Age and Affect in Language: How Emotion Expression on Social Media Varies Across Adulthood

Daniela Teodorescu*, **Jan Philip Wahle**†, **Saif M. Mohammad**‡
*University of Alberta, †University of Göttingen, ‡National Research Council Canada
*dteodore@ualberta.ca, †wahle@uni-goettingen.de, ‡saif.mohammad@nrc-cnrc.gc.ca

Abstract

As we age, the way we experience and express emotions changes. This is because of a number of factors, including: changes in our body, differing types of experiences at different ages, improved emotion regulation strategies, and increasing experience of dealing with affective situations. However, work in psychology points to differing findings on how emotions, typically valence and happiness, changes with age. Psychologists measure happiness and well-being through questionnaires, which can have biases and result in limited data. Thus corpus analyses can provide useful complementary insights. We compile and release a large dataset of social media posts annotated with the age of the author at the time of posting. We refer to it as *AgeCorpus*. Using this dataset, we apply simple and interpretable methods to explore research questions pertaining to how social media posts, especially emotion expression through these posts, varies by age groups. Analyzing the emotions expressed in the posts, we find that the average valence increases until the middle ages, and then decreases; arousal decreases (Reddit)/plateaus with age (Twitter); and dominance follows the inverted U-shape (Reddit)/increases with age (Twitter). For categorical emotions, we find they follow the inverted U-shape on Reddit and increase in intensity with age on Twitter. We hope our dataset enables further research into age related phenomenon, such as well-being and language use.

Keywords: Age, Affect, Emotion, Valence, Arousal, Dominance, Categorical Emotions, Corpus creation, Social Media

1. Introduction

The way in which we perceive and express emotions is dependent on many aspects (e.g., environment, life events, personality dispositions, etc.), and change with age. With age and increased exposure to emotional situations, people's ability to regulate emotions improves and therefore older adults report higher levels of well-being (Urry and Gross, 2010; Gross et al., 1997). The way in which we perceive emotions is important as it directly impacts how we process that emotion information, respond, and act on that decision (Appraisal Theory of Emotion (Arnold, 1960)).

Literature in psychology has found that there are general trends in emotion expression with age. However, these findings do not align and point to differing trends which can be grouped into 3 broad categories: "U-shape" (Blanchflower and Oswald, 2008), "Inverted U-shape", and linear or there is no change with age. The U-shape theory says that generally happiness is highest for those in their early twenties to thirty-five, dips during midlife (e.g., 35-50 years old), and rises again during fifties to seventies, following a U-shaped curve. On the other hand, in the inverted U-shape theories happiness peaks during midlife, and is lowest in 20's to early 30's, and likewise again in older age. Whereas other work points to there being no relationship between emotions and age, with positive and negative emotions being evenly distributed across the life-course.

Understanding how emotions change across the lifespan is critical for answering questions related to what changes (e.g., experience, regulation, expression) and why changes occur across age (e.g., cognition, biology, culture). Answering such questions is important to determine whether such changes are adaptive or problematic. If problematic, are there appropriate interventions that can be taken to help improve health outcomes and well-being and when should they be taken? Such findings can support clinicians provide more tailored care and appropriate interventions when needed. A better understanding the relationship between age and emotion is of interest to numerous fields: psychology (e.g., how does emotion experience, expression and regulation change across the lifespan), mental health (e.g., what does healthy changes look like vs. those associated with a pathology), education (e.g., emotional experiences are key to learning, how can classrooms best support age-related emotional processing), behavioral economics (e.g., how does emotions changes with age impact decisions and consumer trends) and more.

In much of the psychology work on age and emotions, the data studied was obtained from surveys or census asking how happy people feel, or how satisfied they feel. Surveys can have biases (e.g., social pressures to appear happy, population representation biases, etc.), and are timely to collect for large populations. Further, the amount of data that can be collected from surveys is rather limited. In this work, we analyze large amounts of data systematically and efficiently by looking at the emotions expressed in text, specifically social

Proceedings of Computational Affective Science (CAS) @ LREC 2026, pages 175–189
16 May 2026. ©ELRA Language Resources Association (ELRA), 2026

media. While people are not asked to describe how they feel when posting on social media, inadvertently through the words used, they express how they are feeling. Therefore, we can often infer how people feel through the emotions expressed in their posts. Further, the text we write on social media also impacts others – providing another angle on this question. We create a large corpus of social media posts and determine whether the expression of emotion in text varies across age groups, and if so, how does it vary. Further, most if not all of the work in psychology measures one dimension of emotion – valence (the pleasure–displeasure, happiness–unhappiness, or positive–negative dimension). Affect encompasses valence, arousal, and dominance and are thought to be the the primary dimensions of emotion (Osgood et al., 1957; Russell, 2003; Russell and Mehrabian, 1977). Whereas categorical emotions include joy, sadness, fear, etc. such as Ekman's 6 emotions (anger, fear, joy, sadness, disgust, surprise) (Ekman, 1992) and Plutchik's Wheel of Emotions (additionally including anticipation and trust) (Plutchik, 1980). Emotions are multifaceted and it is important to understand how do all types of emotions which we experience every day (e.g., fear, anger) and affect vary across age groups? Each and every one of these dimensions has an impact on overall well-being.

In this paper we introduce *AgeCorpus* – a dataset of social media posts annotated with the age of the author at the time of writing. We use *AgeCorpus* to explore research questions on how emotion varies across age groups. We provide novel insights from two social media platforms, X (formerly known as Twitter) and Reddit, and contribute a large dataset for affective science research. However, there are several ways one could go about studying emotion change with age. One approach is having a group of individuals and following how their emotion expression in text changes as they age through their 20's, 30's, and onwards, allowing us to study *how emotion expressed changes across age* with a consistent cohort. Although such data sources are limited, so this would be a challenge. Therefore, we ask a more specific question - how does emotion expression vary across age groups *during specific periods of time*. This date range, or period of time, is a parameter and can be set to any value e.g., from a few years to a decade. We create a corpus of *contemporary* age utterances, as our dataset covers 2010–2022 for Reddit and 2020–2021 for Twitter. We refer to this dataset as *AgeCorpus*. We add this nuance to our research question as being in your 20's during the 1940's is quite different from being in your 20's during the early 2000's. For simplicity, from here onwards we will refer to our research question as "how does emotion vary across age groups".

Using *AgeCorpus* dataset we investigate the following research questions:

1. What patterns do we see in posting characteristics with age? This can help inform research on well-being such as posting frequency may be related to social withdrawal, or how use of pronouns can be indicative of mental health.
2. Does emotion expression vary by age? How?

We will make our code and dataset publicly available.[1] Our data is also a part of the ABCDE (Affect, Body, Cognition, Demographics, and Emotion) corpus (Wahle et al., 2026) where there are various additional annotations included beyond age, allowing for a wide range of research questions on the relationship between age and other features to be explored.

2. Related Work

2.1. Psychology, Age, and Emotions

Many theories in psychology have explored how well-being and happiness change with age. One branch of theories describe happiness levels as following a "U-shape" with age, whereas others say it follows an "Inverted U-Shape". More so, others say there is simply no relationship with age. Blanchflower and Oswald (2008) introduced the U-shape theory, demonstrating that psychological well-being is at its lowest and mental distress at its highest during midlife. Potential confounding factors such as income, education, marital status, and region have been considered and the U-shape still holds. Similar trends were found regardless of employment status (Clark and Oswald, 1994). Likewise, Deaton (2008) found that mean life satisfaction is U-shaped with age for high income countries, whereas declines with age for middle-income countries but to a lesser extent for low-income countries. Similarly, Kim et al. (2023) noted the U-shape relationship between well-being and age, however it was moderated by personality. On the other hand, some work in psychology points to an inverted U-shape between age and well-being with happiness increasing until 50's/60's and decreasing afterwards (Easterlin, 2006), even after considering extraversion level (Mroczek and Spiro, 2005). Whereas Myers and Diener (1995) claimed that regardless of gender, race, and socioeconomic status, satisfaction with life remained constant with age.

Work in psychology has also studied the relationship between patterns in language use and well-being. For example, the proportion of pronouns can be an indicator of depression (consistently found in a meta-survey (Edwards and Holtzman, 2017)), and the pattern with which emotions change over

[1] https://github.com/dteodore/Age--Affect_AgeCorpus

time (i.e., referred to as *emotion dynamics*) is associated with mental health diagnoses (e.g., more negative affect over time and higher fluctuations of emotional states are associated with anxiety, depression (Houben et al., 2015; Heller et al., 2018)).

2.2. NLP, Age, and Emotions

Past NLP work created age corpora for supporting the development of age-detection models (Chew et al., 2021; Rosenthal and McKeown, 2011; Peersman et al., 2011) and personalized models (e.g., search and recommendations (Tigunova et al., 2020)). Other work, created age corpora to enable research in the social sciences (e.g., Reddit data with demographic and personality labels (Gjurković et al., 2021)) and to mitigate bias in language technologies (e.g., demographic-aware word embeddings (Welch et al., 2020)).

Few works, studied how emotions varies across age groups. Hipson and Mohammad (2020) explored the emotion expressed in poems for grade 1–12 students and found decreased valence during mid-adolescence, and arousal and dominance peaked during adolescence. We focus on how emotion varies across adulthood in the social media domain. Schwartz et al. (2013) found that in Facebook posts there was a positive relationship between positive emotion and age, whereas a negative relationship for negative emotion with age. Lastly, the most related is by Li (2023), who explored whether older people expressed more positive emotions than younger people for 8 discrete emotions with the spoken portion of the BNC (fiction, magazines, newspapers, and academic). Overall, their findings aligned with the U-shaped theory in psychology – older people expressed less negative emotions (except for sadness), but also there were fluctuations in how positive emotions increased with age. In contrast, we create and explore an utterance–age resource for Reddit and Twitter.

3. AgeCorpus Dataset

We create a dataset of textual utterances/posts from social media annotated with the speaker's age at the time of posting. Such a dataset enables a wide array of research such as on how language characteristics vary by age (e.g., post length, metaphor usage, etc.), or for studying aspects of well-being (e.g., loneliness, mental health) across age groups. We focus on social media as a source of text data as it allows us to study large amounts of data and is more accessible compared to other text mediums (e.g., essays, discussion forums). We chose Reddit and Twitter as our platforms as these are two representative social media sources which are rich in text data. Rather than using APIs to scrape these social media platforms, we use existing dumps of social media. Future research could extend this work to multimodal platforms such as Instagram or TikTok. We make our dataset publicly available.[2]

Social media data instances (e.g., Tweets, Reddit posts) often have various metadata information associated with them such as the time of posting, user id, number of likes, reposts, retweets, location, etc. However, the age of the author is not included in the metadata. Some people choose to include their age in their profile or in their social media bio, however this is often noisy data and may be outdated. In our work, we make use of self-declarations of age in posts: e.g., "*I'm so happy to be turning 25, happy birthday to me!*" or "*I'm 30 and relate to this...*". Using the age mentioned in the self-declarations and the time stamp of that post, we can determine the age of the person at the time of the posting. Then we can determine their age at the time they made any other (past or future) target post, by simply determining the difference in time between the target post and the time of the self-declaration post.

People can self-declare their age in many ways, but there are certain templates that are common: e.g., ''*I'm turning 25 years old today!*". We make use of these templates to automatically identify self-declarations.[3] In Appendix A, we show the regexes used to capture these templates. We note that age declarations can appear in two forms: people declare their age at the time of posting; and those that declare their date of birth (e.g., birth year). We make use of high-precision templates since these templates can sometimes capture false positives (e.g., "*When I was 26 years old ...*"). Further, an author can have multiple age declarations, so we increase the precision of the self-declarations using the approach described in the section below.

3.1. Resolving Multiple Age Declarations

We determine the birth year from the self-declaration of age and time of posting. For any user that is below 13 years old, we remove their posts (minimum age to use these social media platforms). For users with several birth year declarations, birth years that are similar (within 2 years) are clustered together. Each cluster is assigned a weight, and the cluster with the highest weight is selected. The final birth year is calculated as the weighted average of that cluster. We remove posts from users with ages determined to be outside of 13–100 (as users older than 100 are unlikely). This methodol-

[2] https://github.com/dteodore/Age--Affect_AgeCorpus

[3] We do not aim to capture all possible ways; just commonly used ways.

Dataset	Time Period	# Users	Total # Utterances	User Avg. # Utterances (Std dev.)
Reddit	2010–2022	1,047,553	33,495,292	31.97 (841.89)
TUSC-City	2020–2021	48,712	7,486,672	153.69 (230.715)

Table 1: The number of users and number of utterances with age annotations obtained from the Reddit and Twitter datasets.

ogy ensures that if multiple ages are declared, the most probable one is selected.

3.2. Reddit Dataset

Reddit data is not easily accessible due to API changes so we use past Reddit Pushshift dumps from 2010 up to and including 2022.[4] We search for posts matching the age declaration templates and identify potential users. Next, we collect all of the users' posts.

3.3. Twitter Dataset

We use an existing Twitter corpus, TUSC-City (Vishnubhotla and Mohammad, 2022), which contains geo-located English Tweets from US and Canada. It consists of posts collected via the free Twitter API by using geo-location and random sampling from 46 American and Canadian cities.

We likewise search for users from TUSC-City using the above mentioned matching templates, and obtain the posts for these users.

The number of users, utterances, and other details of the dataset are in Table 1, and the number of users in each age group is in Appendix C. Most age groups have about 2,000 users from Twitter and about 20,000 users from Reddit, and the smallest age group had 360 users, ensuring a strong enough signal when performing the below experiments.

4. How do Social Media Post Characteristics Change with Age?

Using *AgeCorpus* we explore a series of questions on how social media posting habits and social media posts vary across age groups. In the following section, we explore how emotion word usage varies with age. We remind readers (as mentioned in the Introduction), we will use for short how emotion expression "varies across age groups" although we are referring to more specifically "varies across age groups during the time periods of the dataset" (e.g., 2010–2020 for Reddit and 2020–2021 for Twitter).

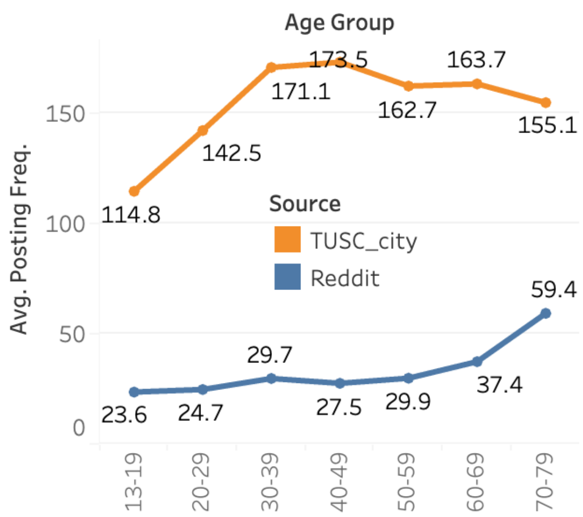

Figure 1: Posting frequency on Reddit and Twitter.

4.1. How does Posting Frequency Change with Age?

We ask how posting habits, such as the frequency of posting, varies across age groups. This is an important question as it is informative for research on well-being (e.g., loneliness) which uses the amount of online activity as a signal. It has been shown that individuals who experience depression, may withdraw from social media use (De Choudhury et al., 2021), and there were significant differences in Facebook posting quantities for individuals diagnosed with depression (Smith et al., 2017).

Experiment Setup: We find the number of posts for each author and then take the average across authors in that respective age group.

Results: We show the results in Figure 1. On Twitter, there is an increase in posting frequency until the 40's, with a decrease afterwards. On Reddit, we see an increasing trend in the frequency of posting across age groups up until the 40's, with a decrease afterwards until the 60's to 70's.

Discussion: These findings contrast those by Chang et al. (2015) who found that age was negatively associated with self-posting on Facebook, and found there is less activity in general for older adults on Facebook (McAndrew and Jeong, 2012).

[4]We do not consider more recent Reddit data as it contains numerous LLM- and bot-generated posts.

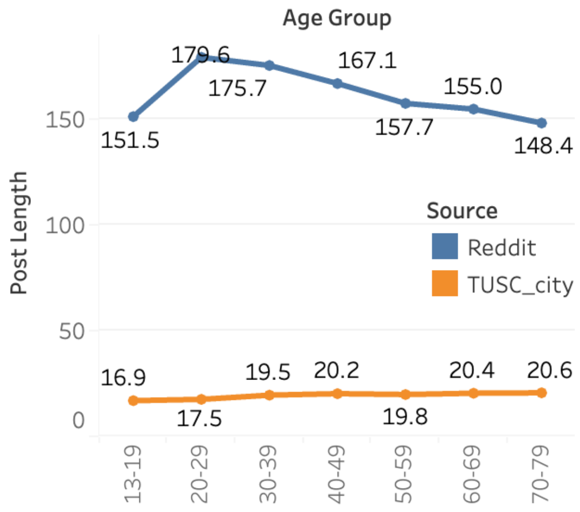

Figure 2: Post length (tokens) by age group.

4.2. How do Post Lengths Change with Age on Reddit and Twitter?

Is it that as we age, we tend to share longer messages online, or is it the opposite? Or, is there no clear trend? These findings provide insights into the writing style of each age group.

Experiment Setup: The experiment is simple – we compute the average length of each post in terms of number of tokens, average the length of posts per author, and compute the average post length per age group. This allows us to see how post length varies across groups and if there are consistent trends across the two platforms.

Results: The results are shown in Figure 2. In the Reddit data we see a marked difference between the teens and 20's, compared to the Twitter data. Generally, there is an increasing trend in post length on TUSC, whereas the Reddit data shows the opposite trend with decreasing post length from the 20's to 70's. Overall, we see more variation in post length on Reddit than Twitter, which could be due to the character limit on Twitter (280 characters).

Discussion: Previous literature found that post length in blogs was longer for older adults (Rosenthal and McKeown, 2011), post length increased on a health forum with age (Shrestha et al., 2016), and the total number of content words in Facebook messages had a positive relationship with age (Schwartz et al., 2013). These findings align with the trends found on Twitter. Whereas other literature found no relationship between social network post length and age (Peersman et al., 2011). The opposing trends between Twitter and Reddit could be explained by how these platforms are used for different purposes.

Age Group	Most popular Subreddits
13–19	teenagers, quzangle, randonaut_reports, AskReddit, buildapc
20–29	AskReddit, NoStupidQuestions, personalfinance, buildapc, NoFap
30–39	funwithwords, Vocabulario, EsperantoFeed, KazakhFeed, CatalanFeed
40–49	stopdrinking, AskReddit, NoStupidQuestions, personalfinance, keto
50–59	SandersAlerts, nba, stopdrinking, MillennialBets, personalfinance
60–69	gw2devtrackalt, Wallstreetsilver, personalfinance, Jokes, teenagers
70-79	WorldNewsinPictures, CMKMArchive, u_WorldNewsinPictures, teenagers, pathofexile

Table 2: The top 5 most popular subreddits per age group in descending order.

4.3. How does Most Popular Subreddits Change with Age?

We are interested in the top subreddits each age group posts in and what changes, if any, we see across age groups. This informs us of the interests and communities that each group engages with.

Experiment Setup: We find the subreddit counts of posts for each age group, and consider the top 5 most popular. We also performed these experiments with up until the top 20 subreddits for each age group, and summarize the results below. We show the top 20 most popular subreddits in Appendix D. This research question is only applicable to the Reddit subset of *AgeCorpus*.

Results: The results are shown in Table 2. We see that in the teens and 20's there is interest in r/teenagers, asking questions online, and hobbies such as building personal computers and an interest in finance. In the 30's people engage in interests with languages, vocabulary and words (e.g., r/funwithwords, r/Vocabulario, r/Word_of_The_Hour). In the 40's and 50's health-conscious interests appear such as reducing drinking and keto diets, and home improvement. In the 60's, 70's other interests such as finances, games, and news appear which could be as people retire they are passing time looking at entertainment and news on their computers. Overall, we see little overlap between groups except when looking at the top 20 subreddits we see sustained interest in r/personalfinance for 20's to 60's, as well asking questions online (e.g., r/AskReddit and r/NoStupidQuestions across all age groups). Looking at the top 20 subreddits we see overall that the teens to 20's is a different time than 30's, as people are in transitory periods of their life seeking advice from r/relationship_advice and r/offmychest, and have interests in gaming (r/gaming, r/leagueoflegends). The teens and 20's is also a time when mental health becomes a prevalent topic (r/depression, r/Anxiety), and it reappears in the 40's (e.g., r/depression, r/ADHD). We also see topics such as r/Menopause appear in the 50's.

Discussion: There is no vast literature exploring which subreddits are popular among each age group, but Chew et al. (2021) found that teenagers commonly post in r/teenagers (age 13–20), whereas older adults (age 21-54) posted more frequently in r/news, which aligns with our findings. Monti et al. (2023) studied topics discussed in r/news and found that generally younger people (teenagers) tended to discuss more about arts, technology, and entertainment, while older adults (30's and 40's) discussed business, politics, and green/environmental topics.

4.4. How does Pronoun Usage Change with Age?

We explore how the pronoun usage in posts on social media varies with age. This informs us of people's focus e.g., themselves, those in direct contact with them, or others. We analyze subject and object pronouns listed in Appendix B.

Experiment Setup: We use a binary flag to indicate whether a post contains a given pronoun or not. Then, we calculate the percentage of posts that have each pronoun across age groups.

Results: In Figure 3, we show the percentage of pronoun usage for subject and object pronouns across the platforms. Across groups, the most common pronoun is *I*, followed by *you*, *me*, *they*, *we* and *them*. Overall, first-person pronoun usage (e.g., *I*, *me*) increases from teens to 20's, and decreases from 30's to 70's. Whereas second-person (e.g., *you*) tends to have an inverse pattern, increasing until 40's and then plateauing. For third-person pronouns (e.g., *we*, *they*) we see the overall trend increasing with age. It appears that for *them*, proportions are relatively constant with age, and just slightly increase for TUSC-City.

Discussion: There is little work comparing how each age group uses first- vs. second- vs. third-person pronouns, however Hendriks et al. (2008) found that elderly adults produced pronouns significantly more often than young adults when describing picture stories and referring to past topics, and more potentially ambiguous pronouns (Hendriks et al., 2014). Recently, the percentage of tweets containing various pronouns has been explored, finding that about 60% of tweets had "I" and 30% had "you" ((Mohammad, 2026), Figure 3). More recent work in the area of pronoun usage has focused on contrasting self-disclosed use of *he*, *she*, and *them* in profile descriptions on social media (Jiang et al., 2023; Lauscher et al., 2022), attitudes towards pronouns (Sendén et al., 2021), pronoun resolution (Andy et al., 2020; Seminck and Amsili, 2017; Yu et al., 2019; Chada, 2019), or errors in translating pronouns (Lauscher et al., 2023; Luong and Popescu-Belis, 2016), rather than looking at overall trends in pronoun usage across age groups.

5. How do Emotions Expressed in Social Media Vary with Age?

We explore how does emotion expression on social media vary across age groups. These findings are important for diverse research questions such as how do characteristics of language use change with age, and in interdisciplinary work such as examining the relationship between emotion in text and mental health. For example, more negative emotion expressed in text has been found to be associated with mental health diagnoses such as depression, ADHD, etc. (Seabrook et al., 2018; De Choudhury et al., 2013, 2021). Further, the way in which emotion expression in text changes over time (e.g., emotional variance) has been linked to mental health diagnoses (Teodorescu et al., 2023a).

We explore two questions about how emotions vary across age groups in the two subsections below. We answer these questions by looking at the three primary dimensions of affect (valence, arousal and dominance), and four important categorical emotions (anger, fear, joy and sadness). We apply common preprocessing to the text by first tokenizing, lowercasing and obtaining the emotion scores. We use the NRC Emotion Lexicon (Mohammad and Turney, 2012) for anger, fear, joy, and sadness, where scores range from 0 (lowest amount of the emotion) – 1 (highest amount of the emotion). We use the NRC Valence Arousal Dominance Lexicon (Mohammad, 2018) for valence, arousal and dominance, where scores range from -1 (lowest V/A/D) to 1 (highest V/A/D).[5] Past work has shown that lexical approaches produce emotion arcs that are highly correlated with the true (gold) arcs (correlations > 0.9) (Teodorescu and Mohammad, 2023).

5.1. How does Average Emotion Change with Age?

Experiment Setup: We calculate the emotion score per post by taking the average of the emotion

[5] http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm https://saifmohammad.com/WebPages/nrc-vad.html

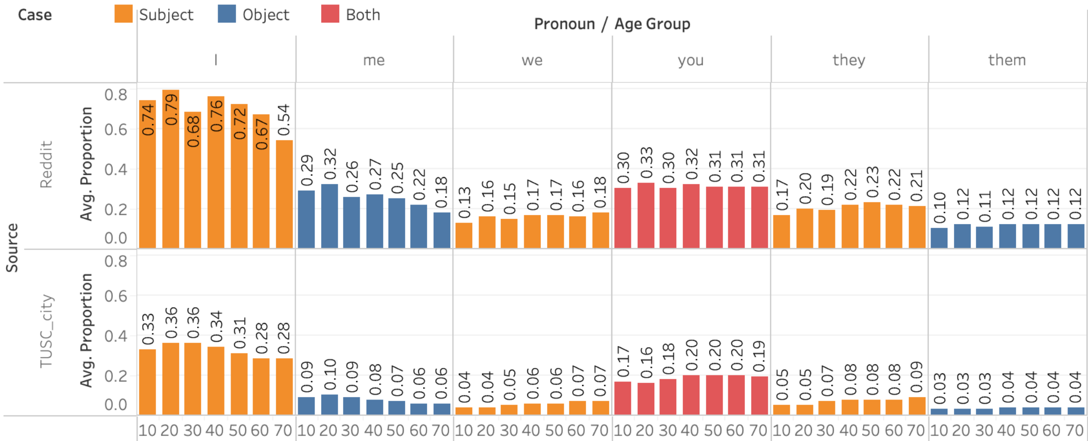

Figure 3: Pronoun usage by age group.

scores for each word in the post. Then, we compute the average emotion score per user by taking the average across their posts. To compare how emotional states vary across age groups, we average the emotion scores across all users per group to analyze the results at the aggregate level.

Results: We plot the results for average valence, arousal, dominance in Figure 4 for TUSC and Reddit, and the average anger, fear, joy, and sadness in Figure 5 for TUSC and Reddit.

Valence, Arousal, Dominance: In both TUSC and Reddit, we see an increase in average valence from teens up until the 30's, and a steady decrease from the 30's till the 70's. We see the reverse pattern for arousal with decreasing scores on Reddit from teen's up until the 70's. For TUSC, we see an increase from teens to 20's, and then arousal intensities plateau until the 70's. This means that younger and older groups are expressing similar levels of arousal in TUSC, whereas there is a consistent decrease in arousal with age on Reddit. For dominance, there are increasing scores with age from teen's to the 70's in TUSC, meaning that language which is more associated with control is used with age. However, there is an inverted U-shape on Reddit, where dominance reaches a peak in the 20's and then decreases with age. Generally, it appears that valence follows the inverted U-shape, arousal decreases (Reddit)/plateaus (TUSC) with age, and dominance follows the inverted U-shape (Reddit)/increases (TUSC) with age. Overall, we see consistent trends between platforms for valence but differing patterns for arousal and dominance.

Discussion: Work in psychology has found various trends in terms of happiness across ages. The majority of the literature points to the U-shape of happiness, and some to the inverted-U or bimodal trends. Our findings support the inverted U-shape, meaning that younger and older adults experience less valence/happiness compared to those in their middle ages. Past work has found that when children wrote poems, there was decreasing valence and increasing arousal and dominance across grades 1–12 (Teodorescu et al., 2023b). No work examines arousal or dominance expression in text across adulthood, so we establish baseline findings for these dimensions.

Anger, Fear, Joy, Sadness: We see inverted U-shape patterns appearing for anger, fear, joy and sadness on Reddit, whereas in TUSC we found increasing intensities for all categorical emotions with age. However, we note that the inverted U-shape on Reddit is not symmetrical e.g., anger and joy is highest for teens than for those in their 70's. In terms of categorical emotions, we find that the negative emotions (i.e., anger, fear, sadness) show similar trends of positive emotions such as valence. Valence was highest for those in midlife, while anger, fear, joy and sadness were highest on Reddit. We expected that valence and joy would have similar trends as both are positive emotions/dimensions and overall they are both have an inverted U-shape (with joy increasing in TUSC).

Discussion: Past work has found more anger and sadness for older adults, however it was noted that sadness only increased for those with decreased perceived control (Blanchard-Fields and Coats, 2008; Wrosch et al., 2018). Haase et al. (2012) point to the adaptive function of anger and sadness at certain stages of life, as higher levels of self-reported sadness in response to a neutral film were associated with higher well-being for older adults. Higher self-reported anger was associated with higher well-being for middle-aged adults (rather than for younger and older adults). Anger is said to be more adaptive to younger adults when they have many resources and are trying to determine their niche in society, and reach goals, whereas sadness is adaptive to older adults when dealing with unachievable goals, diminishing resources, loss, and to elicit support (Haase et al., 2012; Kunzmann and Wrosch, 2024). Other work has found that the intensity of anger, fear, joy and sadness expressed in

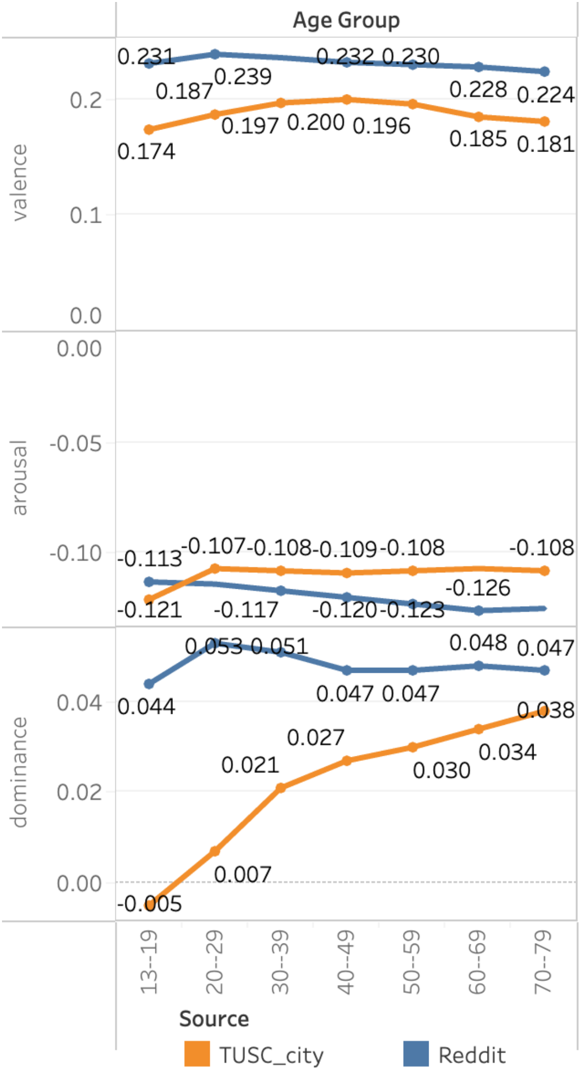

Figure 4: **Average valence, arousal, dominance** across age groups for TUSC-City and Reddit. Note that the lexicon scores for valence, arousal, dominance range from -1 to 1.

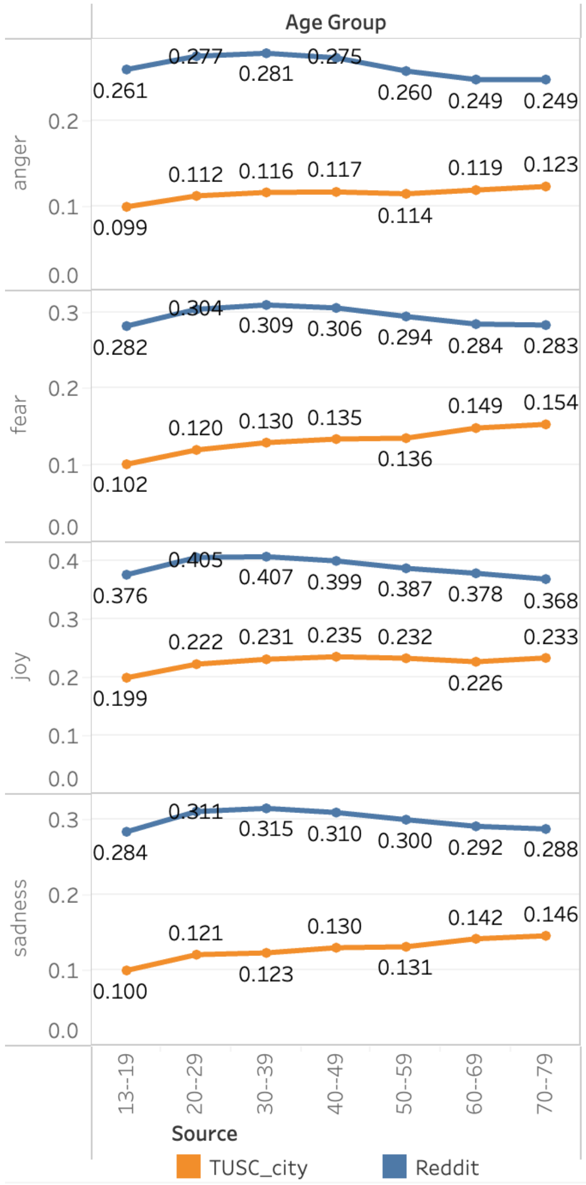

Figure 5: **Average anger, fear, joy, sadness** across age groups for TUSC-City and Reddit.

poems written by children increases across grades 1–12 (Teodorescu et al., 2023b).

5.2. How does Emotional Variance Change with Age?

Experiment Setup: Using the emotion score per post, we calculate the standard deviation across the posts per user. Then we average the standard deviations per age group to compare how emotional variance varies with age.
Results: We show the results in Figure 6 for valence, arousal, and dominance variation on TUSC and Reddit.
Valence, Arousal, Dominance: In Twitter, we see an overall U-shape for valence, arousal and dominance variation, and then a slight drop in the 60's and 70's. On Reddit, we also see a U-shape for valence, arousal and dominance variation across age groups, although teens have markedly higher variation.
Discussion: This is the first work that studies how the variation of emotions changes across age groups, contributing novel findings to affective developmental psychology. Work in psychology has pointed to the relationship between emotional variance and mental health diagnoses (e.g., ADHD, anxiety, depression (Houben et al., 2015; Heller et al., 2018; Seabrook et al., 2018; Teodorescu et al., 2023a)), however little is known about the pattern of emotion variation with age. While we see similar U-shape trends between Reddit and Twitter, teens had high variation on Reddit. This could be due to the nature of the platforms. Reddit allows the space for longer posts, such as stories, where we may expect younger crowds to have more vary-

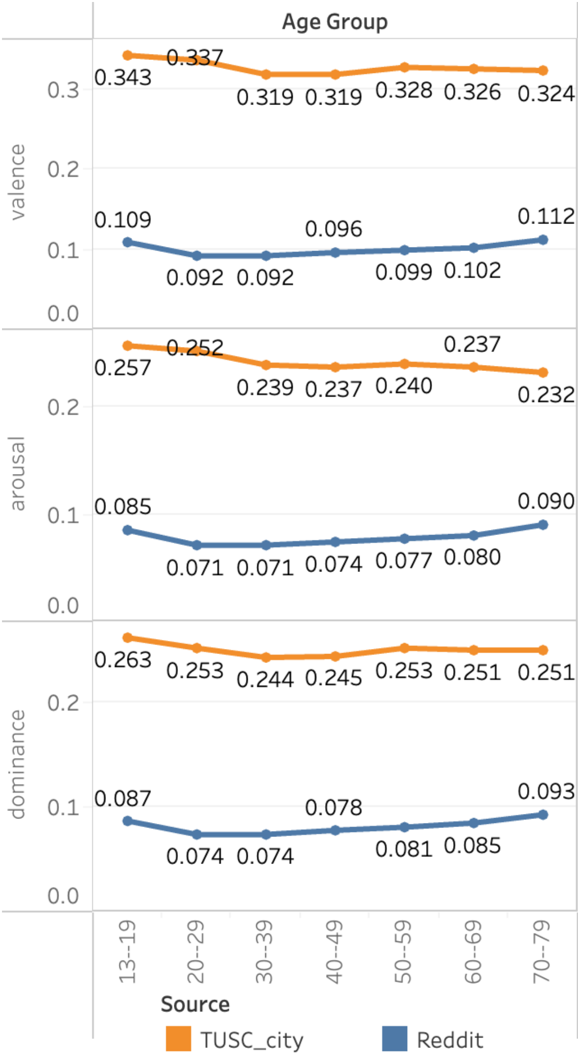

Figure 6: **Valence, arousal, and dominance variation** across age groups for TUSC-City and Reddit. Note that the lexicon scores for valence, arousal, dominance range from -1 to 1.

ing emotional experiences during transitory periods of life. Whereas Twitter is more suitable for shorter updates/posts and generally has an older audience. Future work should explore emotional variation on other platforms and through surveys across age groups to further findings across disciplines.

6. Conclusion

We created a dataset of social media utterances annotated with the author's age at the time of posting. This unique dataset has social media posts from Reddit during 2010–2022 and Twitter from 2020–2021. Such a dataset enables diverse research in the social sciences, psychology, and NLP. With the dataset we explored how emotion expression in text varied across age groups. We found valence followed the inverted-U shape, arousal decreased with age (Reddit)/plateaus (TUSC), and dominance followed the inverted-U shape (Reddit)/increased with age (TUSC). For the categorical emotions, we found all emotions followed an inverted U-shape on Reddit and increased with age in TUSC. We found generally similar trends of U-shaped emotional variance on Reddit and TUSC. Such insight into how online behavior changes with age is informative for research in mental health and well-being. Further, this dataset can be used to explore the relationship between age and many other aspects of human behavior and language, such as emotion regulation, interoception, anxiety, etc. For example, body part mentions in tweets has been linked with poorer health outcomes (Wu et al., 2025) and this relationship could be investigated across ages.

Limitations

This dataset relies on self-disclosure of age on social media. As with all scenarios involving self-disclosure, individuals may falsely report their age due to social pressures, or in order to relate to and engage with certain communities. Additionally, our templates may not fully capture all possible ways to express age, so there may be age declarations not included in our dataset.

Ethical Considerations

There are ethical considerations around creating such an age corpus. The Reddit User Agreement mentions that users agree not to disclose sensitive information of other users and they consent that their posts are publicly available (e.g., through API to other services).

Our research interest is to study emotional development through adulthood at the aggregate/-group level. This has applications in emotional development psychology and in public health (e.g., overall well-being and mental health). However, emotions are complex, private, and central to an individual's experience. Additionally, each individual expresses emotion differently through language, which results in large amounts of variation. Therefore, several ethical considerations should be accounted for when performing any textual analysis of emotions (Mohammad, 2022, 2023). The ones we would particularly like to highlight are listed below:

- Our work on studying emotion word usage should not be construed as detecting how people feel; rather, we draw inferences on the emotions that are conveyed by users via the language that they use.
- The language used in an utterance may convey information about the emotional state (or perceived emotional state) of the speaker, listener, or someone mentioned in the utterance.

However, it is not sufficient for accurately determining any of their momentary emotional states. Deciphering the true momentary emotional state of an individual requires extralinguistic context and world knowledge. Even then, one can be easily mistaken.

- The inferences we draw in this paper are based on aggregate trends across large populations. We do not draw conclusions about specific individuals or momentary emotional states.

Bibliographical References

Anietie Andy, Chris Callison-Burch, and Derry Tanti Wijaya. 2020. Resolving pronouns in Twitter streams: Context can help! In *Proceedings of the Third Workshop on Computational Models of Reference, Anaphora and Coreference*, pages 133–138, Barcelona, Spain (online). Association for Computational Linguistics.

Magda B Arnold. 1960. Emotion and personality. vol. i. psychological aspects.

Fredda Blanchard-Fields and Abby Heckman Coats. 2008. The experience of anger and sadness in everyday problems impacts age differences in emotion regulation. *Developmental psychology*, 44(6):1547—1556.

David G. Blanchflower and Andrew J. Oswald. 2008. Is well-being u-shaped over the life cycle? *Social Science & Medicine*, 66(8):1733–1749.

Rakesh Chada. 2019. Gendered pronoun resolution using BERT and an extractive question answering formulation. In *Proceedings of the First Workshop on Gender Bias in Natural Language Processing*, pages 126–133, Florence, Italy. Association for Computational Linguistics.

Pamara F. Chang, Yoon Hyung Choi, Natalya N. Bazarova, and Corinna E. Löckenhoff. 2015. Age differences in online social networking: Extending socioemotional selectivity theory to social network sites. *Journal of Broadcasting & Electronic Media*, 59(2):221–239. PMID: 31223198.

Robert Chew, Caroline Kery, Laura Baum, Thomas Bukowski, Annice Kim, and Mario Navarro. 2021. Predicting age groups of reddit users based on posting behavior and metadata: Classification model development and validation. *JMIR Public Health Surveill*, 7(3):e25807.

Andrew E. Clark and Andrew J. Oswald. 1994. Unhappiness and unemployment. *The Economic Journal*, 104(424):648–659.

Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013. Social media as a measurement tool of depression in populations. In *Proceedings of the 5th Annual ACM Web Science Conference*, WebSci '13, page 47–56, New York, NY, USA. Association for Computing Machinery.

Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2021. Predicting depression via social media. *Proceedings of the International AAAI Conference on Web and Social Media*, 7(1):128–137.

Angus Deaton. 2008. Income, health, and well-being around the world: Evidence from the gallup world poll. *Journal of Economic perspectives*, 22(2):53–72.

Richard A. Easterlin. 2006. Life cycle happiness and its sources: Intersections of psychology, economics, and demography. *Journal of Economic Psychology*, 27(4):463–482.

To'Meisha Edwards and Nicholas S. Holtzman. 2017. A meta-analysis of correlations between depression and first person singular pronoun use. *Journal of Research in Personality*, 68:63–68.

Paul Ekman. 1992. An argument for basic emotions. *Cognition and Emotion*, 6(3-4):169–200.

Matej Gjurković, Vanja Mladen Karan, Iva Vukojević, Mihaela Bošnjak, and Jan Snajder. 2021. PANDORA talks: Personality and demographics on Reddit. In *Proceedings of the Ninth International Workshop on Natural Language Processing for Social Media*, pages 138–152, Online. Association for Computational Linguistics.

JJ Gross, LL Carstensen, M Pasupathi, J Tsai, CG Skorpen, and AY Hsu. 1997. Emotion and aging: experience, expression, and control. *Psychology and aging*, 12(4):590—599.

Claudia M Haase, Benjamin H Seider, Michelle N Shiota, and Robert W Levenson. 2012. Anger and sadness in response to an emotionally neutral film: evidence for age-specific associations with well-being. *Psychology and aging*, 27(2):305—317.

Aaron S Heller, Andrew S Fox, and Richard J Davidson. 2018. Parsing affective dynamics to identify risk for mood and anxiety disorders. *Emotion*, 19(2):283–291.

Petra Hendriks, Christina Englert, Ellis Wubs, and John Hoeks. 2008. Age differences in adults' use of referring expressions. *Journal of Logic, Language and Information*, 17(4):443–466.

Petra Hendriks, Charlotte Koster, and John C.J. Hoeks. 2014. Referential choice across the lifespan: why children and elderly adults produce ambiguous pronouns. *Language, Cognition and Neuroscience*, 29(4):391–407. PMID: 24771955.

Will Hipson and Saif M. Mohammad. 2020. PoKi: A large dataset of poems by children. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 1578–1589, Marseille, France. European Language Resources Association.

Marlies Houben, Wim Van Den Noortgate, and Peter Kuppens. 2015. The relation between short-term emotion dynamics and psychological well-being: A meta-analysis. *Psychological Bulletin*, 141(4):901–930.

Julie Jiang, Emily Chen, Luca Luceri, Goran Muric, Francesco Pierri, Ho-Chun Herbert Chang, and Emilio Ferrara. 2023. What are your pronouns? examining gender pronoun usage on twitter.

Joo Hyun Kim, Eunsoo Choi, Namhee Kim, and Incheol Choi. 2023. Older people are not always happier than younger people: the moderating role of personality. *Applied Psychology: Health and Well-Being*, 15(1):275–292.

Ute Kunzmann and Carsten Wrosch. 2024. Not all negative emotions are equal - sadness and anger develop differently and their adaptivity is age-graded. *Current Opinion in Psychology*, 55:101766.

Anne Lauscher, Archie Crowley, and Dirk Hovy. 2022. Welcome to the modern world of pronouns: Identity-inclusive natural language processing beyond gender. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 1221–1232, Gyeongju, Republic of Korea. International Committee on Computational Linguistics.

Anne Lauscher, Debora Nozza, Ehm Miltersen, Archie Crowley, and Dirk Hovy. 2023. What about "em"? how commercial machine translation fails to handle (neo-)pronouns. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 377–392, Toronto, Canada. Association for Computational Linguistics.

Jialei Li. 2023. Emotion and ageing in discourse: Do older people express more positive emotions? *Corpus-based Studies across Humanities*, 1(1):23–50.

Ngoc Quang Luong and Andrei Popescu-Belis. 2016. Pronoun language model and grammatical heuristics for aiding pronoun prediction. In *Proceedings of the First Conference on Machine Translation: Volume 2, Shared Task Papers*, pages 589–595, Berlin, Germany. Association for Computational Linguistics.

Francis T. McAndrew and Hye Sun Jeong. 2012. Who does what on facebook? age, sex, and relationship status as predictors of facebook use. *Computers in Human Behavior*, 28(6):2359–2365.

Saif Mohammad. 2018. Obtaining reliable human ratings of valence, arousal, and dominance for 20,000 English words. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 174–184, Melbourne, Australia. Association for Computational Linguistics.

Saif Mohammad. 2023. Best practices in the creation and use of emotion lexicons. In *Findings of the Association for Computational Linguistics: EACL 2023*, pages 1825–1836, Dubrovnik, Croatia. Association for Computational Linguistics.

Saif M. Mohammad. 2022. Ethics sheet for automatic emotion recognition and sentiment analysis. *Computational Linguistics*, 48(2):239–278.

Saif M. Mohammad. 2026. When are we worried? temporal trends of anxiety and what they reveal about us.

Saif M. Mohammad and Peter D. Turney. 2012. Crowdsourcing a word–emotion association lexicon. *Computational Intelligence*, 29(3):436–465.

Corrado Monti, Jacopo D'Ignazi, Michele Starnini, and Gianmarco De Francisci Morales. 2023. Evidence of demographic rather than ideological segregation in news discussion on reddit. In *Proceedings of the ACM Web Conference 2023*, WWW '23, page 2777–2786, New York, NY, USA. Association for Computing Machinery.

Daniel K. Mroczek and Avron Spiro. 2005. Change in life satisfaction during adulthood: Findings from the veterans affairs normative aging study. *Journal of personality and social psychology*, 88(1):189–202.

David G. Myers and Ed Diener. 1995. Who is happy? *Psychological Science*, 6(1):10–19.

Charles E Osgood, George J Suci, and Percy Tannenbaum. 1957. *The measurement of meaning*. University of Illinois Press.

Claudia Peersman, Walter Daelemans, and Leona Van Vaerenbergh. 2011. Predicting age and gender in online social networks. In *Proceedings of the 3rd International Workshop on Search* *and Mining User-Generated Contents*, SMUC '11, page 37–44, New York, NY, USA. Association for Computing Machinery.

Robert Plutchik. 1980. Chapter 1 - a general psychoevolutionary theory of emotion. In Robert Plutchik and Henry Kellerman, editors, *Theories of Emotion*, pages 3–33. Academic Press.

Sara Rosenthal and Kathleen McKeown. 2011. Age prediction in blogs: A study of style, content, and online behavior in pre- and post-social media generations. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 763–772, Portland, Oregon, USA. Association for Computational Linguistics.

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

James A Russell and Albert Mehrabian. 1977. Evidence for a three-factor theory of emotions. *Journal of Research in Personality*, 11(3):273–294.

H. Andrew Schwartz, Johannes C. Eichstaedt, Margaret L. Kern, Lukasz Dziurzynski, Stephanie M. Ramones, Megha Agrawal, Achal Shah, Michal Kosinski, David Stillwell, Martin E. P. Seligman, and Lyle H. Ungar. 2013. Personality, gender, and age in the language of social media: The open-vocabulary approach. *PLOS ONE*, 8(9):1–16.

Elizabeth M Seabrook, Margaret L Kern, Ben D Fulcher, and Nikki S Rickard. 2018. Predicting depression from language-based emotion dynamics: Longitudinal analysis of facebook and twitter status updates. *J Med Internet Res*, 20(5):e168.

Olga Seminck and Pascal Amsili. 2017. A computational model of human preferences for pronoun resolution. In *Proceedings of the Student Research Workshop at the 15th Conference of the European Chapter of the Association for Computational Linguistics*, pages 53–63, Valencia, Spain. Association for Computational Linguistics.

Marie Gustafsson Sendén, Emma Renström, and Anna Lindqvist. 2021. Pronouns beyond the binary: The change of attitudes and use over time. *Gender & Society*, 35(4):588–615.

Prasha Shrestha, Nicolas Rey-Villamizar, Farig Sadeque, Ted Pedersen, Steven Bethard, and Thamar Solorio. 2016. Age and gender prediction on health forum data. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 3394–3401, Portorož, Slovenia. European Language Resources Association (ELRA).

Robert J Smith, Patrick Crutchley, H Andrew Schwartz, Lyle Ungar, Frances Shofer, Kevin A Padrez, and Raina M Merchant. 2017. Variations in facebook posting patterns across validated patient health conditions: A prospective cohort study. *J Med Internet Res*, 19(1):e7.

Daniela Teodorescu, Tiffany Cheng, Alona Fyshe, and Saif Mohammad. 2023a. Language and mental health: Measures of emotion dynamics from text as linguistic biosocial markers. In *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing*, pages 3117–3133, Singapore. Association for Computational Linguistics.

Daniela Teodorescu, Alona Fyshe, and Saif Mohammad. 2023b. Utterance emotion dynamics in children's poems: Emotional changes across age. In *Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, pages 401–415, Toronto, Canada. Association for Computational Linguistics.

Daniela Teodorescu and Saif Mohammad. 2023. Evaluating emotion arcs across languages: Bridging the global divide in sentiment analysis. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pages 4124–4137, Singapore. Association for Computational Linguistics.

Anna Tigunova, Paramita Mirza, Andrew Yates, and Gerhard Weikum. 2020. RedDust: a large reusable dataset of Reddit user traits. In *Proceedings of the Twelfth Language Resources and Evaluation Conference*, pages 6118–6126, Marseille, France. European Language Resources Association.

Heather L. Urry and James J. Gross. 2010. Emotion regulation in older age. *Current Directions in Psychological Science*, 19(6):352–357.

Krishnapriya Vishnubhotla and Saif M. Mohammad. 2022. Tweet Emotion Dynamics: Emotion word usage in tweets from US and Canada. In *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, pages 4162–4176, Marseille, France. European Language Resources Association.

Jan Philip Wahle, Gipp Bela Vishnubhotla, Krishnapriya, and Saif M. Mohammad. 2026. Affect, body, cognition, demographics, and emotion: The abcde of text features for computational affective science. In *Proceedings of the 1st Workshop on Computational Affective Science (CAS 2026)*, Palma de Mallorca, Spain. European Language Resources Association (ELRA).

Charles Welch, Jonathan K. Kummerfeld, Verónica Pérez-Rosas, and Rada Mihalcea. 2020. Compositional demographic word embeddings. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 4076–4089, Online. Association for Computational Linguistics.

Carsten Wrosch, Meaghan A Barlow, and Ute Kunzmann. 2018. Age-related changes in older adults' anger and sadness: The role of perceived control. *Psychology and aging*, 33(2):350—360.

Sophie Wu, Jan Philip Wahle, and Saif M. Mohammad. 2025. The language of interoception: Examining embodiment and emotion through a corpus of body part mentions. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, pages 23375–23399, Suzhou, China. Association for Computational Linguistics.

Xintong Yu, Hongming Zhang, Yangqiu Song, Yan Song, and Changshui Zhang. 2019. What you see is what you get: Visual pronoun coreference resolution in dialogues. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 5123–5132, Hong Kong, China. Association for Computational Linguistics.

A. Patterns to Match Age Declarations

In Table 3, we show the patterns used to find age declarations in the seed posts.

B. Subject and Object Pronouns

We study the occurrences of the following pronouns shown in Table 4 to determine if there are patterns across age categories.

C. Number of Users per Age Group

In Table 5, we show the number of users in each age group in the *AgeCorpus*. This data was used to perform the experiments.

D. Top 20 Most Popular Subreddits per Age Group

In Table 6, we show the 20 most popular subreddits by age group. We describe the findings in Section 4.3.

<table>
<tr><th>Regex</th><th>Example</th></tr>
<tr><td><code>\bI(?:\s+am|’m)\s+(\d{1,2})\s+years?\s+old\b</code></td><td>I am 25 years old
I’m 30 year old</td></tr>
<tr><td><code>\bI(?:\s+am|’m)\s+(\d{1,2})</code>
<code>(?=\s*(?:$|[,.!?;:\-]|(?:and|but|so|yet)\s))</code></td><td>I am 25.
I’m 30, and ...</td></tr>
<tr><td><code>\bI(?:\s+was|\s+am|’m)\s+born\s+in\s+</code>
<code>(19\d{2}|20(?:0\d|1\d|2[0-4]))\b</code></td><td>I was born in 1998
I am born in 2005</td></tr>
<tr><td><code>\bI(?:\s+was|\s+am|’m)\s+born\s+in\s+’(\d{2})\b</code></td><td>I was born in ’98
I’m born in ’05</td></tr>
<tr><td><code>\bI\s+was\s+born\s+on\s+</code>
<code>(?:\d{1,2}(?:st|nd|rd|th)?\s+)?</code>
<code>(?:January|February|March|April|May|June|</code>
<code>July|August|September|October|November|December|</code>
<code>Jan|Feb|Mar|Apr|May|Jun|Jul|Aug|Sep|Sept|Oct|Nov|Dec)</code>
<code>\s+ (?:\d{1,2}(?:st|nd|rd|th)?,?\s+)?</code>
<code>(19\d{2}|20(?:0\d|1\d|2[0-4]))\b</code></td><td>I was born on 15 March 1998
I was born on March 15th, 1998</td></tr>
<tr><td><code>\bI\s+was\s+born\s+on\s+\d{1,2}</code>
<code>[/\-]\d{1,2}[/\-](19\d{2}|20(?:0\d|1\d|2[0-4]))\b</code></td><td>I was born on 03/15/1998
I was born on 15-03-1998</td></tr>
</table>

Table 3: Regexes used to identify users from both the Reddit and Twitter datasets.

Subject	Object
I	me
you	you
he	him
she	her
it	it
we	us
they	them

Table 4: Subject and object pronouns we studied occurrences of in the dataset.

Age Group	#Posts	
	TUSC-City	Reddit
13–19	94,857	9,281,055
20–29	406,027	15,455,426
30–39	462,238	6,151,757
40–49	360,946	1,220,498
50–59	278,809	449,230
60–69	218,008	246,398
70–79	97,432	160,465

Table 5: The number of posts across the age groups in each subset of the *AgeCorpus*. “City” refers to TUSC-City subset of the dataset. We use this data in our experiments.

Age Group	Most popular Subreddits
13–19	teenagers, quzangle, randonaut_reports, AskReddit, buildapc, NoFap, NoStupidQuestions, Advice, GlobalOffensiveTrade, leagueoflegends, pcmasterrace, teenagersnew, depression, techsupport, trees, alt_source_bot_log, tipofmytongue, offmychest, relationship_advice, gaming
20–29	AskReddit, NoStupidQuestions, personalfinance, buildapc, NoFap, depression, Advice, techsupport, leagueoflegends, offmychest, relationship_advice, tipofmytongue, raisedbynarcissists, AskDocs, legaladvice, trees, Fireteams, r4r, gaming, Anxiety
30–39	funwithwords, Vocabulario, EsperantoFeed, KazakhFeed, CatalanFeed, personalfinance, GalicianFeed, Lessico, SwedishFeed, Traductions, Word_of_The_Hour, HindiFeed, AskReddit, PortugueseFeed, Sprache, NoStupidQuestions, BCSFeed, WelshFeed, DestinyTheGame, BengaliFeed
40–49	stopdrinking, AskReddit, NoStupidQuestions, personalfinance, keto, HomeImprovement, raisedbynarcissists, Jokes, legaladvice, techsupport, Isletforum, DestinyTheGame, buildapc, NoFap, depression, getdisciplined, tipofmytongue, AskDocs, ADHD, sysadmin
50–59	SandersAlerts, nba, stopdrinking, MillennialBets, personalfinance, exmormon, Jokes, AskReddit, NoStupidQuestions, keto, volunteer, legaladvice, open_bots_test, techsupport, Menopause, leagueoflegends, The_Donald, buildapc, raisedbynarcissists, HomeImprovement
60–69	gw2devtrackalt, Wallstreetsilver, personalfinance, Jokes, teenagers, TargetedEnergyWeapons, anychat, stopdrinking, NoStupidQuestions, zerohedge, ketoscience, amcstock, AskReddit, wow, exmormon, techsupport, LibraryofBabel, playcodformoney, leagueoflegends, buildapc
70-79	WorldNewsinPictures, CMKMArchive, u_WorldNewsinPictures, teenagers, pathofexile, makingascenemag, bestchange, CattleExchange, nba, PollsAndSurveys, BollyBlindsNGossip, Jokes, u_classeAS2002, u_From_Him_For_Him, NoStupidQuestions, guns, AskReddit, Gunsforsale, exjw, FashionWinter

Table 6: The top 20 most popular subreddits per age group in descending order.