Parenting During COVID-19: A Sentiment Analysis of Reddit Data

Kaitlin P. Ward

University of Michigan

SurvMeth 727

Github repository: https://github.com/kaitlinpward/ServMeth727_Reddit

Introduction

The coronavirus disease 2019 (COVID-19) pandemic has caused significant disruption and disturbance in the lives of individuals across the globe. Specifically, COVID-19 has affected over 213 countries and territories (World Health Organization, 2020) and has caused over 290,000 deaths in the U.S. alone (Centers for Disease Control and Prevention, 2020). Since mid-March of 2020, Americans have engaged in social distancing, social isolation, and lockdown, which has led to drastic decreases in in-person social interactions (Douglas, 2020). These social distancing and lockdown requirements had a significant impact on the economy, leading to 15% of U.S. adults reporting job losses due to COVID-19 as of September 2020 (Pew Research Center, 2020). Further, children and young adults were required to switch to remote learning, which has disproportionally affected the learning, nutrition, and drop-out rates among students from vulnerable and oppressed communities (UNESCO, 2020). Clearly, the COVID-19 pandemic is having a detrimental effect on the physical and mental health of many.

Due to the added responsibilities of caring for children and maintaining a household, parents may be a particularly vulnerable population that exists during COVID-19. The economic strain and unemployment associated with COVID-19 may make it more difficult for parents to meet their children's physical and emotional needs (Conger & Conger, 2002; Neppl et al., 2016). Indeed, recent survey research suggests that, after the onset of COVID-19, parents reported high rates of anxiety, financial worries, and parenting stress (Lee & Ward, 2020). Further, with at-home learning beginning in March 2020 and now extending into 2021, parents are reporting difficulties in providing at-home education, particularly those parents who meet criteria for anxiety and depression (Lee et al., 2020). More research is needed in order to determine the short- and long-term impacts that COVID-19 will have on parents.

Thus far, the majority of research that examines parenting during COVID-19 has been conducted with survey data. While survey data is certainly important for understanding human responses to national crises, self-presentation biases may prevent survey researchers from obtaining authentic and accurate experiences from parents (Lavrakas, 2008). In regard to reporting their emotional wellbeing, parents frequently feel pressure to live up to societal standards of motherhood or fatherhood, which may prevent them from fully disclosing their emotional distress in a survey distributed by prominent research institutions (Foucault, 2005; Lavrakas, 2008). Analyzing social media data may provide researchers with a more accurate representation of how parents are adjusting and coping with COVID-19.

Research suggests that social media interaction—specifically anonymous social media interaction—can help parents have more intimate and open conversations with others (Bernstein et al., 2011). Reddit is a social networking website where individuals use pseudonyms (i.e., usernames), making individual activity largely anonymous. Reddit users can also create temporary "throwaway" accounts (separate from their username account) where they can post anonymously. Prior empirical work that utilizes Reddit data to study parenting suggests that parents utilize username accounts to discuss topics common to parenthood such as naming a child, work-life balance, and vaccinations (Ammari et al., 2018) and use throwaway accounts to discuss more stigmatizing topics (Ammari et al., 2019). Thus, analyzing parents' Reddit posts may help researchers better understand parents' authentic responses to COVID-19. To date, no study has utilized Reddit data to examine parental emotional wellbeing during COVID-19.

The purpose of this study was to examine parents' sentiment and emotional wellbeing during COVID-19 using Reddit data—specifically, posts from the Mommit and Daddit subreddits. As of November 2020, the Mommit subreddit has over 160,000 members, and the

Daddit subreddit has over 180,000 members. To provide a general sense of how positive or negative Reddit posts were, polarity scores of mothers and fathers are examined. To provide insight into parents' expression of emotions, a sentiment analysis is conducted using the NRC Word-Emotion Association Lexicon (EmoLex). Emolex identifies 8 emotions which are argued to be the most prototypical emotions (Plutchik, 1980): anger, sadness, trust, surprise, anticipation, joy, fear, and disgust. In order to determine whether parents' emotional expression during COVID-19 is unique, Reddit data from 2019 was also examined and compared to the data collected in 2020.

Research Questions

1) What is the sentiment and emotional wellbeing among mothers and fathers during the first 6 months of the COVID-19 pandemic?

2) Does the sentiment from 2020 significantly differ from the sentiment found from the same time period in 2019?

Method

Participants and Procedure

Reddit posts from the Mommit and Daddit subreddits were scraped using Python and the PushShift application programming interface (API). Date ranges were specified as March 16, 2020-September 16, 2020. March 16 was chosen as the start date due to the fact that this was the date wherein the U.S. White House issued social distancing guidelines (The White House, Office of the Press Secretary, 2020). Data from March16, 2019-September 16, 2019 were also collected to serve as a pre-COVID-19 comparison. Data are publicly available and completely anonymous; therefore, our data did not meet criteria for Internal Review Board (IRB) oversight. The data were downloaded as a .csv file and imported into R. The final 2019 sample included 3,267

Mommit posts and 7,274 Daddit posts, and the 2020 sample included 4,864 Mommit posts and 7,484 Daddit posts.

Data Cleaning and Analysis

To clean the data, numbers were converted to text; punctuation was removed; all words were transformed to lowercase; and all words were stemmed. Average polarity scores among Mommit and Daddit subreddits were calculated for 2019 and 2020. Independent samples t-tests using Bonferonni corrections were conducted to determine whether polarity differed across years and between Mommit and Daddit subreddits. Polarity was then plotted across time to examine overall trends in polarity across the 6 months in 2019 and 2020. To inspect whether changes in polarity in 2020 visually correlated with changes in new COVID-19 cases, Mommit and Daddit subreddit polarity was graphed against COVID-19 cases (data for COVID-19 cases were pulled from the New York Times GitHub repository). Sentiment analysis was conducted using the NRC Word-Emotion Association Lexicon (i.e., EmoLex). Emolex elucidates percentages of 8 emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. Percentages of each sentiment were calculated among Mommit and Daddit subreddits for 2019 and 2020. Then, sentiments were graphed across time to determine whether general trends in emotions changed across time.

<div align="center">

Results

</div>

Average Polarity

Independent samples t-tests suggested the average polarity for the Mommit subreddit did not differ between 2019 and 2020 ($t = -1.80[7067.6]$, $p = .071$); likewise, average polarity of the Daddit subreddit did not differ between 2019 and 2020 ($t = 0.75[14,673]$, $p = .450$). However, the average polarity from the Daddit subreddit was higher than the Mommit subreddit in 2019 (t

= 9.19[6,213], $p < .001$) and 2020 (t = 7.66[9.870], $p < .001$). The plot depicting average polarity across time (see Figure 1) confirmed that Mommit subreddits had a lower average polarity than Daddit subreddits. Additionally, the polarity across years did not appear to differ greatly, with some notable exceptions: Mommit subreddit polarity in March 2020 was considerably lower than Mommit subreddit polarity in March 2019; Daddit subreddit polarity in April 2020 was lower than Daddit subreddit polarity in April 2019; Mommit subreddit polarity greatly increased between March and April 2020; and, across both years, Mommit subreddit polarity followed a curvilinear from May-September, with polarity decreasing from May-July, and increasing again in August and September.
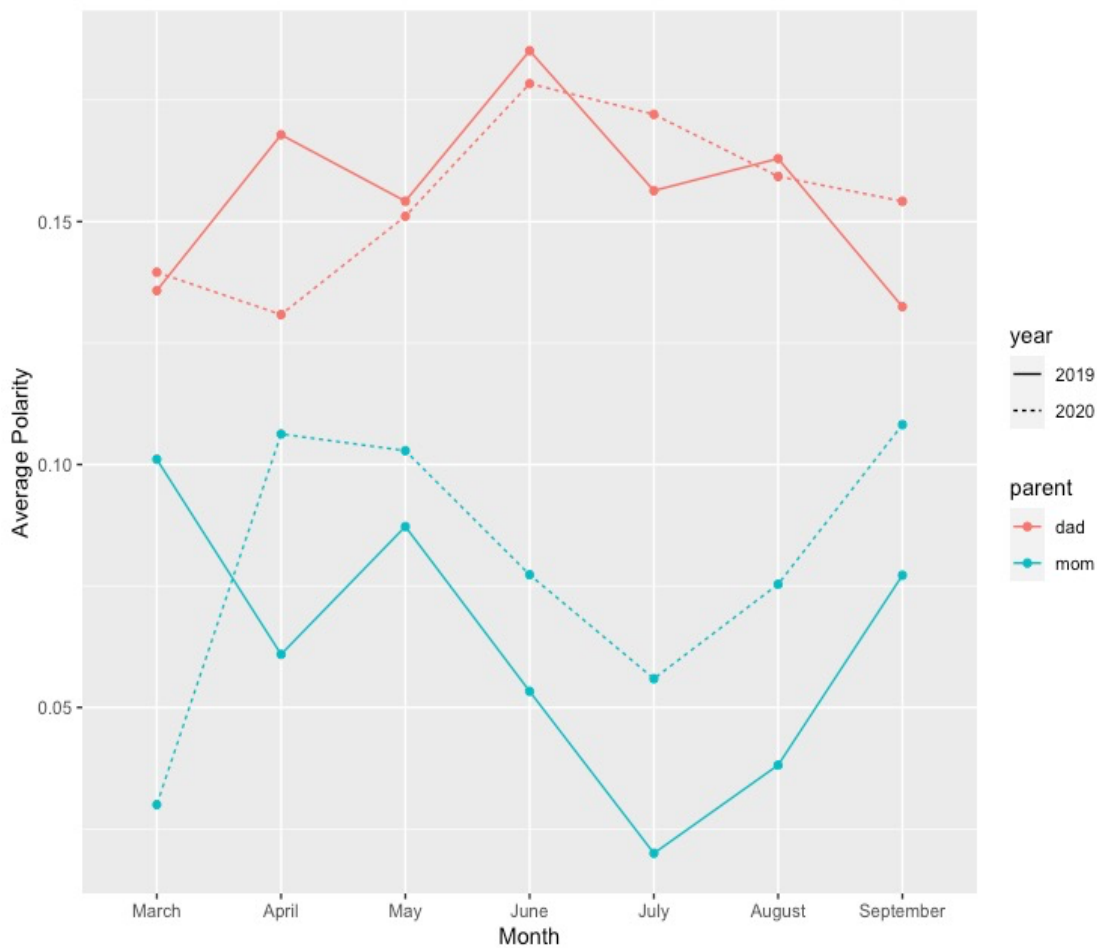

Figure 1. Mommit and Daddit subreddit polarity scores across time.

The graphs that examine Mommit and Daddit subreddit polarity against COVID-19 cases can be found in Figure 2. Daddit polarity did not appear to be associated with COVID-19 cases; Mommit's polarity was the lowest in July, which was also when the highest number of new COVID-19 cases was reported.
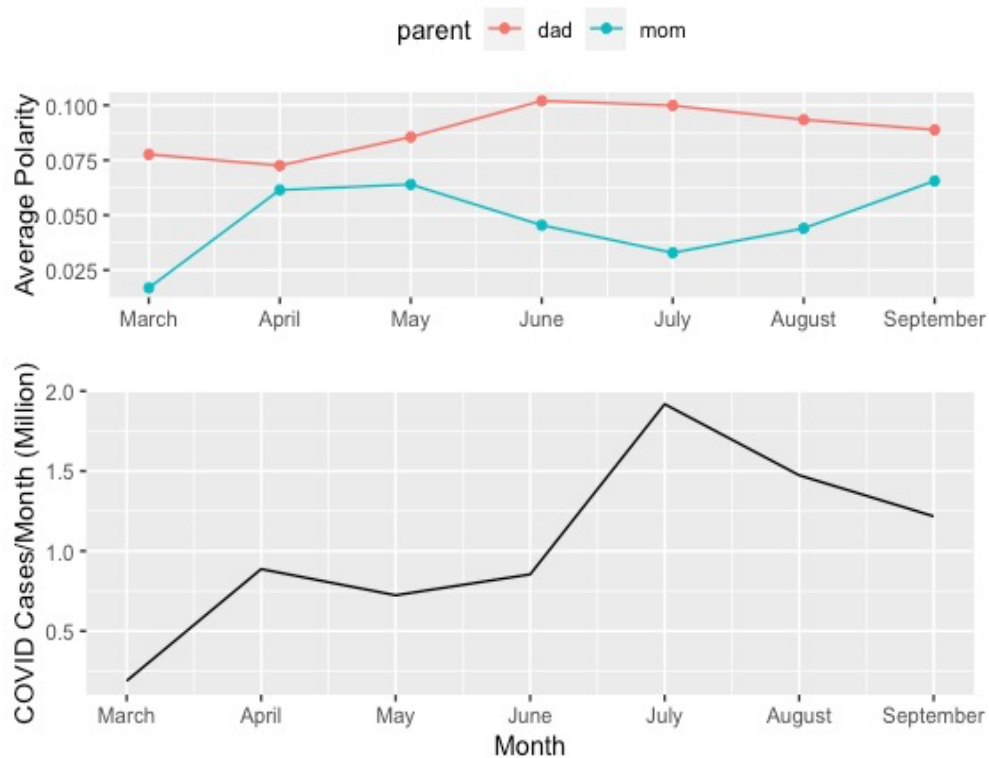

Figure 3. Mommit and Daddit subreddit polarity graphed against number of new COVID-19 cases in 2020.

The percentage of sentiments for 2019 and 2020 can be found in Figures 3 and 4, respectively. Irrespective of month, Mommit and Daddit sentiments appeared very similar to one another and across years. However, Mommit subreddits appeared to have a higher percentage of sadness in 2019 (Daddit: 10.3%, Mommit: 12.3%) and 2020 (Daddit: 10.6%, Mommit: 12.1%). Mommit appeared to have a lower percentage of trust in 2019 (Daddit: 18.3%, Mommit: 17.4%) and 2020 (Daddit: 18.6%, Mommit: 17.4%). For both subreddits, there was a higher percentage

of anticipation, joy, and trust in 2020 compared to 2019; and there was a lower percentage of

anger, disgust, and fear in 2020 compared to 2019.

Figure 3. Daddit subreddit sentiment percentages in 2019 and 2020.
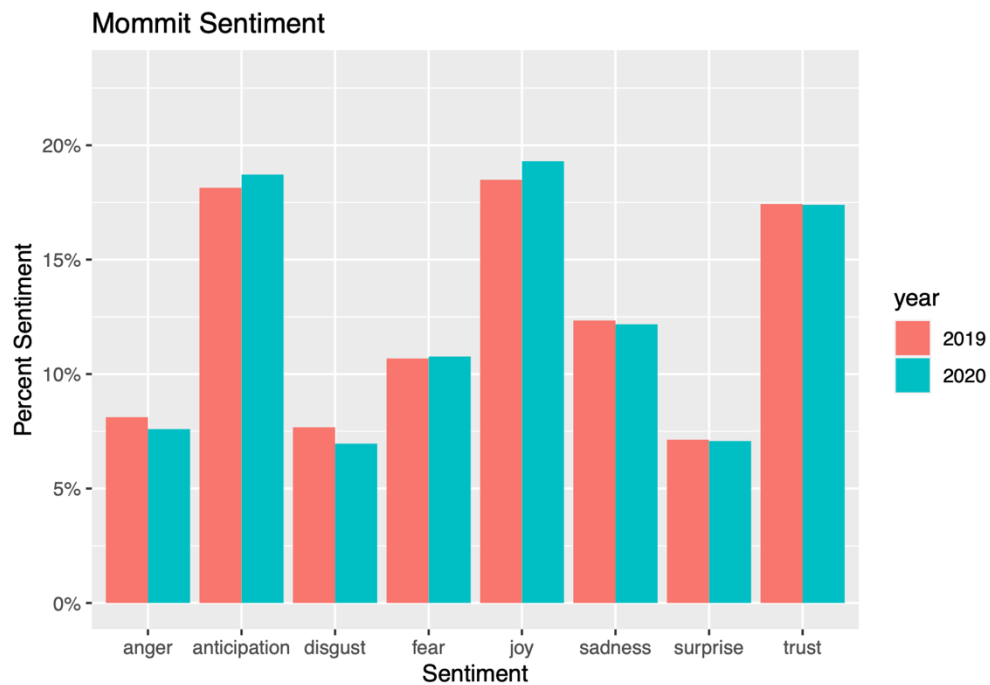

Figure 4. Mommit subreddit sentiment percentages in 2019 and 2020.

Changes in Mommit subreddit sentiment over time in 2019 and 2020 can be found in Figures 5 and 6, respectively. Regarding Mommit sentiment, in general, sentiment trends across 2019 looked very similar to those found in 2020. However, some notable exceptions occurred: whereas anticipation decreased from March-April 2019 (from 19.0% to 17.9%), anticipation increased from March-April 2020 (from 17.8% to 19.5%). Additionally, whereas anger did not change from August-September 2019 (from 8.2% to 7.8%), anger increased from August-September 2020 (from 7.6% to 8.4%). Finally, whereas a linear trend of trust was found from July-September 2019, the trend of trust remained flat from July-September 2020.
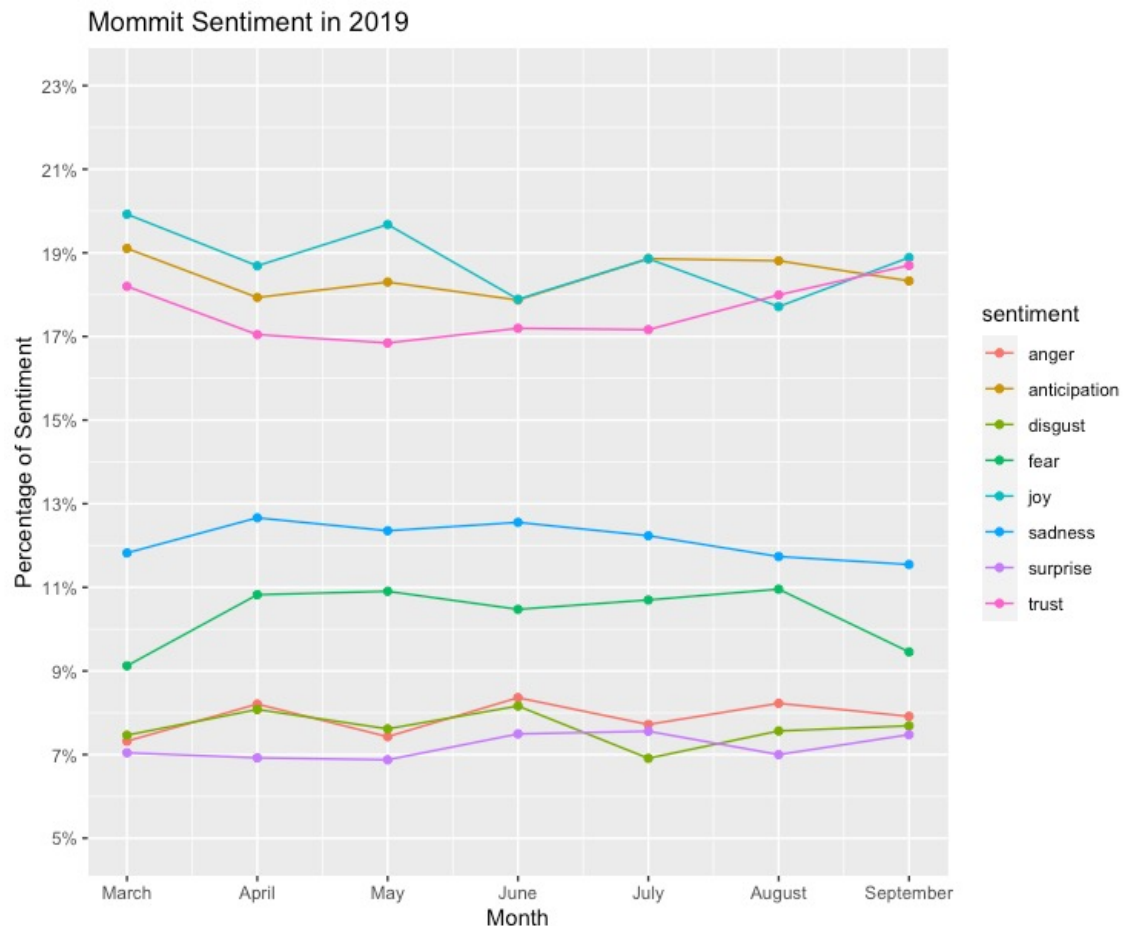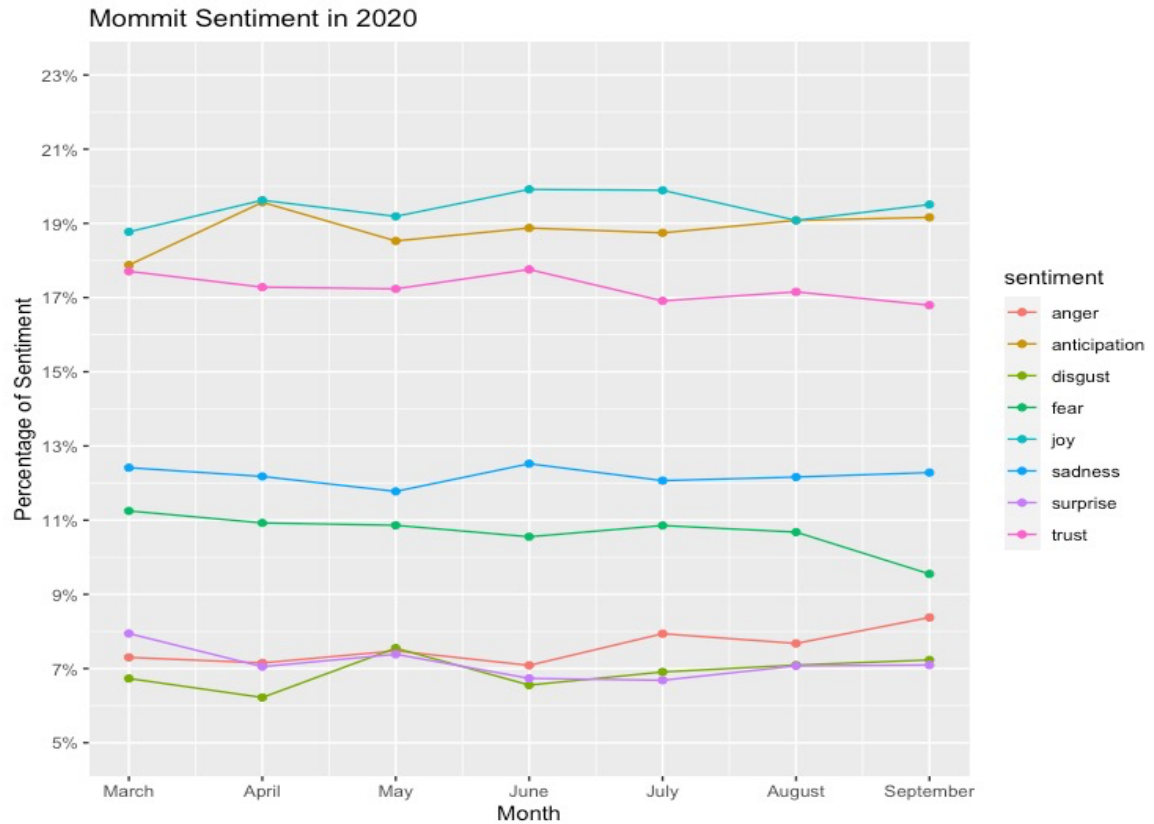

Figure 5. Mommit subreddit sentiment across time in 2019

Figure 6.

Mommit subreddit sentiment across time in 2020.

Changes in Daddit subreddit sentiment over time in 2019 and 2020 can be found in Figures 7 and 8, respectively. Similar to the trends found in the Mommit subreddit, whereas trust increased from August-September 2019 (from 17.9% to 19.1%), trust did not change from August-September 2020 (staying at 18.3%). Also similar to the Mommit subreddit, whereas anger did not change from August-September 2019 (staying at 7.8%), anger increased from August-September 2020 (from 7.3% to 8.0%). However, on average, trends in 2019 sentiment looked very similar to the trends found in 2020.
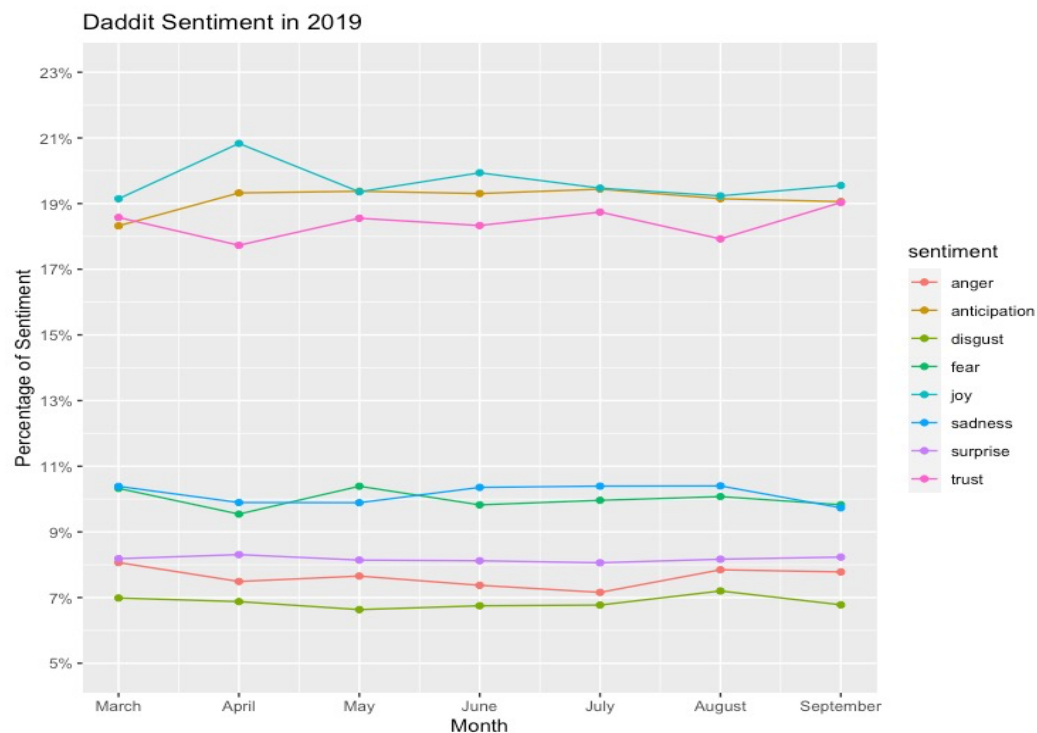
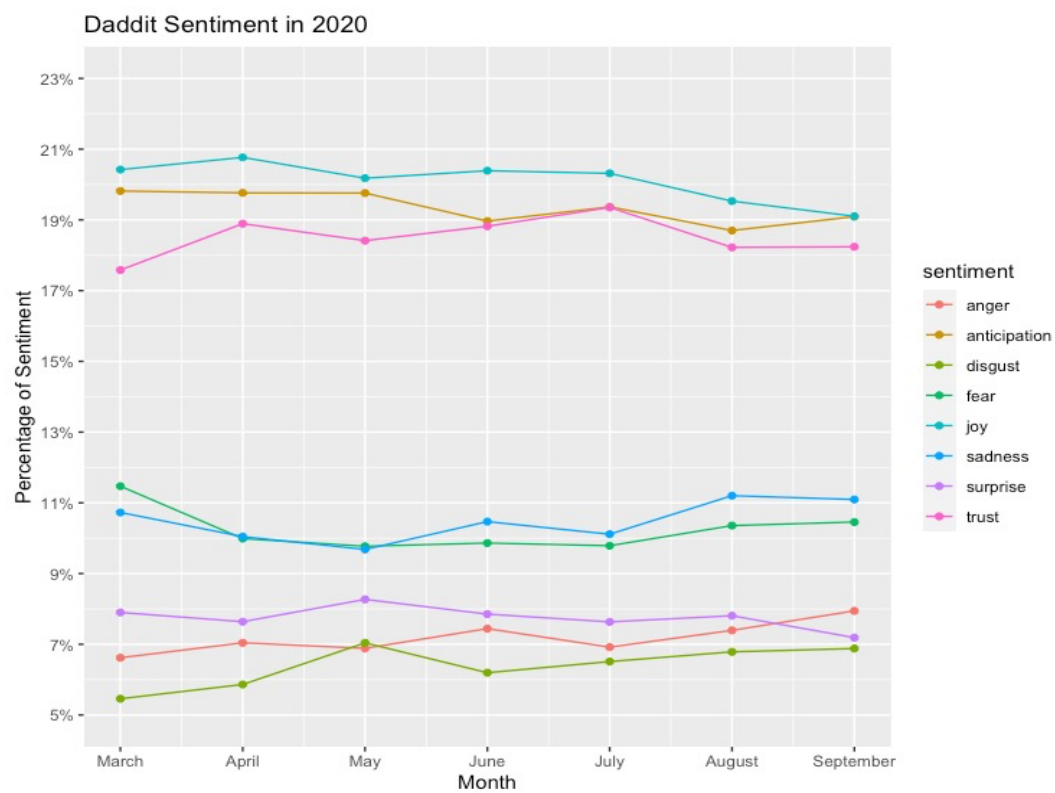*Figure 7*. Daddit subreddit sentiment across time in 2019

Figure 8. Daddit subreddit sentiment across time in 2020

Discussion

The purpose of this study was to examine the sentiment and emotional wellbeing among mothers and fathers during the first six months (i.e., March 16, 2020-September 16, 2020) of the COVID-19 pandemic. Data from the same six months in 2019 were collected to serve as a pre-COVID-19 comparison. Findings suggest that, on average, Mommit subreddits had lower polarity compared to Daddit subreddits, meaning that mothers were posting content that was more negative than fathers. This result may reflect gender norms surrounding posting online. For example, research suggests mothers turn to anonymous online sources such as Reddit to avoid judgment for not living up to the ideal of motherhood (Foucault, 2005); therefore, mothers may utilize Reddit as a mechanism for discussing more negative information than fathers. Additionally, in 2019 and 2020, we found that Mommit subreddit posts decreased in polarity from May-July. This may reflect the increased burden placed on mothers when children are home from school for the summer. When graphing 2020 polarity scores against new COVID-19 cases, mother's polarity was very low (i.e., more negative) in the month of July, when new COVID-19 cases were the highest. However, mothers' polarity also decreased in July of 2019; therefore, it is unclear whether mothers' low polarity in July 2020 were associated with the rising of COVID-19 cases, or whether this dip took place due to other factors.

Results from the sentiment analysis revealed that, overall, mothers' and fathers' sentiment percentages were quite similar to one another. However, based on graphical data, mothers showed a higher percentage of sadness and a lower percentage of trust compared to fathers. Again, this may provide insight into the norms and purposes of mothers' Reddit usage, with mothers using Reddit to express more melancholic and critical thoughts and feelings compared to fathers. When sentiments were graphed across time, some differing trends emerged

when 2019 was compared to 2020. Specifically, for mothers, anticipation increased from March 2020-April 2020 (this trend did not occur in 2019). As stay-at-home orders began to emerge in mid-March, this trend may reflect mothers' initial adjustment to the COVID-19 pandemic.

Additionally, for mothers, whereas trust showed a linear trend from July-September 2019, trust did not change from July-September 2020. Importantly, these months in 2020 were the months when U.S. experienced the largest numbers of new COVID-19 cases. Lastly, among mothers and fathers, anger increased from August-September 2020 (these trends did not occur in 2019). This lack of change in trust and increase in anger may be reflective of some distrust and irritation regarding how the COVID-19 pandemic was handled. However, future research using more sophistical analyses (such as topic modeling) would be needed to explore this supposition.

Limitations

This study should be interpreted in light of its limitations. Although Reddit data provides insight into the possibly candid and authentic expressions of parents, Reddit data does not provide information on gender identity or other important sociodemographic information. Therefore, this study assumes that Mommit users are mothers and Daddit users are fathers, although this may not be the case. Additionally, this study is primarily descriptive in nature, with percentage differences in sentiments observed based on descriptive and graphical data; therefore, causal and firm directional conclusions cannot be made. Further, our sentiment analysis did not control for the number of words or number of sentiments, which may bias the results. Additionally, while this study provides a rough approximation of parents' emotional expressions in the Mommit and Daddit subreddit, more advanced analyses would need to be conducted in order to explore the content and nature of parents' posts. One promising technique could be

Latent Dirichlet Allocation (LDA) topic modeling, which has been used previously on Reddit data (e.g., Ammari et al., 2018).

Conclusions

Due to economic downturns, school closures, and social isolation, parents are under great financial and emotional pressure during the COVID-19 pandemic. More research is needed on how parents are adjusting to COVID-19. This study suggests that mothers and fathers may be have been feeling less trusting and angrier in late-summer 2020 compared to 2019. Additionally, mothers were much more negative in March 2020 compared to March 2019, and mothers' anticipation increased from March 2020-April 2020 (which did not occur in 2019). This suggests that mothers may have experienced initial distress due to the COVID-19 pandemic. Future research should continue to monitor the emotional wellbeing of parents as the COVID-19 pandemic continues.

References

Ammari, T., Schoenbeck, S., & Romero, D. M. (2019). Self-declared throwaway accounts on Reddit: How platform affordances and shared norms enable parenting disclosure and support. *Computer Supported Cooperative Work, 3*(135), 1-30.

Ammari, T., Schoenebeck, S., & Romero, D. M. (2018). Pseudonymous parents: Comparing parenting roles and identities on the Mommit and Daddit subreddits. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, 489,* 1-13. https://doi.org/10.1145/3173574.3174063

Bernstein, M. S., Monroy-Hernandez, A., Harry, D., Andre, P., Panovich, K., & Vargas, G. G. (2011). 4chan and /b/: An analysis of anonymity and ephemeraility in a large online community. In *ICWSM*, 50-57. Retrieved from https://www.media.mit.edu/publications/4chan-and-b-an-analysis-of-anonymity-and-ephemerality-in-a-large-online-community-2/

Centers for Disease Control and Prevention (2020). *United States COVID-19 Cases and Deaths by State.* Retrieved from https://covid.cdc.gov/covid-data-tracker/#cases_casesper100klast7days

Conger, J. C., & Conger, K. J. (2002). Resilience in Midwestern families: Selected findings from the first decade of a prospective, longitudinal study. *Journal of Marriage and Family, 64*, 261-373. doi: 10.1111/j.1741-3737.2002.00361.x

Douglas, Y. (2020, April 29). The costs of social isolation: Loneliness and COVID-19. *Psychiatry Advisor*. www.psychiatryadvisor.com/home/topics/general-psychiatry/costs-of-social-isolation-loneliness-covid19/

Foucault, B. E. (2005). Designing technology for growing families. *Technology*: 1.

Lavrakas, P. J. (2008). *Encyclopedia of survey research methods* (Vols. 1-0). Thousand Oaks, CA: Sage Publications, Inc. https://doi.org/10.4135/9781412963947

Lee, S. J., & Ward, K. P. (2020, March). Stress and parenting during the coronavirus pandemic. University of Michigan Parenting in Context Research Lab. https://bit.ly/2D3Fv2I

Lee, S. J., Ward, K. P., Chang, O. D., & Downing, K. M. (2020). Parenting activities and the transition to home-based education during the Covid-19 pandemic. *Children and Youth Services Review*. https://doi.org/10.1016/j.childyouth.2020.105585

Neppl, T. K., Senia, J. M., & Donnellan, M. B. (2016). Effects of economic hardship: Testing the family stress model over time. *Journal of Family Psychology, 30*, 12-21. doi:10.1037/fam0000168

Pew Research Center (2020). Economic fallout from COVID-19 continues to hit lower-income Americans the hardest. Retrieved from https://www.pewsocialtrends.org/2020/09/24/economic-fallout-from-covid-19-continues-to-hit-lower-income-americans-the-hardest/#:~:text=Fully%2015%25%20of%20adults%20report,they%20are%20currently%20not%20employed.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Emotion: Theory, research, and experience, 1*(3), 3–33.

The White House, Office of the Press Secretary. (2020, March 16). *The President's Coronavirus Guidelines for America: 15 Days to Slow the Spread.* Retrieved from https://www.justice.gov/doj/page/file/1258511/download

UNESCO (2020). *Adverse consequences of school closures*. Retrieved from https://en.unesco.org/covid19/educationresponse/consequences

World Health Organization (2020). *Coronavirus disease 2019 (COVID-19) Situation Report 93.*

Retrieved from https://www.who.int/docs/default-source/coronaviruse/situation-

reports/20200422-sitrep-93-covid-19.pdf?sfvrsn=35cf80d7_4