

Viewpoint Analysis of Autism-Related Comments in Reddit During COVID-19

Narges Azizifard^a, Lidia Pivovarova^b and Eetu Mäkelä^c

University of Helsinki, Finland
fi *fi*

Keywords: Autism, COVID-19, Reddit, Social Media.

Abstract: Major events, such as the COVID-19 health crisis, have different effects on various groups, with vulnerable populations like individuals with Autism Spectrum Disorder (ASD) being especially affected. Social media platforms capture the unique experiences of these individuals and their families, offering a wide range of perspectives and voices. This study utilizes text analysis methods to examine Reddit discussions concerning both autism and COVID-19. Through the analysis of these comments, we identify key themes, including challenges in education, employment, and family life, as well as conspiracy theories and propaganda surrounding vaccination. Our findings shed light on the struggles faced by individuals with autism during the lockdown and highlight their coping strategies. Additionally, the study reveals significant variations in sentiment and emotions across different themes within the comments, providing deeper insights into how these experiences are expressed and understood in online communities.

1 INTRODUCTION

Significant events, like the COVID-19 outbreak, impact various groups differently and can be particularly challenging for vulnerable populations, such as individuals with conditions like Autism Spectrum Disorder (ASD). In the digital age, they have the opportunity to share their unique experiences online through social media. At the same time, a variety of views and opinions on these people is also presented in social media. Thus, social media allows for studying a multidimensional discussion of certain events and conditions, and for revealing communities discussing same themes from different angles. Various language analysis techniques, such as topic modelling, sentiment analysis, etc. help to organize collections of social media posts and comments into subsets that differ in their viewpoint. In addition, in many of social media platforms, discussions are organized in a form of groups and channels, that provide an additional facet of analysis.

In this paper we focus on analysis of ASD-related discussions in Reddit during the COVID outbreak. To investigate the impact of COVID-19 and lockdown

on individuals with ASD or people who are close to them, we find all comments regarding autism in COVID-related subreddits, and all comments from autism-related subreddits linked to COVID time. This research aims to uncover patterns, sentiments, and emotions, providing insights into autism during the COVID-19 lockdown. Overall, our goal is to highlight the challenges that arise in such situations, which often result in shock and confusion.


We utilized a variety of NLP techniques, such as including topic modeling and sentiment analysis to identify recurring themes in the discourse and examine how different signals detected by these tools are distributed across communities and themes.


We discovered significant differences among discussed topics particularly in the emotions and sentiments expressed in the comments. We combine data-driven analysis with close reading and demonstrate that the differences found by automatic tools are indeed reflected in different attitudes and discussion culture variety among topics which we extracted.


To sum up, our objective is to explore the following research questions (RQs):

RQ1: What are the most frequently discussed topics in the discourse? What types of issues they address?

RQ2: Do users exhibit the different emotional reactions to the different extracted topics?

^a  <https://orcid.org/0000-0002-7525-4561>

^b  <https://orcid.org/0000-0002-0026-9902>

^c  <https://orcid.org/0000-0002-8366-8414>

RQ3: Are there any differences in community reaction on the comments within each topic?

The main contributions are the following:

- To the best of our knowledge, this is the first attempt to provide a quantitative analysis of subreddit comments regarding ASD during the COVID-19 lockdown.
- Our analysis encompasses a wide range of subreddit communities without restricting to specific subreddit, community types, genders, or age groups. This ensures that the filtered comments represent a diverse cross-section of the autistic population and minimize bias.
- Our approach offers a novel insight for a more comprehensive perception of unrecorded issues related to ASD, especially during public disasters such as global health crisis and lockdowns. This can lead to prevention, resolution, and control of public health effects.

2 RELATED WORK

The utilization of social media as a medium for disseminating health information has gained considerable attention in recent years. While some families of children with ASD prefer to obtain health information from established, local resources, others turn to social media for support and information. This trend highlights the necessity of understanding how health-related issues are represented in online discussions.

Social media platforms provide users with the opportunity to discuss health-related information and advocacy (Whisenhunt Saar et al., 2023). Reddit, in particular, is popular due to its accessibility and user-generated content but remains relatively unexplored in empirical research compared to larger platforms like Twitter and Facebook (Yeung et al., 2022). Moreover, a recent review indicates that much of the knowledge about Reddit content is derived from computational textual analysis, and it advocates for further contextualization of this data (Proferes et al., 2021).

Reddit is a pseudonymous social media platform, where users create accounts that do not directly reveal their identities but may still include personalized information. For instance, the subreddits that an account follows indicate the user's interests, although this information is not immediately recognizable to individuals unfamiliar with the account holder. Reddit's pseudonymous nature facilitates intimate discussions among users, who do not fear to disclose identifiable public information (Palonis, 2021). An increasing body of research examines Reddit as a forum for

Table 1: Basic statistics of filtered dataset.

Subreddit	# comments	# users
<i>AskReddit</i>	9,622	8,066
<i>conspiracy</i>	3,708	2,145
<i>wallstreetbets</i>	3,088	2,483
<i>DebateVaccines</i>	2,302	485
<i>insanepeoplefacebook</i>	2,083	1758
<i>vaxxhappened</i>	2,069	1,500
<i>memes</i>	2,114	1,880
<i>AmItheAsshole</i>	2,016	1,048
<i>politics</i>	1,884	1,561
<i>autism</i>	1,817	1,273
<i>worldnews</i>	1,755	1,482
<i>news</i>	1,654	1,377
<i>Coronavirus</i>	1,547	1,283
<i>AntiVaxxers</i>	1,292	552
<i>TrueAntiVaccination</i>	1,147	277
<i>insaneparents</i>	1,109	984
<i>facepalm</i>	1,108	982
<i>aspergers</i>	1,032	743
total	41,875	28,270

discussing topics related to ASD. These discussions generally occur within specific subreddits, i.e. dedicated forums that address various aspects of autism (Whisenhunt Saar et al., 2023).

Qualitative methods are also frequently used to analyze Reddit comments related to autism. For example, in (Thom-Jones et al., 2024), the authors employed a qualitative approach to study how individuals discuss and reflect on their experiences of motherhood within a Reddit community dedicated to autistic parents. This method allowed them to explore the nuanced perspectives and personal experiences shared by community members, providing a deeper understanding of the unique challenges and insights associated with parenting as an autistic individual.

Among those empirically investigating Reddit content related to ASD (Bellon-Harn et al., 2022) explored the representation of applied behavior analysis (ABA) interventions within Reddit discussions through topic modeling. Additionally, the study in (Whisenhunt Saar et al., 2023) analyzed ABA discourse on Reddit using inductive manual coding of contextualized thread content.

Other studies investigate how user use internal voting system to shape their communities by promoting content they find useful, relevant, and trustworthy, while downvoting and thus suppressing content they consider irrelevant, false, or self-serving (Leavitt and Robinson, 2017; Richterich, 2014).

3 METHODOLOGY

3.1 Data Collection and Filtering

Data retrieval was conducted using the Reddit Application Programming Interface (API).

All comments and subreddit titles were converted to lowercase. To retain comments related to both COVID-19 and autism, we applied the following filtering criteria:

- We considered Reddit comments from March 2020 to October 2021 to ensure they were written during the COVID-19 lockdown, as documented on Wikipedia¹.
- Next, after searching related phrases on Google and the Reddit website, we filtered comments that contained at least one of the following COVID-related terms: “covid”, “corona”, “c19”, “c-19”, “sars-cov-2”, “sars-cov2”, “sarscov2”, “china_flu”, “china-flu”, “chinaflu”, “wuhan_flu”, “wuhan-flu”, “wuhanflu”, “lock-down”, “quarantine”, “vaxx”, or “vaccine”, and at least one of “asperg” or “autis” for autism-related comments.

To analyze the filtered comments, we excluded non-English posts. For preparing the data for further analysis, we performed data cleaning using Natural Language Processing (NLP) techniques to remove unnecessary information from the collected tweets, including URLs, special characters (e.g., @, /, ...), and stop-words (e.g., in, the, ...). After these steps, a total of 111,507 comments remained for analysis.

We focused on subreddits with more than 1000 comments. Among these subreddits, we excluded the “csci040temp” after determining that it was a temporary subreddit active only during October and November 2020. The final dataset consisted of around 42 thousand comments from 18 subreddits.

3.2 Data Analysis

3.2.1 Explanatory Analysis

In Table 1 we present the basic statistics of our dataset. We also performed an analysis of the temporal distribution of comments within our selected subreddits from the official start of quarantine to its conclusion. As shown in Figure 1, at the start of the quarantine, there was an increase in comments related to our themes in most of the subreddits, although these discussions eventually decreased (except conspiracy

¹https://en.wikipedia.org/wiki/COVID-19_lockdowns

subreddit). The initial surge was largely driven by the fear and uncertainty people felt during the early stages of the lockdown, as the situation was unprecedented and many were unsure how to cope. The significant amount of news, particularly on social media, further contributed to the heightened anxiety. Many users also debated vaccines, the anti-vaxx movement, and the pros and cons of vaccination. This trend reflects a growing need for individuals to seek guidance and information regarding autism-related issues during the lockdown. In the following section, we explore the themes of these discussions to gain a deeper understanding of the issues raised.

3.2.2 Topic Modeling

We utilized Latent Dirichlet Allocation (LDA) to answer the first RQ, i.e. to reveal the most frequently discussed topics during the COVID-19 lockdown, as well as the challenges faced by people with autism and those who support them. We applied LDA model using different number of topics, from 2 to 20 and choose the best model based on coherence metric (Röder et al., 2015)².

3.2.3 Emotion and Sentiment Analysis

To answer RQ2 and uncover differences between topics and the users’ reactions to the issues they face, we performed emotion and sentiment analyses on filtered comments for each topic separately.

For sentiment analysis, we employed VADER³ (Valence Aware Dictionary and sEntiment Reasoner), a lexicon and rule-based tool specifically designed to detect sentiments expressed in social media.

For emotion analysis we utilized Plutchik’s model of emotions (Plutchik, 1980), which identifies eight primary emotions—anger, fear, sadness, disgust, surprise, anticipation, trust, and joy—from user comments. To extract emotions based on Plutchik’s framework, we employed the NRC emotion lexicon (Mohammad and Turney, 2013), which comprises a list of English words associated with each primary emotion, as well as two sentiments: negative and positive. The annotations within the lexicon were manually curated through crowdsourcing. To detect text emotions, we count a number of tokens from each lexicon category in comments and then average them for each topic.

We did not take the negative and positive sentiments from the NRC lexicon into account in this analysis, as this is analysed separately using VADER.

²We used Gensim implementation, <https://radimrehurek.com/gensim/>

³<https://pypi.org/project/vaderSentiment/>

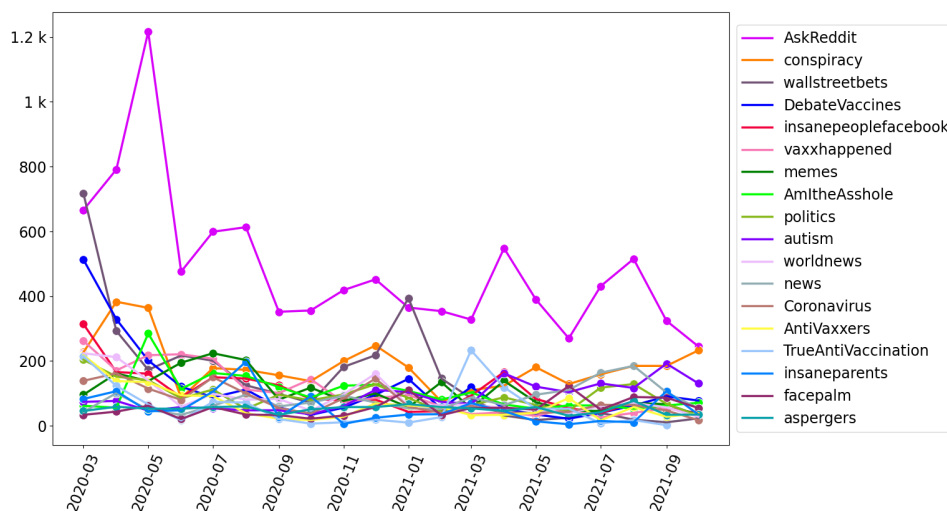


Figure 1: Temporal distribution of comments in subreddits.

Table 2: Categorical variable of vote score.

Category name	Score range
Slightly upvoted	2–10
Moderately upvoted	11–50
Greatly upvoted	51–200
Very greatly upvoted	201–1000
Extremely upvoted	Greater than 1000
Slightly downvoted	-10–0
Moderately downvoted	-50 to -11
Greatly downvoted	-200 to -51
Very greatly downvoted	-1000 to -201
Extremely downvoted	Less than -1000

Table 3: Prominent words for each topic.

Topic 0	Topic 1	Topic 2
vaccine	doctor	people
autism	fact	kid
study	money	work
children	government	school
cause	health	parent
claim	mask	mom
evidence	disease	family
anti_vaxxer	science	live
risk	opinion	job
support	issue	hard

3.3 Vote Scores

To address RQ3, we analyzed the comment scores to identify any changes in user voting behavior across different topics, applying the categories defined in (Davis and Graham, 2021). The Reddit vote score is determined by subtracting the number of downvotes from the number of upvotes. This score is then classified into one of ten categories, reflecting whether a comment is slightly, moderately, greatly, very greatly, or extremely upvoted (or downvoted)(Davis and Graham, 2021) as shown in Table 2.

3.4 Results

3.4.1 Topic Modeling

Table 3 presents a selection of words associated with each identified topic. We further analysed each topic through a close reading of the comments related to each topic, focusing on the dominant keywords asso-

ciated with each topic. The primary discussions centered around the following themes:

Topic 0:

- **Issues regarding vaccination and its potential harms:** This topic accounts for 47 percent of all comments in the dataset. Some users discuss claims and studies suggesting that vaccines may have side effects, such as causing autism in children or increasing the risk of having autistic offspring in the future. Consequently, some individuals are concerned about why others take the perceived risk of using vaccines. Additionally, it is argued that vaccines can lead to death and certain diseases. Therefore, discussions on this topic tend to either support vaccinations or align with the views of anti-vaxxers.

Topic 1:

- **Conspiracy theories and propaganda regarding vaccination and wearing mask:** Topic 1 comprises 31 percent of all comments in the

dataset; it also discusses whether to trust vaccinations, particularly in the context of conspiracy theory, propaganda, and governmental perspectives. Many users argue that medical research and science affirm the safety of vaccines, while those who deny science and promote conspiracy theories, anti-science propaganda, and anti-government narratives claim that vaccines cause autism, death, and disease. Numerous comments ridiculed these claims by comparing the statement "vaccines cause autism" to assertions like "5G technology is deadly", "the Earth is flat", and "COVID is a hoax". Additionally, there was discussion about the influence of authorities in shaping political opinions. Some users argued that governments, pharmaceutical companies and doctors have plans to profit from vaccines by promoting their brands as safe and not causing autism or disease. However, other users disagreed, expressing appreciation for the efforts and hard work of healthcare workers. Similar discussions arose around wearing masks, with some users criticizing anti-maskers who claimed that mask mandates were driven by financial motives and intended to prevent a market crash due to COVID-19.

Topic 2: This topic makes up 22 percent of the comments in the dataset, and discuss several themes related to effects that the crisis causes in everyday life.

- **Parental issues related to COVID-19 lockdown:** This topic explores how the COVID-19 has affected family dynamics, particularly focusing on how parents, especially single mothers, have struggled with managing autistic children exhibiting aggressive behavior due to quarantine measures and the loss of recreational activities, and challenges to cover expenses while struggling to meet the needs of their autistic children. Many autistic individuals or parents of autistic children have been unable to speak with therapists or access other services due to COVID-19, adversely affecting their mental health.
- **Issues related to incorrect behavior of parents:** Numerous autistic users voiced concerns about how their parents blamed and punished them for their condition. Due to lockdown restrictions and social distancing rules, they could not access affordable shared living spaces and were forced to live with their aggressive parents. Many parents also pressured their autistic children not to receive vaccines, fearing it would worsen their condition.
- **Issues related to work:** Some users discussed losing their jobs and the challenges of employment as autistic individuals, given the high un-

employment rate within the autistic community. Single mothers, in particular, face significant challenges as they must work more to cover expenses while struggling to meet the needs of their autistic children.

- **Issues related to wearing mask:** The final issue raised in the comments was the difficulty autistic children face in wearing masks and adhering to other rules, such as social distancing at school. This highlights the sensory challenges associated with mask-wearing and physical contact.
- **Positive but still negative point:** Some autistic users expressed relief from the isolation, noting that not attending school and taking exams remotely reduced their stress. They also highlighted that being away from environments where they are often mocked, especially by anti-vaxxer groups, provided a sense of peace.

To enhance our analysis and determine the specific subreddits where each topic is most prevalent, we identified the subreddits associated with each topic and examined the distribution of comments pertaining to each topic. As illustrated in Figure 2, the largest concentration of Topic 0, which includes discussions that either support vaccinations or align with anti-vaccination sentiments, is found in the subreddits "DebateVaccines" and "TrueAntiVaccination". This is followed by the "AntiVaxxers" subreddit, which also primarily features debates centered around vaccines. The content in these subreddits reflects a diverse range of opinions, from those advocating for vaccination to those opposing it, highlighting the polarized nature of the vaccine debate.

Topic 1, which revolves around vaccine and mask debates related to conspiracy theories and propaganda, is most prevalent in the "wallstreetbets" subreddit, with significant discussions also occurring in the "politics" and "worldnews" subreddits. As mentioned earlier, many comments within this topic suggest that vaccines and masks are perceived by some as tools for financial exploitation and control, reflecting a broader skepticism towards public health measures and governmental motives.

Finally, Topic 2 primarily focuses on individuals coping with autism, caring for autistic family members, addressing parental challenges, dealing with aggressive parenting, facing difficulties with mask-wearing, navigating workplace problems, and managing mental health issues. This topic is most prominent in the "AmItheAsshole" subreddit, which aims to provide advice based on personal stories and conflicts shared by users. Additionally, this topic is highly relevant in the "aspergers" and "autism" subreddits, which are dedicated to discussions about autism and

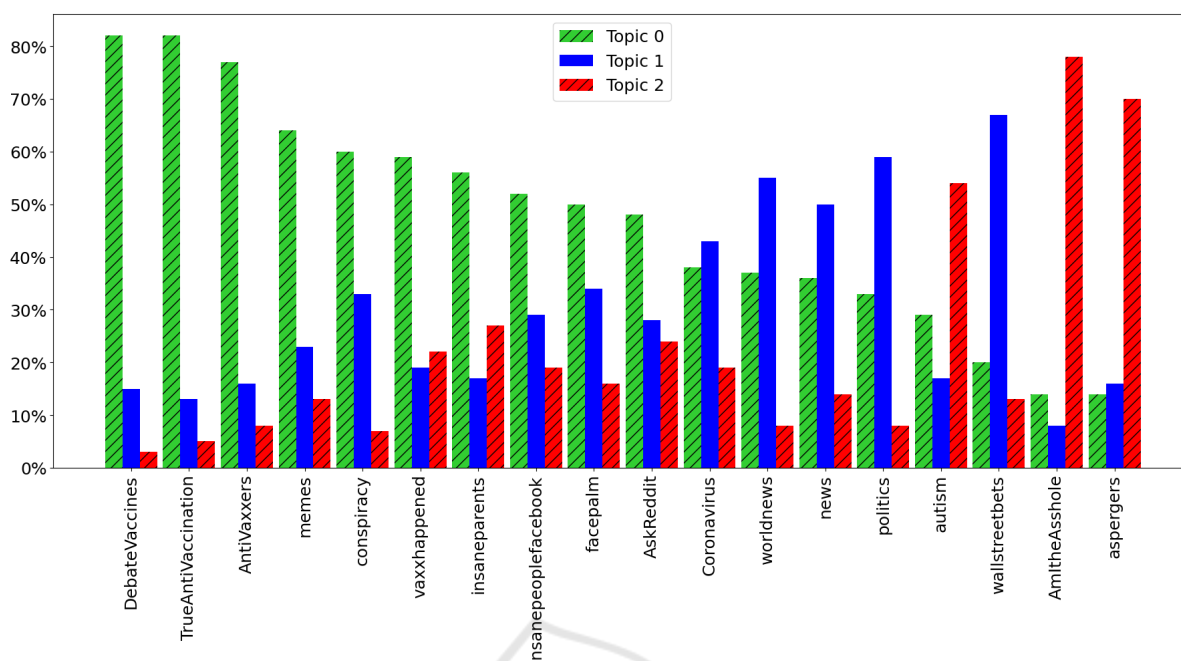


Figure 2: Distribution of topics per each subreddit, sorted by the prominence of Topic 0.

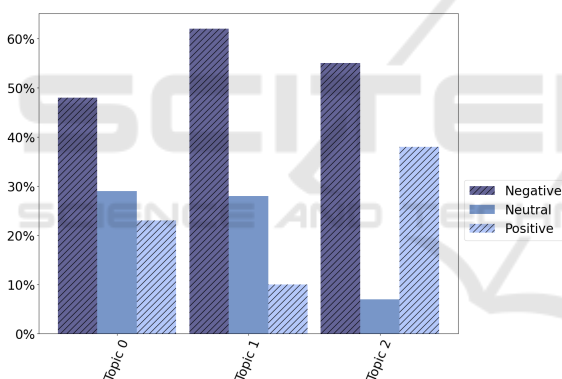


Figure 3: Sentiment distribution across topics.

the experiences of autistic individuals; these subreddits serve as support communities for those seeking guidance, understanding and shared experiences.

3.4.2 Sentiment Analysis

Figure 3 displays the results of the sentiment analysis, in relation to each of the topics. From the data presented, it is evident that the comments associated with Topic 1 tend to be the most negative and the least positive overall. As mentioned in the previous section, users engaged in this topic often discuss how powerful entities manipulate false and misleading claims, crafting conspiracy theories and propaganda to influence public opinion on vaccination and mask-wearing. These discussions suggest that these actors are motivated by profit, exploiting the situation

for financial gain while disregarding the well-being and health of the public.

Conversely, Topic 2 contains the highest number of positive comments and the fewest neutral ones. As we previously discussed, this topic primarily centers on the challenges faced by autistic individuals and their families during the lockdown, as well as the coping strategies they employed. These discussions tend to be more optimistic and supportive, reflecting a positive outlook rather than a neutral stance.

The discussions in Topic 0 remain relatively neutral, not leaning strongly toward one side or the other. They generally revolve around the debate on whether or not to trust vaccines, with participants expressing either anti-vaccine sentiments or support for vaccination. As a result, this topic showcases a range of user perspectives, reflecting the diverse stances within the conversation.

3.4.3 Emotion Analysis

The results of the emotion analysis are shown in Figure 4. To visualize the emotional fingerprint of the corpora and the text annotated with Plutchik’s model, we utilized "PyPlutchik" (Semeraro et al., 2021), a specialized tool designed for this purpose.

Once again, the comments associated with Topic 2 exhibit the highest levels of emotional expression, followed by those in Topics 1 and 0. This heightened emotional intensity in Topic 2 can be attributed to the nature of the comments, where individuals share per-

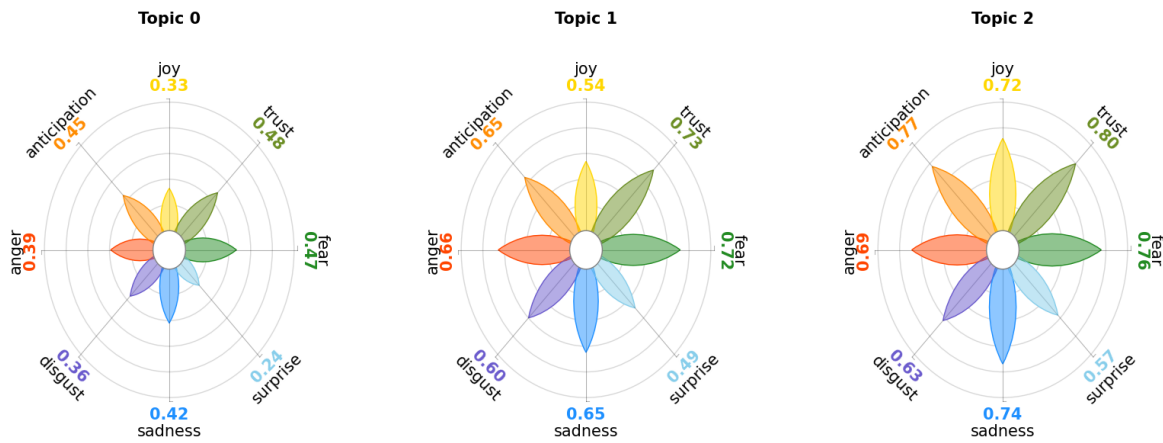


Figure 4: Plutchik's emotions of Topics.

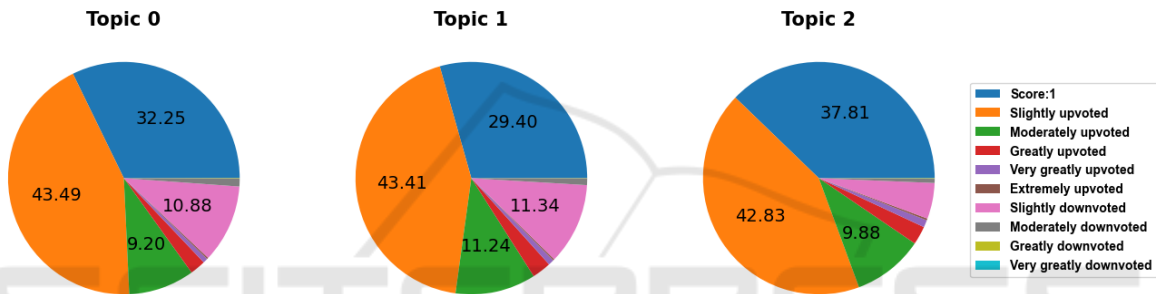


Figure 5: Distribution of score categories among topics.

sonal experiences and seek support in coping with their challenges, leading to a richer expression of emotions compared to the other topics.

At the same time, it is hard to find any single emotion is more prominent in any of the topics. All three contain the whole spectrum of emotions, although with different intensity.

3.4.4 Vote Scores of Comments

We created vote score categories for each comment in our dataset and grouped the comments based on their respective topics. The results are shown in Figure 5, where we show a proportion for each vote score category for each topic. We did not include percentage labels for categories where the percentage was below 5%. A large number of comments had a score of 1, which do not fall into any specific category and neither supports nor opposes the content; therefore, we assigned these comments to a separate category labeled "score 1".

As it was mentioned earlier, the scores reflect other users' reactions to a comment rather than the sentiment of the comment author. After reviewing the scores, we discovered that there were no comments in the "Extremely downvoted" category across

any of the topics, and only one comment was in the "Very Greatly downvoted" category. Additionally, as shown in Figure 5, the percentage of comments in the "Slightly upvoted" category is higher than in other categories, indicating that more comments are promoted rather than suppressed. Regarding the "Moderately upvoted," "Greatly upvoted," "Very greatly upvoted," and "Extremely upvoted" categories, there are not significant differences across the three topics.

Overall, our analysis revealed that the percentage of downvoted categories in Topic 2, which primarily focused on issues faced by autistic individuals and those who interact with them, was lower compared to the other topics. This lower rate of downvotes indicates that users perceived these comments as more valuable, insightful, and relevant. The topics discussed in this category were likely seen as meaningful and resonated more strongly with the community, suggesting a higher level of engagement and agreement with the content.

4 CONCLUSION

Online communication is important for many people, and it can be a part of coping strategies during crisis. At the same time, it is important to discriminate between different communities, and to be able to point people to right places on the WEB where they can get support. Viewpoint analysis may potentially help in that regard.

In this paper we presented a first attempt to apply data-driven analysis on Reddit comments devoted to autism during COVID-19 outbreak. While the previous work studying autism in Reddit were focused on some pre-selected subreddits and employs mostly qualitative analysis, we combined quantitative analysis with close reading. We have shown that language analysis techniques, such as topic modelling and sentiment analysis, can reveal viewpoint differences in reddit communities.

We observed various themes in the comments, such as vaccination and disinformation, family issues, conspiracy theories, and propaganda. Comments that focus on the challenges faced by autistic individuals and their families during quarantine, as well as their coping strategies, tend to be more emotional and less neutral. These comments also display more positive sentiment and are generally more optimistic and supportive. Additionally, they tend to receive fewer downvotes compared to other types of comments.

Based on the discussions related to various topics, sentiments, emotions, and vote scores, we can observe the complex and multi-dimensional nature of online conversations. These discussions reveal how different themes and emotional tones interact with user engagement, highlighting the diverse perspectives and reactions within online communities. This complexity reflects the broader structure of online discourse, where multiple factors contribute to shaping the flow and tone of the conversation, making it both dynamic and multifaceted. In general, the analysis presented in this study offers valuable insights into the challenges faced by autistic individuals and their families during the outbreak. These findings can aid communities, particularly children, parents, and individuals with disorders, in better managing autism. Furthermore, the study provides critical information for government and authorities to enhance their policy-making processes regarding support and services for communities with disorders, including ASD, during critical situations such as global health crisis.

REFERENCES

- Bellon-Harn, M., Manchaiah, V., and Boyd, R. (2022). Applied behavior analysis as treatment for autism spectrum disorders: Topic modeling and linguistic analysis of reddit posts. *Frontiers in Rehabilitation Sciences*, 2:1–10.
- Davis, J. and Graham, T. (2021). Emotional consequences and attention rewards: the social effects of ratings on reddit. *Information, Communication & Society*, 24:1–18.
- Leavitt, A. and Robinson, J. (2017). Upvote my news: The practices of peer information aggregation for breaking news on reddit.com. *Proceedings of the ACM on Human-Computer Interaction*, 1:1–18.
- Mohammad, S. M. and Turney, P. D. (2013). Crowdsourcing a word–emotion association lexicon. *Computational intelligence*, 29(3):436–465.
- Palonis, B. (2021). *Understanding the Themes of Disability Discourse Through Reddit Comments*. Rochester Institute of Technology.
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Theories of emotion*, 1:3–31.
- Proferes, N., Jones, N., Gilbert, S., Fiesler, C., and Zimmer, M. (2021). Studying reddit: A systematic overview of disciplines, approaches, methods, and ethics. *Social Media+ Society*, 7(2):20563051211019004.
- Richterich, A. (2014). ‘karma, precious karma!’karmawhoring on reddit and the front page’s econometrisation. *Journal of Peer Production*, 4.
- Röder, M., Both, A., and Hinneburg, A. (2015). Exploring the space of topic coherence measures. In *Proceedings of the eighth ACM international conference on Web search and data mining*, pages 399–408.
- Semeraro, A., Vilella, S., and Ruffo, G. (2021). Pyplutchik: Visualising and comparing emotion-annotated corpora. *PLOS ONE*, 16(9):e0256503.
- Thom-Jones, S., Melgaard, I., and Gordon, C. (2024). Autistic women’s experience of motherhood: A qualitative analysis of reddit. *Journal of Autism and Developmental Disorders*, pages 1–11.
- Whisenhunt Saar, K., Saar, M., Bellon-Harn, M., Harn, W., Boyd, R., and Manchaiah, V. (2023). Upvote or downvote aba for autism? content and support in reddit posts. *Journal of Consumer Health on the Internet*, 27(2):119–138.
- Yeung, A. W. K., Tosevska, A., Klager, E., Eibensteiner, F., Tsagkaris, C., Parvanov, E. D., Nawaz, F. A., Völkl-Kernstock, S., Schaden, E., Kletecka-Pulker, M., et al. (2022). Medical and health-related misinformation on social media: bibliometric study of the scientific literature. *Journal of medical Internet research*, 24(1):e28152.