust**	**fear**	**happiness**	**sadness**	**surprise**
anger	confusion	disgust	fear	amusement	embarrassment	curiosity
annoyance			nervousness	excitement	disappointment	realization
disapproval				joy	grief	surprise
				love	remorse	
				optimism	sadness	

Note: This table presents the mapping used in our main analysis to group the 28 fine-grained GoEmotions into the eight broader emotion categories used in our main analysis. We exclude the following fine-grained emotions from the joy category: admiration, approval, caring, desire, gratitude, pride, relief.

TABLE B.3
Macro F1 Scores

Prediction	**Macro F1**	**Number of Examples**	**Exclusion Rate**
Panel A: Text			
Primary Emotion	0.289	588	35.0%
Emotion (Binarized)	0.467	588	11.7%
Is Human	0.862	564	0.7%
Gender	0.802	554	3.4%
Age	0.680	588	7.8%
Race	0.531	532	7.3%
Panel B: Image			
Primary Emotion	0.480	459	22.9%
Emotion (Binarized)	0.640	459	5.7%
Is Human	0.941	459	0.9%
Gender	0.821	459	2.6%
Age	0.782	368	3.8%
Race	0.685	305	12.5%
Skin Color	0.636	333	10.2%

Notes: This table reports macro F1 scores, which are obtained by computing the F1 score for each class and averaging over classes. The number of examples changes because raters may find demographic characteristics to be not applicable (i.e., the race of a non-human character). Examples are excluded if all three human raters disagree so that there is no modal rater. The high exclusion rate for text and image emotion labeling reflects the relative difficulty of this task.

TABLE B.4
Alternative Annotator Test

Prediction	**Avg Advantage**	**Pass Rate** ($q = 0.10$)	**Pass Rate** ($q = 0.05$)
Panel A: Text			
Primary Emotion	-9.3%	0.333	0.333
Emotion (Binarized)	-8.5%	0.333	0.333
Is Human	-0.8%	1.000	1.000
Gender	5.1%	1.000	1.000
Age	5.2%	1.000	1.000
Race	-11.3%	0.333	0.333
Panel B: Image			
Primary Emotion	-3.9%	0.500	0.833
Emotion (Binarized)	-6.3%	0.167	0.667
Is Human	2.3%	1.000	1.000
Gender	-1.1%	1.000	1.000
Age	-7.2%	0.333	0.500
Race	1.3%	1.000	1.000
Skin Color	-4.8%	0.500	0.500

Notes: This table reports the results from the alternative annotator test. The top panel shows the results for the text predictions, and the bottom panel shows the results for the image predictions. Binarized emotion is constructed from the primary emotion label by binning into Happy/Calm and Non-Happy/Calm categories (along with a No Emotion/Unsure category).

C Methods Appendix

C.A Text Analysis

LMM Prompt Used for Emotion Identification Our primary text model is OpenAI's GPT-4o-mini, version 2024-07-18, accessed through the API in April 2025 with temperature set to 1.0 and the system prompt: "You are a knowledgeable and unbiased judge." For each page of OCR-extracted text, we submitted a standardized prompt which instructed the model to classify each character along with their age, gender, race, whether or not they were human, and any emotions expressed by that character. Emotion labels were recorded under three taxonomies: the six Ekman emotions, our eight-category MiiE taxonomy, and the 28-category GoEmotions taxonomy, each with a confidence score. The prompt restricted responses to predefined answer sets, allowed "unsure" and "not applicable" values for ambiguous cases, required "no emotion" for characters without expressed affect, and instructed the model not to assign emotions or attributes unsupported by the text. Hallucinations of emotions beyond simple morphological variants account for less than 1% of all observations, indicating high prompt adherence. The full prompt can be viewed at `https://github.com/miielab/emotions_paper`. This gives us a final dataset containing character-page level emotions.

Alternative Text Analysis Approach: Using a Lexicon-Based Approach To create our lexicon, we primarily use the synsets of WordNet. A synset is a collection of cognitive synonyms that convey the same concept. Some words have a single synset, while others have multiple. For example, the word "anger" is part of the following synsets: 1) a feeling that is oriented toward some real or supposed grievance; 2) the state of being angry; and 3) belligerence aroused by a real or supposed wrong (personified as one of the deadly sins). Hence, with an initial emotion-related search term, we can access a comprehensive network of semantic relations.

We constructed an initial set of WordNet inputs, extract their synonyms, and manually review the results. Following this, we engineer a prompt to elicit more emotional synonyms from an LLM. Our final lexicon maps these synonyms to our 8 categorical emotions (confusion, anger, sadness, surprise, calm, happiness, fear, and disgust). In the process, we lemmatize all words using SpaCy and, through manual inspection, ensure that all categories are mutually exclusive and each keyword unique. Lemmatization is chosen in order to simplify the lexicon construction at the cost of less flexibility over word choice. Table B.1 shows a sample of the dictionary. The full lexicon can be viewed at `https://github.com/miielab/emotions_paper`.

In total, our lexicon contains 1,706 entries. More specifically, breaking this down into the number of synonyms for each emotion, we have 253 anger, 256 calm, 191 confusion, 224 disgust, 173 fear, 204 happiness, 246 sadness, and 159 surprise entries.

The sample of textbooks and children's literature is then processed by splitting it into sentences using SpaCy and then cleaned by removing certain punctuation and digits, by making it lowercase and by lemmatizing it. This gives us 3,361,503 sentences across 1,306 books. Subsequently, we apply a count vectorizer to each sentence, allowing us to identify if a sentence contains one or more keywords and the number of times it appears. This gives us a final dataset containing sentence level emotions.

Alternative Approach: Using an Emotions AI-Based Approach This approach uses a RoBERTa-based model trained on the GoEmotions dataset created by Demszky et al. (2020). It is a multi-label classification model that takes a document as an input and returns 28 'probability' float outputs, one for each of the 28 emotion labels.

We use the same sample of 3,361,503 cleaned sentences as in the lexicon approach, but in this case do not lemmatize. We then feed these sentences into the model, receive the output, and assign each sentence the emotion with the highest predicted probability. Sentences labeled as "neutral" are dropped, as we are only concerned with text that has emotional content. This parallels our lexicon method, where our denominator consists only of emotion-related words. Next, we map each of the 27 emotions to 7 aggregate emotions guided by a correspondence released by Demszky et al. (2020). Ideally, we would like to have aligned the output to our MiiE Lab 8, but the model outputs did not allow us to easily identify "calm." This gives us a final dataset containing sentence level emotions.

In comparison to the lexicon-based approach, this method takes into account the entire context of sentence when making a prediction. The base model was pretrained on a large volume of texts and learned the contextual representation of words based on their surroundings. Hence, it is able to reduce semantic ambiguity otherwise introduced by analyzing each word individually. However, we lose some of the transparency our lexicon-based method provides through insight into the words with which emotions are being conveyed.

C.B Image Analysis

We quantify emotional representation in illustrated pages of children's books using a two-stage, fully automated pipeline. We implement the pipeline on Google Cloud Vertex AI using Gemini 2.5 Pro Preview, model version `gemini-2.5-pro-preview-03-25`.

First, each scanned page is processed using a lightweight Gemini classifier. This classifier assigns a binary label (either `TEXT_ONLY` or `CONTAINS_IMAGE`) using a low temperature setting of 0.1 and a response limit of 256 tokens to ensure consistent binary classification. Pages identified as `TEXT_ONLY` are excluded from further analysis, whereas pages labeled `CONTAINS_IMAGE` pages advance to the next stage in the pipeline.

Next, for our illustrated pages, the model outputs character level data (e.g., character bounding boxes, demographic attributes, and emotion labels), using a temperature 0.0, a 10,000-token limit. Importantly, the character level output includes three distinct emotion codings (`emotion_6_options`, `emotion_10_options`, `emotion_28_options`).

Because the image prompt allowed for open-ended race predictions, we standardized the model output before using it in the analysis. In particular, the re-categorized variable for race (race_original) was constructed according to the following classification of the Gemini model's output.

- Not applicable: Occurs when the character is classified as non-human
- Black: if race_original contains any of: black, african, afro, ashanti, mende

- Latine: if race_original contains any of: latin, cuban, rican, hispanic, mexica, mestizo, caboclo, mestiza
- Asian: if race_original contains any of: asian, afghan, chinese, hmong, indian, japanese, korean, mongol, tibetan, vietnamese, cambodian, filipino, kashmiri, kurdish, pashtun
- White: if race_original contains any of: white, greek, egyptian, european, german, russia, spanish, armenian, balkan, hungarian, italian, kosovan, mediterranean, norse, roman, celtic, jewish, ashkenazi
- Indigenous if race_original contains any of: indigenous, native, inuit, aboriginal, aben, aztec, guarani, hawaiian, inuk, inupiaq, kwakiutl, maya, navajo, olmec, penan, polynesian, saami, samoan, sioux, yup, zapotec, pacific islander, melanesian
- Unsure: if race_original contains any of: unclear, unknown, unspecified, unsure, ambiguous
- Other: if race_original contains any other classification