Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Combining Natural Language Processing with Patient-Reported Outcome Measures scores to investigate the impact of pandemic regulations on anxiety in children

Michiel A. J. Luijten
m.luijten@amsterdamumc.nl

Amsterdam UMC Location University of Amsterdam, Emma Children's Hospital

Chris Gibbons
Oracle (United States)

Conrad J. Harrison
University of Oxford

Hedy A. Oers
Amsterdam UMC Location University of Amsterdam, Emma Children's Hospital

Josjan Zijlmans
Amsterdam UMC Location University of Amsterdam, Emma Children's Hospital

Jacintha M. Tieskens
Leiden University Medical Center

Hekmat Alrouh
Vrije Universiteit Amsterdam

Emma M. Broek
Leiden University Medical Center

Janna N. Boer
University of Groningen, University Medical Center Groningen (UMCG)

Lotte Haverman
Amsterdam UMC Location University of Amsterdam, Emma Children's Hospital

Tinca J. C. Polderman
Amsterdam UMC Location University of Amsterdam, Emma Children's Hospital

Article

Keywords:

Posted Date: November 13th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-7869448/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on October 8th, 2026. See the published version at https://doi.org/10.1038/s41598-026-73632-2.

Abstract

Background

The COVID-19 pandemic and the governmental regulations taken during the pandemic had a profound impact on the experienced anxiety of children. However, there is a gap in research that explains how and which regulations had an impact. In this study we combine quantitative and qualitative data through structural topic modelling (STM) and sentimental analysis.

Methods

Children and adolescents (N = 3,821, aged 8–18 years old) completed an open-ended question; "*How are the corona-regulations for you?*", and the PROMIS Anxiety questionnaire, bi-annually from April 2020 until March 2023 (7 measurement occasions). This yielded 6,672 open-ended responses, with anxiety classified as (sub)clinical for PROMIS T-scores > 50.6. We applied STM to the open-ended responses to identify relevant topics, using dichotomized Anxiety as covariate. We used sentiment analysis to assess the affective state of responses and obtain polarity scores for each response (ranging from − 1 (negative) to + 1 (positive)). We identified the topics of importance to children during the pandemic and assessed the affective state of these responses per topic in the whole sample and split by presence of (sub)clinical anxiety.

Results

Eight topics emerged: 1) Adaptation/Resilience, 2) Distress, 3) Adherence to social distancing, 4) Limited social and family contact 5) Restrictions in activities/boredom, 6) Future perspectives, 7) Homebound (school closures), and 8) Lack of celebratory events. Compared to their peers, children with (sub)clinical anxiety reported significantly more negative sentiment, particularly regarding adaptation/resilience (-0.19 vs. -0.10, Cohen's D = 0.24) and distress due to lockdown (-0.06 vs. 0.02, Cohen's D = 0.20). They also expressed stronger emotional language when discussing these topics.

Conclusions

We successfully identified topics relating to the governmental regulations that are associated with (sub)clinical anxiety during the COVID-19 pandemic. This study shows that children with (sub)clinical anxiety scores experienced the pandemic period more negatively, and may have more problems with coping and adapting to lockdown measures.

Introduction

The COVID-19 pandemic has had a large impact on the life of children throughout the world. Fear of infection, uncertainties about the future, and social isolation affected the social and mental health of children. Previous studies have shown that children reported increased internalizing problems, such as anxiety and depressive symptoms throughout the pandemic [1–3]. A recent systematic review and meta-analysis summarized findings of a multitude of studies and showed that the majority of studies reported deterioration in mental health during the pandemic. Moreover, during the pandemic a significant number of children reported (sub)clinically relevant anxiety and depressive symptoms that affected their daily life significantly [1].

While many studies have investigated changes in mental health during the pandemic [4–6], only few studies provide insight into how or why the pandemic has had such a large impact on the mental health of children and adolescents. It is hypothesized that governmental regulations during the pandemic had an impact on the psychosocial development of children and adolescents, due to limited social contact during the pandemic. This may have hindered forming relationships with peers, which are of vital importance to develop social skills and a social network at this age [7]. The presence of the virus, the uncertainty of its development and the effects of the virus itself may also have played a role in developing anxiety and depressive symptomatology [7]. While there were no specific explanations given for mental health problems, several factors have been identified that were related to worsened mental health in children during the pandemic in quantitative studies. These are not just limited to characteristics of the child (age, gender) or regulations aimed at the child (school closures, limited social contact)[2,3], but also changes within the family dynamic due to regulations, such as the loss of income and work of the parent(s) [8].

Qualitative studies can provide more background to the observed mental health deterioration of children during the pandemic. By asking children what they experienced during the pandemic in interviews, focus groups, diaries or open-ended surveys, and subsequently performing thematic analyses, we can gain insights into what themes were most important for children during the pandemic. The few qualitative studies so far [9] reported that the majority of children experienced a negative impact on their life related to themes such as school, leisure time and family/social bonding. Although valuable, generalization of results of this type of study is often difficult as sample sizes are usually small and selective, resulting in possible selection bias. In addition, the relation of the qualitative results (e.g., themes), with specific mental health outcomes, such as anxiety and depression, remains unclear.

Open-ended survey questions that are collected in large samples can solve these problems. However, the analysis of such large quantities of open text is burdensome and identifying clear topics or themes can be challenging. Topic modeling is a text-mining technique that allows for assessing a thematic structure within a large collection of unstructured texts. It has become increasingly popular to efficiently analyze large amounts of qualitative data, such as notes from electronic health records [10] or (sub)clinical social work notes [11]. Structural topic modeling (STM) is a recently popularized topic model for large scale open-ended survey data, that can outperform other topic models. In addition, it allows for the combination of qualitative data with quantitative data [12–14].

In this study we employed STM to combine qualitative data from an open-ended survey question on the impact of governmental COVID-19 regulations with quantitative scores [15] of anxiety in a large sample of children and adolescents from the general population. In addition, we applied a lexicon-based approach called sentiment analysis to assess the 'affective state' of the open-ended responses that shows whether a topic was negatively or positively experienced.

By combining these methods, we specifically aimed to: 1) identify important topics related to the pandemic (measures) as experienced by children and adolescents, 2) investigate the affective states of these topics 3) examine how the topics and the affective states differed between children with and without self-reported (sub)clinical anxiety.

Methods

Participants and Measures

Children and adolescents (8–18 years old) were recruited through an online panel agency to attain a representative sample of the Dutch general population (for more details see van Oers et al. [16]). Children completed online questionnaires including the open-ended question and the Patient-Reported Outcomes Measurement Information System (PROMIS) Anxiety computerized adaptive test (CAT). We received approval for data collection from the appropriate ethics committees and all children and their parents provided informed consent. The study was conducted in accordance with the ethical standards outlined in the 1964 Declaration of Helsinki and its later amendments.

Open-ended question

With a biannually assessed open-ended question during and after the COVID-19 pandemic (April 2020 – April 2023, seven times in total) we asked children how the pandemic regulations affected their life. The question was formulated as: "*How are the corona-regulations for you?*". For post-pandemic measures (November/December 2022/March/April 2023) we asked how they previously experienced the pandemic regulations: "*How were the corona-regulations for you?"*.

PROMIS Anxiety

The PROMIS pediatric Anxiety v2.0 item bank contains 15 items that use a 7-day recall period and are scored on a five-point Likert scale; 1 (never), 2 (almost never), 3 (sometimes), 4 (often), 5 (almost always). The instrument has been recommended in the DSM-V-TR [17] as an online cross-cutting assessment of pediatric anxiety. The item bank results in a single T-score for each participant. The item bank has been previously validated for assessing anxiety in the Netherlands [15] and Dutch cut-offs have been previously established, where T-scores > 50.6 were defined as scores within the (sub)clinical range (> 1SD)) of anxiety [15]. Based on this cut-off we divided the responses in a 'no (sub)clinical anxiety' and '(sub)clinical anxiety' group.

Analyses

Open-ended data and text-preprocessing

In total there were 7605 open-ended responses from 3821 unique respondents. Prior to processing, the text was spellchecked manually, as children tend to misspell words in a manner that cannot be corrected automatically. As exclusion criteria we removed open-ended answers below a minimum sentence length of three words (e.g., "Stupid."). For preprocessing we used stemming from the stm-package in R [13] and a list of Dutch stop words from the Snowball stemmer [18]. Numbers and punctuations were removed and the minimum word length for words to be included in the analyses was set to at least three letters. Furthermore, topics were only included if they occurred at least ten times. For the majority of the responses, accompanying PROMIS Anxiety scores were available (N = 6672, 87.7%), which could thus be used as a covariate for the structural topic model.

Structural Topic Model

An important aspect of developing a suitable topic model is defining and testing hyper-parameters. The number of topics to be estimated is vital in this regard. By testing topic models with a varying number of topics, the optimal number of topics is identified. To assess the effectiveness of the chosen *K* number of topics, the multinomial dispersion of residuals of the model and held-out-likelihood were assessed, and semantic coherence and similarity/exclusivity were taken into consideration. First, higher residuals indicate that a larger number of topics are required, therefore lower residuals are preferred [19]. Second, held-out-likelihood is an estimation of the probability of topics occurring in data that was not used (set to 25%) to build the model (i.e., 25% of the responses are withheld from the initial development of the model to check whether similar topics are found in the new data). A higher held-out likelihood tells us that the model predicts topics in unused data better and is therefore more suitable. Third, semantic coherence and exclusivity tend to decrease when allowing more topics. As structural topic models with covariates do not allow for computation of coherence and exclusivity, we focused on the residuals and held-out likelihood and selected the number of topics where the residuals no longer showed signs of improvement, to maintain a parsimonious and interpretable model.

To find the optimal number of topics a range from 3 to 12 topics were tested, based on the aforementioned criteria. The dichotomous variable for anxiety was used as covariate in all analyses to investigate which terms (i.e. words) within a topic were primarily used by children with (sub)clinical anxiety scores and how the choice of terms to describe the same topic differed between children with and without (sub)clinical anxiety. As topics may be dominated by a

specific subgroup, the percentage of children with (sub)clinical anxiety occurring in the 200 most representative responses was calculated. A response in this case is one open-ended survey answer. Analyses were performed in R [20], using the STM package [13].

Once topics are found they need to be interpreted. To reduce the subjectiveness of the topic descriptions, seven professionals (psychologists, methodologists, and linguists) were asked to describe the topic in a short sentence based on a list of the terms associated with the topic and the 200 responses most associated with that specific topic. We defined sufficient consensus as over 50% of experts reaching a similar gist on topic content and two authors (TJP & MAJL) summarized this into one topic description.

Sentiment analyses

Once the topics were defined, we assessed the affective states of responses in order to investigate whether topics were discussed positively or negatively. With sentiment analysis, we calculated the polarity score of responses using a lexicon-based approach based on the pattern.nl module in Python [21]. Polarity scores of responses range from − 1 (most negative response) to + 1 (most positive response). First, we calculated polarity scores for all responses and then split the responses and accompanying scores by the anxiety covariate. Next, we assessed the previously calculated polarity scores for each topic separately (based on the top 200 responses), also split by anxiety covariate. We compared the polarity scores of children with (sub)clinical anxiety with the scores of children without (sub)clinical anxiety by independent T-tests, where we considered a Cohen's $D \geq 0.2$ as a small effect. We assigned a positive (> 0.01), negative (<-0.01) or neutral (-0.01 to 0.01) label to summarize the polarity of a response. These labels were used to calculate the percentage of positive/negative responses of all data and per topic (of the top 200 responses), also split by covariate. The amount of positive or negative responses provides information on how a topic is experienced.

Results

Structural Topic Model

Descriptives of responses

The 7,605 responses had an average number of words of 7.1 (range 1-395). After text preprocessing, 3,401 responses remained that contained three or more words and had the accompanying dichotomized anxiety scores.

Number of topics

The number of topics were selected based on the residuals and held-out-likelihood (see Fig. 1). Residuals did not show any improvement after eight topics and also the held-out-likelihood did not show any major improvement, therefore we selected eight topics as the optimal number of topics. Semantic coherence and exclusion metrics tend to decrease at a larger number of topics, which was also taken into consideration when selecting eight topics.

Final Topic Model

The final terms belonging to each topic can be seen in Table 1. We specifically selected the FREX terms (FREX; frequent and exclusive) to identify and interpret topics from terms, as they offer the most suitable (exclusive) descriptors of the emerging themes. Consensus was reached for all topics and the terms were summarized into the following topics; 1) Adaptation/Resilience, 2) Distress due to lockdown, 3) Adherence to social distancing measures, 4) Limited social and family contact (bonding), 5) Restrictions in activities/boredom, 6) Future perspectives, 7) Homebound (school closures) and 8) Lack of celebratory events. To illustrate the content of different topics two quotes were selected from the top twenty most representative responses, one for children with (sub)clinical anxiety and one for children without (sub)clinical anxiety.

Table 1
Final topics, terms, associated quotes of the structural topic model and the associated polarity scores based on sentiment analysis.

Topics	Terms	Quotes[1]	% Anxiety[2]	Polarity scores[3]				% Positive (polarity > 0.01)			% Negative (polarity < -0.01)		
				A	B	Effect size D	Total	A	B	Total	A	B	To
All	All	All	41.1%	-0.12	-0.06	-0,14	-0.09	32%	28%	30%	51%	48%	49
1: Adaptation/Resilience	with, sometimes, well, difficulty, especially, difficult, mother, falls, problems, problem, heavy, lessons, tricky, understand, learn	1A. "Sometimes it's difficult. Sometimes suddenly a different teacher. But otherwise, it's okay."	39.5%	-0.19	-0.10	-0,24	-0.14	33%	30%	31%	59%	49%	53
		1B. "It's okay now, there are actually no more rules. Home schooling wasn't fun, I did miss my friends. But I didn't have any further problems with it."											
2: Distress due to lockdown	fortunately, annoying, sits, home, over, notice, easy, follow, hope, afraid, to, makes, tricky, full, goes	2A. "Very difficult, I'm still afraid of a new lockdown. And I don't want to become as gloomy and sad as during the last lockdown when I really couldn't cope at times. Luckily, it's much better now. But I'm disappointed that many things, like school camp and the Christmas celebration in the church with school, can't go ahead. And I'm annoyed that not all my friends and classmates get vaccinated."	40.0%	-0.06	0.02	-0,20	-0.02	37%	44%	42%	44%	40%	42
		2B. "Home schooling but started an internship two weeks ago. It was difficult to find a place, but now I can thankfully go to my internship and I'm no longer stuck at home all the time."											
3: Adherence to social distancing measures	stupid, rule, mask, people, distance, keep, understand, other, wear, hands, meter, wash, come, sick, inside	3A. "Keeping distance, the idea that some people aren't being careful and don't follow the measures or don't listen."	40.5%	-0.10	-0.09	-0,03	-0.10	25%	19%	22%	49%	48%	49
		3B. "It's an adjustment, but at some point, it becomes routine, like wearing a mask in public transport, keeping 1.5 meters distance, disinfecting/washing hands."											
4: Limited social and family contact (bonding)	meeting, allowed, soccer, sport, grandma, miss, grandpa, see, see,	4A. "I miss my social contacts… I even miss not being able to go to school. No more sports, not allowed to visit my grandparents. Little distraction. Not	39.0%	-0.13	-0.15	0,07	-0.15	24%	20%	22%	47%	51%	50

Topics	Terms	Quotes[1]	% Anxiety[2]	Polarity scores[3]				% Positive (polarity > 0.01)			% Negative (polarity < -0.01)		
				A	B	Effect size D	Total	A	B	Total	A	B	To
	play, family, can, alone, to, not anywhere	seeing my boyfriend. I feel quite sad about it."											
		4B. "Annoying. I have fewer football trainings and matches. Many free periods at school due to cancellations. And I see fewer friends because we're not allowed to be in large groups anymore."											
5: Restrictions in activities/boredom	burden, very, find, fine, had, (not) very, found, clear, bit, completely, really, so much, almost, corona, mom	5A. "I miss school and my teachers and classmates very much. I'm afraid to go outside and am worried that my mom, dad, grandma, and grandpa might get sick. I cried a lot during the Passion because of everything related to corona. My mom and dad were very sweet and comforted me."	43.5%	-0.08	-0.03	-0,13	-0.05	41%	44%	43%	55%	52%	54
		5B. "I don't think I was really affected by it much. Except for the lack of parties and new encounters. But I could continue doing almost everything I used to do."											
6: Future perspectives	things, (not) happy, (not) fun, just, normal, all, gone, (not) boring, again, nothing, long, necessary, lasts, stay, strict	6A. "I am fed up and want my normal life back. Just working out, just going to school, just being able to hang out with my friends and do fun things outside."	36.5%	0.027	0.049	-0,07	0.041	41%	48%	46%	51%	45%	47
		6B. "I don't like any of this, it's been going on for so long. But everything else is fine, hopefully, everything will be more fun again soon."											
7: Homebound (school closures)	friend, school, contact, (not) nice, may, often, teacher, gladly, exam, work, hug, homework, bored, best, lesson	7A. "I really want to party with all my friends, but that's not possible now. I want to go out, I would have liked to do my final exams, and I would love to go on vacation, but unfortunately, that's probably not going to happen."	50.0%	-0.13	-0.11	-0,06	-0.12	27%	27%	27%	47%	53%	50
		7B. "I can't go to school. And I can't hang out with friends. And I have to do schoolwork at home. I can't hug my grandparents."											

Topics	Terms	Quotes[1]	% Anxiety[2]	Polarity scores[3]				% Positive (polarity > 0.01)			% Negative (polarity < -0.01)		
				A	B	Effect size D	Total	A	B	Total	A	B	To
8: Lack of celebratory events	go, little, outside, together, life, closed, time, pity, such as, less, birthday, going out, vacation, confusing, celebrate	8A. "It's a shame I couldn't have a party for my birthday, but I do cough into my elbow."	44.0%	-0.16	-0.09	-0,18	-0.12	26%	32%	30%	54%	48%	51
		8B. "Annoying, because we couldn't go on vacation and missed out on camp. And I couldn't celebrate my birthday in a big way."											

[1] = Each topic contains two quotes – one of **children with (sub)clinical anxiety (A)** and one for **children without (B)**. [2] = Percentage of children with (sub)clinical anxiety in the top 200 most representative documents.

The majority of the topics focused on specific aspects brought forward by pandemic regulations, and were related to social aspects and social development (e.g., participating in extracurricular activities, visiting family, playing with friends) and school (participation/performance). The topic “homebound”, which has a focus on school performance, has a higher percentage of children with (sub)clinical anxiety (50.0%) than what would be expected based on the percentage of (sub)clinical anxiety in the entire sample (41.1%). This indicates that children with (sub)clinical anxiety more often reported issues from being homebound. Investigating the documents belonging to this topic shows that children missed their friends, schoolmates and their sports club, as well as having difficulties with school closure, homeschooling and completing homework at home. Regarding online schooling, children mentioned both positive and negative aspects. For some, online school was calming, for others it was distracting and demotivating. Generally, all topics shared positives and negative terms.

Children with (sub)clinical anxiety more often used terms such as “difficult”, “hard”, whereas children without (sub)clinical anxiety more often used words such as “stupid”, “annoying” and “too bad/disappointing”. Children with (sub)clinical anxiety also more often mentioned issues surrounding school (“school”, “lessons”) and the regulations (“mask”, “hands”) in relationship to the difficulties they were experiencing regarding the governmental regulations. Specifically in the topic on social distancing measures, children with (sub)clinical anxiety more often mentioned specific measures (“washing”, “distance”, “facemask”) than children without (sub)clinical anxiety, who more often just mentioned “rules”.

Sentiment analysis

The sentiment analysis shows that overall children mainly had negative responses (polarity = -0.09, 49% of responses negative) to the pandemic regulations (see Fig. 2). All topics, except Future Perspectives, were experienced primarily negatively (polarity range − 0.09 to -0.15) with higher proportions of negative than positive responses (range positive; 27–46%, range negative; 42–54%). Overall, children with (sub)clinical anxiety scores were more negative in general (-0.12 vs -0.06, Cohen’s *D* = -0.14) on all topics (in terms of polarity), however the proportion of positive and negative responses differed between topics. Differences between children with (sub)clinical anxiety and children without (sub)clinical anxiety with a small effect size ($D \geq 0.20$) were found on Adaptation/Resilience (-0.19 vs -0.10 polarity; Cohen’s *D* = -0.24, 59% vs. 49% negative responses) and Distress due to lockdown (-0.06 vs 0.02 polarity; Cohen’s *D* = -0.20). For the latter the direction of the average polarity scores (i.e., positive or negative) differed between groups.

Discussion

In this study we combined qualitative open-ended survey data with quantitative outcomes to assess differences in how children with (sub)clinical anxiety experienced the COVID-19 pandemic compared to children who did not report (sub)clinical anxiety scores. Overall, for all children, social aspects and social development were most impacted by the pandemic regulations, as shown by how often these were mentioned as topics and how negatively they were experienced. Children were particularly affected by social distancing measures and school closures, which may have resulted in social isolation and the accompanying emotional distress.

In the topics that emerged, children with (sub)clinical anxiety more often mentioned negative aspects of the pandemic. Children with (sub)clinical anxiety also reported more (specific) problems with the pandemic regulations, such as washing hands, wearing masks can keeping 1.5m distance and more often used heftier (sentimental) terms. This was also visible in the sentiment analysis, as polarity scores for children with (sub)clinical anxiety were more negative than responses from children without (sub)clinical anxiety. All topics were primarily described with negative affect, except the topic future perspectives, which was described more positively (e.g., “hopefully it will be better soon”). The largest differences between the two groups were found on the topics of resilience/adaptation and distress due to lockdown, where children with (sub)clinical anxiety were more negative. This may indicate that children with (sub)clinical anxiety had more difficulties with adaptation to the lockdown. These children also discussed future perspectives less often and the association was more negative than for children without (sub)clinical anxiety. This could be indicative of difficulties with coping, which has been previously linked to mental health problems during the pandemic [22,23]. Specifically, studies suggest that avoidance coping strategies resulted in higher anxiety and depressive scores during the pandemic [23]. A systematic review reported the negative effects of avoidance coping strategies on mental health, and specifically internalizing problems, of children and adolescents [24], which aligns with the topics and terms reported in our study.

The themes that were found also align with a previous thematic analysis that our group performed on data collected in April 2020 on the initial lockdown period [25]. All themes we identified during the initial lockdown are included in the topics from this study. Two of the current topics were not previously mentioned by children: adaptation/resilience and future perspectives. This might be explained by the initial lack of knowledge about the virus and which adaptations were needed, and the fact that these topics mainly apply to long-term situations while during our previous study it was unclear how long the pandemic would last.

In the past, most STM papers have focused on using only qualitative data from open-ended surveys or open text fields [12–14]. By using STM with the inclusion of a quantitative anxiety measure, we successfully identified themes that provide context for (sub)clinical anxiety scores of children during the pandemic. It is important however, that the open-ended question posed relates to the measured domain of the PROM directly. If the quantitative measure included in the topic model is not explanatory or does not relate to the open-ended questions, the terms assigned to the different levels of such a measure will not contribute to a better interpretation.

This study is one of the first studies that has applied STM to open-ended survey data of children and adolescents in the context of mental health. Responses from children and adolescents require additional preprocessing steps and the vocabulary of children has to be taken into account when deciding on minimal word and sentence length. In addition, the lexicon-based sentiment analysis approach on the topics that were found provides a deeper insight into how the emerged topics were experienced.

Due to the relatively small sample size for these specific types of analyses, we were unable to investigate the relationships with other covariates (or interaction effects). For example, including sex in addition to the dichotomized anxiety measure may have provided additional insights into how both sexes experienced the pandemic differently. Similarly, we were unable to investigate more extreme (low and high anxiety) groups. However, this may be a valuable approach for future studies as it may lead to a clearer separation and interpretation of topics and terms used by children with severe anxiety. In addition, future studies regarding the influence of the COVID-19 pandemic on mental health could expand upon the current study by including time in the analyses. This provides insight into whether specific topics were of importance in certain periods of the pandemic. In combination with sentiment analysis changes in polarity scores over time can be examined.

Conclusions

In this study we applied STM in anxiety data of children and adolescents. Important themes of the lockdown were identified, which showed that children with (sub)clinical anxiety scores experienced the COVID regulations more negatively and possibly struggle with adaptation/resilience and distress during the lockdown. We recommend combining PROMs with open-ended survey questions in future studies, as it provides unique information that may not be obtained otherwise. It allows us to assess the association between topics extracted from open-ended survey responses to a quantitative, validated outcome.

Statements and Declarations

Authors Contributions

ML, CG, CB, LH and TP conceptualized and designed the study. ML, HvO, LH were involved indata acquisition. ML, CG, CB, JZ, JT, HA, EB, JB, TP helped with the interpretation of the data. ML drafted and revised the initial manuscript. All authors read, reviewed, provided feedback and approved the final manuscript.

Ethics approval and consent to participate

We received approval for data collection from the appropriate ethics committees and all children and their parents provided informed consent. The study was conducted in accordance with the ethical standards outlined in the 1964 Declaration of Helsinki and its later amendments. Informed consent was obtained from all participants and their parents.

Data Availability

Anonymized data may be made available upon reasonable request. R code to perform the structural topic analyses and the Python script for performing sentiment analysis are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no competing interests.

Funding Statement

This work was funded by the Netherlands Organisation for Health Research and Development (ZonMW) (grant: 10430372310005).

References

1. Panda, P. K. *et al.* Psychological and Behavioral Impact of Lockdown and Quarantine Measures for COVID-19 Pandemic on Children, Adolescents and Caregivers: A Systematic Review and Meta-Analysis. *J Trop Pediatr* **67** (2021). https://doi.org/10.1093/tropej/fmaa122
2. Fischer, K. *et al.* Internalizing problems before and during the COVID-19 pandemic in independent samples of Dutch children and adolescents with and without pre-existing mental health problems. *Eur Child Adolesc Psychiatry* **32**, 1873–1883 (2023). https://doi.org/10.1007/s00787-022-01991-y

3. Zijlmans, J. *et al.* The effects of COVID-19 on child mental health: Biannual assessments up to April 2022 in a clinical and two general population samples. *JCPP Adv* **3**, e12150 (2023). https://doi.org/10.1002/jcv2.12150
4. Guzman Holst, C. *et al.* Examining Children and adolescent mental health trajectories during the COVID-19 pandemic: Findings from a year of the Co-SPACE study. *JCPP Adv* **3**, e12153 (2023). https://doi.org/10.1002/jcv2.12153
5. Korczak, D. J. *et al.* The trajectory of depression and anxiety among children and adolescents over two years of the COVID-19 pandemic. *Psychiatry Research* **339**, 116101 (2024). https://doi.org/https://doi.org/10.1016/j.psychres.2024.116101
6. Bauducco, S. *et al.* Adolescents' trajectories of depression and anxiety symptoms prior to and during the COVID-19 pandemic and their association with healthy sleep patterns. *Scientific Reports* **14**, 10764 (2024). https://doi.org/10.1038/s41598-024-60974-y
7. González-Calvo, G., Varea, V. & García-Monge, A. Children's Experiences of Lockdown and Social Distancing in the Covid-19 Pandemic. *Journal of Family Issues* **44**, 2422–2445 (2022). https://doi.org/10.1177/0192513X221094038
8. Luijten, M. A. J. *et al.* The impact of lockdown during the COVID-19 pandemic on mental and social health of children and adolescents. *Qual Life Res* **30**, 2795–2804 (2021). https://doi.org/10.1007/s11136-021-02861-x
9. Dewa, L. H. *et al.* The impact of COVID-19 on young people's mental health, wellbeing and routine from a European perspective: A co-produced qualitative systematic review. *PLOS ONE* **19**, e0299547 (2024). https://doi.org/10.1371/journal.pone.0299547
10. Rijcken, E. *et al.* Topic Modeling for Interpretable Text Classification From EHRs. *Frontiers in Big Data* **5** (2022). https://doi.org/10.3389/fdata.2022.846930
11. Sun, S., Zack, T., Williams, C. Y. K., Sushil, M. & Butte, A. J. Topic modeling on clinical social work notes for exploring social determinants of health factors. *JAMIA Open* **7**, ooad112 (2024). https://doi.org/10.1093/jamiaopen/ooad112
12. Roberts, M. E. *et al.* Structural Topic Models for Open-Ended Survey Responses. *American Journal of Political Science* **58**, 1064–1082 (2014). https://doi.org/https://doi.org/10.1111/ajps.12103
13. Roberts, M. E., Stewart, B. M. & Tingley, D. stm: An R Package for Structural Topic Models. *Journal of Statistical Software* **91**, 1 – 40 (2019). https://doi.org/10.18637/jss.v091.i02
14. Roberts, M. E., Stewart, B. M. & Airoldi, E. M. A Model of Text for Experimentation in the Social Sciences. *Journal of the American Statistical Association* **111**, 988–1003 (2016). https://doi.org/10.1080/01621459.2016.1141684
15. Klaufus, L. H. *et al.* Psychometric properties of the Dutch-Flemish PROMIS(®) pediatric item banks Anxiety and Depressive Symptoms in a general population. *Qual Life Res* **30**, 2683–2695 (2021). https://doi.org/10.1007/s11136-021-02852-y
16. van Oers, H. A. *et al.* Changes in child and adolescent mental health across the COVID-19 pandemic (2018–2023): Insights from general population and clinical samples in the Netherlands. *JCPP Advances* **n/a**, e12213 (2023). https://doi.org/https://doi.org/10.1002/jcv2.12213
17. Edition, F. Diagnostic and statistical manual of mental disorders. *Am Psychiatric Assoc* **21**, 591–643 (2013).
18. Porter, M. F. (2001).
19. Taddy, M. in *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics* Vol. 22 (eds D. Lawrence Neil & Girolami Mark) 1184––1193 (PMLR, Proceedings of Machine Learning Research, 2012).
20. R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, Austria, 2023).
21. Gatti, L. & van Stegeren, J. Improving Dutch sentiment analysis in Pattern. *Computational Linguistics in the Netherlands Journal* **10**, 73–89 (2020).
22. Vallejo-Slocker, L., Sanz, J., García-Vera, M. P., Fresneda, J. & Vallejo, M. A. Mental Health, Quality of Life and Coping Strategies in Vulnerable Children During the COVID-19 Pandemic. *Psicothema* **34**, 249–258 (2022). https://doi.org/10.7334/psicothema2021.467
23. Zhang, Q., Zhou, Y. & Ho, S. M. Y. Active and avoidant coping profiles in children and their relationship with anxiety and depression during the COVID-19 pandemic. *Sci Rep* **12**, 13430 (2022). https://doi.org/10.1038/s41598-022-15793-4
24. Theberath, M. *et al.* Effects of COVID-19 pandemic on mental health of children and adolescents: A systematic review of survey studies. *SAGE Open Medicine* **10**, 20503121221086712 (2022). https://doi.org/10.1177/20503121221086712
25. Zijlmans, J. *et al.* Mental and Social Health of Children and Adolescents With Pre-existing Mental or Somatic Problems During the COVID-19 Pandemic Lockdown. *Front Psychiatry* **12**, 692853 (2021). https://doi.org/10.3389/fpsyt.2021.692853

Figures

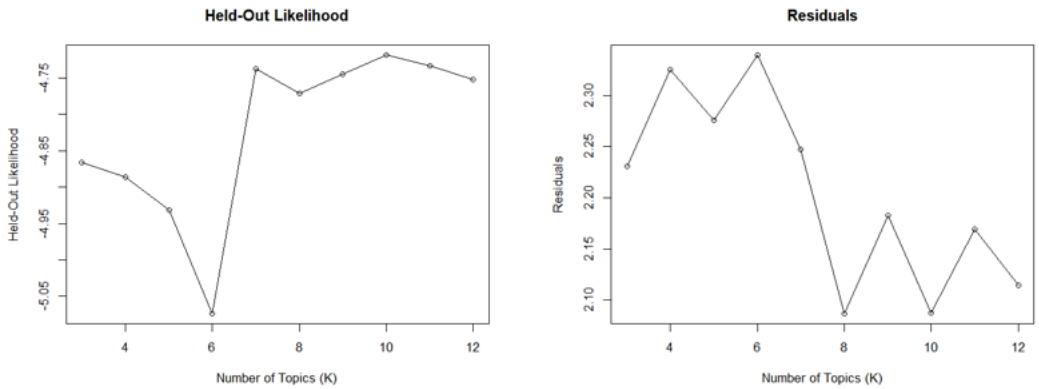


Figure 1

Held-out likelihood and residuals of *K* number of topics.

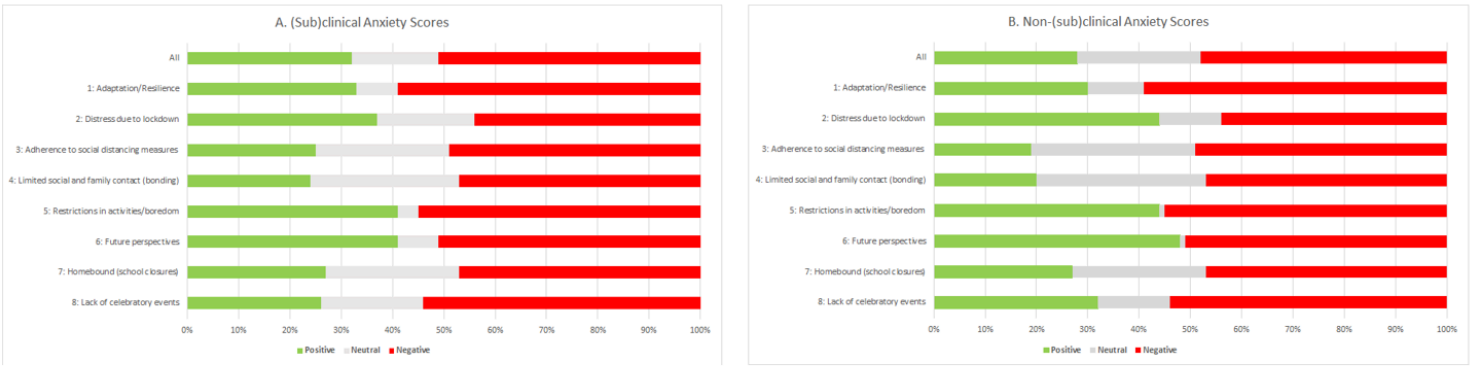


Figure 2

Percentages of positive, neutral and negative responses per topic of children with and without (sub)clinical anxiety scores.