AMERICAN PSYCHOLOGICAL ASSOCIATION

Emotion

ISSN: 1528-3542

2025, Vol. 25, No. 5, 1137–1150
https://doi.org/10.1037/emo0001491

Emotion Word Production Tasks Grant Insight Into the Development of Emotion Word Organization and Accessibility

Chantal A. Valdivia-Moreno[1], Stephanie F. Sasse[2], Hilary K. Lambert[3], Katie A. McLaughlin[2], Leah H. Somerville[2], and Erik C. Nook[1]
[1] Department of Psychology, Princeton University
[2] Department of Psychology, Harvard University
[3] McLean Hospital/Harvard Medical School, Boston, Massachusetts, United States

Children are often instructed to "use their words" to communicate their emotions, which requires them to quickly access words that best describe their feelings. Adults vary in their ability to bring both nonemotion and emotion words to mind (two capacities called *verbal fluency* and *emotion fluency*). However, no studies have examined how emotion fluency emerges across development, despite the fact that mastering emotion language is an important developmental task. A cross-sectional sample of participants aged 4–25 years ($N =$ 194) generated as many fruit words as possible in 60 s (to measure verbal fluency) and as many emotion words as possible in 60 s (to measure emotion fluency). Emotion fluency was highly correlated with verbal fluency, and both showed similar increases across age, plateauing in late adolescence. Participants produced more negative emotion words than positive or neutral words, and these proportions were invariant across age. Network analyses shed light on the emergence of semantic networks underlying emotion organization across age. Finally, age of acquisition, valence, dominance, concreteness, and word length were significantly associated with the order in which emotion words came to participants' minds, suggesting that these dimensions are associated with the accessibility of emotion concepts. Interestingly, the influence of these dimensions on the order of emotion word production was invariant across age. Results from this study illustrate the developmental emergence of emotion fluency and provide new insight into the key dimensions that are associated with which emotion words rapidly come to mind.

Keywords: affect, child development, emotion fluency, valence, verbal fluency

Supplemental materials: https://doi.org/10.1037/emo0001491.supp

We often ask children to "use their words" to identify their emotions, which requires them to generate and select emotion words that best describe their current feelings. This practice is increasingly common in educational settings, where children are taught to broaden their emotional vocabularies (Nathanson et al., 2016). Indeed, naming and communicating emotions are foundational skills that we employ in everyday life (Rimé, 2009), and current theories suggest that language is central to emotional processes (Lindquist & Gendron 2013). Emotion vocabulary in particular may provide the building blocks for understanding and communicating emotion concepts throughout development (Hoemann et al., 2019; Nook et al., 2017), and growing evidence suggests that language facilitates the development of emotion regulation (Fields-Olivieri et al., 2024). Nevertheless, it remains unclear how the capacity to generate emotion words develops across age and what features of emotion words might shape which words most

This article was published Online First January 16, 2025.
Kristen A. Lindquist served as action editor.
Chantal A. Valdivia-Moreno https://orcid.org/0009-0007-2240-5737
Deidentified data and analysis scripts for this project are publicly available on the Open Science Framework at https://osf.io/h3gtp/. The authors declare no conflicts of interest. This work was supported by a Harvard University seed grant to Leah H. Somerville, a National Institutes of Mental Health, National Institutes of Health grant to Katie A. McLaughlin (R01-MH103291), and a National Science Foundation Graduate Research grant to Erik C. Nook (DGE1144152).

The authors thank C. R. M. Bolden, A. Dews, E. Fearey, M. Garrad, T. Gogue-Garcia, K. Kent, A. Sareen, M. Sirak, T. Stacy, C. Stavish, and C. Uhrig for assistance with data collection and P. Mair and members of Harvard University's Institute for Quantitative Social Science for guidance on statistical analyses. The authors also thank M. Zettersten for guidance on semantic network analyses.

Chantal A. Valdivia-Moreno played a lead role in conceptualization, data curation, formal analysis, visualization, and writing–original draft and an equal role in writing–review and editing. Stephanie F. Sasse played a supporting role in conceptualization, data curation, and writing–review and editing. Hilary K. Lambert played a supporting role in conceptualization, data curation, and writing–review and editing. Katie A. McLaughlin played a supporting role in conceptualization, supervision, and writing–review and editing and an equal role in funding acquisition. Leah H. Somerville played a supporting role in conceptualization, supervision, and writing–review and editing and an equal role in funding acquisition. Erik C. Nook played a lead role in supervision and writing–review and editing, a supporting role in conceptualization, formal analysis, and visualization, and an equal role in data curation and funding acquisition.

Correspondence concerning this article should be addressed to Chantal A. Valdivia-Moreno, Department of Psychology, Princeton University, 40 Woodlands Way, Princeton, NJ 08544, United States. Email: cvaldivia@princeton.edu

readily come to mind. Here, we examine the development of *emotion fluency* (i.e., how rapidly people can generate emotion words) and examine the order of emotion word production to assess which dimensions appear to make emotion words more or less accessible.

A handful of prior studies have examined emotion fluency in adults, seeking to chart the key predictors of emotion fluency and its potential relations with mental health. These studies typically assess emotion fluency by having participants say aloud or write as many emotions as they can think of in a minute. This task is an adaptation of established methods of assessing *verbal fluency* (i.e., the ability to rapidly produce *any* words, typically assessed by asking participants to generate as many words as possible in 60 s that belong to a specific semantic category or that begin with a specific letter). These prior studies consistently show that emotion fluency and verbal fluency are positively correlated in adults ($.30 < r\text{s} < .54$; Abeare et al., 2017; Camodeca et al., 2021; Hegefeld et al., 2023). Consequently, emotion fluency appears to be strongly related simply to how rapidly people can generate *any* words. Studies have also tested whether emotion fluency might be related to mental health in adults, as facility with emotion words could provide greater efficacy managing one's emotions (Barrett et al., 2001; Erbas et al., 2019; Fields-Olivieri et al., 2024; Hoemann et al., 2021; Kalokerinos et al., 2019). For instance, work by Vine et al. (2020) suggests that increased use of negative emotion words in one's everyday vocabulary might stem from increased familiarity with experiencing negative emotions. However, research thus far has shown inconsistent correlations between emotion fluency and psychopathology (Abeare et al., 2017; Camodeca et al., 2021), with a recent preregistered study demonstrating no significant correlations between emotion fluency and depression symptoms, emotion regulation, or alexithymia in adults (Hegefeld et al., 2023). As such, current evidence suggests that emotion fluency may primarily index verbal abilities, rather than mental health.

Crucially, these studies showing that emotion fluency is not significantly correlated with psychopathology have only been conducted in adults, leaving the relation between emotion fluency and psychopathology uncharted in youth, whose emotion concepts are still developing. Moreover, emotion fluency is positively associated with children's socioemotional word use in a naturalistic context (Sun et al., 2024), suggesting that emotion fluency tracks how people use emotion words in daily life. Additionally, establishing normative trajectories of emotion fluency is a critical first step in understanding how deviations from these trajectories predict emotional health and well-being (Cicchetti, 1993). These motivate additional study of emotion fluency in youth. Specifically, a developmental framework can provide a stronger test of whether emotion fluency is simply a subset of verbal fluency by comparing their shifts across age, and merely charting these normative trajectories can allow future researchers to test whether deviations relate to important social and psychological outcomes.

Verbal abilities in general—and verbal fluency in particular—increase through early development (Regard et al., 1982; Riva et al., 2000; Sauzéon et al., 2004). Consequently, if emotion fluency and verbal fluency are indeed tightly coupled, they might also increase in tandem across age. If so, demonstrating that these two processes share overlapping developmental trends would add an additional point of evidence that verbal fluency is central to emotion fluency. Indeed, given data showing that verbal fluency and emotion fluency are closely related (Abeare et al., 2017; Camodeca et al., 2021; Hegefeld et al., 2023), our primary hypothesis is that they will follow similar developmental patterns across age. Nonetheless, an alternative hypothesis is that emotion fluency and verbal fluency may decouple. This could occur if the development of general vocabulary does not progress at the same rate as emotion vocabulary, such that children gain greater skill in accessing some clusters of words (e.g., animals or foods) faster than emotion words. We consequently test these competing hypotheses.

Prior research in adults has also consistently found that participants tend to produce more negative words than positive or neutral emotion words during emotion fluency tasks, a tendency known as *negative valence bias* (Hegefeld et al., 2023; Schrauf & Sanchez, 2004; Zammuner, 2011). A likely explanation of this observation is that English has more negative emotion words than positive or neutral words (Jackson et al., 2023), although it may also be related to the classic literature showing that people typically attend more to negative than positive information (Rozin & Royzman, 2001). From a developmental perspective, negative valence bias may be especially elevated in childhood and decrease linearly across age, as prior research shows that children tend to seek out and attend to negative information more than other age groups (Grisanzio et al., 2021). Alternatively, the relationship between age and negative valence bias could be quadratic, such that negative valence bias peaks in adolescence, as is found for other clinical and affective processes that involve elevated focus on negative affect (e.g., risk of psychopathology; Fairchild, 2011; McGrath et al., 2023; and contrahedonic motivation; Riediger et al., 2009). Finally, the relationship could be null, such that one's negative valence bias does not significantly vary across age. A null relationship would imply that as emotion fluency increases, people expand the number of negative emotion words they produce in equal proportion to positive and neutral words. Adjudicating between these hypotheses provides greater insight into how the valence of emotion word accessibility varies across age. Establishing normative differences in negative valence bias is a necessary first step toward evaluating whether over- or underproduction of negative emotions might be related to psychopathology in development. Additionally, this analysis lets us examine whether children or adolescents show systemic biases in which emotion words they know and can access under time pressure. Are youth more, less, or equivalently focused on negative words as adults?

Beyond these questions regarding the development of emotion fluency and valence bias, emotion fluency tasks offer an opportunity to examine the order in which emotion words spontaneously come to mind and thereby learn about the underlying semantic structures that organize emotion words. Studies in cognitive psychology have shown that people produce words in fluency tasks by moving through *semantic networks* (Abwender et al., 2001; Mills & Phillips, 2023; Sauzéon et al., 2004). Participant responses on memory retrieval and verbal fluency tasks often "cluster" around semantically related words. For instance, a participant may produce the string *cow*, *pig*, and *goat* (three words belonging to the category of farm animals) before moving to another semantically related string, like *cat*, *dog*, and *hamster* (belonging to the category of pets; Zemla & Austerweil, 2017). Semantically relevant groups might also emerge during emotion fluency tasks, such that participants produce a string of words that share some semantic features (e.g., negative emotions) and then move on to emotion words that share other features (e.g., positive emotions). Thus, one could exploit the order of words people produce in this task to examine the dimensions that shape the accessibility of emotion language and concepts.

Affective scientists and social psychologists have long sought an understanding of how people organize emotions at the conceptual level. One leading framework, called the circumplex model, proposes that two key dimensions organize affective states: valence (i.e., positivity vs. negativity) and arousal (i.e., deactivating/sedating to activating/energizing; Posner et al., 2005; Russell, 1980). Russell and Mehrabian (1977) also identified dominance (i.e., extent to which an emotion reflects submissiveness or control; Mohammad, 2018) as a third dimension organizing affective states. Children as young as 3.5 organize emotional stimuli within a circumplex model, although children appear to place more attention to valence than other dimensions (Barrett, 2004; Nook et al., 2017; Posner et al., 2005; Widen, 2013, Widen & Russell, 2010). As such, both youth and adults may produce words in the emotion fluency task by moving through a semantic network organized by valence, arousal, and dominance.

In contrast to these affective dimensions, properties of the words themselves may shape which emotion words first come to mind. For example, people might start with the words they learned earliest (e.g., happy, sad) and then move to words learned later (e.g., anxious, jubilant; Baron-Cohen et al., 2010). In fact, Widen and Russell (2010) reported that children's emotion categories start broad (i.e., "good" vs. "bad") and become more differentiated with the acquisition of early emotion words, such as happy, sad, and angry. Thus, age of acquisition (i.e., the age at which participants learned a word) might be a dominant factor shaping the accessibility of emotion words. Indeed, words learned first are more readily accessible during word training and other word production tasks (Brysbaert & Biemiller, 2017; Kuperman et al., 2012). Concreteness may also influence which words are more readily available, as concrete words are easier to remember and recall than abstract words (Brysbaert et al., 2014). Shorter emotion words (e.g., mad, scared) may also be more readily accessible than longer words (e.g., apprehensive, overwhelmed; Kuperman et al., 2012). Finally, words that are used more frequently in everyday life could also be "primed" and thus more easily come to mind than words heard less frequently (Kuperman et al., 2012).

Here, we provide two novel analyses of emotion fluency data to (a) shed new light on the developing semantic network of emotion words and (b) identify dimensions that appear to make words more easily accessed. First, we deploy network analyses (Zemla et al., 2020) to visualize clusters of emotion words that reliably appear in close proximity to each other. This method allows us to gain qualitative insight into emotion words that are semantically associated both across all participants and within specific age groups. Second, we use regression methods to test valence, arousal, dominance, concreteness, age of acquisition, word length, and word frequency as predictors of the order in which words are generated in the task. Again, we perform this analysis both at the sample level to see which dimensions shape word order for people regardless of age, and we also test whether the potency of specific dimensions in predicting word order varies with age. These analyses reveal which dimensions underlie emotion word accessibility and assess whether they are static or variable across age.

In this study, we administered verbal fluency and emotion fluency tasks to participants between the ages of 4 and 25. We examined age-related variation in verbal fluency, emotion fluency, negative valence bias, and semantic networks. Using established lexical norms, we also tested which dimensions (valence, arousal, dominance, age of acquisition, concreteness, word length, and word frequency) predict the order in which participants produce emotion words. Documenting normative developmental patterns of emotion fluency lays crucial groundwork for understanding the downstream implications of typical and atypical development in these processes (Cicchetti, 1993). This study also uses development as a wedge to gain further clarity on the overlap between emotion fluency and verbal fluency by testing whether they develop in tandem or diverge across age. Though they are closely correlated in adulthood (Abeare et al., 2017; Camodeca et al., 2021; Hegefeld et al., 2023), lesser familiarity with or access to emotion concepts in early development could lead to divergent age trajectories. Finally, we exploit these data to access the rich cognitive, semantic, and linguistic factors that shape the accessibility of emotion words (Mills & Phillips, 2023; Zemla et al., 2020), both in adults and across development.

Method

Participants

This study includes data from participants who were recruited for a study examining several components of emotion development (Grisanzio et al., 2021; Jenness et al., 2021; Nook et al., 2017, 2018, 2020; Weissman et al., 2020). No publications from this study include the emotion fluency or verbal fluency data analyzed here. A total of 203 participants completed these tasks. This target sample size provided adequate power to detect small-to-medium relations between age and dependent variables (i.e., 95% power for $r = .25$). Six participants who did not understand directions, did not cooperate, or were distracted during the fluency tasks were excluded from our analyses, and audio files for three additional participants were not saved. As such, analyses presented here represent a final sample size of $N = 194$ participants who ranged in age from 4.13 to 25.91 years ($M_{\text{age}} = 14.55$; $SD_{\text{age}} = 5.83$). This sample reflected a slight female majority (99 females [51%] and 95 males [49%]) and a wide range of income-to-needs ratios (0.07–13.32; $M_{\text{income-to-needs}} = 5.05$; $SD_{\text{income-to-needs}} = 3.14$; 9.28% below poverty line, 25 participants did not report data required to compute income-to-needs ratio). Participants reported their race as American Indian/Alaska Native: 0.52%; Asian: 8.76%; Black/African American: 11.34%; more than one race: 11.86%; native Hawaiian/Pacific Islander: 0.52%; White: 63.92%; and other: 1.55%, and 1.55% declined to provide information. Most participants did not identify as Hispanic/Latino (89.18%). All participants reported being native English speakers, were recruited from areas near Harvard University and the University of Washington, and were compensated for their time. All procedures were approved by the Committee on the Use of Human Subjects at Harvard University and the University of Washington Institutional Review Board.

Procedure

All data from this study were collected at a single in-person laboratory session in 2015 and 2016. As a measure of verbal fluency, participants were given 60 s to name aloud as many fruits as possible (Rosselli et al., 2000; Sokołowski et al., 2020). Then, as a measure of emotion fluency, participants were given 60 s to name as many emotions as possible (Abeare et al., 2017; Camodeca et al., 2021). These tasks were not counterbalanced to allow younger participants to begin with a more familiar category (i.e., fruits) compared to emotions. Participants' responses to both conditions were recorded as audio files and later transcribed.

Data Processing

Verbal Fluency

We computed each participant's verbal fluency score by counting the total number of valid and distinct fruits they produced in the verbal fluency task. Participants did not receive points for fruit-related words that were not themselves fruits (e.g., applesauce, salad). Participant responses were also synonymized such that varieties of fruits (e.g., bosc pear, asian pear, and d'anjou pear) were collapsed into a single category (e.g., pear). This process ensured that participants received 1 point per distinct fruit and all subsequent varieties of the same fruit were counted as repetitions (see the Supplemental Materials for more details regarding the coding process). We provide our full dictionary on the Open Science Framework (OSF) at https://osf.io/h3gtp/.

Emotion Fluency

As an advance beyond prior work, we generated a coding system for scoring answers from the emotion fluency task and categorizing these words by valence. We provide a thorough description of these methods in the Supplemental Materials, and our full dictionaries are available on the OSF at https://osf.io/h3gtp/.

We first computed emotion fluency scores by counting the total number of distinct and "valid" emotion words produced in the emotion fluency task (i.e., words that referred to common-sense emotions, rather than nonemotional states). Given that the field has not identified a strict boundary between what is and is not an "emotion" (Gendron et al., 2020), these distinctions were determined iteratively by the first and anchor authors. First, all participant responses were sorted into three categories: emotion words, borderline emotion words, and nonemotion words. Emotion words (92.71% of total utterances) were defined as (a) words that could be used to describe how one is feeling and/or (b) words that are colloquially acknowledged as emotions (e.g., happy, mad, sad). Borderline emotion words (4.46% of total utterances) were defined as words that do not meet the previous criteria but nonetheless (a) could be used to describe how one is feeling and (b) label an affective or physiological state (e.g., aware, childish, hungry). Nonemotion words (2.82% of total utterances) were defined as words that do not describe common-sense feelings, affective states, or physiological states (e.g., achieved, schizophrenic, yelling). Given that borderline words comprised only 4.46% of total utterances and borderline word production did not vary systematically across age (see the Supplemental Materials), we ultimately treated them as nonemotion words (i.e., they did not count toward emotion fluency scores, they were not coded by valence, and they were not included in word order analyses; see below). Consequently, emotion fluency scores for each participant were produced by counting the number of valid and distinct emotion words they produced (i.e., no points were received for borderline words, nonemotion words, or repetitions of emotion words). Additional information on emotion fluency scoring is provided in the Supplemental Materials.

Valence Categories

Emotion words were then categorized by valence (i.e., positive, negative) so that we could examine the proportion of words that fell into each valence category, as well as whether these proportions varied by age. Given that people occasionally refer to their own emotions using labels that are not clearly positive or negative (e.g., neutral, nothing), we coded some emotion words as neutral. Valence categorizations were informed by large-scale coding norms from the National Research Council Canada Valence Arousal Dominance (NRC VAD) lexicon, which provides human ratings of valence, arousal, and dominance for more than 20,000 English words (Mohammad, 2018). The NRC VAD provides normative ratings for each of these dimensions that vary from 0 to 1, where higher scores indicate more positive valence, higher arousal, and higher dominance. Mohammad (2018) reported split-half reliability scores of 0.95 for valence, 0.90 for arousal, and 0.90 for dominance. Though we used NRC VAD to guide our categorizations, some manual rescoring was needed to categorize participants' responses with common-sense valence categories. See the Supplemental Materials for further details on this coding process.

To examine whether there was any age-related variation in negative valence bias, we calculated the proportion of negative, neutral, and positive emotion words produced relative to each participant's emotion fluency score. Similar to Hegefeld et al. (2023), we computed each participant's negative valence bias by calculating the proportion of emotion words they produced that fell in the negative category.

Predictors of Word Order

To test which dimensions made emotion words more accessible (i.e., were associated with earlier word order), we created a word-level data set that aligned each word produced in the emotion fluency task with relevant lexical parameters. For each participant, we recorded the placement of each word relative to all other words produced in the emotion fluency task (which we call *word order*). Note that this order was only analyzed for valid emotion words, but order was assigned in relation to all words a participant produced, including borderline or nonemotion words. For example, if a participant produced the words "angry" (emotion), "itchy" (nonemotion), and "sad" (emotion), we recorded the order placement of all three words such that angry would receive a word order placement of one, itchy would be two, and sad would be three. However, analyses only examined relations between lexical parameters and emotion words (i.e., "itchy" would not be analyzed in the prior example).

To examine whether qualities of emotion words predicted the order they were produced in the emotion fluency task, we compiled relevant qualities for each emotion word using seven lexical parameters, mostly drawn from large norming studies. These lexical parameters included valence (i.e., extent to which an emotion is positive or negative; Mohammad, 2018), arousal (i.e., the degree of activation or deactivation of an emotion; Mohammad, 2018), dominance (i.e., extent to which an emotion reflects submissiveness or control; Mohammad, 2018), age of acquisition (i.e., averages of self-reported age at which participants learned this word; Kuperman et al., 2012), concreteness (i.e., degree to which an emotion refers to a physical, observable, or perceptible entity; Brysbaert et al., 2014), word length (i.e., number of letters), and word frequency (i.e., how often a word appears in the English language; Brysbaert & New, 2009). A small number of valid emotion words participants generated in the task did not exist in these norms ($n = 38$ words, 3% of utterances). Whenever possible, we used the closest available synonyms (e.g., "jubilant" for "jubilance") to draw relevant data (see the Supplemental Materials). We also conducted a supplemental analysis to demonstrate that these substitutions did not influence key results (Supplemental Table S1).

Analytical Approach

We first produced descriptive statistics (e.g., mean, standard deviation) of verbal fluency and emotion fluency scores. We also conducted a paired-samples t test to compare means between emotion fluency and verbal fluency scores. Subsequent analyses were designed to assess (a) age-related variation in verbal fluency and emotion fluency, (b) correlations between verbal fluency and emotion fluency, (c) age-related variation in negative valence bias, (d) network organization of emotion word production, and (e) which linguistic dimensions predict the order in which emotion words were produced.

Testing Age-Related Variation in Emotion Fluency and Verbal Fluency

We first tested whether age was related to verbal and emotion fluency using two Pearson's correlations. We hypothesized that both verbal and emotion fluency scores would positively correlate with age, given the overall increase in verbal abilities across age (Anglin et al., 1993). However, visualization of these data suggested nonlinear relationships with age, as emotion fluency and verbal fluency appeared to increase across childhood and adolescence before seemingly reaching a plateau. First, we examined Akaike information criterions (AICs) for each model and used the *lmtest* package in R (Zeileis & Hothorn, 2002) to conduct a likelihood-ratio test, assessing whether a spline model provided a better fit for the data, even after penalizing for additional complexity. Then, we followed prior methods (Nook et al., 2020) to quantify the age of this plateau. Specifically, we used the *mgcv* package to conduct a smoothing spline regression that essentially constructs a regression equation that is nonlinear but does not overfit the data (Version 1.8.41; Wood, 2003, 2017). We then calculated the first derivative of the line of best fit for each spline (Simpson, 2014) and identified the age at which the rate of change in these dependent variables approached 0 (i.e., slowed to 0.2 words/year). This cutoff was selected in line with methods from prior work (Nook et al., 2020) and provides an adequate visual fit with the slope.

To test whether there was a significant difference in the ages at which verbal fluency and emotion fluency plateaued, we employed nonparametric bootstrapping methods (as in Nook et al., 2020). Using the *boot* package in R, we conducted 10,000 bootstrapped simulations (Version 1.3.30; Canty & Ripley, 2024; Davison & Hinkley, 1997). Within each simulation, we (a) conducted spline analyses of verbal fluency and emotion fluency, (b) used the derivative method described above to identify the age of each variable's plateau, and (c) computed pairwise differences between the two plateau ages. We then examined the 95% bias-corrected and accelerated confidence intervals of the difference between the emotion fluency and verbal fluency plateau points. If this range did not include 0, we concluded that verbal fluency and emotion fluency plateau ages were significantly different (DiCiccio & Efron, 1996). Plateaus could not be identified for 2.55% of verbal fluency bootstraps and 0.36% of emotion fluency bootstraps. These simulations were excluded from confidence interval computations.

Testing Relations Between Verbal Fluency and Emotion Fluency

We conducted a Pearson's correlation to test for an association between verbal fluency and emotion fluency. As in prior work with adults (Hegefeld et al., 2023), we hypothesized that verbal fluency and emotion fluency would be positively correlated with each other.

Examining Valence

Using the *ez* package, we conducted a one-way within-subjects analysis of variance to compare emotion fluency scores across each valence category (Version 4.4.0; Lawrence, 2016). Then, we conducted a Pearson's correlation between age and negative valence bias (i.e., the proportion of words they produced that were negatively valenced) to examine whether negative valence bias varied by age. We also used the *lmtest* package in R (Zeileis & Hothorn, 2002) to conduct a likelihood-ratio test, ensuring that a linear model provided the best fit for the data.

Examining Network Organization of Emotion Word Production

We used the Semantic Network and Fluency Utility (SNAFU; Zemla et al., 2020) program to produce network visualizations of emotion fluency responses. In brief, this program uses computational tools to identify "clusters" of words that participants tend to produce in close proximity to each other. We applied the Community Network (CN) approach to participants' emotion fluency responses (after removing nonemotion words, borderline words, and repetitions and synonymizing emotion words, as described above). The CN method first counts the co-occurrences of emotion words (i.e., the number of times participants produced two emotion words within two utterances of each other, as we set our window to 2). We also set the "threshold" (i.e., minimum number of co-occurrences to merit analysis) to 2, meaning co-occurrences of 0–1 are ignored. An algorithm then determines which co-occurrence rates meet a statistical threshold that is beyond chance (see Zemla & Austerweil, 2018, for details). Those that meet this threshold are marked as "edges" in a network diagram. We used SNAFU's CN algorithm to produce a network from all participants' responses, and we also produced networks for participants split into four age groups that spanned ~5 years each: early and middle childhood (Ages 4–9; $n = 38$), late childhood (Ages 9–12.99; $n = 44$), adolescence (Ages 13–17.99; $n = 49$), and adulthood (Ages 18–25.99; $n = 61$). Sample sizes here only include participants who produced >0 emotion words. Figures displaying network results were produced in R using the *igraph* package (Version 2.0.3; Csardi & Nepusz, 2006).

Examining the Predictors of Word Order

We then examined whether our seven lexical parameters (i.e., valence, arousal, dominance, age of acquisition, concreteness, word length, and word frequency) could significantly predict word order (i.e., one for the first word produced, two for the second word, etc.) in the emotion fluency task. Because these data were nested (i.e., participants produced several words, each with their own word order and corresponding lexical parameters), we used analytic approaches that accounted for this multilevel structure. As a primary analysis, we used the *lme4* (Version 1.1.27.1; Bates et al., 2015), *lmerTest* (Version 3.1.3; Kuznetsova et al., 2017), *parameters* (Version 0.20.2; Lüdecke et al., 2020), *and effect size* (Version 0.8.3; Ben-Shachar et al., 2020) packages to conduct a mixed-effects model testing which parameters were related to word order. In this model, the order in which each emotion word was produced was modeled as the outcome variable, each lexical parameter was included as a fixed effect, and subject identification was included as a random effect to nest words within participants. This analysis tests which parameters are related to word order, even after controlling for all other parameters (see the Supplemental Materials). We then examined whether age moderated any associations between word order and the seven parameters by conducting mixed-effects models that tested whether age interacted with each parameter in predicting word order. This analysis also included subject identification as a random effect to nest words within participants.

Transparency and Openness

All data and analytic code can be accessed at https://osf.io/h3gtp/. Statistical analyses were performed using RStudio Version 4.2.1 (R Core Team, 2022). Analyses for this study were not preregistered.

Results

Verbal Fluency and Emotion Fluency Scores

Verbal fluency scores ranged from 0 to 26 ($M = 12.11$, $SD = 5.17$), and emotion fluency scores ranged from 0 to 23 ($M = 9.71$, $SD = 4.07$). Results from a paired-samples t test indicated that, on average, participants produced more fruit words than emotion words, $t(193) = 8.47$, $p < .001$, 95% CI [1.84, 2.96], $d = -.52$. Across our sample, however, more unique emotion words were generated ($n = 219$) than fruit words ($n = 64$). This posed a paradox: There were a greater number of emotion words than fruit words generated across participants, but each participant tended to produce more fruit words than emotion words in 60 s. We reasoned this could arise if there were a larger number of words that are easily accessed as examples of fruits than emotions, allowing people to rapidly sample fruit words. To test this hypothesis, we examined the proportion of participants who produced each fruit and emotion word in our data set (Supplemental Figure S4). In line with our reasoning, there were more "popular" fruit words (i.e., words said by >50% of participants) than emotion words, and there were more "rare" emotion words (i.e., words said by <5% of participants) than fruit words.

Verbal Fluency and Emotion Fluency Increase Across Age and Plateau in Late Adolescence

Pearson's correlations showed that age was positively correlated with both verbal fluency, $r(192) = .69$, $p < .001$, 95% CI [.61, .76], and emotion fluency, $r(192) = .59$, $p < .001$, 95% CI [.49, .67]. However, these relationships did not appear linear, so we applied spline models to fit the data (Figure 1A and 1B). Results from a likelihood-ratio test indicated that the spline model was a significantly better fit than a linear model for both verbal fluency ($AIC_{linear} = 1066.34$; $AIC_{spline} = 1033.40$; $\chi^2 = 36.73$, degrees of freedom = 1.90, $p < .001$) and emotion fluency ($AIC_{linear} = 1017.68$; $AIC_{spline} = 996.28$; $\chi^2 = 25.02$, degrees of freedom = 1.81, $p < .001$). Results revealed that verbal fluency and emotion fluency increased across childhood and adolescence before plateauing in early adulthood. Verbal fluency plateaued at age 19.41 years, 95% CI [17.15, 22.62], and emotion fluency plateaued at 18.08 years, 95% CI [15.80, 20.97]. Bootstrapped comparisons of the difference between these plateau points suggested that they did not differ significantly from each other, 95% CI of the difference in plateau points [−1.53, 5.87].

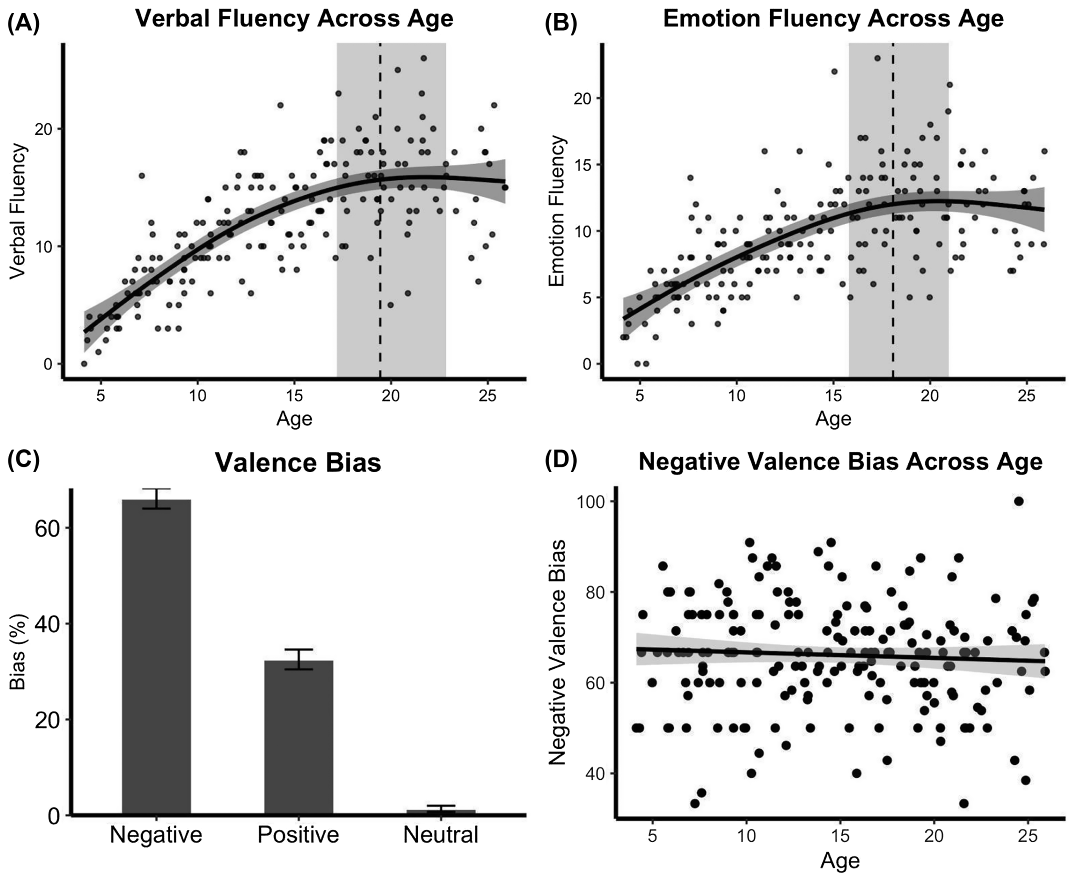
Verbal Fluency and Emotion Fluency Are Strongly Correlated

As hypothesized, verbal fluency and emotion fluency were strongly and positively correlated, $r(192) = .66$, $p < .001$, 95% CI [.57, .73]. Age did not significantly moderate the correlation between verbal fluency and emotion fluency, $\beta = -.07$, $p = .180$. As such, the strength of the relationship between verbal and emotion fluency did not significantly change from the ages of 4 to 25.

Negative Valence Bias Does Not Significantly Differ by Age

A within-person one-way analysis of variance showed that the number of negative, positive, and neutrally valenced emotion words that participants produced differed significantly, $F(2, 386) = 622.07$, $p < .001$ (Figure 1C). Follow-up paired t tests showed that participants produced significantly more negative ($M = 66\%$) than positive ($M = 33\%$) emotion words, $t(193) = 17.16$, $p < .001$, $d = -1.35$. Participants also produced significantly more positive than neutral words, $M = 1\%$; $t(193) = -21.71$, $p < .001$, $d = -2.19$. Contrary to our hypothesis, however, a correlation analysis indicated that negative valence bias did not significantly increase or decrease across age, $r(190) = -.81$, $p = .420$, 95% CI [−.20, .08] (Figure 1D). To test for potential nonlinear relations with age, we fit a thin plate smoothing spline. However, a likelihood-ratio test indicated that the spline did not significantly improve model fit beyond the null ($AIC_{null} = 1505.16$; $AIC_{spline} = 1505.67$; $\chi^2 = 4.10$, degrees of freedom = 2.30, $p = .13$), suggesting no linear or nonlinear relations with age.

Figure 1
Emotion Fluency and Verbal Fluency Task Results



Note. (A) Verbal fluency (i.e., the number of fruit words participants produced in 60 s) scores across age. Verbal fluency increased across age and reached a plateau at age 19.41. (B) Emotion fluency (i.e., the number of emotion words participants produced in 60 s) across age. Emotion fluency increased across age and reached a plateau at age 18.08. In A and B, the dark gray bands represent 95% confidence intervals, the vertical dashed line represents the plateau point, and the light gray column represents the bootstrapped 95% confidence interval of the plateau point. (C) Average proportion of responses that fell within each valence category in the emotion fluency task. Participants produced more negative words than positive or neutral words. (D) Negative valence bias did not significantly vary by age.

Descriptive Analysis: Average Placement of Frequent Emotion Words

As a first descriptive snapshot of word order, we selected the most commonly produced emotion words (i.e., only words that were produced more than 40 times across all participants) and computed the average "word order" for each of these words. We then visualized these orders (Figure 2). Interestingly, "happy/happiness" and "sad/sadness" were commonly produced at the very beginning of the emotion fluency task, followed by negatively valenced words like anger/angry, mad, fear/ful, disgust/ed/ing, and scared. In fact, 68% of participants' first responses in this task were happy/happiness and sad/sadness (in either order).

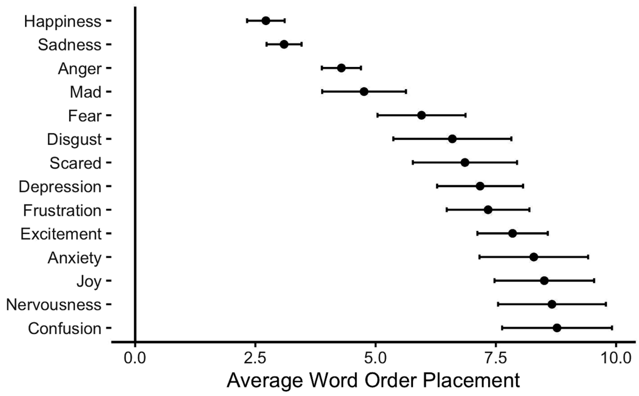
Visualizing the Network Organization of Emotion Words

CN analyses of all participants' data using the SNAFU program are presented in the left panel of Figure 3. This sample-level network reveals 14 clusters that provide interesting insight into semantic associations between emotion words. In line with results above and below, early-acquired "basic level" emotion words cluster together (anger/fear/disgust/sadness/happiness/mad), with a bridge to more complex emotion words (upset/hurt/exhausted and grumpy/boredom; Rosch et al., 1976). Another large cluster appears to group "fear family" and "surprise-family" emotion words, along with a bridge to words related to anxious misery/psychopathology: startled/afraid/scared/frightened/surprise/concerned/bewildered/shock/apprehensive/excitement/anxiety/depression/melancholy/nervousness (Craske et al., 2009). Interestingly, some clusters contain words that are highly semantically associated (e.g., joy/exuberance and shame/guilt, sympathy/empathy, stress/distress/overwhelmed), but other clusters group antonyms (love/hate, misery/contentment, comfortable/uncomfortable). Finally, the envy/elation cluster seems to reflect phonetic (not semantic) associations (e.g., envy/elation).

Conducting CN analyses within age groups is further illustrative (right of Figure 3). The overall number of words, edges, and clusters

Figure 2

Average Word Order Placement for Frequent Emotion Words



Note. Frequently produced emotion words (i.e., those produced >40 times) are displayed with their average word order placement in the emotion fluency task. For instance, happy/happiness and sad/sadness were on average the earliest words produced by participants.

grows across age, reflecting increases in emotion word comprehension and emotion fluency. However, some nuanced age-related patterns appear. First, a happiness/sadness cluster exists even in the youngest participants, and this cluster grows to include anger by adolescence. This is in line with results below showing that these two words are often the first to be spoken by participants, regardless of age. Second, anger, disgust, and fear were associated with different words across age, sometimes clustering together, but sometimes not. Specifically, adolescents appeared to form an "anger family" cluster (irritation/annoyance/frustration/upset), and adults had two distinct "fear family" clusters, one more focused on anxiety and the other on fear. Last, an "anxious misery/psychopathology" cluster (anxiety/depression/nervousness/excitement) was unique to adolescents (Craske et al., 2009).

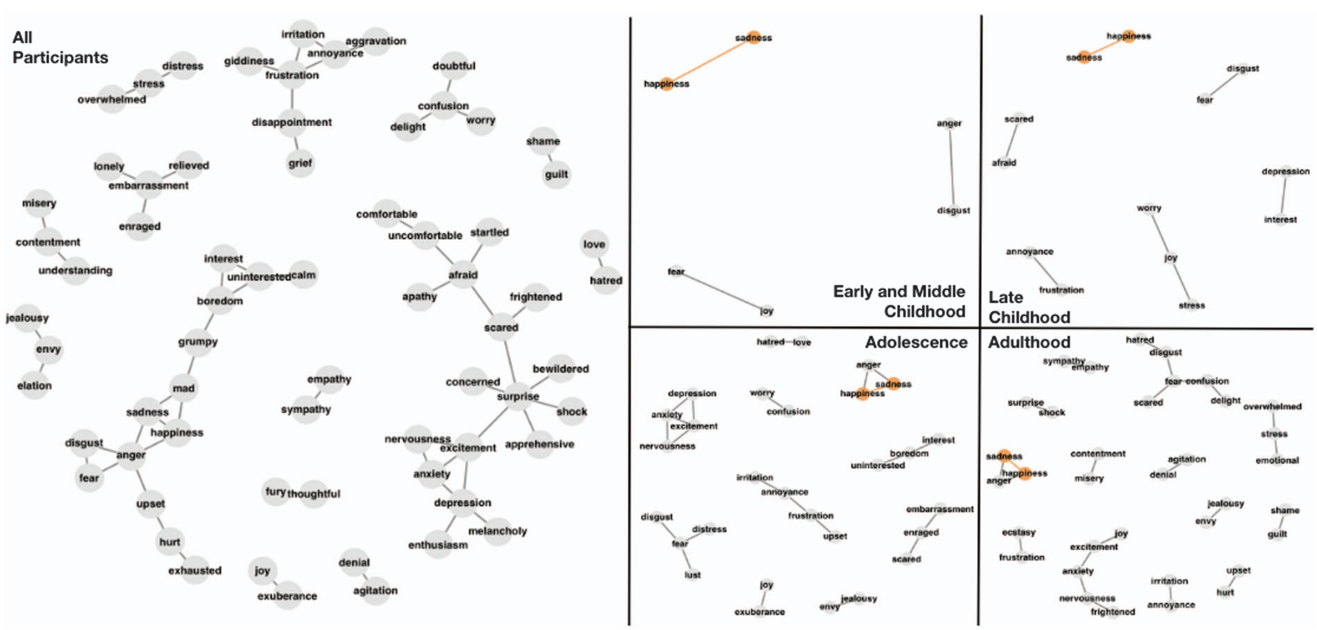
Testing Associations Between Word Order and Lexical Parameters

We used a mixed-effects model to test which lexical parameters were associated with word order when controlling for all other parameters. This analysis revealed that age of acquisition, valence, dominance, concreteness, and word length (listed in order of strength) were significant predictors of word order (Figure 4; Table 1). Words produced later in the emotion fluency task tended to be learned later in life, more positive, less dominant, less concrete, and longer. A supplemental analysis confirmed that the significance of these parameters did not change when we excluded words for which we used lexical data for a best-fitting substitute when an exact match was not available (e.g., "jubilant" for "jubilance"; see the Supplemental Table 1).

Finally, mixed-effects models showed that age did not moderate any associations between word order and lexical parameters ($ps > .463$), suggesting that the strength of each of these parameters in predicting word order does not significantly vary across age. To protect against the possibility that these results (especially the relationship between word order and age of acquisition) may have emerged because our sample included young participants, we conducted another supplementary mixed-effects model that

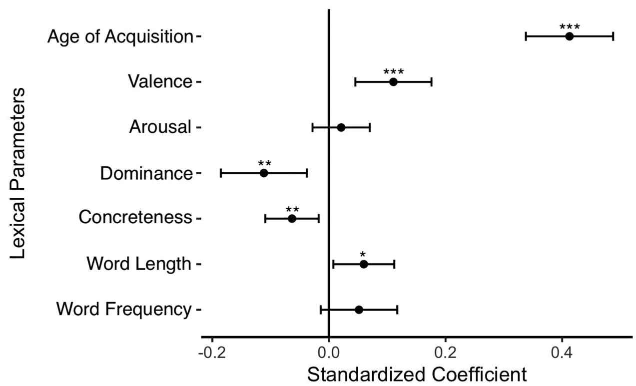
Figure 3

Network Analyses Displaying Clusters of Emotion Words That Emerge Based on Co-Occurrence of Emotion Words



Note. Left panel displays results from analyses of all participants' data. Subsequent panels display results for participants in four age bins: early and middle childhood (4–9), late childhood (9–12.99), adolescence (13–17.99), and adulthood (18–25.99). Happiness-sadness nodes are highlighted in orange, as an edge occurred between these nodes in all age groups. Words that were spoken by participants but had no edges connecting them to other words are not displayed.

Figure 4
Lexical Parameters and Their Relations With Word Order



Note. Standardized betas and 95% confidence intervals are presented from a mixed-effects model testing which of seven lexical parameters predict the order in which participants produced emotion words. Age of acquisition, valence, dominance, concreteness, and word length significantly predicted word order, such that words produced later in the emotion fluency task tended to be learned later in life, more positive, less dominant, less concrete, and longer.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

restricted analyses only to adult participants aged 18 and older (see the Supplemental Materials). In this analysis, age of acquisition, valence, and dominance survived as predictors of word order, but concreteness and word length were no longer significant (Supplemental Figure S5, Table S2). However, age of acquisition remained the *strongest* predictor of word order, assuaging the concern that age of acquisition emerged as a significant predictor of word order only because the study included youth participants. We also reasoned that the loss in significance for concreteness and word length was due to reduced power for detecting these small effects when restricting the sample to 63 adults.

Discussion

In this study, we explored several research questions regarding the development of verbal fluency and emotion fluency. We found that verbal and emotion fluency increased across age and plateaued in late adolescence around ages 18–19 years. As expected, verbal and emotion fluency were positively and strongly correlated. Participants produced more negative than positive or neutral emotion words, and this negative valence bias did not vary across age. Network approaches revealed a conserved happiness-sadness cluster across age but also shed light on the role of semantic similarity, phonemic similarity, and oppositionality in structuring emotion word generation. Finally, valence, dominance, age of acquisition, concreteness, and word length significantly predicted word order in the emotion fluency task even when controlling for all other dimensions, and the strength of these relationships did not significantly vary with age. In all, results from this study suggest that emotion fluency is a particular subset of verbal fluency and not a distinct kind of fluency. Nevertheless, the order in which words are produced during emotion fluency tasks can reveal key dimensions organizing emotion concepts.

In order to analyze behavioral responses in the emotion fluency and verbal fluency tasks, we had to first score and categorize the words participants produced in these tasks. Previous work has articulated dimensions for defining what is/is not an emotion word (Ortony et al., 1987). However, our field still grapples with the central and culturally variant question of "what is an emotion" (Gendron et al., 2020), and there is no published scoring guide for this task. Consequently, we developed our own publicly available coding dictionaries that determine (a) word validity in each of the tasks and (b) valence categorizations for valid emotion words (OSF at https://osf.io/h3gtp/). We approached this task iteratively and only collapsed categories or excluded responses after verifying that they were produced very rarely. We made this choice both because excluding these words is thus not likely to majorly influence results (as they were produced so rarely) and because the rarity with which borderline words were produced supports the notion that these words are not commonly thought of as "emotion words" (the key criterion of our coding scheme). Following this coding process, we observed similar overall task results as prior studies of adults: Emotion and verbal fluency were strongly correlated, and people tended to produce more negative emotion words than other valences (Abeare et al., 2017; Camodeca et al., 2021; Hegefeld et al., 2023).

Our developmental sample provides novel insight into how these faculties vary across age. Interestingly, we found that both verbal fluency and emotion fluency followed similar age-related trajectories, increasing across childhood and adolescence, before plateauing around 18–19 years of age. In fact, bootstrapping analyses suggested that the plateau points for emotion and verbal fluency were not significantly different. These findings add to emerging evidence that one's verbal fluency abilities strongly predict emotion fluency. However, prior studies primarily demonstrated these similarities using "letter fluency" tasks (i.e., producing as many words that start with the letter "j" in 1 min; Hegefeld et al., 2023), and here we show that these similarities emerge with a more closely matching "category fluency" task (i.e., producing as many fruit words as possible in 60 s). Though it was theoretically possible for the age-related patterns of verbal and emotion fluency to dissociate, we did not observe that to be the case. Consequently, emotion fluency seems to be a specific instance of verbal fluency. Nonetheless, it remains an open question whether differences could emerge in clinical samples. For instance, emotion concepts could become so inaccessible in clinically severe populations that emotion fluency

Table 1
Mixed-Effects Model Testing the Relationship Between Lexical Parameters and Emotion Word Order

Lexical parameter	β	95% CI	*p*
Age of acquisition	.41	[.34, .49]	<.001***
Valence	.11	[.05, .18]	<.001***
Arousal	.02	[−.03, .07]	.402
Dominance	−.11	[−.19, −.04]	.003***
Concreteness	−.06	[−.11, −.02]	.007**
Word length	.06	[.01, .11]	.025*
Word frequency	.05	[−.01, .12]	.124

Note. β values indicate standardized β values, and CI indicates confidence intervals. In the order presented above, higher β values of each lexical parameter indicate that a word was acquired later in life, more positively valenced, more arousing, more dominant, more concrete, longer, and more frequently used in English. CI = confidence interval.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

becomes decoupled from verbal fluency. Interestingly, our observed plateau point for verbal fluency is in slight tension with prior studies showing that children reached stability or adult-like performance in semantic verbal fluency tasks by 10–12 years of age (Regard et al., 1982; Sauzéon et al., 2004). However, neither study examined verbal fluency past adolescence, nor did they apply spline methods to visualize results. As such, our study extends on these prior findings and suggests that verbal fluency may plateau later in development than previously thought.

Participants generated fewer emotion words than fruit words in 60 s, but across our sample, more unique emotion words were generated than fruit words. Although our study was not designed with the intention of directly comparing these nuances, we probed this discrepancy for potential explanations. Visualizations suggested that this may be due to larger consensus on fruits than emotions, allowing participants to generate fruit words more rapidly (Supplemental Figure S4). There are several explanations for this difference. First, fruit words are more concrete than emotion words and thus are likely more accessible due to this overall greater ease in visualizing them and bringing them to mind (Brysbaert et al., 2014). Second, there are likely societal forces producing greater consensus around exemplars of fruits than emotions. For instance, in the American schooling system, children often practice songs naming fruits, but there are not parallel songs for emotions. Third, there is likely overall greater consensus on the boundaries of the category of fruit than emotions. This difference in consensus reflects lay intuitions about fruits being more "natural" kinds with clear group boundaries versus emotions as culturally constructed categories with more idiosyncratic boundaries (Berent et al., 2020; Lindquist & Gendron 2013). Again, in American classrooms, children learn to distinguish fruits from vegetables, but even emotion scholars do not have clear definitions of what is or is not an emotion (Gendron et al., 2020). All these processes could result in greater fluency (though lesser diversity) for generating fruits compared to emotions.

We also found that negative valence bias did not vary significantly across age. Theoretically, this suggests that the proportions of negative, positive, and neutral words that people rapidly access remain stable from ages 4 to 25 years. When coupled with data showing that the number of emotion words that people know increases with age (Baron-Cohen et al., 2010, Nook et al., 2020; Widen & Russell, 2010), this finding suggests that people expand their access to positive and negative emotion words in roughly even proportions as they age. This need not occur out of necessity, as the proportion of negative or positive words that people know or can rapidly generate could vary for different age groups (e.g., children could know many more negative than positive or neutral words, and this proportion could shrink with age). In fact, we considered alternative hypotheses whereby (a) children show heightened access to negative words due to their willingness to approach normatively "negative" stimuli (Grisanzio et al., 2021) and (b) adolescents show an elevated negative valence bias due to normative peaks in risk for psychopathology and contrahedonic motivation in this developmental stage (Fairchild, 2011; McGrath et al., 2023; Riediger et al., 2009). However, our findings did not support either alternative hypothesis and instead suggest that youth and adults are sampling from negative and positive words in similar proportions.

These basic insights into the development of emotion fluency set the stage for future research that addresses both the ontogeny and implications of emotion fluency. What familial, biological, environmental, or cultural factors might shape someone's level of emotion fluency, and what psychological and neural changes might explain the development of these abilities across age? Furthermore, both theory and empirical data argue that facility with emotion language is important to mental health (Erbas et al., 2019; Nook et al., 2018), suggesting that being able to generate and apply emotion words could be important to youth well-being. As noted in the introduction, emotion fluency was *not* related to depression symptoms or emotion regulation success in a community adult sample (Hegefeld et al., 2023). However, this picture may differ in youth or in more severe clinical samples. Specifically, finding that the lion's share of variance in emotion fluency boils down to verbal fluency can guide the field's thinking in how relations between emotion fluency and well-being must unfold. Either this relationship must operate through general word production difficulties or it must emerge through a (rare) selective impairment in emotion word production. This can motivate more pointed research designs, such as (a) examining more severe patient populations that are known to experience verbal fluency impairments (Bertschy et al., 2023) and testing whether emotion fluency has predictive power in explaining mental health symptoms beyond verbal fluency in these populations or (b) testing whether emotion fluency but not verbal fluency explains mental health difficulties in people with semantic aphasia for emotion words (Lindquist et al., 2015; Souter et al., 2022). Thus, this finding narrows in the theoretical space in which this measure of emotion language can track mental health symptoms.

Beyond examining emotion fluency variables themselves, we borrowed from the cognitive literature (Mills & Phillips, 2023; Sauzéon et al., 2004; Zemla & Austerweil, 2017) and tested whether analyzing the *order* in which participants thought of emotion words could reveal the dimensions underlying emotion word representation. Network analyses revealed clusters of emotion words that tended to be produced in close proximity by participants. Happiness and sadness were consistently in close proximity, regardless of age. This again underscores the foundational importance of these two early-acquired emotion words as highly accessible exemplars of emotions, even in very young minds. These network analyses also supported semantic associations between "fear family" and "anger family" emotions (Keltner et al., 2023). However, clustering results also revealed that participants associated *antonyms* (e.g., love/hate) and that they used phonemic qualities (such as the first letter of envy and elation) to access emotion words in this task. These findings are in line with other verbal fluency tasks (Abwender et al., 2001), again pointing to general overlaps between emotional and nonemotional verbal fluency tasks. A final note is the appearance of "anxious misery/psychopathology" words in adolescence, a finding that may reflect emerging awareness of these constructs in this period of heightened risk for psychopathology (Kessler et al., 2005).

Turning to results of our mixed-effects models, we hypothesized that dimensions of the circumplex model could structure word order (e.g., people would move along the dimensions of valence, arousal, and dominance as they search semantic space for emotion words; Posner et al., 2005; Russell, 1980; Russell & Mehrabian, 1977). Results partially supported this notion, as valence and dominance did predict word order: People tended to say negative and more dominant emotion words before positive and less dominant emotion words. Interestingly, though, arousal was not related to word order, suggesting that it does not structure accessibility of emotion words in tasks like these. However, the size of the effect of valence and dominance was much smaller than a more "cognitive" dimension: age of acquisition. Consequently, the first emotion words that people think of tend to be those that they learned earliest in life, and the strongest factor that reflects how people access emotion words is age of acquisition, not valence or other affective dimensions. This parallels results for nonemotion fluency tasks, again suggesting that this task taps a more cognitive or semantic framework for emotion accessibility. In fact, 68% of participants began with the words "happy" and "sad," reflecting early attention to the "good/bad" emotion dichotomy (Bridges, 1930; Nook et al., 2017; Widen & Russell, 2010). People also tended to think of more concrete and shorter words before they thought of more abstract and longer words, as has also been found in other word naming and memory retrieval tasks (Brysbaert et al., 2014; Kuperman et al., 2012). Finally, all these effects emerge above and over all other predictors (e.g., word frequency, word length).

Within our full sample, age did not moderate the strength of the associations between these dimensions and word order. These results suggest that whether you are a child, an adolescent, or an adult, the first words that come to mind when pressed to think of emotions are those that you learned earliest and those that are more negative, less dominant, less concrete, and shorter. A supplementary analysis demonstrated that age of acquisition, valence, and dominance remained significant predictors of word order when analyzing data only from adult participants. As such, even adults (not just youth) start naming emotion words they learned earliest.

Together, these word order results reveal something interesting about the way people organize emotion concepts. When people are prompted to generate emotion words, they gravitate first to happiness and sadness. This finding echoes early theories from Bridges (1930), arguing that happiness and sadness are foundational emotions from which all others stem, and Widen and Russell (2010), who show that toddlers tend to categorize emotional faces on this good/bad dichotomous scale. In fact, recent studies show that children attend more to a positivity/negativity dimension of emotion than other dimensions, suggesting that happiness/sadness and good/bad dichotomies are foundational to segmenting emotion concepts (Nook et al., 2017). Moreover, happiness and sadness are often the first emotion words and concepts acquired in early childhood, further supporting the idea that children may use happiness/sadness as anchors to organize new emotions (Ridgeway et al., 1985). One curious implication of these findings is that emotion differentiation (i.e., one's ability to separate similarly valenced emotions using precise terms) might emerge based on the extent to which people move beyond the most accessible terms of "happy" and "sad." As such, some people might only think of emotions using these rather than more specific terms because it is effortful to move beyond these early-acquired and foundational emotion words. Consequently, these findings prompt future research testing whether searching and applying words beyond happy or sad are indeed effortful, and if so, whether this explains individual differences in emotion differentiation.

Constraints on Generality

The present study has several strengths. First, it uses a wide age range to examine broad normative shifts in our dependent variables all the way from childhood through young adulthood. Second, we use cutting-edge statistical approaches and use supplemental analyses to examine whether word substitutions or specific age ranges influenced results. Third, we integrate cognitive, linguistic, affective, and developmental approaches to provide new insight into the accessibility of emotion language across development. Nevertheless, our study also has several limitations that should be addressed in future work.

First, a cross-sectional sample limits causal conclusions regarding the influence of age per se on our dependent variables. Cohort effects or other third variables could possibly shape these results. Future longitudinal studies could rule out these alternative explanations. Additionally, word production in both fluency tasks could be impacted by the fact that responses were verbalized instead of written. Although we intuited that asking participants to write in this task would likely produce age-related confounds (as children likely write slower than adults, and the youngest children in our sample could not yet write), alternative methods of assessing fluency or access to emotion words could grant deeper insight into the phenomena at hand.

Participants always completed the verbal fluency task before the emotion fluency task, possibly priming a certain approach to the emotion fluency task or inducing some fatigue that may have contributed to the finding that participants produced more fruit words than emotion words. Our motivation for keeping task order consistent was to give children the opportunity to start with a more familiar category for their fluency task (i.e., fruits) rather than the potentially more challenging category of emotions. Nonetheless, we agree that the consistent order should be kept in mind when interpreting results and that future studies should consider counter-balancing task orders to provide a firmer comparison of these measures.

We also studied a majority White population of English-speaking participants recruited in the United States, limiting our inferences about emotion words in other populations, countries, and languages. Further, we would like to acknowledge that (a) our categorization of emotion words and (b) intuition about the role that emotion words play in mental health are inherently a reflection of our linguistic and cultural context (North American English). For instance, our decision to exclude body-centered emotion words (e.g., hungry) from our list of "valid" emotion words may not be appropriate for all cultures (Dzokoto, 2010; Hoemann et al., 2024). Consequently, future researchers should consider their local contexts when adapting or interpreting our categorization of emotions by crowdsourcing normative classifications and employing intercoder reliability methods.

Last, though we posit that these tasks might imply how people generate words to label their own and others' emotions, our findings cannot be generalized to these specific settings, as they were conducted in the absence of mood inductions and the explicit goal to label one's own affect. As such, future work could strengthen our understanding of the implications of tasks like these by prompting participants to generate a string of emotion words in response to personally relevant or emotional stimuli. Such investigations would further reveal the downstream implications of emotion fluency and its development.

Conclusion

To conclude, the present study examined age-related variation in verbal fluency, emotion fluency, and negative valence bias (i.e., the tendency to produce more negative than positive or neutral emotion words). Using established lexical norms, we also tested whether the order in which participants generate emotion words reveals the underlying organizational structure of emotion concepts. Results from this study suggest that emotion fluency is a subtype of verbal fluency, whereby their age-related trajectories similarly plateau in late adolescence. Further, the order in which words are produced in the emotion fluency task reflects key dimensions of emotion concepts.

References

Abeare, C. A., Freund, S., Kaploun, K., McAuley, T., & Dumitrescu, C. (2017). The Emotion Word Fluency Test (EWFT): Initial psychometric, validation, and physiological evidence in young adults. *Journal of Clinical and Experimental Neuropsychology*, *39*(8), 738–752. https://doi.org/10.1080/13803395.2016.1259396

Abwender, D. A., Swan, J. G., Bowerman, J. T., & Connolly, S. W. (2001). Qualitative analysis of verbal fluency output: Review and comparison of several scoring methods. *Assessment*, *8*(3), 323–338. https://doi.org/10.1177/107319110100800308

Anglin, J. M., Miller, G. A., & Wakefield, P. C. (1993). Vocabulary development: A morphological analysis. *Monographs of the Society for Research in Child Development*, *58*(10), v–165. https://doi.org/10.2307/1166112

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

Barrett, L. F. (2004). Feelings or words? Understanding the content in self-report ratings of experienced emotion. *Journal of Personality and Social Psychology*, *87*(2), 266–281. https://doi.org/10.1037/0022-3514.87.2.266

Barrett, L. F., Gross, J., Christensen, T. C., & Benvenuto, M. (2001). Knowing what you're feeling and knowing what to do about it: Mapping the relation between emotion differentiation and emotion regulation. *Cognition and Emotion*, *15*(6), 713–724. https://doi.org/10.1080/02699930143000239

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Ben-Shachar, M. S., Lüdecke, D., & Makowski, D. (2020). effectsize: Estimation of effect size indices and standardized parameters. *Journal of Open Source Software*, *5*(56), Article 2815. https://doi.org/10.21105/joss.02815

Berent, I., Feldman Barrett, L., & Platt, M. (2020). Essentialist biases in reasoning about emotions. *Frontiers in Psychology*, *11*, Article 562666. https://doi.org/10.3389/fpsyg.2020.562666

Bertschy, G., Martz, E., Weibel, S., & Weiner, L. (2023). Psychopathological dissection of bipolar disorder and ADHD: Focusing on racing thoughts and verbal fluency. *Neuropsychiatric Disease and Treatment*, *19*, 1153–1168. https://doi.org/10.2147/NDT.S401330

Bridges, K. M. B. (1930). A genetic theory of the emotions. *The Pedagogical Seminary and Journal of Genetic Psychology*, *37*, 514–527.

Brysbaert, M., & Biemiller, A. (2017). Test-based age-of-acquisition norms for 44 thousand English word meanings. *Behavior Research Methods*, *49*(4), 1520–1523. https://doi.org/10.3758/s13428-016-0811-4

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, Instruments & Computers*, *41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, *46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Camodeca, A., Walcott, K., Hosack, A., & Todd, K. Q. (2021). Preliminary evidence for the Emotion Word Fluency Test as a unique semantic fluency measure. *Psychological Assessment*, *33*(2), 195–200. https://doi.org/10.1037/pas0000965

Canty, A., & Ripley, B. (2024). *boot: Bootstrap R (S-Plus) functions* (R package Version 1.3-30) [Computer software].

Cicchetti, D. (1993). Developmental psychopathology: Reactions, reflections, projections. *Developmental Review*, *13*(4), 471–502. https://doi.org/10.1006/drev.1993.1021

Craske, M. G., Rauch, S. L., Ursano, R., Prenoveau, J., Pine, D. S., & Zinbarg, R. E. (2009). What is an anxiety disorder? *Depression and Anxiety*, *26*(12), 1066–1085. https://doi.org/10.1002/da.20633

Csardi, G., & Nepusz, T. (2006). "The igraph software package for complex network research." *InterJournal, Complex Systems*, *1695*, 1–9. https://www.scirp.org/reference/referencespapers?referenceid=2637714

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge University Press. https://doi.org/10.1017/CBO9780511802843

DiCiccio, T. J., & Efron, B. (1996). Bootstrap confidence intervals. *Statistical Science*, *11*(3), 189–212. https://doi.org/10.1214/ss/1032280214

Dzokoto, V. (2010). Different ways of feeling: Emotion and somatic awareness in Ghanaians and Euro-Americans. *Journal of Social, Evolutionary, and Cultural Psychology*, *4*(2), 68–78. https://doi.org/10.1037/h0099299

Erbas, Y., Ceulemans, E., Blanke, E. S., Sels, L., Fischer, A., & Kuppens, P. (2019). Emotion differentiation dissected: Between-category, within-category, and integral emotion differentiation, and their relation to well-being. *Cognition and Emotion*, *33*(2), 258–271. https://doi.org/10.1080/02699931.2018.1465894

Fairchild, G. (2011). The developmental psychopathology of motivation in adolescence. *Developmental Cognitive Neuroscience*, *1*(4), 414–429. https://doi.org/10.1016/j.dcn.2011.07.009

Fields-Olivieri, M. A., Kim, Y., Jennings, K. J., & Cole, P. M. (2024). The role of language in the development of emotion regulation. In M. A. Bell (Ed.), *Child development at the intersection of emotion and cognition* (2nd ed., pp. 31–50). American Psychological Association. https://doi.org/10.1037/0000406-003

Gendron, M., Hoemann, K., Crittenden, A. N., Mangola, S. M., Ruark, G. A., & Barrett, L. F. (2020). Emotion perception in Hadza hunter–gatherers. *Scientific Reports*, *10*(1), Article 3867. https://doi.org/10.1038/s41598-020-60257-2

Grisanzio, K. A., Sasse, S. F., Nook, E. C., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2021). Voluntary pursuit of negatively valenced stimuli from childhood to early adulthood. *Developmental Science*, *24*(2), Article e13012. https://doi.org/10.1111/desc.13012

Hegefeld, H. M., Satpute, A. B., Ochsner, K. N., Davidow, J. Y., & Nook, E. C. (2023). Fluency generating emotion words correlates with verbal measures but not emotion regulation, alexithymia, or depressive symptoms. *Emotion*, *23*(8), 2259–2269. https://doi.org/10.1037/emo0001229

Hoemann, K., Gendron, M., Crittenden, A. N., Mangola, S. M., Endeko, E. S., Dussault, È., Barrett, L. F., & Mesquita, B. (2024). What we can learn about emotion by talking with the Hadza. *Perspectives on Psychological Science*, *19*(1), 173–200. https://doi.org/10.1177/17456916231178555

Hoemann, K., Nielson, C., Yuen, A., Gurera, J. W., Quigley, K. S., & Barrett, L. F. (2021). Expertise in emotion: A scoping review and unifying framework for individual differences in the mental representation of emotional experience. *Psychological Bulletin*, *147*(11), 1159–1183. https://doi.org/10.1037/bul0000327

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

Jackson, J. C., Lindquist, K., Drabble, R., Atkinson, Q., & Watts, J. (2023). Valence-dependent mutation in lexical evolution. *Nature Human Behaviour*, *7*(2), 190–199. https://doi.org/10.1038/s41562-022-01483-8

Jenness, J. L., Lambert, H. K., Bitrán, D., Blossom, J. B., Nook, E. C., Sasse, S. F., Somerville, L. H., & McLaughlin, K. A. (2021). Developmental variation in the associations of attention bias to emotion with internalizing and externalizing psychopathology. *Research on Child and Adolescent Psychopathology*, *49*(6), 711–726. https://doi.org/10.1007/s10802-020-00751-3

Kalokerinos, E. K., Erbas, Y., Ceulemans, E., & Kuppens, P. (2019). Differentiate to regulate: Low negative emotion differentiation is associated with ineffective use but not selection of emotion-regulation strategies. *Psychological Science*, *30*(6), 863–879. https://doi.org/10.1177/0956797619838763

Keltner, D., Brooks, J. A., & Cowen, A. (2023). Semantic space theory: Data-driven insights into basic emotions. *Current Directions in Psychological Science*, *32*(3), 242–249. https://doi.org/10.1177/09637214221150511

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 593–602. https://doi.org/10.1001/archpsyc.62.6.593

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lawrence, M. A. (2016). *ez: Easy analysis and visualization of factorial experiments* (R Package Version 4.4-0) [Computer software]. https://rdrr.io/cran/ez/

Lindquist, K. A., & Gendron, M. (2013). What's in a word? Language constructs emotion perception. *Emotion Review*, *5*(1), 66–71. https://doi.org/10.1177/1754073912451351

Lindquist, K. A., Satpute, A. B., & Gendron, M. (2015). Does language do more than communicate emotion? *Current Directions in Psychological Science*, *24*(2), 99–108. https://doi.org/10.1177/0963721414553440

Lüdecke, D., Ben-Shachar, M., Patil, I., & Makowski, D. (2020). Extracting, computing and exploring the parameters of statistical models using R. *Journal of Open Source Software*, *5*(53), Article 2445. https://doi.org/10.21105/joss.02445

McGrath, J. J., Al-Hamzawi, A., Alonso, J., Altwaijri, Y., Andrade, L. H., Bromet, E. J., Bruffaerts, R., de Almeida, J. M. C., Chardoul, S., Chiu, W. T., Degenhardt, L., Demler, O. V., Ferry, F., Gureje, O., Haro, J. M., Karam, E. G., Karam, G., Khaled, S. M., Kovess-Masfety, V., … the WHO World Mental Health Survey Collaborators. (2023). Age of onset and cumulative risk of mental disorders: A cross-national analysis of population surveys from 29 countries. *The Lancet Psychiatry*, *10*(9), 668–681. https://doi.org/10.1016/S2215-0366(23)00193-1

Mills, T., & Phillips, J. (2023). Locating what comes to mind in empirically derived representational spaces. *Cognition*, *240*, Article 105549. https://doi.org/10.1016/j.cognition.2023.105549

Mohammad, S. (2018). Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. *Proceedings of the 56th annual meeting of the association for computational linguistics* (Vol. 1: Long Papers, pp. 174–184). https://doi.org/10.18653/v1/P18-1017

Nathanson, L., Rivers, S. E., Flynn, L. M., & Brackett, M. A. (2016). Creating emotionally intelligent schools with RULER. *Emotion Review*, *8*(4), 305–310. https://doi.org/10.1177/1754073916650495

Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2017). Increasing verbal knowledge mediates development of multidimensional emotion representations. *Nature Human Behaviour*, *1*(12), 881–889. https://doi.org/10.1038/s41562-017-0238-7

Nook, E. C., Sasse, S. F., Lambert, H. K., McLaughlin, K. A., & Somerville, L. H. (2018). The Nonlinear development of emotion differentiation: Granular emotional experience is low in adolescence. *Psychological Science*, *29*(8), 1346–1357. https://doi.org/10.1177/0956797618773357

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

Ortony, A., Clore, G. L. & Foss, M. A. (1987). The referential structure of the affective lexicon. *Cognitive Science*, *11*(3), 341–364. https://doi.org/10.1207/s15516709cog1103_4

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*, *17*(3), 715–734. https://doi.org/10.1017/S0954579405050340

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

Regard, M., Strauss, E., & Knapp, P. (1982). Children's production on verbal and non-verbal fluency tasks. *Perceptual and Motor Skills*, *55*(3), 839–844. https://doi.org/10.2466/pms.1982.55.3.839

Ridgeway, D., Waters, E., & Kuczaj, S. A. (1985). Acquisition of emotion-descriptive language: Receptive and productive vocabulary norms for ages 18 months to 6 years. *Developmental Psychology*, *21*(5), 901–908. https://doi.org/10.1037/0012-1649.21.5.901

Riediger, M., Schmiedek, F., Wagner, G. G., & Lindenberger, U. (2009). Seeking pleasure and seeking pain: Differences in prohedonic and contra-hedonic motivation from adolescence to old age. *Psychological Science*, *20*(12), 1529–1535. https://doi.org/10.1111/j.1467-9280.2009.02473.x

Rimé, B. (2009). Emotion elicits the social sharing of emotion: Theory and empirical review. *Emotion Review*, *1*(1), 60–85. https://doi.org/10.1177/1754073908097189

Riva, D., Nichelli, F., & Devoti, M. (2000). Developmental aspects of verbal fluency and confrontation naming in children. *Brain and Language*, *71*(2), 267–284. https://doi.org/10.1006/brln.1999.2166

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382–439. https://doi.org/10.1016/0010-0285(76)90013-X

Rosselli, M., Ardila, A., Araujo, K., Weekes, V. A., Caracciolo, V., Padilla, M., & Ostrosky-Solís, F. (2000). Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied Neuropsychology*, *7*(1), 17–24. https://doi.org/10.1207/S15324826AN0701_3

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296–320. https://doi.org/10.1207/S15327957PSPR0504_2

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*(6), 1161–1178. https://doi.org/10.1037/h0077714

Russell, J. A., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, *11*(3), 273–294. https://doi.org/10.1016/0092-6566(77)90037-X

Sauzéon, H., Lestage, P., Raboutet, C., N'Kaoua, B., & Claverie, B. (2004). Verbal fluency output in children aged 7–16 as a function of the production criterion: Qualitative analysis of clustering, switching processes, and semantic network exploitation. *Brain and Language*, *89*(1), 192–202. https://doi.org/10.1016/S0093-934X(03)00367-5

Schrauf, R. W., & Sanchez, J. (2004). The preponderance of negative emotion words in the emotion lexicon: A cross-generational and crosslinguistic study. *Journal of Multilingual and Multicultural Development*, *25*(2–3), 266–284. https://doi.org/10.1080/01434630408666532

Simpson, G. (2014). *First derivative of a spline term* [Computer software]. GitHub Repository. https://gist.github.com/gavinsimpson/e73f011fdaaab4bb5a30

Sokołowski, A., Tyburski, E., Sołtys, A., & Karabanowicz, E. (2020). Sex differences in verbal fluency among young adults. *Advances in Cognitive Psychology*, *16*(2), 92–102. https://doi.org/10.5709/acp-0288-1

Souter, N. E., Wang, X., Thompson, H., Krieger-Redwood, K., Halay, A. J., Lambon, M., Michel Thiebaut de Schotten, R., & Jeffries, E. (2022). Mapping lesion, structural disconnection, and functional disconnection to symptoms in semantic aphasia. *Brain Structure & Function*, *227*, 3043–3061. https://doi.org/10.1007/s00429-022-02526-6

Sun, S., Plate, R. C., Jones, C., Rodriguez, Y., Katz, C., Murin, M., Pearson, J., Parish-Morris, J., & Waller, R. (2024). Childhood conduct problems and parent–child talk during social and nonsocial play contexts: A naturalistic home-based experiment. *Scientific Reports*, *14*(1), Article 1018. https://doi.org/10.1038/s41598-024-51656-w

Vine, V., Boyd, R. L., & Pennebaker, J. W. (2020). Natural emotion vocabularies as windows on distress and well-being. *Nature Communications*, *11*(1), Article 4525. https://doi.org/10.1038/s41467-020-18349-0

Weissman, D. G., Nook, E. C., Dews, A. A., Miller, A. B., Lambert, H. K., Sasse, S. F., Somerville, L. H., & McLaughlin, K. A. (2020). Low emotional awareness as a transdiagnostic mechanism underlying psychopathology in adolescence. *Clinical Psychological Science*, *8*(6), 971–988. https://doi.org/10.1177/2167702620923649

Widen, S. C. (2013). Children's interpretation of facial expressions: The long path from valence-based to specific discrete categories. *Emotion Review*, *5*(1), 72–77. https://doi.org/10.1177/1754073912451492

Widen, S. C., & Russell, J. A. (2010). Differentiation in preschooler's categories of emotion. *Emotion*, *10*(5), 651–661. https://doi.org/10.1037/a0019005

Wood, S. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society. Series B. Methodological*, *65*(1), 95–114. https://doi.org/10.1111/1467-9868.00374

Wood, S. (2017). *Generalized additive models: An introduction with R* (2nd ed.). Chapman and Hall/CRC. https://doi.org/10.1201/9781315370279

Zammuner, V. (2011). People's active emotion vocabulary: Free listing of emotion labels and their association to salient psychological variables. In A. Esposito, A. Vinciarelli, K. Vicsi, C. Pelachaud, & A. Nijholt (Eds.), *Lecture notes in computer science: Vol. 6800. Analysis of verbal and nonverbal communication and enactment: The processing issues* (pp. 449–460). Springer. https://doi.org/10.1007/978-3-642-25775-9_40

Zeileis, A., & Hothorn, T. (2002). Diagnostic checking in regression relationships. *R News*, *2*(3), 7–10. https://CRAN.R-project.org/doc/Rnews/

Zemla, J. C., & Austerweil, J. L. (2017). Modeling semantic fluency data as search on a semantic network. *Proceedings of the 39th annual conference of the cognitive science society* (pp. 3646–3651). Cognitive Science Society.

Zemla, J. C., & Austerweil, J. L. (2018). Estimating semantic networks of groups and individuals from fluency data. *Computational Brain & Behavior*, *1*(1), 36–58. https://doi.org/10.1007/s42113-018-0003-7

Zemla, J. C., Cao, K., Mueller, K. D., & Austerweil, J. L. (2020). SNAFU: The semantic network and fluency utility. *Behavior Research Methods*, *52*(4), 1681–1699. https://doi.org/10.3758/s13428-019-01343-w

Received October 26, 2023
Revision received November 11, 2024
Accepted November 19, 2024 ■