

# How Do Parents Use Online Communities? An Analysis of the r/Parenting Subreddit

Anna Lieb, Athena Chapekis, Sono Shah, Aaron Smith

Pew Research Center, Washington, DC, USA  
alieb@pewresearch.org

## Abstract

As today's parents navigate the challenges of raising a child, many are turning to online spaces where fellow parents share their experiences. To better understand parents' use of online communities, we conduct a two-part analysis of data from (1) the r/Parenting subreddit and (2) a nationally representative survey of U.S. parents. Based on 29,295 posts and 853,209 comments collected from r/Parenting, we find that about half of posts express negative emotions — yet the vast majority of comments in reply to these posts are supportive. We also surface themes from three salient topics that appear in posts on the subreddit. Survey data from a sample of 3,054 parents of children ages 12 and under complements the Reddit analysis by providing insight on how parents use online communities in day-to-day life. We find that about two-thirds of U.S. parents say they visit online parenting communities, and that mothers, young parents, parents with young children, and parents with lower incomes are especially likely to do so. Overall, we leverage both data sources to characterize who participates in online parenting communities, what they discuss there, and how parents experience these online spaces.

## Introduction

Parents face a variety of pressures including financial challenges, children's health & safety, managing children's digital device usage, societal expectations, and balancing time between work and home (Anderson, Faverio, and Park 2024; Braga 2022; Leppert 2024; Murthy 2024). As they navigate these challenges, many parents consult online parenting communities, such as social media and discussion forums.

Previous research suggests that younger parents, parents of children with special needs, and families facing barriers to accessing other forms of support are among the most likely to use online parent communities and benefit from them (Baker, Sanders, and Morawska 2017; Ammari, Schoenebeck, and Morris 2014). Analyses of these online discussions have found that they are wide-ranging and often center around asking questions, offering support, and sharing information about child care and managing parental stress (Francisco 2025; Ibrahim et al. 2024; McSorley et al. 2022; Mertens et al. 2024). But these communities are not a monolith: Sepahpour-Fard and Quayle (2022) found that

topics of discussion can differ between moms and dads, while a topical analysis by Gao, Jang, and Yang (2021) focuses on discussions among parents of younger children.

Some researchers suggest that the dynamics of these online parenting communities were shaped by the Covid-19 pandemic. During the pandemic, parents reported higher levels of perceived stress and experienced interruptions in their support networks (Brown et al. 2020; Yakub et al. 2024; Murthy 2024). Multiple studies suggest that online parenting communities were an important source of social support for parents at the time (Hooper, Schweiker, and Kerch 2022; Francisco 2025). To the extent that online communities can boost parents' social connectedness, they are an opportunity for increasing support for parents and improving outcomes for parents and their children (Plesko et al. 2024). On the other hand, these communities have the potential to pressure parents to conform to social norms or fear judgment from other parents (Robbeets et al. 2024; Clark et al. 2023).

Using six months of data from the r/Parenting subreddit, we analyze multiple factors including the emotional valence of posts, supportiveness of comments, and themes from three salient topics that appear on the subreddit. We supplement the Reddit analysis with a nationally representative survey of U.S. parents to better understand the role of online communities. This study aims to answer: **(RQ1)** What issues do parents discuss on r/Parenting, and what sentiments do they express in these discussions? **(RQ2)** How do U.S. parents use online communities in their day-to-day life, and how do they make them feel?

## Methods

This analysis uses two data sources: a representative survey of U.S. parents and data from the r/Parenting subreddit.

### Reddit data collection

The r/Parenting subreddit is the largest community for parents on Reddit, with more than 8 million members as of November 2025. Reddit posts, comments, and engagement metrics were collected from r/Parenting using requests to the official Reddit Data API. We ran the Reddit data collection pipeline on a continuous hourly schedule from Jan. 17 to July 17, 2025 to comply with API rate limits while constructing a complete picture of activity on the subreddit. Each hour, we collected the following:

- All submissions posted to r/Parenting in the past hour, including the submission author, title, and body text. We also collected metadata such as the time it was posted and the tags (or “flairs”) applied to the post.
- Comments on all submissions posted to r/Parenting 48 hours earlier, including the comment text, author and upvote count. We also collected metadata such as the time it was posted and whether the comment was a reply to the original post or a reply to another comment.
- Engagement metrics for submissions collected 48 hours earlier, including upvote score.<sup>1</sup>

Data robustness checks indicate that our pipeline successfully retrieved posts and comments at a consistent rate throughout the data collection period, with the exception of a brief sitewide Reddit outage on Feb. 21. In total, we collected 29,295 posts and 853,209 comments throughout the six-month data collection period.<sup>2</sup> Based on language detection with the `stringi` R package (Gagolewski 2022), over 99% of posts in r/Parenting are written in English. The subreddit is most active during the daytime in American time zones: the most common times to post or comment were between 10 a.m. and 2 p.m. U.S. Eastern time.

On Oct. 16, 2025, about nine months after the start of our data collection period, we made a request to the Reddit Data API for every submission in our dataset to check whether the submission status was active, deleted or removed. We found that posts on r/Parenting often get deleted or removed, as just 62% of the posts we collected were still accessible on the site. About 29% were deleted by the post author, while 9% were removed by the subreddit moderators. However, all data from deleted posts were still included in our analysis, since we saved content from posts as we collected them.

### Analysis of emotions and topics on r/Parenting

To better understand discourse on r/Parenting, we identified the emotions expressed and topics mentioned in each post in our dataset. Researchers worked with qualitative coding interns over multiple iterative coding rounds to develop a codebook that describes these categories and accurately reflects themes that commonly appear on r/Parenting.

We defined binary labels for three topics: family finances, children’s technology use, and the division of parenting duties. Among many topics that appeared often in the data, we focused on these because they are relevant to otherwise salient conversations about challenges that parents face today (Murthy 2024). The topic labels were defined based on mentions of each topic in a given post, but they do not necessarily indicate the primary focus of the post. Thus, posts may be labeled as containing more than one topic. We also categorized the emotional valence of each post as positive, negative, or neutral. Positive emotions include gratitude, pride, joy or amusement, while negative emotions include anger,

<sup>1</sup>Upvote score is the number of upvotes minus downvotes, along with “fuzzing” adjustments by a Reddit algorithm whose details are not publicly known. In this analysis, we only use upvote score in combination with comment count to identify popular posts.

<sup>2</sup>Due to the Reddit Data API terms, our Reddit dataset will not be shared publicly. Contact the authors for specific data inquiries.

sadness, disgust, stress or fear. The neutral label includes posts that express unclear, mixed or no emotions.

After finalizing the labels and definitions, two qualitative analysis interns labeled a sample of 700 Reddit posts. Conflicts between annotator labels were resolved by the research team to create a ground truth dataset. Then, we used GPT-4.1 mini to classify all 29,295 Reddit posts in the full dataset. For each post, we prompted the model with a description of each label and the Reddit post text.<sup>3</sup> Inter-rater reliability and model performance metrics are available in Table 1.

### Comment sentiment analysis

Discourse on r/Parenting largely takes place in the comments section. To learn more about community support provided by commenters on r/Parenting, we analyzed comments in our dataset that reply directly to a post. We chose to only analyze these top-level comments, rather than comments replying to other comments, to avoid measuring lengthy threads that are only adjacently related to the original post. Of the 853,209 comments in our dataset, 475,311 were valid top-level comments. This includes comments that reply directly to the original post and are not authored by the original post author or a subreddit moderator.

As with the post-level labels, the comment supportiveness label definitions were developed over multiple iterative rounds of qualitative coding to accurately reflect sentiments that appear on r/Parenting. A comment was classified as “supportive” if it thoughtfully replies to the original post, offers helpful advice, or otherwise contributes positively to the community discussion. By this definition, “supportive” does not require that the comment agrees with the post; for example, constructive criticism or respectful disagreement counts as supportive if it contributes positively to the conversation or replies helpfully to the post. Comments that have a disrespectful tone, criticize the original post or author, are off-topic, or are difficult to interpret would be classified as “not supportive.” We intended to include a “critical” category, but after manually reviewing several hundred comments, our research team failed to identify a sufficient number of these cases to confidently develop a “critical” label distinct from the “not supportive” category. We believe this speaks to the relative rarity of toxicity and criticism in r/Parenting.

To create a ground truth dataset for comment classification, two qualitative coding interns labeled a sample of 1,005 top-level comments as “supportive” or “not supportive.” The qualitative coders reached 92% agreement on the comments ground truth dataset (Cohen’s  $\kappa = 0.624$ ). In the case of conflicts between coders, a third member of the research team selected a winning label. Then, we used GPT-4.1 mini to classify all 475,311 top-level comments in the dataset. Although pre-trained non-LLM sentiment classifiers exist, this LLM-based approach enabled us to use our custom “supportiveness” construct rather than the positive/negative scale found in out-of-the box sentiment models. This is especially important for a community where, based on our exploratory

<sup>3</sup>Classification prompts, including full definitions of topic and emotion labels, are available at [pewresearch.org/data-labs/2025/11/12/appendix-b-post-and-comment-classification-prompts/](https://pewresearch.org/data-labs/2025/11/12/appendix-b-post-and-comment-classification-prompts/)

Label: Definition	IRR (Cohen’s $\kappa$ )	Annotator agreement	F1 score	Precision / Recall
<b>Emotion:</b> Identifies the emotions expressed by the parent authoring the post as positive, negative or neutral.	0.620	77%*	0.825	0.833 / 0.827
<b>Family finances topic:</b> Identifies posts that mention family finances or costs related to raising a child.	0.703	91%	0.882	0.916 / 0.873
<b>Technology use topic:</b> Identifies posts that mention children’s use of digital technology.	0.751	92%	0.877	0.916 / 0.867
<b>Unequal division of labor topic:</b> Identifies posts that mention an unequal division of child care between parents, including delegating responsibilities or navigating co-parenting.	0.664	91%	0.820	0.880 / 0.799

Table 1: Label definitions, human annotator agreement metrics (IRR and % agreement), and GPT-4.1 mini classification performance metrics (F1 score and precision/recall) for Reddit post labels (n = 700 posts, 2 annotators). Precision, recall, and F1 are weighted averages by class. \*For emotion, the main source of disagreement was between the “neutral” label and other labels. Completely opposing label conflicts (i.e., “positive” and “negative”) were rare, occurring in less than 2% of posts in the sample.

rounds of qualitative analysis, sarcasm and commiseration are popular forms of support. The model prompt included annotation instructions and a description of the supportiveness labels, along with the Reddit post and comment text to be classified.<sup>4</sup> Based on the ground truth dataset, the model achieved 87% accuracy, with a weighted average F1 score of 0.876, precision of 0.886 and a recall of 0.868.

### Survey of U.S. parents

For the survey component of this analysis, we surveyed 3,054 parents who have children ages 12 and under from May 13 – 26, 2025. Respondents were recruited through random sampling of U.S. residential addresses, giving nearly all U.S. adults a chance of selection (Harter et al. 2016). Potentially eligible respondents gave consent for survey results to be used for research and were asked a series of screening questions to confirm whether they were the parent or guardian of any child 12 or under. All respondents were offered an incentive for their participation, ranging from \$5 to \$15 depending on whether the respondent belongs to a part of the population that is harder or easier to reach.

The survey data is weighted in a process that accounts for sampling and nonresponse. Each panelist begins with a base weight that reflects their probability of recruitment. Then, the weights were adjusted to align with 15 estimated benchmarks for the population of parents of children ages 12 and under: age, gender, education, race/ethnicity, years living in the U.S. (among foreign born), volunteerism, voter registration, frequency of internet use, religious affiliation, party affiliation, census region, metropolitan status, and number of children in various age ranges. Income tiers are assigned based on the median income of all panel members, adjusted for household size and cost-of-living differences by geography. In this paper, sampling errors and significance tests take into account the effect of weighting. The margin of sampling

<sup>4</sup>The comment sentiment classification prompt, including a detailed definition of our supportiveness measure, is available at [pewresearch.org/data-labs/2025/11/12/appendix-b-post-and-comment-classification-prompts/](https://pewresearch.org/data-labs/2025/11/12/appendix-b-post-and-comment-classification-prompts/). In total, all LLM usage for this research cost about \$300.

error for the full sample of 3,054 respondents at a 95% confidence level is  $\pm 2.2$  percentage points. Sample sizes and sampling errors for subgroups are available upon request. Survey data containing PII is stored securely on internal servers and is accessible only to core survey researchers. Anonymized data will be available after an embargo period at [pewresearch.org/datasets](https://pewresearch.org/datasets).<sup>5</sup>

## Results & Discussion

Our results leverage both the Reddit analysis and the survey of U.S. parents to characterize who participates in online parenting communities, what they discuss there, and how parents experience these online spaces.

**Who visits online parenting communities?** We find that about two-thirds of parents visit online parenting communities, including majorities of parents in nearly every major demographic group. Certain groups, including mothers, younger parents, newer parents, and parents with lower incomes are most likely to do so. This suggests that digital spaces mirror trends in parenting practices at large, such as the positive link between socioeconomic status and parenting practice (Ayoub and Bachir 2025) and unequal division of child care between mothers and fathers (Raley, Bianchi, and Wang 2012). Point estimates for these groups are available in Figure 1. Differences by race and education level were not statistically significant ( $p > 0.05$ ).

The prevalence of online community use among new parents is consistent with our analysis of tags (also known as flairs) on r/Parenting. The most common tags were for discussions of toddlers and young children, with about 20% of posts tagged “Toddler 1-3 years” and a similar share (23%) tagged “Child 4-9 years.” The next most common age group tag was for infants 2-12 months, making up 9% of posts. Tags for older ages (10–12 or 13+ years) were less common.

<sup>5</sup>A detailed copy of the questionnaire is available at [pewresearch.org/wp-content/uploads/sites/20/2025/11/pl\\_2025.11.12\\_reddit-parenting\\_questionnaire.pdf](https://pewresearch.org/wp-content/uploads/sites/20/2025/11/pl_2025.11.12_reddit-parenting_questionnaire.pdf).

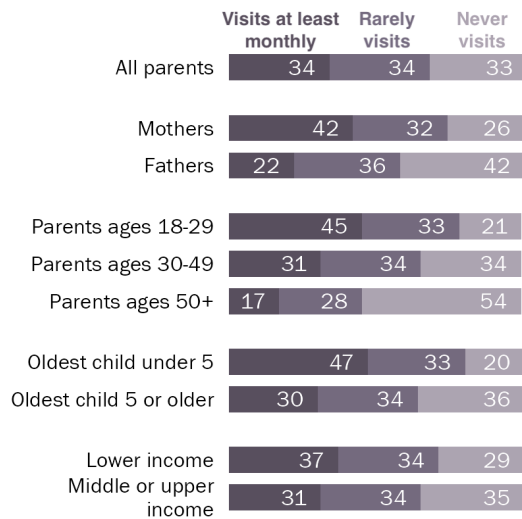


Figure 1: Percent of U.S. parents of a child age 12 or younger who say they visit places online where parents talk about their experiences raising children monthly, rarely, or never (n=3,054). Differences by gender, parent age, age of oldest child, and income level are significant at  $p < 0.05$ .

**Participation in online community.** During the six-month study period, over 110,000 unique accounts authored a post or a comment on r/Parenting. The vast majority of these accounts (94%) left a comment, while only 20% wrote a post. Those who comment tend to do so repeatedly, with 59% of commenters writing multiple comments over the six months. By comparison, only 16% of post authors submitted multiple times in the same period. Commenting is highly concentrated among a small share of users: The top 5% of commenters produced around half (52%) of all comments.

Our survey results put community participation into further perspective, indicating that “lurker” behavior is common in online parenting communities. We found that among parents who ever visit online parenting communities, 40% say they have ever posted or commented about their experiences as a parent, and just 5% say they do this often. Mothers are more likely than fathers to engage in the discussion: 46% of mothers who visit online parenting communities say they ever post or comment compared to 30% of fathers, mirroring gender inequality found in offline child care duties (Raley, Bianchi, and Wang 2012).

**Topic analysis.** Discussions on r/Parenting address a wide range of topics, but we analyzed three topics that appear frequently in user submissions: unequal division of child care responsibilities between parents, children’s use of technology, and the financial challenges of raising children. In total, 44% of all posts in our sample mention one or more of these issues. Themes for each of these topics were surfaced with a distinctive terms analysis based on pointwise mutual information calculations (Bestvater and Shah 2022), along with qualitative analysis of a sample of posts in each topic.<sup>6</sup>

<sup>6</sup>See distinctive terms at [pewresearch.org/data-labs/2025/11/12/appendix-a-distinctive-terms-by-topic-in-r-parenting-posts/](https://pewresearch.org/data-labs/2025/11/12/appendix-a-distinctive-terms-by-topic-in-r-parenting-posts/)

About 23% of all posts mentioned unequal division of parenting duties. Frequent themes in these posts include stay-at-home parenthood, co-parenting between separated parents, and balancing domestic work with a career outside the home. Meanwhile, children’s technology use was mentioned in 18% of all posts. Posts mentioning this topic are distinguished by discussions of devices and screen time, TV, social media like YouTube and TikTok, and popular video games for young people like Minecraft or Roblox. Mentions of financial challenges of raising children appeared in 13% of all posts. Common themes in these posts include budgeting, costs and affordability, and how to navigate significant expenses like rent, tuition, mortgages and paying off debts.

**Emotional valence & community support.** About half of r/Parenting posts (48%) convey negative emotions like anger, sadness, disgust, stress or fear. A similar share (47%) express neutral or mixed emotions. Posts with mainly positive emotions like gratitude, pride, joy or amusement are much rarer, making up just 5% of posts we analyzed. Some topics are more likely to express negativity than others. For instance, 59% of popular posts (i.e. posts in the 90th percentile by number of comments and upvote score) express primarily negative emotions. And of the topics we measured, posts that mention the division of parenting responsibilities are the most negative. Around two-thirds of all posts mentioning this topic (68%) express negative emotions.

Although negative emotions are common in r/Parenting posts, we found that comments are overwhelmingly supportive. Of the top-level comments that we collected over this six-month period, 88% offered support. And upon qualitative review of the 12% of comments that were not labeled as supportive, we found that they mostly expressed neutral or mixed sentiments. Outright critical comments were extremely rare. In fact, we initially intended to classify critical comments, but we could not find enough examples to build a validation dataset. This is a promising sign of the supportiveness of online parenting communities, although a toxic minority can have an outsized impact (Kumar et al. 2023).

Survey findings suggest similarly positive parent experiences with online communities. Among those who visit online parenting communities, 63% say these spaces make them feel more informed about being a parent and about half (48%) say online communities make them feel more connected to other parents. But these experiences are not universally positive: 38% say parenting communities make them feel more overwhelmed because of all the information they need to know. And small but notable shares say these communities make them feel less informed about being a parent (11%) and less connected to other parents (17%).

**Future work.** Future studies can work to compare these results to other forms of community on Reddit by analyzing the dynamics of subreddits that are organized around different identity groups, shared experiences, or interests. Additionally, future mixed methods work combining surveys with web data would benefit from linking individual respondents with their respective web data. This would enable researchers to make more meaningful inferences about online experiences and their impacts in offline life.

## References

- Ammari, T.; Schoenebeck, S.; and Morris, M. 2014. Accessing Social Support and Overcoming Judgment on Social Media among Parents of Children with Special Needs. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1): 22–31.
- Anderson, M.; Faverio, M.; and Park, E. 2024. How Teens and Parents Approach Screen Time. *Pew Research Center*.
- Ayoub, M.; and Bachir, M. 2025. A Meta-analysis of the Relations Between Socioeconomic Status and Parenting Practices. *Psychological Reports*, 128(6): 4470–4494.
- Baker, S.; Sanders, M. R.; and Morawska, A. 2017. Who Uses Online Parenting Support? A Cross-Sectional Survey Exploring Australian Parents' Internet Use for Parenting. *Journal of Child and Family Studies*, 26(3): 916–927.
- Bestvater, S.; and Shah, S. 2022. Analyzing text for distinctive terms using pointwise mutual information. *Pew Research Center*.
- Braga, D. 2022. One-in-four U.S. parents say they've struggled to afford food or housing in the past year. *Pew Research Center*.
- Brown, S. M.; Doom, J. R.; Lechuga-Peña, S.; Watamura, S. E.; and Koppels, T. 2020. Stress and parenting during the global COVID-19 pandemic. *Child Abuse & Neglect*, 110(Pt 2): 104699.
- Clark, S. J.; Woolford, S. J.; Freed, G. L.; Singer, D. C.; and Gebremariam, A. 2023. Sharing on parenting: Getting advice through social media. Technical report, C.S. Mott Children's Hospital, University of Michigan Health.
- Francisco, S. C. 2025. Hey Mama, We Hear You: An Analysis of Online Parenting Support Conversations on Reddit. *Families in Society*, 106(4): 1207–1228.
- Gagolewski, M. 2022. stringi: Fast and Portable Character String Processing in R. *Journal of Statistical Software*, 103(2).
- Gao, Y.; Jang, J.; and Yang, D. 2021. Understanding the Usage of Online Media for Parenting from Infancy to Preschool At Scale. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, 1–12. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-8096-6.
- Harter, R.; Battaglia, M. P.; Buskirk, T. D.; Dillman, D. A.; English, N.; Fahimi, M.; Frankel, M. R.; Kennel, T.; McMichael, J. P.; McPhee, C. B.; Montaquila, J.; Yancey, T.; and Zuckerberg, A. L. 2016. AAPOR Report: Address-based Sampling.
- Hooper, A.; Schweiker, C.; and Kerch, C. 2022. Social support in a parenting Facebook group during the COVID-19 pandemic. *Family Relations*, 10.1111/fare.12804.
- Ibrahim, S.; Ang, J. R. X.; Petsolari, M.; Michelson, R.; Dong, Y.; Theofanopoulou, N.; Van Kleek, M.; Davis, K.; and Slovák, P. 2024. Understanding Online Parental Help-Seeking and Help-Giving in Early Childhood: The Design Challenges of Supporting Complex Parenting Questions. *Proc. ACM Hum.-Comput. Interact.*, 8(CSCW1): 199:1–199:36.
- Kumar, D.; Hancock, J.; Thomas, K.; and Durumeric, Z. 2023. Understanding the Behaviors of Toxic Accounts on Reddit. In *Proceedings of the ACM Web Conference 2023*, WWW '23, 2797–2807. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-9416-1.
- Leppert, R. 2024. 5 facts about child care costs in the U.S. *Pew Research Center*.
- McSorley, L.; Deighton-Smith, N.; Budds, K.; and Wang, X. 2022. “Hang in there mama!” The role of online parenting forums in maternal mental health. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 16(4).
- Mertens, E.; Ye, G.; Beuckels, E.; and Hudders, L. 2024. Parenting Information on Social Media: Systematic Literature Review. *JMIR Pediatrics and Parenting*, 7: e55372.
- Murthy, V. H. 2024. Parents Under Pressure: The U.S. Surgeon General's Advisory on the Mental Health & Well-Being of Parents. *Office of the U.S. Surgeon General*.
- Plesko, C. M.; Yu, Z.; Tobin, K.; Richman, R.; and Gross, D. 2024. A mixed methods study of parents' social connectedness in a group-based parenting program in low-income communities. *The American Journal of Orthopsychiatry*, 94(1): 1–14.
- Raley, S.; Bianchi, S. M.; and Wang, W. 2012. When Do Fathers Care? Mothers' Economic Contribution and Fathers' Involvement in Child Care. *AJS; American Journal of Sociology*, 117(5): 1422–1459.
- Robbeets, C.; Bastien, M.; Jacques, J.; Campion, B.; Roberti-Lintermans, M.; François, A.; and Merla, L. 2024. Exploring Parents' Everyday Experiences With Digital Media: Barriers and Opportunities for Digital Inclusion. *Media and Communication*, 12: 8172.
- Sepahpour-Fard, M.; and Quayle, M. 2022. How Do Mothers and Fathers Talk About Parenting to Different Audiences? Stereotypes and Audience Effects: An Analysis of r/Daddit, r/Mommit, and r/Parenting Using Topic Modelling. In *Proceedings of the ACM Web Conference 2022*, WWW '22, 2696–2706. New York, NY, USA: Association for Computing Machinery. ISBN 978-1-4503-9096-5.
- Yakub, N. A.; Ali, S. F.; Farid, N. D. N.; Shari, N. I.; and Panatik, S. A. 2024. A systematic review comparing pre-post COVID-19 pandemic parenting style and mental health-related factors. *Journal of Affective Disorders Reports*, 17: 100826.

## Paper Ethics Checklist

1. For most authors...
  - (a) Would answering this research question advance science without violating social contracts, such as violating privacy norms, perpetuating unfair profiling, exacerbating the socio-economic divide, or implying disrespect to societies or cultures? **Yes, all Reddit data was publicly accessible and analysis was conducted on an aggregate scale; no user IDs or text from posts or comments was shared. Survey was conducted with informed consent from participants.**

- (b) Do your main claims in the abstract and introduction accurately reflect the paper’s contributions and scope? **Yes**
- (c) Do you clarify how the proposed methodological approach is appropriate for the claims made? **Yes, we explain how Reddit data and survey data are complementary.**
- (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? **Yes, our survey was nationally representative and results are weighted to better represent U.S. parents. The time and scope of Reddit data collection was clearly indicated.**
- (e) Did you describe the limitations of your work? **Yes, in the Methods and Future work sections.**
- (f) Did you discuss any potential negative societal impacts of your work? **No, we do not foresee noteworthy potential negative impacts.**
- (g) Did you discuss any potential misuse of your work? **No, we do not foresee significant potential misuses.**
- (h) Did you describe steps taken to prevent or mitigate potential negative outcomes of the research, such as data and model documentation, data anonymization, responsible release, access control, and the reproducibility of findings? **Yes, Reddit data will not be released and survey data will be released without PII.**
- (i) Have you read the ethics review guidelines and ensured that your paper conforms to them? **Yes**
2. Additionally, if your study involves hypotheses testing...
- (a) Did you clearly state the assumptions underlying all theoretical results? **NA**
- (b) Have you provided justifications for all theoretical results? **NA**
- (c) Did you discuss competing hypotheses or theories that might challenge or complement your theoretical results? **NA**
- (d) Have you considered alternative mechanisms or explanations that might account for the same outcomes observed in your study? **NA**
- (e) Did you address potential biases or limitations in your theoretical framework? **NA**
- (f) Have you related your theoretical results to the existing literature in social science? **NA**
- (g) Did you discuss the implications of your theoretical results for policy, practice, or further research in the social science domain? **NA**
3. Additionally, if you are including theoretical proofs...
- (a) Did you state the full set of assumptions of all theoretical results? **NA**
- (b) Did you include complete proofs of all theoretical results? **NA**
4. Additionally, if you ran machine learning experiments...
- (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? **Yes, the model version and prompts are provided. However, the validation dataset will not be shared due to Reddit Data API terms.**
- (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? **Yes, prompts are provided at a URL address.**
- (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? **No, our LLM classification procedure does not demand this type of random seed testing given our applied use case.**
- (d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? **Yes, the cost of OpenAI API requests was noted.**
- (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? **Yes, we share standard model performance metrics and inter-rater reliability for qualitative coding.**
- (f) Do you discuss what is “the cost“ of misclassification and fault (in)tolerance? **Yes, we discuss the type of disagreements and share precision/recall to help readers understand the type of error that we measured. We do not further discuss specific costs of misclassification, since we did not identify any high-cost consequences.**
5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets, **without compromising anonymity...**
- (a) If your work uses existing assets, did you cite the creators? **Yes, all survey data is our own and Reddit data is attributed to Reddit.**
- (b) Did you mention the license of the assets? **Yes, in the methods section.**
- (c) Did you include any new assets in the supplemental material or as a URL? **Yes, for some materials. Survey data will be made available at the given URL after an embargo period, but Reddit Data API terms prohibit sharing the data publicly.**
- (d) Did you discuss whether and how consent was obtained from people whose data you’re using/curating? **Yes, we explicitly mention obtaining consent as is consistent with all surveys from the American Trends Panel (see more at [pewresearch.org/the-american-trends-panel/](https://www.pewresearch.org/the-american-trends-panel/)).**
- (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? **Yes, we mention that released survey data is anonymized. Reddit data may contain potentially offensive content or PII, but this data will not be released.**
- (f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR? **NA**
- (g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset? **NA**

6. Additionally, if you used crowdsourcing or conducted research with human subjects, **without compromising anonymity**...
- (a) Did you include the full text of instructions given to participants and screenshots? [Yes, in the provided URL for the detailed questionnaire.](#)
  - (b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? [Risks were not discussed directly in the paper, as this was a low-risk survey. However, more details about the survey implementation are available at the questionnaire URL and at \[pewresearch.org/the-american-trends-panel/\]\(http://pewresearch.org/the-american-trends-panel/\).](#)
  - (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [Yes, survey compensation was specified in the Methods section.](#)
  - (d) Did you discuss how data is stored, shared, and de-identified? [Yes, in the Methods section.](#)