

Mental Health Impact of the COVID-19 Pandemic on College Students: A Quasi-Experimental Study on Social Media

Koustuv Saha¹, Bhaskar Kotakonda², Munmun De Choudhury²

¹ Siebel School of Computing and Data Science, University of Illinois Urbana-Champaign

² School of Interactive Computing, Georgia Institute of Technology
ksaha2@illinois.edu, bkotakonda3@gatech.edu, munmund@gatech.edu

Abstract

Given the limited understanding of how the COVID-19 pandemic impacted mental health on college campuses, this paper examines the evolution of the mental health of college students since the onset of the pandemic. We conducted a large-scale study on over 1.2M posts on 173 U.S. college subreddits over 17 months. In particular, we adopted a quasi-experimental approach to examine how the different stages of the pandemic (isolation period, normalization period, and vaccination period) impacted changes in social media discussions of college students. We measured the temporal shifts in the symptomatic mental health expressions and topics of discussion on college subreddits. We find that while the expressions of depression, anxiety, stress, and suicidal ideation significantly increased in the isolation period. Interestingly, these expressions gradually subsided in the normalization period, only to resurface in the vaccination period. We also find unique occurrences of discussion across social, academic, health, and COVID-19-induced topics. Our findings reveal that despite the fragility of college students' mental health in the face of crisis, college students show resilience with sufficient time. We discuss the implications of our work in terms of building tools for real-time comprehension of college students' mental health, and in designing timely and tailored mental health support for college students.

1 Introduction

The COVID-19 pandemic caused numerous challenges throughout our lives and societies (Mishra et al. 2020; Ciotti et al. 2020). College campuses were not immune to the profound impacts of the pandemic, including unprecedented changes and disruptions in the usual college experience which led to heightened stressors among students (Pokhrel and Chhetri 2021). Among the enduring consequences of COVID-19 and immediate health concerns, one of the most significant has been its impact on mental health (Miller 2020; Talevi et al. 2020). This impact has been particularly pronounced on college campuses, where mental health is already a burgeoning concern and mental health care resources are under-utilized (Eisenberg et al. 2009).

As the pandemic unfolded, college students underwent significant transformations, navigating through disruptions in academic routines, abrupt shifts to online learning, physical isolations, reductions in on-campus activities, and the pervasive uncertainties surrounding health, finances, and careers. Although research has revealed the concurrent psychological impacts of COVID-19 on college students (Wang

et al. 2020; Kim et al. 2022), the long-term mental health effects on college students' mental health remain unknown. College students are in a critical developmental stage, where mental health challenges can have lasting effects on academic performance, career trajectories, and personal wellbeing (Eisenberg et al. 2009). An understanding of long-term impacts can equip universities to design better support systems and interventions that not only address immediate concerns, but also promote resilience and recovery from mental health in the years following a crisis. Moreover, this understanding can be valuable for preparing campuses to handle future disruptions effectively, ensuring students' holistic development is safeguarded during and after such events.

Our present understanding of the psychological impacts of COVID-19 relies heavily on surveys and healthcare usage data, which are typically reactive in nature, have biases, and are only collected at a later point in time. Because these assessments are static and in snapshots, they are susceptible to missing out on fine-grained and real-time fluctuations in psychological signals over time (Tourangeau et al. 2000). This implies the need for more scalable and real-time methods to understand the mental health impacts of COVID-19. In addition, understanding the multifaceted ways in which COVID-19 influenced the mental health of college students is crucial for developing tailored support systems and interventions to address their unique needs during and after the pandemic. This can empirically help inform the preparedness on college campuses for crises and external events that can impact the mental health of college students.

We aim to address above gap by leveraging social media data from online college communities. Social media platforms enable the collection of unobtrusive and naturalistic data of people's self-expressions both in real-time and in scale. In particular, social platforms are ubiquitous and widely used, and, in particular within the demographic of college students, are integral spaces to share experiences and expressions, build community, and seek social support and social capital (Ellison, Steinfield, and Lampe 2007). Prior work has revealed the potential of social media data to study college students' mental health (Saha et al. 2022). Essentially, social media data can enable us to understand how discussions and mental health issues evolve in response to changing circumstances of the pandemic, such as fluctuating case numbers, regulations, and vaccine administration. Accordingly, we target the following research questions:

RQ1: How did the pandemic impact the symptomatic mental health expressions of college students?

RQ2: How did the pandemic impact the topics of discus-

sions of college students?

We build on Ashokkumar and Pennebaker (2021)'s research on the different phases of the pandemic—*Isolation*, *Normalization*, and *Vaccination* period, and examined how the college students' mental health evolved through these phases. We collected social media (Reddit) discussions on 173 U.S. college subreddits over a 17-month period during the pandemic between March 2020 and August 2021, or the *Treatment* period. We also obtained pre-pandemic period's data between March 2018 and August 2019, which was our comparison or *Control* dataset to help mitigate temporal confounds in our analyses. Therefore, on an overall dataset of over 1.2M posts, we adopted a quasi-experimental approach to examining the impacts of the pandemic on the subreddit discussions.

For RQ1, we measured symptomatic mental health expressions through classifiers of depression, anxiety, stress, and suicidal ideation. For RQ2, we built BERT-based topic models (BERTTopics) and measured topical occurrences. We conducted interrupted time series (McDowall et al. 2019) and causal impact analyses (Brodersen et al. 2015). We find that the *Isolation* period did not show significantly different mental health concerns compared to the *Control* period, but *Normalization* and *Vaccination* periods had significantly higher mental health concerns than the corresponding *Control* period's data. Looking into the *Treatment* dataset deeper, our findings unveil a noteworthy pattern. As the pandemic started (*Isolation* period begun), there was a sudden and significant surge in symptomatic mental health expressions, which gradually subsided during the *Normalization* period, only to increase in the *Vaccination* period. Additionally, we identify distinctive discussions spanning social, academic, health, and COVID-19-induced topics.

Our study provides a detailed, real-time analysis of how the different phases of the pandemic impacted the mental health discourse among college students. By understanding these multifaceted impacts, we can better inform tailored support systems and interventions that address the unique needs of this demographic. Therefore, our research not only complements existing studies but also extends them by offering a more targeted and immediate analysis of mental health trends within the college student population during the pandemic. Our study not only sheds light on the immediate impacts of crises on the mental health of college students, but also lays the groundwork for a more comprehensive and forward-looking approach to crisis management on college campuses and educational settings. By understanding the temporal, interconnected, and resilient nature of students' experiences, we can better equip college campuses to navigate future challenges with support, adaptability, and a commitment to holistic student wellbeing.

Ethics and Reflexivity. Our work uses publicly accessible social media discussions on online college communities on Reddit (college subreddits), and did not require any direct interactions with the individuals, so, it did not qualify for ethics board approval. However, we are committed to the ethics of the research and followed practices to secure the individuals' privacy in our dataset. We recognize the sensitivities of our study in revealing the identities of college students and campuses. We used de-identified data for analyses, and this paper only presents paraphrased quotes to reduce traceability, yet provide context in readership. Our research findings were contextualized and interpreted through several conversations with college administrators, staff, and students. The authors' diverse experiences—as college student, col-

lege applicant, and faculty member during the COVID-19 period—provided unique perspectives. Additionally, multiple authors experienced major life events and career transitions during this time, which further shaped our interpretations, highlighting how our positionality helped us contextualize and interpret the findings.

2 Related Work

COVID-19 Pandemic and Mental Health Impacts On March 11, 2020, the World Health Organization (WHO) declared COVID-19 a “global pandemic” considering the then over 100K infected cases in over 100 countries (Cucinotta and Vanelli 2020). The coronavirus outbreak led to a pandemic that has globally caused over 770M cases and over 7M deaths, reported by WHO as of January 2024 (World Health Organization 2020). The pandemic led to rapid and significant rises in psychological distress and mental health concerns (Basu 2020). As the pandemic progressed, the fear and anxiety about the disease and various societal and personal factors, exacerbated the existing mental illness of individuals (McGinty et al. 2020). The disease mitigating efforts and regulations such as social distancing and the lack of in-person events led to increased symptoms of anxiety, depression, and feelings of fear, abandonment, loneliness, and stigmatization (Miller 2020). Kontoangelos et al. 2020 conducted a review of the mental health effects of the COVID-19 pandemic. Researchers have also looked at various forms of digital data (Torous et al. 2020), including smartphone sensing (Nepal et al. 2022) and social media (Ashokkumar and Pennebaker 2021; Saha et al. 2020; Guntuku et al. 2020). Our work draws motivation from these studies to leverage social media data to examine how the mental health symptomatic expressions progressed during the pandemic.

Mental Health on College Campuses Per the 2018 AUCCD survey, most frequent concerns for college counseling centers around the world are anxiety (58.9%), depression (48.0%), stress (46.9%), relationship problems (29.5%), family concerns (29.0%), suicidal thoughts (28.4%), academic performance difficulties (28.2%), sleep disturbance (19.1%), social isolation or loneliness (18.5%), significant previous mental health treatment history (16.5%), and adjustment to a new environment (15.8%) (LeViness et al. 2017). While these numbers are already significant, mental health concerns are under-reported for college students (Eisenberg et al. 2009). These numbers are largely based on what the students seek support and care for, yet, given the stigma surrounding mental health concerns, actual numbers are likely to be much higher (Eisenberg et al. 2009).

Besides the data on the utilization of mental health services, mental health data on college campuses (and also generally) are largely derived from surveys through psychometric instruments. While these forms of data are very informative and accurate in snapshots, they are sensitive to self-report and retrospective recall biases. Moreover, such measurements can only be conducted periodically, posing difficulties in understanding the dynamic evolution of college students' mental wellbeing over time. In the specific context of COVID-19, prior work leveraged smartphone sensing to examine the mental health impact of the COVID-19 pandemic on college students (Nepal et al. 2022). While passive sensing techniques capture rich and dense behavioral signals, they suffer from limitations of scale and compliance and are expensive to implement. Parallely, psycholinguistics research reveals how people's wellbeing can be assessed

from their language use (Pennebaker et al. 2014), motivating how social media can serve as a passive sensor for wellbeing of college students (Bagroy, Kumaraguru, and De Choudhury 2017; Saha and De Choudhury 2017).

Social Media, Crisis, and Mental Health Social media data has enabled studying psychological impacts of external events, crises (Palen 2008; Saha and De Choudhury 2017), wars and conflicts (Mark et al. 2012; De Choudhury, Monroy-Hernandez, and Mark 2014), natural disasters (Starbird et al. 2010), and terror attacks (Lin and Margolin 2014). Other studies have leveraged social media to examine psychological reactions following gun violence in schools, such as the Sandy Hook and Virginia Tech Shooting (Glasgow, Fink, and Boyd-Graber 2014; Palen 2008). Social media platforms serve as a safe haven for self-expression, support, and solidarity during crises (Mark et al. 2012; Starbird et al. 2010).

The COVID-19 pandemic led to significant changes in people’s lives and mental health, including that on college campuses (Melcher et al. 2023; Wang et al. 2020; Copeland et al. 2021; Chrikov et al. 2020). Prior work also studied the mental health impacts of the pandemic through social media data, revealing that mental health related concerns gradually decreased after a sudden huge increase in prevalence at the onset of the pandemic (Saha et al. 2020). Our work builds on the above body of literature, by conducting a large-scale study on social media to examine the mental health impacts of the COVID-19 pandemic.

Causal Inference Approaches on Social Media and Observational Data Causal-inference approaches are adopted to examine the effect of an (external or internal) change or intervention. In settings where controlled trials or experimental approaches may be infeasible or unethical, observational study design can be adopted. Observational studies are designed so as to minimize confounders and provide stronger evidence than purely correlational approaches (Imbens and Rubin 2015). These approaches have garnered interest in computational social science research, including using social media data (Keith, Jensen, and O’Connor 2020; Saha et al. 2021). Prior research has explored shifts in online suicidal ideation (De Choudhury et al. 2016), mental health coping stories (Yuan et al. 2023), and county health estimates using Twitter data (Culotta 2014). These approaches have been adopted on social media to examine the effects of COVID-19 on mental health (Saha et al. 2020; Jha et al. 2021; Vowels et al. 2023). In the context of college students, research has adopted causal approaches in social media data on the effects of on-campus gun violence (Saha and De Choudhury 2017), public service announcements after student deaths (Saha et al. 2018), hateful speech (Saha et al. 2019), and alcohol use (Kiciman, Counts, and Gasser 2018) on college campuses. Methodologically, our work draws on the above causal-inference literature, particularly the potential outcomes framework (Rubin 2005), interrupted time series (McDowall, McCleary, and Bartos 2019), and causal impact estimation (Brodersen et al. 2015) to mitigate the confounds and unveil temporal effects of the pandemic on the mental health of college students.

3 Data

Gathering college campus-wise social media data

To study the mental health of college students, the data for our study comes from Reddit. Reddit is a social media

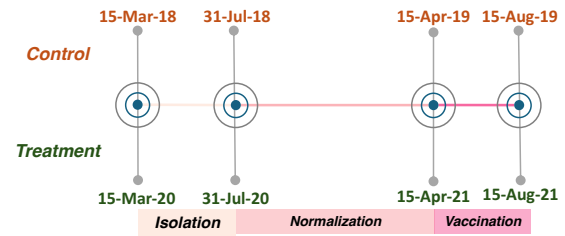


Figure 1: A schematic representation of the timeline of *Treatment* and *Control* datasets, labeled with periods.

platform, which, due to its pseudonymous design, enables candid self-disclosure of stigmatized topics such as mental health (De Choudhury and De 2014). Reddit is widely used and popular among the youth and college student demographic (Statista 2021), and has college-specific communities (or college subreddits). Prior work has established that geographically localized communities like college subreddits are reliable proxies to infer the college students’ mental health (Bagroy, Kumaraguru, and De Choudhury 2017; Saha and De Choudhury 2017; Saha et al. 2022). This body of research has noted that college subreddits can be representative of the student body of their corresponding college campuses. In these communities, students discuss and share about their academic, personal, and many other aspects of campus life. Therefore, college subreddits are unique in their dynamic, with the majority of discussions typically driven by currently enrolled on-campus students. Graduated or former students are often less active in these communities, making the conversation predominantly focused on present campus life. Relatedly, Saha et al. (2022) highlighted the construct validity in how ongoing discussions can be reflective of on-campus mental health consultations.

For the purpose of this study, we identified the top 173 colleges in the U.S. based on the U.S. News rankings and the number of students in the subreddit. For each college, we additionally obtained other identifiers like *city*, *state*, *name of the official subreddit*, *number of users*, *tuition*, *geography*, etc by consulting the respective official webpages, Wikipedia, and by using Reddit’s subreddit search feature. Then, we retrieved the posts corresponding to each college subreddit over a principled set of selected time periods around the COVID-19 pandemic, which we describe below.

Time Periods

Drawing from the causal inference literature (Kiciman, Counts, and Gasser 2018), we needed to separate the data before and after the event in focus occurs to mitigate the effects of confounding or latent variables in the sample space. In our case, to study the comparative effect of COVID-19 on the students, we identified the *Treatment* and *Control* periods corresponding to the duration affected by COVID and the duration before that. Figure 1 summarizes the timeline of our datasets, and we describe these below.

Treatment Period. Within the period post the onset of the COVID-19 pandemic, we identified three phrases in line with prior research (Ashokkumar and Pennebaker 2021; Zhunis et al. 2022). These phases align with the timeline of COVID-19 events in the real world. These phases are:

- **Isolation Phase** (15-Mar-20 to 31-Jul-20) consists of the initial official declaration of the pandemic as a national

Period	Treatment		Control	
	Duration	#Posts	Duration	#Posts
Isolation	15-Mar-20—31-Jul-20	199,241	15-Mar-18—31-Jul-18	78,103
Normalization	1-Aug-20—15-Apr-21	393,780	1-Aug-18—15-Apr-19	226,383
Vaccination	16-Apr-21—15-Aug-21	207,624	16-Apr-19—15-Aug-19	103,718
Overall	15-Mar-20—15-Aug-21	800,645	15-Mar-18—15-Aug-19	408,204

Table 1: Summary of *Treatment* and *Control* datasets.

emergency in the U.S. Major aspects of life moved to the digital mode with states announcing strict lockdown orders restricting all kinds of physical movement. This phase initiated a shift to online classes and remote activities on college campuses.

- **Normalization Phase** (1-Aug-20 to 15-Apr-21) is the period post the initial surge in attention about COVID-19 but the aftereffects of the virus and the discourse around it still existed. Policies like remote education and social distancing became popular among college students, and discussion and monitoring of health issues (e.g., COVID testing or gauging infection trends) became routine—essentially a *new normal* was setting in.
- **Vaccination Phase** (15-Apr-21 to 15-Aug-21) is the duration when the vaccines began being administered and the subsequent discourse around the virus and the vaccine re-entered public discourse.

Control Period. For each of the above identified periods, we also needed to collect data pertaining to a suitable control period that would help us offset the other factors affecting the student posts like time of the year, academic timelines, or other major events, including circumstances responsible for altered student mental health (e.g., breaks or exams). In addition to this, the control period should also take care of the seasonal changes that may affect the mental health of individuals. Prior work approached similar problems by identifying *Control* period as the identical time period (months) from the previous year for *Treatment* period (Saha and De Choudhury 2017; Saha et al. 2020). However, given that our *Treatment* dataset spanned for more than a year, a one-year gap between *Treatment* and *Control* periods would have led to the *Control* period including post-pandemic data (e.g., Apr-21 in *Treatment* would then match to Apr-20 in *Control*—however, Apr-20 is post-pandemic and is actually a part of the *Treatment* period). Rather, we ensured that both the *Control* and *Treatment* datasets had identical lengths and were cleanly pre- and pandemic, respectively, by setting *Control* period’s dataset to be two years prior to the *Treatment* period for any respective date. Therefore, the overall *Control* period was chosen to be from 15-Mar-18 to 15-Aug-19, i.e. two years prior to the *Treatment* period for any respective date. Within the *Control* period, the placebo phases can be separated similarly to the above:

- **Isolation Phase** (15-Mar-18 to 31-Jul-18)
- **Normalization Phase** (1-Aug-18 to 15-Apr-19)
- **Vaccination Phase** (15-Apr-19 to 15-Aug-19)

Data Collection and Data Processing

Based on the above-identified periods, we collected our dataset using the Pushshift API V4.0. The Pushshift API (at the time of this research) is a tool designed by the moderators of the *r/datasets* subreddit to provide advanced search functionalities for the comments and the post submissions on any given subreddit. Of the two endpoints, we utilized

the “submissions” endpoint to collate the data for a given subreddit and duration. This paper uses the term “posts” as a unified term referring to both submissions and comments.

Following the above data collection process, we queried the 173 college subreddits in the specific time periods on Pushshift API, and additionally checked for any deleted posts or missing values in the query responses. This led us to a total of 1,208,849 posts—800,645 belonged to the *Treatment* period, and 408,204 belonged to the *Control* period. This shows almost double the activity (96% growth) during the same duration in the two time periods—aligning with prior work on how Reddit activities significantly spiked during the pandemic (Veselovsky and Anderson 2023). Table 1 summarizes the dataset.

We examined if the dataset in our study is comparable per subreddit across the various time periods. Therefore, we aggregated the data per subreddit in different periods for both *Control* and *Treatment* datasets, and obtained the proportion of data per 173 subreddit. Figure A1 (a&b) shows the density distribution of proportion of posts per subreddit over the different phases in both *Control* and *Treatment* datasets. We observe no significant visual differences across the three phases within the *Control* and *Treatment* periods—this was further confirmed by Kruskal-Wallis tests—($H=0.01$, $p=0.99$) for *Control* period and ($H=0.31$, $p=0.85$) for *Treatment* period—revealing no statistical significance in differences across phases in both the periods. We also conducted similar analyses for the dataset per state across the time periods, where we aggregated the data per state in different periods. Appendix Figure A1 (c&d) shows the density distributions where again we find very similar trends across all the phases, further confirmed by Kruskal-Wallis tests—($H=0.01$, $p=0.99$) for *Control* period and ($H=0.22$, $p=0.90$) for *Treatment* period—revealing no statistical significance. These observations confirm that posting activity in different subreddits showed uniform trends and was not sensitive to varying events that might have occurred in different colleges or geographical locations during the pandemic.

Further, we obtained the correlation between the quantity of posts per subreddit and the college size, where we found a high correlation (Pearson’s $r=0.55$, $p<0.001$). We also obtained the correlation between the quantity of the posts per state and population of the state as per 2020 U.S. census (data.census.gov), which also revealed a very high correlation (Pearson’s $r=0.94$, $p<0.001$). The high and statistically significant correlation is somewhat indicative of the representativeness of subreddit’s data with offline origin (college campus and state)—also aligning with prior research (Bagroy, Kumaraguru, and De Choudhury 2017).

4 RQ1: Evolution of Mental Health Expressions

Measuring Mental Health Expressions

Measurement Approach Following prior work (Saha et al. 2022), we postulate that we can assess college students’ symptomatic mental health expressions based on their social media posts. We leveraged machine learning classifiers developed in prior work (Saha et al. 2019, 2022) to identify symptomatic mental health expressions of depression, anxiety, stress, and suicidal ideation on college subreddit datasets. These classifiers were developed using transfer machine learning techniques and are n -gram-based ($n=1,2,3$) binary SVM models. The positive labels of the training data were sourced from subreddits of *r/depression*,

$r/anxiety$, $r/stress$, and $r/SuicideWatch$ for depression, anxiety, stress, and suicidal ideation respectively. The negative classes came from the top 20 subreddits on Reddit’s homepage unrelated to mental health, such as $r/AskReddit$, r/aww , and $r/movies$. These classifiers were expert-validated with the DSM-5 clinical framework to reveal high-performing accuracy of >90% (Saha et al. 2019). Further, the construct validity of these classifiers was again by these authors in Saha et al. 2022 on college subreddit dataset, to reveal that the classifiers’ predictions of symptomatic mental health expressions were indicative of on-campus mental health consultations. As a cautionary note, these classifiers are not diagnostic of someone’s mental health conditions; rather, they are only indicative of symptomatic mental health expressions on social media language.

We employed the above classifiers on our dataset to obtain the symptomatic mental health expressions. For a given post and a classifier output, a class label=1 indicates the presence of that symptomatic expression, and a class label=0 denotes its absence. We obtained the daily occurrence of each symptomatic expression per time series, and then normalized these daily occurrences over the total frequency of posts per day. To measure the relative significance of COVID-19 on the mental health of the student community, we contrasted the *Treatment* period time series by superimposing with the corresponding *Control* period time series for each symptomatic mental health expression. We calculated the z -scores of the combined *Treatment* and *Control* series for each phase and symptomatic expression. We obtained effect size (Cohen’s d) and evaluated statistical significance in differences using independent sample t -tests. We conducted Kolmogorov-Smirnov (KS) tests, which essentially test against the null hypothesis that the distributions of *Treatment* and *Control* datasets are drawn from the same distributions. For both t -tests and KS-tests, we conducted Bonferroni correction to adjust for the multiple hypotheses tests.

Time Series Observations Table 2 summarizes the relative differences of the z -score time series obtained for the three identified periods—*Isolation*, *Normalization*, and *Vaccination*, along with the statistical tests. We observe that while there is not a major difference in the symptomatic expression occurrences in the *Isolation* phase of the *Treatment* period compared to the *Control*, the *Normalization* and *Vaccination* periods show a distinctive difference in all the symptomatic expressions. These findings reveal that the symptomatic expressions of students’ mental health increased after the start of the pandemic.

In comparison to the *Control* period, within the *Isolation* phase, the *Treatment* period shows slightly lower anxiety (-0.9%) and stress (-0.2%), and higher suicidal ideation (3%). This can be attributed to the fact that mental health discourse among college students are typically during these months (during pre-pandemic) (Saha et al. 2022), and that the mental health disruptions due to the pandemic may have taken a while to affect the college students compared to a regular academic year, as also noted by Das Swain et al. (2024). In *Normalization* phase, *Treatment* period’s expressions were higher—depression by 15%, anxiety by 12%, stress by 9%, and suicidal ideation by 15%. The *Vaccination* phase also shows similar trends of higher occurrences in *Treatment* than *Control* period—depression by 14%, anxiety by 11%, stress by 8%, and suicidal ideation by 4%.

Next, Figure 2 plots the occurrences of symptomatic mental health expressions over time. These plots superimpose the *Treatment* and *Control* dataset’s values in comparable

Expression	Relative Difference%	t -test	KS-test	Cohen’s d
<i>Isolation Period</i>				
Depression	-1.24	-6.23	0.15	-0.75
Anxiety	-0.86	-3.60***	0.24***	-1.10
Stress	-0.21	-1.91	0.18**	-0.69
Suicidal Ideation	2.99	6.18***	0.41***	1.17
<i>Normalization Period</i>				
Depression	15.40	10.70***	0.41***	0.95
Anxiety	12.40	8.08***	0.35***	0.72
Stress	8.98	8.75***	0.43***	0.78
Suicidal Ideation	15.46	28.30***	0.79***	2.52
<i>Vaccination Period</i>				
Depression	13.86	3.17**	0.24**	0.41
Anxiety	10.93	2.13*	0.19*	0.28
Stress	7.86	3.70***	0.24***	0.48
Suicidal Ideation	3.65	11.43***	0.60***	1.49

Table 2: Relative difference in mental health expressions between the *Treatment* and *Control* periods and statistical significance as per t -test, KS-test, and Cohen’s d . Significance (p -values) reported after Bonferroni correction as per * $0.01 < p < 0.05$, ** $0.001 < p < 0.01$, *** $p < 0.001$.

time periods. We analyze the z -score trends of each of the expressions. Consider the case of the depression graph; the slopes of the curves in each of the *Treatment* period phases gradually decreased from 0.017 to 0.004 and 0.001. Therefore, within the *Treatment* period, the expression of depression in the college subreddits leveled out over the course of the pandemic. The same can also be observed in the case of anxiety (0.017 to 0.005 to 0.004) and stress (0.016 to 0.006 to 0.003). Another interesting observation is the relatively flat curve for the normalization period in each of the trends—depression (0.004), anxiety (0.005), stress (0.003), and suicidal ideation (0.004) shows that there is little variation in these expressions during these phases, implying back to normalcy. This aligns with prior work on how people’s social media expressions get closer to normalcy after a crisis (Lin and Margolin 2014), and aligns with principles of habituation from behavior science research (Rankin et al. 2009).

Examining Causal Impact

Causal Impact Analysis Causal Impact Analysis is used to study the effects of an intervention by an external factor(s) (Brodersen et al. 2015). For our study, this is especially helpful in understanding the causal impact of interventions at the periphery of each identified COVID-19 phase. We drew on Brodersen et al. to measure the causal impact as the difference in the expected and observed time series data.

To study the effect of an intervention, both expected and observed time series were given, and the two periods corresponding to the “pre-intervention” and “post-intervention” were identified using the Causal Impact model (Brodersen et al. 2015; Saha et al. 2024). Then, we used Bayesian structural time series modeling to fit a model based on the observed data. This fitted model was used to predict the counterfactual, or what the response in the post-intervention period might have looked like had the intervention not taken place. We divided the time series data into cause and effect series, and studied the three interventions in our study—*Isolation*, *Normalization*, and *Vaccination*, corresponding to the three phases (Section 3). Table 3 summarizes the pre- and post- intervention periods for the three phases. Our ensuing results report the relative change in the predicted time series data as well as the probability of the causal effect.

Causal Impact Findings Table 4 summarizes causal impact metrics—the relative effect and posterior probability of

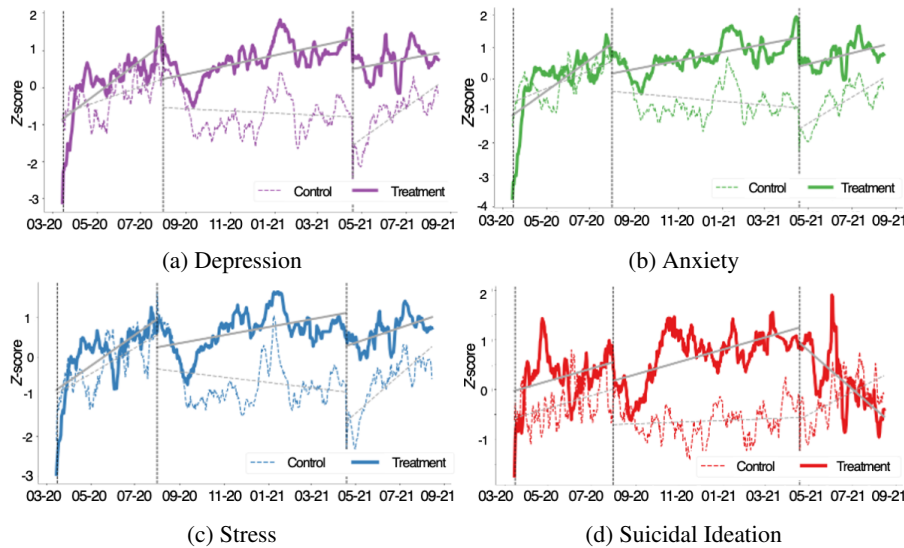


Figure 2: Evolution of mental health expressions during the pandemic. The trend lines (gray) in both the *Treatment* and *Control* periods are also shown in addition to the z -score trends—denoting the linear fit over that duration.

Phase	Intervention	Pre-Intervention Period	Post-Intervention Period
<i>Isolation</i>	Mar 15, 20	Dec 15, 19–Mar 15, 20	Mar 16, 20–Jul 31, 20
<i>Normalization</i>	Jul 31, 20	Mar 16, 20–Jul 31, 20	Aug 1, 20–Apr 15, 21
<i>Vaccination</i>	Apr 15, 21	Aug 1, 20–Apr 15, 21	Apr 16, 21–Aug 15, 21

Table 3: Time periods around interventions (COVID-19 phases) for causal impact analysis (Brodersen et al. 2015).

Expression	<i>Isolation</i>		<i>Normalization</i>		<i>Vaccination</i>	
	Effect	P.P.	Effect	P.P.	Effect	P.P.
Depression	3.18	100	-10.70	99	4.76	98
Anxiety	4.79	100	-8.65	99	6.23	98
Stress	1.77	99	-12.11	97	5.82	96
Suicidal Ideation	5.63	100	-5.45	100	2.89	100

Table 4: Causal impact estimation metrics of relative effect (%) and posterior probability (P.P.%) of causal effect.

causal effect. We plot the causal impact estimation in Figure A4, A5, and A6 for *Isolation*, *Normalization*, and *Vaccination* phase respectively. In addition, the posterior probability measure signifies the confidence with which we can ascertain the significance of the intervention. Overall, it is interesting that all interventions lead to statistically significant changes in their respective series (>95%).

Causal Impact for Isolation phase. The *Isolation* intervention shows a marked increase in the cumulative effect for all four expressions, indicating that the intervention of *Isolation* resulted in a greater expression of these symptoms. The highest change in the predicted effect is for suicidal ideation (5.63%) with 100% posterior probability (Table 4).

Causal Impact for Normalization phase. For the *Normalization* intervention, interestingly, the overall cumulative effects dipped, i.e., a true sense of normalcy with regard to reduced levels of depression, anxiety, stress, and suicidal ideation are observed. We also note that the posterior probabilities are all >97%, indicating a strong causal impact of the intervention.

Causal Impact for Vaccination phase. Finally, an increased

trend corresponding to the symptomatic expressions similar to the *Isolation* phase is observed in the vaccination intervention graphs of Figure A6. This can be attributed to the other COVID-19 variants and the multiple waves of infection, as well as the anxiety corresponding to the vaccines (Kumar et al. 2022). The highest change is observed in the cumulative changes in the anxiety values (6.23%).

5 RQ2: Evolution of Topics of Discussions Building and Evaluating Topic Models

For RQ2, we examined the relative change in the topics of discussion among the student community in the various stages of the COVID-19 pandemic. We adopted topic modeling—a popularly used technique to extract the common themes or topics in a corpus of text (Blei, Ng, and Jordan 2003). On our college subreddit dataset, we employed topic modeling, specifically BERTopic (Grootendorst 2022).

Topic Modeling using BERT BERTopic leverages NLP algorithms like transformers and a class-based term frequency-inverse document frequency score to create dense clusters with easily interpretable topics for the entire corpus (Grootendorst 2022). BERTopic provides flexibility with different custom-trained models for the clustering and embeddings of topics, enabling us to incorporate semantic similarities in obtaining topics in our dataset. We initialized the BERTopic model with the n -gram ($n=1,2,3$) as this could capture our topics with sufficient linguistic context and comprehensiveness. We used the UMAP algorithm for dimensionality reduction, HDBSCAN for clustering, and TF-IDF based topical representations to conduct our topic modeling (A provides details). Using this approach, we obtained topics from all posts in our entire dataset.

Obtaining Optimal Number of Topics An important aspect of topic modeling—an inherently unsupervised approach—is obtaining the optimal number of topics. For this, we evaluated the degree of semantic similarity between the top representative words in each cluster and by assessing the number of topics for which this intra-cluster semantic

Theme	Topic Keywords
Parking	parking, park, pass, permits, garage, ticket, passes, parked, cars
Outdoor Events	tickets, ticket, selling, sale, stadium, tailgate, tours, tailgating, bought, sold
Social Hangouts	friends, meet, lonely, dating, friend, social, homeless, group, clubs, meeting
Housing	roommate, roommates, rent, room, apartment, house, bedroom, lease, rooms, beds
Online Classes	online, asynchronous, synchronous, webreg, courses, webassign, async, schedule, instruction, sync
College Admits	decision, decisions, appeal, admissions, acceptance, applied, admission, received, accepted, selection
Outdoor Dining	dining, food, restaurants, tassel, lunch, taco, restaurant, cafe, tacos, mcdonald
Dropping Courses	drop, dropping, dropped, adddrop, deadline, credits, add, credit, hours, zoomers
Physical Health	yoga, workout, fitness, meditation, exercises, pilates, stretch, stretches, massage, exercise
Vaccination	vaccine, immunization, vaccination, vaccines, vaccinated, covid, quarantine, covid19, vaccinations, immunizations
Indoor Games	chess, poker, games, nintendo, playlist, gaming, game, jeopardy, bingo, chesscom
Masks	masks, mask, face, vaccinated, coverings, masking, vaccination, nose, vaccine, covering
Mental Health	counseling, therapy, therapist, therapists, depression, suicide, counselor, counselors, psychiatrist, depressed
Grief	cry, sad, death, depressed, died, crying, die, depression, weeps, funeral
Financial Aid	aid, financial, loan, disbursement, subsidized, disbursed, federal, pay, offered

Table 5: Summary of major topical themes.

similarity is maximized while the inter-cluster similarity is minimized. So, we obtained the coherence scores for a range of topics where we found that the maximum value of the coherence score is obtained for the $k=200$ number of topics (see Appendix Figure A2). Therefore, we used the BERTopic with 200 topics for the remainder of the analysis.

Topical Occurrences After running a BERTopic model with 200 topics, two coauthors conducted thematic analyses, by manually inspecting the keywords in each topical cluster and assigning interpretable labels to these topics, drawing on a similar approach from prior work (Zhou et al. 2022). Then, we measured the relative frequency of the topics by aggregating the day-wise occurrences of topics, normalized by the volume of day-wise posts. Table 5 summarizes the top occurring topical themes and representative keywords (the Appendix provides the entire list of 200 topics). Next, we measured the relative changes in the occurrences between the corresponding *Treatment* and *Control* periods and within-*Treatment* between different phases. Table 6 summarizes the relative frequency of the major occurring topics, representative keywords, and relative occurrence changes. Appendix Figure A3 plot the time series trends of relative frequencies of the top topics in the *Treatment* and *Control* periods.

Examining Causal Impact on Topical Occurrences

We conducted causal impact analyses similar to Section 4 to examine how each of the *Isolation*, *Normalization*, and *Vaccination* interventions causally impacted the topical usage in the college subreddits. This analysis also helps us ascertain which topics’ occurrences were impacted with statistical significance at the phase transitions. Table 7 shows the topics’ relative effect and posterior probabilities.

Isolation Intervention on Topics. We observe a significant impact in the topics of *parking*, *housing*, *dropping courses*,

Topic	Between Treatment and Control			Within Treatment	
	ΔI (%)	ΔN (%)	ΔV (%)	$\Delta I \rightarrow N$ (%)	$\Delta N \rightarrow V$ (%)
Parking	-4,848	-3,753	-1,754	-48	10,256
Outdoor Events	-7,597	-6,943	-1,354	7,590	10,129
Social Hangouts	-1,982	1,876	1,503	7,011	-1,178
Housing	-2,943	-1,419	334	-2,654	12,364
Online classes	118,410	52,225	30,823	-5,011	-3,129
College Admits	523	3,985	3,386	4,881	-2,693
Outdoor dining	-4,167	-2,416	-628	4,976	614
Dropping Courses	12,402	5,291	4,948	-962	-2,850
Physical Health	3,656	9,152	9,558	4,099	711
Vaccination	7,497	163,432	81,656	25,786	9,094
Indoor Games	-1,157	-1,855	43	3,008	-1,761
Masks	91,626	55,177	208,061	2,308	4,213
Mental Health	-1,247	1,485	5,451	9,114	-2,483
Sadness & Grief	1,360	4,125	-796	4,394	-4716%
Financial Aid	3,188	2,595	-1,345	-6,259	10,275

Table 6: Changes in Relative frequency of topics over the COVID-19 phases of *Isolation* (I), *Normalization* (N), and *Vaccination* (V). Δ represents the change in the phase, expressed as percentage (%). Within the *Treatment* period, we measure the shifts in topical usage across consecutive phases, i.e., I to N, and N to V.

Topic	Isolation		Normalization		Vaccination	
	Effect	P.P.	Effect	P.P.	Effect	P.P.
Parking	45.41	100	-19.32	94.11	45.71	100
Outdoor Events	40.07	99.6	64.02	100	-12.72	85.01
Social Hangouts	8.77	85.81	27.39	100	-11.68	91.51
Housing	63.74	100	-44.41	100	41.76	100
Online classes	22.23	79.02	188.71	100	-80.37	100
College Admits	-10.41	84.92	9.26	79.62	-19.58	97.2
Outdoor dining	-15.95	93.71	16.22	97.1	-6.60	75.22
Dropping Courses	-24.85	97.4	29.24	99.6	-40.78	100
Physical Health	-3.05	59.64	38.91	99.90	-29.38	99.9
Vaccination	94.16	100	-56.30	100	48.50	98.7
Indoor Games	-24.96	98.90	37.49	100	-26.84	99.30
Masks	124.15	100	-10.14	87.55	-32.51	63.88
Mental Health	28.92	95.48	45.37	100	-9.50	84.2
Sadness & Grief	74.59	100	50.80	100	-40.22	92.73
Financial Aid	12.41	63.84	29.42	97.5	-18.40	80.64

Table 7: Causal impact metrics on the evolution of topical occurrences of 15 major topics. Posterior Probabilities (P.P.) are reported as percentages (%).

international travel, and *masks*.

Normalization Intervention on Topics For the *Normalization* intervention, we find a significant impact in the occurrences of *social life*, *online classes*, *dining*, *fitness*, *vaccine*, *test kits*, *board games*, *masks*, and *sadness and grief*.

Vaccination Intervention on Topics For the *Vaccination* intervention, we find a significant impact in the topical occurrences of *housing*, *online classes*, *admissions*, and *board games*. This phase brought vaccines into the equation and increased the stress surrounding them. In addition, some of the after-effects of the initial isolation phase were also seen in the form of health and grief.

Qualitative Examination of the Topical Themes

Following up on the analyses in the previous paragraphs, we qualitatively examined the topics, and identified four broader themes in the discourse of college students during the COVID-19 pandemic. We describe the findings of our qualitative analysis on the themes of discussions below:

On-Campus and Social Topics This topic spans discussions on social events, housing, meeting friends, outdoor dining, parking, etc. The *Isolation* phase saw discussions on the **lack of social life**. Students were restricted to their

homes and screens, and the initial reactions included reassessing their social lives, feelings of loneliness, and finding solidarity, e.g., “How well are you guys keeping up with your friends from college? Do you guys feel your friendships are largely affected by covid?” However, discussions on social hangouts emerged significantly with the onset of the *Normalization* phase and, subsequently, the *Vaccination* phase. Through various stages of the pandemic, with multiple restrictions on movement and social distancing, discussions relying on on-campus or physical presence saw a decreased occurrence. Topics such as parking, dining, and social events, given their dependency on regulations set by campus administration, saw confusion in policy changes, such as parking permits, concert refunds, etc.

Parking. Before the *Isolation* phase, parking discussions were mainly inquiry-related and informative in nature, such as, “Where’s the best spot to park on college ave without getting ticketed?”. However, after the onset of the pandemic, significant shifts occurred, including confusion and inquiries about the changed regulations on campus parking, such as, “I just got my parking refund of only \$0. They didn’t even warn us about the value of permits still decreasing while the campus isn’t open for lectures—this is ridiculous!”

Housing. The housing discussions before the pandemic concerned describing rental places with details on mobility to nearby areas. The selling point for seeking tenants revolved around the social attractions and accessibility to events, activities, and eateries nearby, e.g., “I am releasing my room in south campus for the spring 2020 semester. The other 3 roommates are very clean and awesome. We decorated the place, and it is within a 5-minute walk to most campus buildings, Chipotle, Cava, Lotsa, etc.” However, as the pandemic progressed, there were shorter and cautionary posts, focusing more on off-campus housing and social distancing practices, e.g., “Anyone still looking for a roommate? I’m looking for a place close to campus. The safer places are pricey, so I thought I’d get the roommates to share the burden. I won’t be going out other than classes, so don’t worry about me getting COVID.”

In the *Normalization* phase, there was a distinctive increase in the student posts mentioning COVID topics similar to social distancing and hygiene in the listings. The host made it a point to address COVID-19 and related protocols in the initial stages. For example, “My two roommates and I are trying to find another person to sign with us. We take covid very seriously and do our best to stay safe.”

However, the *Vaccination* phase saw the revival of the social life in the listing details and the status of “fully vaccinated” as a proxy for the safety of the host and surroundings, as well as prospective rentals. The vaccination status likely improved their chances of getting a place, such as, “Hey everyone, I am looking for anyone planning on canceling their housing for 2021-22. I am vaccinated, very clean, and respectful of my roommates. I would be happy being friends with my roommates, but I can keep to myself if that’s what my roommates seek.”

Dining. The *Isolation* phase saw limited discussions on dining, except some about food delivery and supporting small businesses. While some of these discussions continued in the *Normalization* phase, there was also frustration about the new protocols and the conflicting opinions of the new normal. The limited options related to the dining options after the students came back to the campus were discussed, such as, “Why are all the dining halls closed except the one in the east, which only serves two options?”

Board Games. The initial *Isolation* phase saw requests for online board and card games such as chess, ludo, and poker, there was a change in these requests during the *Normalization* phase. Students either sought a personal connection over these topics or expressed the need for a community with this interest. For example, “Online classes make friendships so hard, anyone wanna join a gaming group?” While, the *Vaccination* phase continued seeing the need for online support and community; the focus shifted towards more in-person and physical settings. Students were willing to meet in-person and also host smaller groups to meet in person for board game nights and other events, indicating a transition from online to in-person interactions. For instance, “My fiancée and I started school right before COVID hit. It has been hard to make new friends during the age of Zoom University. We have a Discord server to hang out and play video games with friends. We’d also love to meet new friends.”

Academic Topics. The academic topics consisted of discussions on college admissions and the dropping of courses. In terms of the relative frequencies of occurrences, Figure A3 (f-h) shows a cyclical trend and increase in frequency in line with the college admission cycles and semester deadlines. In the fall of 2020 (around August 2020), there was immense confusion and worry regarding admission decisions and deferrals. Students were not only concerned about the nature of the classes going into the next semester but also inquired about their options for deferral.

Online Classes. This was a key topic as the concept of online classes spiked with the advent of the pandemic. During the *Isolation* phase, discussions included speculations about the reopening of college, as most students underestimated the timeline for normalcy to kick in. There were also strong opinions about coping up with online classes as opposed to thriving in them, such as, “Do you guys think classes will be back to normal by fall 2020?. Will it continue online?” Once the *Normalization* phase started, the discussions became more nuanced with discussions on remote learning, flexibility, and safety protocols coming into play. By this time, the students not only gathered sufficient experiences of both pros and cons of remote learning, they also expressed various opinions on this. The onset of the *Vaccination* phase saw a divergence in opinions regarding the revival of in-person classes. Students who opined in favor of in-person classes mentioned the vaccine as a reason for the normalcy in the state of affairs, such as, “Most of my fall classes are in person, and I’m one of the few that wants to come to campus for fall, or at least I’m comfortable with current covid conditions, given the vaccine.” In contrast, other students were concerned about in-person classes because of health concerns despite the vaccine, such as, “I want online classes. More than half a million people died in the US because of covid. But people still want in-person classes, not even hybrid. Don’t forget we are still in the middle of PANDEMIC.”

Admissions. The *Normalization* phase saw a surge in the posts on adjusting to the new policies. For example, the cancellation of the SAT for several admission decisions was discussed, and how that affected their renewed application. An individual posted, “I got a 1200 on SAT, which I am very unhappy with. I can’t take it anymore because everything is closed due to COVID. With a lower-end GPA, do I risk it and submit the 1200 for my application?” However, this trend was less prevalent in the *Vaccination* phase and shifted to the post-*Normalization* topics like active campus participation. Post the vaccines, students again began focusing on

topics related to normalcy, like student elections and active campus participation, such as, “I’m interested in rushing a sorority this fall, but I’m worried about rushing as a sophomore. Last semester was only covid. Would it be difficult getting into a sorority as a sophomore?”

We also find that during the pandemic, some colleges included a mandatory essay on “COVID-19” during the admissions process for assessing how the pandemic affected the students. Finally, the academic rigor coupled with the stress of attending classes from home during the pandemic brought additional challenges. For those planning to take a break or drop some courses for a reduced course load, the processes and the repercussions were not clear. Many such students sought support on the college subreddits.

Health & Wellbeing Topics. The college subreddits enabled the students to express their physical and mental health. Post the isolation phase and especially in the vaccination phase, health-related discussions significantly surged in the *Treatment* compared to the *Control* period (Figure A3) (i,m). Discussions on physical health gained more prominence over the course of the pandemic. We observe that students resorted to remote classes for yoga and fitness. The focus shifted from in-person to remote and again back towards in-person settings over the pandemic phases. In fact, we saw conversations about a more hybrid or community-based need for fitness, especially for gymming and sports, where students recognized the shortcomings of online methods for achieving their fitness goals. They sought the student communities to find accountability partners. On mental health, the students openly expressed discomfort with the current mode of classes as well as sought options for online therapy. Their queries also sometimes sought information on managing the finances of treatment and therapy—if these would be covered under the tuition or insurance.

Fitness. Due to restrictions, the *Isolation* phase saw many posts focused on online and socially distanced options related to fitness and mental health. Whatever could be demonstrated and performed online, went the zoom classes route to mobilize the community and were quick to realize the extra importance of this topic in the isolation phase, for example the meditation clubs of the college actively advertised new enrollments. For instance, “Hey guys, I am a regular member of the meditation club, and we meet twice. I just learned that they are holding online meditation sessions for us.” In the *Normalization* phase, fitness-driven discussion shifted to a more collaborative and community-based topic. There was a pronounced change to online as the new normal. Students not only saw online activities and community-building as a new normal, but they also got comfortable with the mix of online and offline, with most posts targeting both kinds of audiences for their accountability, e.g., keywords such as “fitness buddies” emerged more frequently.

Sadness & Grief. The main concern related to this topic was that of the students missing their graduation or losing a loved one. Having spent a significant time of their youth on campus, students took to the subreddits to express their angst and lack of closure due to the COVID crisis, e.g., “I’m fucking sad. Is anyone else extremely sad that there’s no graduation? But I’m just upset at the situation.”

COVID-Induced Topics. These topics explicitly emerged during the pandemic, such as regarding masks, vaccines, and international travel-related regulations and restrictions.

International Travel. The pandemic affected international students in the US. Several aspects of travel regulations,

visa requirements, academic uncertainties, and the fear of infection were discussed. This topic also drew interest from prospective students planning to pursue their higher studies in the US and current students who were unclear about the future of classes and their specific regulations. Several posts sought support and connections from students in similar situations, e.g., “What happens to international students in the US? One rule concerns online classes and a limit to what can be taken. Can we get an update regarding that? I am sorry I am so anxious over this?”

Masks. This was a popular topic of discussion that evolved into multiple perspectives as the pandemic spread. As the pandemic started, this topic saw confusion and surprises related to pricing and forced policies, and also garnered interest from students who feared being judged for wearing a mask, for e.g., “Is wearing a mask a good idea on campus? My family is very concerned, and my mom asks me to wear a mask on campus. However, I wonder whether I will be judged.” Interestingly, many opinions on masks in the *Normalization* phase revolved around students expressing their right to not wear the mask, e.g., “Why are face masks still required on campus? Cases are dropping dramatically, and so many of us have already had covid or gotten the vaccine. At this point, it’s no longer about health and safety; it’s now only about power and control. It’s time to take back our damn freedom as students.”

Vaccine. Vaccine discussions were centered around how the new normal might look related to the post-vaccinated world. In the *Isolation* phase, students took the vaccine as a certain event, and the focus of this topic seemed to be around speculations of a post-vaccine world. However, the *Normalization* phase saw discussions on the logistics of the vaccine delivery and the science behind it. The discussions revolved around the various institutional channels for administering the vaccines as well as posts on the viability of its working, e.g., “Any info on when will students be offered vaccine through the university? It might be in the summer, but what about students who aren’t registered during the summer?”

COVID-19 Testing. Students were initially concerned with the availability and the booking of testing kits, especially using existing health insurance programs. In some initial phases, there was also a question of whether and where they were available, e.g., “I keep seeing posts about people getting their covid testing kits. How do you get the kits? I ask coz I want to head to my off-campus apartment around 10th so that I can meet with my advisor on campus a bit later.” In the *Normalization* period, there were discussions surrounding the random testing rule in several campuses, such as, “I am currently taking all of my classes remotely at home but was sent an email to be tested for covid within 48 hours. Is anyone else who is taking classes remotely selected for it?”

6 Discussion and Conclusion

Our study leveraged college subreddit data of 173 U.S. college campuses to assess the mental health effects of the COVID-19 pandemic on college students. We conducted a quasi-experimental study to examine how social media expressions over a 17-month period post-pandemic (*Treatment*) compared against a similar pre-pandemic period (*Control*). We conducted causal impact analyses on symptomatic mental health expressions of depression, anxiety, stress, and suicidal ideation, and on topics of expressions using BERTopic models. We measured the shifts in social media expressions over three phases of the pandemic (*Isolation*,

Normalization, and *Vaccination* phases). We find that mental health expressions spiked in the *Isolation* phase, gradually normalized over the *Normalization* phase, but again increased in the *Vaccination* phase, and that discussions spanning social, academic, and health, as well as COVID-induced topics, shifted over the course of the pandemic.

Our work builds on and contributes to the literature on social media, college students' mental health (Saha and De Choudhury 2017; Saha et al. 2022), and the COVID-19 pandemic (Guntuku et al. 2020; Saha et al. 2020). We note that our study is a first of its kind in looking into a much longer time-frame (e.g., we also look into the *Vaccination* phase, which was unexplored prior to our work). Methodologically, a quasi-experimental approach grounded in causal inference (Rubin 2005) strengthens our study by mitigating confounds, complemented by qualitative interpretations for deeper insights. In addition, linguistic analyses provide deeper insight beyond lexical and psycholinguistic (LIWC) expressions through symptomatic mental health classifiers. We also examined topics specifically relevant to the mental wellbeing of college students, providing more fine-grained insights compared to Ashokkumar and Pennebaker (2021), which takes a broader, geography-based approach. The insights contribute to an empirical understanding of how the pandemic impacted college students' mental health and how they coped with the changing circumstances as the pandemic progressed. The novelty of our paper lies in its specific focus on the college student demographic, longer timeframe, and methodological framework. The finer insights provided by our study can be more actionable for better preparedness and informed decision-making in future crises that affect college students.

Practical and Design Implications

The computational approaches developed in this work can be adopted in several other contexts and settings, e.g., how COVID-19 or another crisis or external intervention affects the wellbeing of a community. Additionally, these approaches can be leveraged in designing tools to help monitor and manage mental health on college campuses. For instance, these tools can apply the methods outlined in this paper to measure campus mental health at scale and in real-time, providing insights that support timely and tailored interventions. This can be in the form of dashboards (Yoo and De Choudhury 2019) through which campus administrators can monitor real-time mental health pulse and concerns of college students. Our work emphasizes the importance of recognizing temporal and protracted effects—not all mental health consequences manifest immediately; some may unfold gradually over time. In our case, we observed the mental health symptomatic expressions shifted from an increase (in *Isolation* phase) to a decrease (in *Normalization* phase) and again an increase (in *Vaccination* phase). These shifts in temporal dimension underscore the need for ongoing support mechanisms to address the evolving mental health needs of college students post-crisis.

While the COVID-19 pandemic may have “ended” in some regard, this does not mark the end of mental health challenges for college campuses. It is imperative to consider the enduring relevance of our findings in the context of potential future crises, recognizing that pandemics or crises are not isolated occurrences. Further, the spillover effects of crises, as observed in the case of COVID-19, are noteworthy. Beyond its status as a public health crisis, COVID-19 had far-reaching consequences, impacting students' learning

outcomes and overall college experiences—as revealed in our findings on how the topics of discussions varied through the pandemic. These observations underscore the interconnectedness of several aspects of students' lives and the need for comprehensive crisis management strategies that go beyond the immediate health implications.

Our findings highlight the fragility of student mental health in the face of crises—supporting and complementing prior research about college students' mental concerns during the pandemic through surveys (Liu et al. 2020; Wang et al. 2020) and interviews (George and Thomas 2021). Our study made observations on the longer-term evolution and effects. We found that even though the *Isolation* phase saw a significant spike compared to the immediate few pre-pandemic months—the students' mental health concerns were not significantly different from other academic years. This aligns with prior research that college students' mental health needs are not just pandemic-specific but also a general trend in other academic years (Eisenberg et al. 2009). This is especially significant as it highlights a case of when chronic stress was acutely impacted by stressors induced by the pandemic. Such instances are known to be very detrimental to the mental health (Schneiderman et al. 2005). For example, based on our observations, during the initial phase of an emerging public health threat, interventions could prioritize anxiety-reducing communications and resources. During an isolation phase, support systems can focus on mitigating loneliness and fostering virtual community engagement. The normalization phase can then involve strategies to help students adapt to a new routine while addressing any lingering psychological impacts.

Our findings signal the emergence of psychological resilience over time—mental health concerns slowly subsided during the *Normalization* period, and usual conversations resurfaced over time. This shows how college students navigated through the changing circumstances, such as remote classes, lack of in-person events, and financial and personal hurdles—as also evident from our topic analyses. Given the critical importance of community resilience to crisis events (Sharma and Yukhymenko-Lescroart 2022), this work informs an opportunity for college campuses to consider incorporating resilience as a formal learning outcome—i.e., how can college campuses foster resilience in students, recognizing it as an integral component of their holistic development? This also raises the question of what it would mean for colleges to accommodate resiliency as part of learning outcomes actively. Integrating programs and initiatives that explicitly aim to nurture and enhance resilience could be a transformative approach, preparing students academically and emotionally for future challenges.

Salimi et al. (2023) noted the importance of expanding campus services during the pandemic—our study can help identify the specific stressors that students face during crises—crucial for developing targeted interventions and support systems. Equally important is recognizing aspects that provide students with strength and joy. For instance, we found the need for social connectedness and activities that help students engage beyond academics, such as board games, outdoor dining, and social events. In fact, while social isolation is detrimental to mental health (Loades et al. 2020), digital technologies can help reduce students' isolation (Kelly et al. 2021). This provokes us to think about whether we can integrate these positive elements into resiliency programs, offering students strategies to navigate future crises with a sense of empowerment and wellbeing.

This research reveals the key role played by social media as a platform for college students to stay connected and seek support during the pandemic. This highlights the need to continue designing and improving these platforms to serve students' needs better. We have already seen on-line learning and remote socialization technologies growing in the last few years (Hoefler et al. 2022). These platforms can also incorporate peer and expert support by involving counselors and administrators in these communities. For example, students were concerned about the uncertainties and the lack of clarity of information. In addition to the personal experiences shared by the students, if the moderators could sort and share the questions with college stakeholders, this would benefit the whole community. Again, involving college stakeholders will also make them aware of the student concerns and proactively act upon them to mitigate them.

Limitations and Future Directions

We acknowledge that our study has limitations, which also suggest promising future directions. For instance, our data suffers from self-selection bias—we are only able to study those who express themselves on these platforms. Prior work revealed that college subreddit data is representative of college campuses (Bagroy et al. 2017) and the construct validity with the mental health of college communities (Saha et al. 2022). Again, while we adopt a causal-inference approach and mitigate the temporal confounds, we cannot claim “true causality”—there is a lack of “true counterfactual”, or how the community would have behaved in the same time period without the occurrence of a pandemic. Also, our work is cross-sectional and gives insights into how the community’s mental health evolved over the course of the pandemic. Future work can conduct longitudinal studies on several individuals to see how the shifts in mental health. These studies can also incorporate complementary sources of behavioral and mental wellbeing data through surveys, ecological momentary assessments, and passive sensing (Nepal et al. 2022). Further, we cannot claim the clinical validity of our study—which can be overcome in future research by incorporating clinical datasets as well as expert inputs. Our work used symptomatic mental health classifiers, which are based on using Reddit data from communities focusing on mental health and a diversity of communities. While these classifiers have been validated through human experts in prior work (Saha et al. 2019), they are not fool-proof. For example, (Andy 2021) noted individuals could post mental health concerns on generic subreddits—such expressions can lead to inaccurate predictions in the classifiers. Therefore, future research can also incorporate more sophisticated machine learning algorithms, and study other sources of data beyond Reddit, which are popular among college students (e.g., SnapChat, TikTok, etc.).

References

Andy, A. 2021. Understanding user communication around loneliness on online forums. *PloS one*, 16(9): e0257791.

Ashokkumar, A.; and Pennebaker, J. W. 2021. Social media conversations reveal large psychological shifts caused by COVID-19's onset across US cities. *Science advances*, 7(39): eabg7843.

Bagroy, S.; Kumaraguru, P.; and De Choudhury, M. 2017. A Social Media Based Index of Mental Well-Being in College Campuses. In *CHI*.

Basu, T. 2020. The coronavirus pandemic is a game changer for mental health care. *Humans and Technology*, 20.

Blei, D. M.; Ng, A. Y.; and Jordan, M. I. 2003. Latent dirichlet allocation. *JMLR*, 3(Jan): 993–1022.

Brodersen, K. H.; Gallusser, F.; Koehler, J.; Remy, N.; and Scott, S. L. 2015. Inferring causal impact using Bayesian structural time-series models.

Chrikov, I.; Soria, K. M.; Horgos, B.; and Jones-White, D. 2020. Undergraduate and graduate students' mental health during the COVID-19 pandemic.

Ciotti, M.; Ciccozzi, M.; Terrinoni, A.; Jiang, W.-C.; Wang, C.-B.; and Bernardini, S. 2020. The COVID-19 pandemic. *Critical reviews in clinical laboratory sciences*, 57(6): 365–388.

Copeland, W. E.; McGinnis, E.; Bai, Y.; Adams, Z.; Nardone, H.; Devadanam, V.; Rettew, J.; and Hudziak, J. J. 2021. Impact of COVID-19 pandemic on college student mental health and wellness. *Journal of the American Academy of Child & Adolescent Psychiatry*, 60(1): 134–141.

Cucinotta, D.; and Vanelli, M. 2020. WHO declares COVID-19 a pandemic. *Acta bio medica: Atenei parmensis*, 91(1): 157.

Culotta, A. 2014. Estimating county health statistics with Twitter.

Das Swain, V.; Ye, J.; Ramesh, S. K.; Mondal, A.; Abowd, G. D.; De Choudhury, M.; et al. 2024. Leveraging Social Media to Predict COVID-19–Induced Disruptions to Mental Well-Being Among University Students: Modeling Study. *JMIR Formative Research*.

De Choudhury, M.; and De, S. 2014. Mental health discourse on reddit: Self-disclosure, social support, and anonymity. In *ICWSM*.

De Choudhury, M.; Kiciman, E.; Dredze, M.; Coppersmith, G.; and Kumar, M. 2016. Discovering shifts to suicidal ideation from mental health content in social media. In *CHI*.

De Choudhury, M.; Monroy-Hernandez, A.; and Mark, G. 2014. Narco emotions: affect and desensitization in social media during the mexican drug war. In *CHI*, 3563–3572. ACM.

Eisenberg, D.; Downs, M. F.; Golberstein, E.; and Zivin, K. 2009. Stigma and help seeking for mental health among college students. *Medical Care Research and Review*, 66(5): 522–541.

Ellison, N. B.; Steinfield, C.; and Lampe, C. 2007. The benefits of Facebook “friends”: Social capital and college students' use of online social network sites. *JCMC*.

Geburu, T.; Morgenstern, J.; Vecchione, B.; Vaughan, J. W.; Wallach, H.; Iii, H. D.; and Crawford, K. 2021. Datasheets for datasets. *Communications of the ACM*.

George, G.; and Thomas, M. R. 2021. Quarantined effects and strategies of college students—COVID-19. *AEDS*.

Glasgow, K.; Fink, C.; and Boyd-Graber, J. L. 2014. “Our Grief is Unspeakable”: Automatically Measuring the Community Impact of a Tragedy. In *ICWSM*.

Grootendorst, M. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv preprint arXiv:2203.05794*.

Guntuku, S. C.; Sherman, G.; Stokes, D. C.; Agarwal, A. K.; Seltzer, E.; Merchant, R. M.; and Ungar, L. H. 2020. Tracking mental health and symptom mentions on Twitter during COVID-19. *J. Gen. Intern. Med.*

Hoefler, G.; Massachi, T.; Xu, N. G.; Nugent, N.; and Huang, J. 2022. Bridging the Social Distance: Offline to Online Social Support during the COVID-19 Pandemic. *PACM HCI*, (CSCW2).

Imbens, G. W.; and Rubin, D. B. 2015. *Causal inference in statistics, social, and biomedical sciences*. Cambridge.

Jha, I. P.; Awasthi, R.; Kumar, A.; Kumar, V.; and Sethi, T. 2021. Learning the mental health impact of COVID-19 in the United States with explainable artificial intelligence: observational study. *JMIR mental health*.

Keith, K.; Jensen, D.; and O'Connor, B. 2020. Text and Causal Inference: A Review of Using Text to Remove Confounding from Causal Estimates. In *ACL*.

Kelly, R. M.; Cheng, Y.; McKay, D.; Wadley, G.; and Buchanan, G. 2021. “it's about missing much more than the people”: how students use digital technologies to alleviate homesickness. In *CHI*.

Kiciman, E.; Counts, S.; and Gasser, M. 2018. Using Longitudinal Social Media Analysis to Understand the Effects of Early College Alcohol Use. In *ICWSM*, 171–180.

- Kim, H.; Rackoff, G. N.; Fitzsimmons-Craft, E. E.; Shin, K. E.; Zainal, N. H.; Schwob, J. T.; Eisenberg, D.; Wilfley, D. E.; Taylor, C. B.; and Newman, M. G. 2022. College mental health before and during the COVID-19 pandemic: Results from a nationwide survey. *COTR*.
- Kontoangelos, K.; Economou, M.; and Papageorgiou, C. 2020. Mental health effects of COVID-19 pandemic: a review of clinical and psychological traits. *Psychiatry investigation*, 17(6): 491.
- Kumar, N.; Corpus, I.; Hans, M.; Harle, N.; Yang, N.; McDonald, C.; Sakai, S. N.; et al. 2022. COVID-19 vaccine perceptions in the initial phases of US vaccine roll-out: an observational study on reddit. *BMC public health*.
- LeViness, P.; Bershady, C.; and Gorman, K. 2017. The association for university and college counseling center directors annual survey.
- Lin, Y.-R.; and Margolin, D. 2014. The ripple of fear, sympathy and solidarity during the Boston bombings. *EPJ Data Science*.
- Liu, J.; Zhu, Q.; Fan, W.; Makamure, J.; Zheng, C.; and Wang, J. 2020. Online mental health survey in a medical college in China during the COVID-19 outbreak. *Frontiers in psychiatry*.
- Loades, M. E.; Chatburn, E.; Higson-Sweeney, N.; Reynolds, S.; Shafran, R.; Brigden, A.; Linney, C.; McManus, M. N.; Borwick, C.; and Crawley, E. 2020. Rapid systematic review: the impact of social isolation and loneliness on the mental health of children and adolescents in the context of COVID-19. *J. Am. Acad. Child Adolesc. Psychiatry*.
- Mark, G.; Bagdouri, M.; Palen, L.; Martin, J.; Al-Ani, B.; and Anderson, K. 2012. Blogs as a collective war diary. In *CSCW*.
- McDowall, D.; McCleary, R.; and Bartos, B. J. 2019. *Interrupted time series analysis*. Oxford University Press.
- McGinty, E. E.; Presskreischer, R.; Han, H.; and Barry, C. L. 2020. Psychological Distress and Loneliness Reported by US Adults in 2018 and April 2020. *JAMA*.
- Melcher, J.; Lavoie, J.; Hays, R.; D’Mello, R.; Rauseo-Ricupero, N.; Camacho, E.; Rodriguez-Villa, E.; Wisniewski, H.; Lagan, S.; Vaidyam, A.; et al. 2023. Digital phenotyping of student mental health during COVID-19: an observational study of 100 college students. *J. Am. Coll. Health*.
- Miller, G. 2020. Social distancing prevents infections, but it can have unintended consequences. *Science*.
- Mishra, N. P.; Das, S. S.; Yadav, S.; Khan, W.; Afzal, M.; Alarifi, A.; Ansari, M. T.; Hasnain, M. S.; Nayak, A. K.; et al. 2020. Global impacts of pre-and post-COVID-19 pandemic: Focus on socio-economic consequences.
- Nepal, S.; Wang, W.; Vojdanovski, V.; Huckins, J. F.; Dasilva, A.; Meyer, M.; and Campbell, A. 2022. COVID student study: A year in the life of college students during the COVID-19 pandemic through the lens of mobile phone sensing. In *CHI*.
- Palen, L. 2008. Online social media in crisis events. *Educause Quarterly*.
- Pennebaker, J. W.; Chung, C. K.; Frazee, J.; Lavergne, G. M.; and Beaver, D. I. 2014. When small words foretell academic success: The case of college admissions essays. *PLoS one*, 9(12): e115844.
- Pokhrel, S.; and Chhetri, R. 2021. A literature review on impact of COVID-19 pandemic on teaching and learning. *Higher education for the future*.
- Rankin, C. H.; Abrams, T.; Barry, R. J.; Bhatnagar, S.; Clayton, D. F.; Colombo, J.; Coppola, G.; Geyer, M. A.; Glanzman, D. L.; Marsland, S.; et al. 2009. Habituation revisited: an updated and revised description of the behavioral characteristics of habituation. *Neurobiol. Learn. Mem.*
- Rubin, D. B. 2005. Causal inference using potential outcomes: Design, modeling, decisions. *J. Am. Stat. Assoc.*
- Saha, K.; Chandrasekharan, E.; and De Choudhury, M. 2019. Prevalence and psychological effects of hateful speech in online college communities. In *ICWSM*.
- Saha, K.; and De Choudhury, M. 2017. Modeling stress with social media around incidents of gun violence on college campuses. *PACM HCI, (CSCW)*.
- Saha, K.; Gupta, P.; Mark, G.; Kiciman, E.; and De Choudhury, M. 2024. Observer Effect in Social Media Use. In *CHI*.
- Saha, K.; Liu, Y.; Vincent, N.; Chowdhury, F. A.; Neves, L.; Shah, N.; and Bos, M. W. 2021. Adverting matters: Examining user ad consumption for effective ad allocations on social media. In *CHI*.
- Saha, K.; Sugar, B.; Torous, J.; Abrahao, B.; Kiciman, E.; and De Choudhury, M. 2019. A Social Media Study on the Effects of Psychiatric Medication Use. In *ICWSM*.
- Saha, K.; Torous, J.; Caine, E. D.; and De Choudhury, M. 2020. Psychosocial effects of the COVID-19 pandemic: large-scale quasi-experimental study on social media. *JMIR*.
- Saha, K.; Weber, I.; and De Choudhury, M. 2018. A Social Media Based Examination of the Effects of Counseling Recommendations After Student Deaths on College Campuses. In *ICWSM*.
- Saha, K.; Yousuf, A.; Boyd, R. L.; Pennebaker, J. W.; and De Choudhury, M. 2022. Social media discussions predict mental health consultations on college campuses. *Scientific reports*.
- Salimi, N.; Gere, B.; Talley, W.; and Iriooogbe, B. 2023. College students mental health challenges: Concerns and considerations in the COVID-19 pandemic. *Journal of College Student Psychotherapy*.
- Schneiderman, N.; Ironson, G.; and Siegel, S. D. 2005. Stress and health: psychological, behavioral, and biological determinants. *Annu. Rev. Clin. Psychol.*, 1: 607–628.
- Sharma, G.; and Yukhymenko-Lescroart, M. A. 2022. Life purpose as a predictor of resilience and persistence in college students during the COVID-19 pandemic. *CSR*.
- Starbird, K.; Palen, L.; Hughes, A. L.; and Vieweg, S. 2010. Chatter on the red: what hazards threat reveals about the social life of microblogged information. In *CSCW*, 241–250. ACM.
- Statista. 2021. Percentage of U.S. adults who use Reddit as of February 2021, by age group.
- Talevi, D.; Succi, V.; Carai, M.; Carnaghi, G.; Faleri, S.; Trebbi, E.; di Bernardo, A.; Capelli, F.; and Pacitti, F. 2020. Mental health outcomes of the CoViD-19 pandemic. *Rivista di psichiatria*.
- Torous, J.; Myrick, K. J.; Rauseo-Ricupero, N.; et al. 2020. Digital mental health and COVID-19: using technology today to accelerate the curve on access and quality tomorrow. *JMIR mental health*.
- Tourangeau, R.; Rips, L. J.; and Rasinski, K. 2000. *The psychology of survey response*.
- Veselovsky, V.; and Anderson, A. 2023. Reddit in the Time of COVID. In *ICWSM*.
- Vowels, L. M.; Vowels, M. J.; Carnelley, K. B.; Millings, A.; and Gibson-Miller, J. 2023. Toward a causal link between attachment styles and mental health during the COVID-19 pandemic. *Br. J. Clin. Psychol.*
- Wang, X.; Hegde, S.; Son, C.; Keller, B.; Smith, A.; and Sasangohar, F. 2020. Investigating mental health of US college students during the COVID-19 pandemic: Cross-sectional survey study.
- World Health Organization. 2020. WHO Coronavirus Disease (COVID-19) Dashboard.
- Yoo, D. W.; and De Choudhury, M. 2019. Designing dashboard for campus stakeholders to support college student mental health. In *PervasiveHealth*.
- Yuan, Y.; Saha, K.; Keller, B.; Isometsä, E. T.; and Aledavood, T. 2023. Mental health coping stories on social media: a causal-inference study of Papageno effect. In *TheWebConf*.
- Zhou, J.; Saha, K.; Lopez Carron, I. M.; Yoo, D. W.; Deeter, C. R.; De Choudhury, M.; and Arriaga, R. I. 2022. Veteran Critical Theory as a Lens to Understand Veterans’ Needs and Support on Social Media. *PACM HCI, (CSCW1)*.
- Zhunis, A.; Lima, G.; Song, H.; Han, J.; and Cha, M. 2022. Emotion bubbles: Emotional composition of online discourse before and after the COVID-19 outbreak. In *TheWebConf*.

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**
 - (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**
 - (d) Do you clarify what are possible artifacts in the data used, given population-specific distributions? **Yes**
 - (e) Did you describe the limitations of your work? **Yes**
 - (f) Did you discuss any potential negative societal impacts of your work? **Yes**
 - (g) Did you discuss any potential misuse of your work? **Yes**
 - (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**
 - (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)? **No. The algorithms for detecting symptomatic mental health expressions are sensitive and can be misused. Therefore, we will not be releasing these models publicly. However, with proper data sharing protocols, we are open to sharing the models with researchers.**
 - (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? **Yes**
 - (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? **NA**
 - (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)? **No**
 - (e) Do you justify how the proposed evaluation is sufficient and appropriate to the claims made? **Yes**
 - (f) Do you discuss what is "the cost" of misclassification and fault (in)tolerance? **NA**
5. Additionally, if you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
 - (a) If your work uses existing assets, did you cite the creators? **Yes**
 - (b) Did you mention the license of the assets? **NA**
 - (c) Did you include any new assets in the supplemental material or as a URL? **NA**
 - (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? **Yes**
 - (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? **Yes**
 - (f) If you are curating or releasing new datasets, did you discuss how you intend to make your datasets FAIR (see ?)? **NA**
 - (g) If you are curating or releasing new datasets, did you create a Datasheet for the Dataset (see Gebru et al. (2021))? **NA**
 6. Additionally, if you used crowdsourcing or conducted research with human subjects...
 - (a) Did you include the full text of instructions given to participants and screenshots? **NA**
 - (b) Did you describe any potential participant risks, with mentions of Institutional Review Board (IRB) approvals? **NA**
 - (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? **NA**
 - (d) Did you discuss how data is stored, shared, and de-identified? **NA**

Ethical Statement

This research employs data-driven approaches for studying the mental health impacts of crises, such as the COVID-19 pandemic on college students. Therefore, given the sensitivity of the problem and the population, the ethical implications of this research are multifaceted and should be carefully considered. The machine learning models of identifying symptomatic mental health expressions on social media can be misused and misappropriated to cause harm to college students. For example, these assessments can be used by targeted advertising, commercial benefits, and insurance companies. We also caution against misinterpretation of these models for diagnostic purposes of mental health conditions of college students. Again, developing tools based on our computational framework would require additional consideration of college students' diverse backgrounds and experiences. It is important to acknowledge how different cultural contexts may influence differences in social media expressions, as well as mental health and coping mechanisms.

It is also important to avoid perpetuating stereotypes and biases about different demographic groups and their concerns, which can be extracted through these approaches.

A Appendix

Topic Modeling

We describe the specific steps employed in topic modeling with BERTopic as below:

Document Embedding: First, we transformed our data into vector embeddings of semantic space, using UKPLab sentence transformers optimized for semantic similarity.

Dimensionality Reduction: The sentence embeddings obtained above are high dimensional which can raise issues with clustering algorithms. Therefore, we used the UMAP algorithm for dimensionality reduction, maintaining both the local and global relationships.

Document Clustering: The next step involves the main clustering, which we conducted using the HDBSCAN algorithm, which is a density-based clustering technique. We used the automatic topic reduction feature.

Bag of Words: The HDBSCAN algorithm outputs the clusters in different densities and shapes, ruling out a centroid-based topical representation. Therefore, we apply a bag-of-words approach based on the count of occurrences in entire cluster of the documents. This representation is normalized to take into account the clusters of different sizes.

Topic Representations: Finally, we obtain the topical representations using the TF-IDF approach, which obtains the relative importance of a term in a cluster contributing to the topic importance. For example, the score of a term i in cluster C is computed as $tf_i \cdot \log(1 + \frac{A}{f})$, where tf_i is the frequency of i in C , f is the frequency of i in the entire corpus, and A is the average number of words per cluster.

Causal Impact Estimations

We plot the causal impact estimation in Figure A4, A5, and A6 for *Isolation*, *Normalization*, and *Vaccination* phase respectively. For a given period and symptomatic expression, there are three individual subplots—1) the predicted model on the top, 2) the point effects in the middle, and 3) the cumulative effect at the bottom. The black line and the blue dotted line in the predicted model plot denote the actual and predicted values of the time series, i.e., the black line denotes the actual values given as observed in the time series post-intervention, and the blue dotted line denotes the values as predicted if the intervention had not taken place. The difference between these two series is plotted in the point effects graph in the middle to observe these differences further. Finally, the cumulative sum of these differences is plotted in the bottom plot.

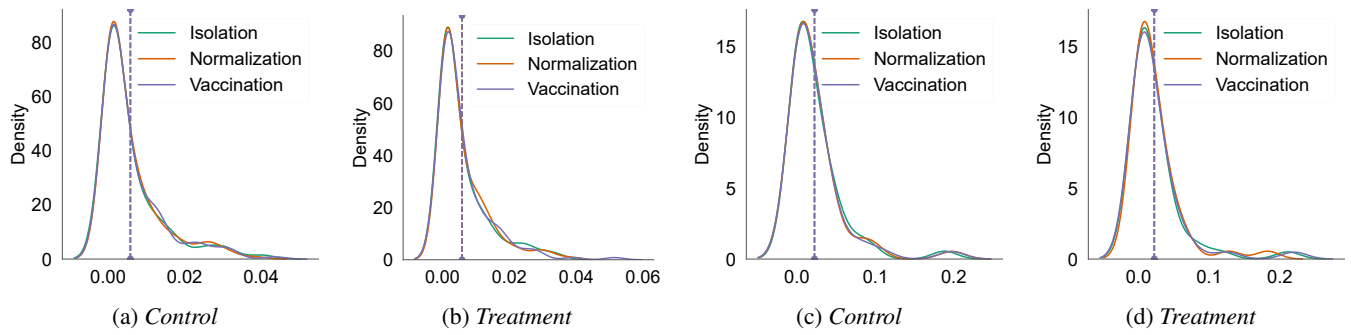


Figure A1: Density distribution of the proportion of data contributed per in the *Control* and *Treatment* periods at (a&b) **sub-reddit**-level and (c&d) **state**-level.

h

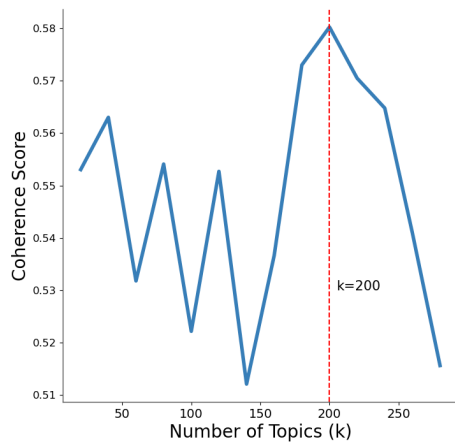


Figure A2: Optimal number of topics ($k=200$) for the BERTopic Model

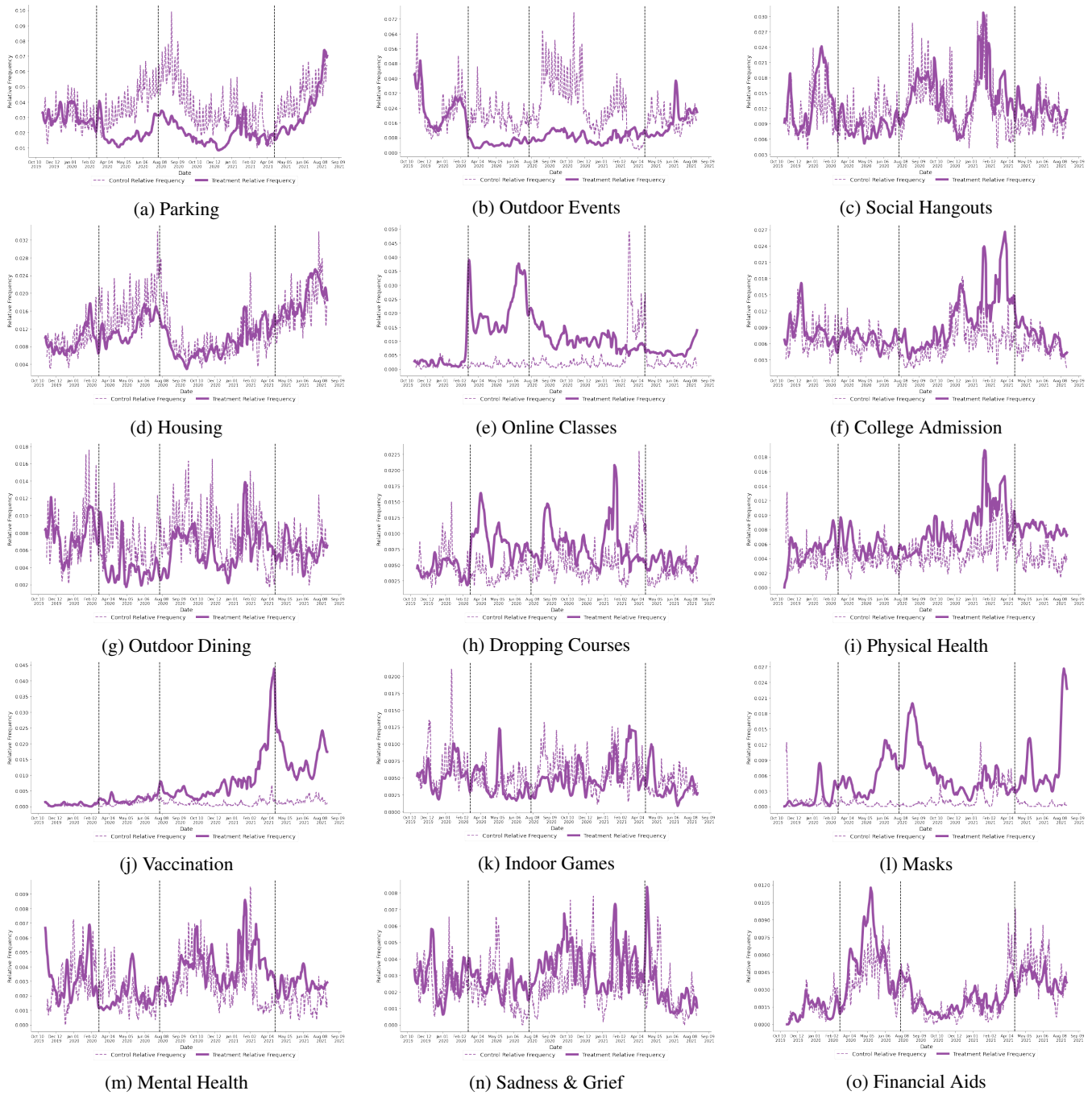
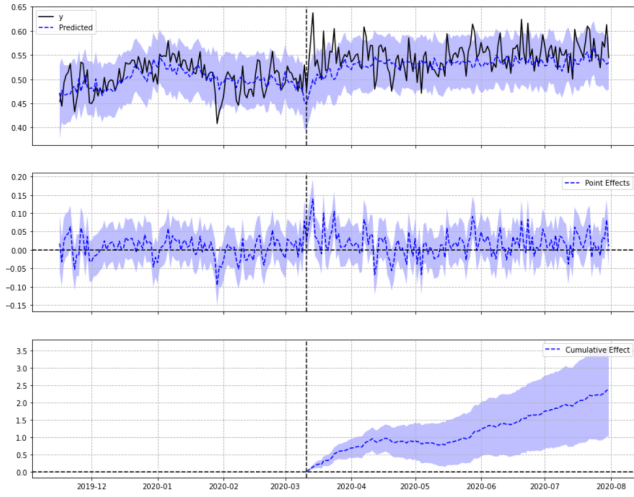
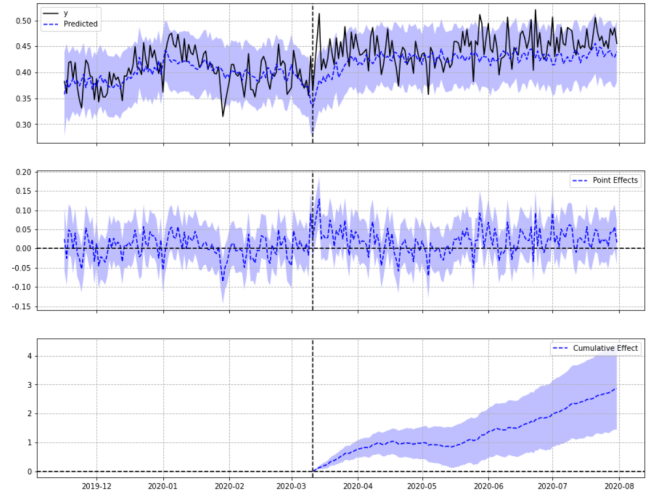


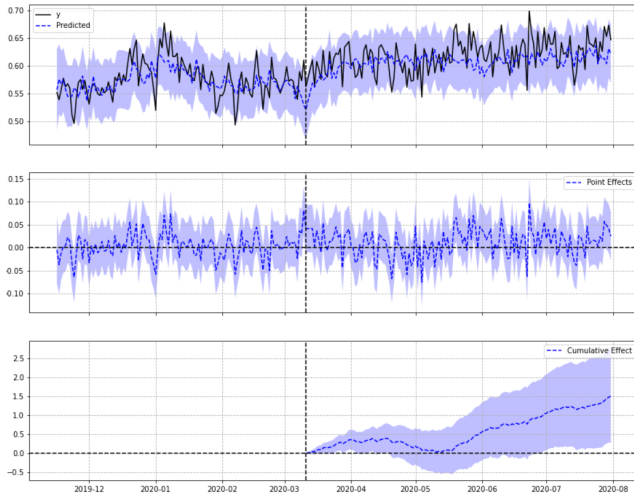
Figure A3: Relative Frequency trends over different phases



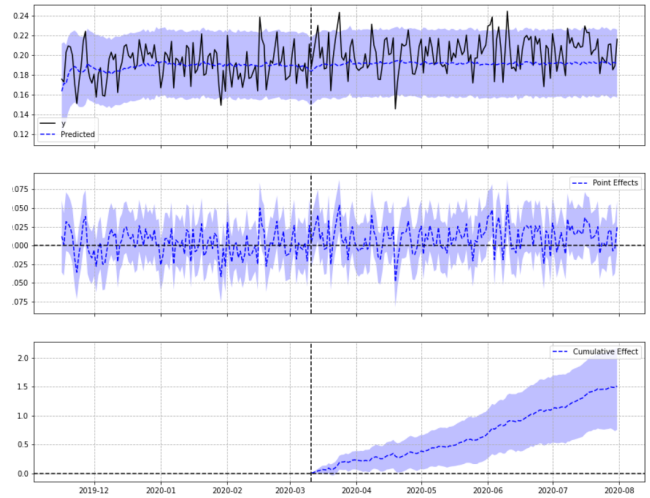
(a) Depression



(b) Anxiety

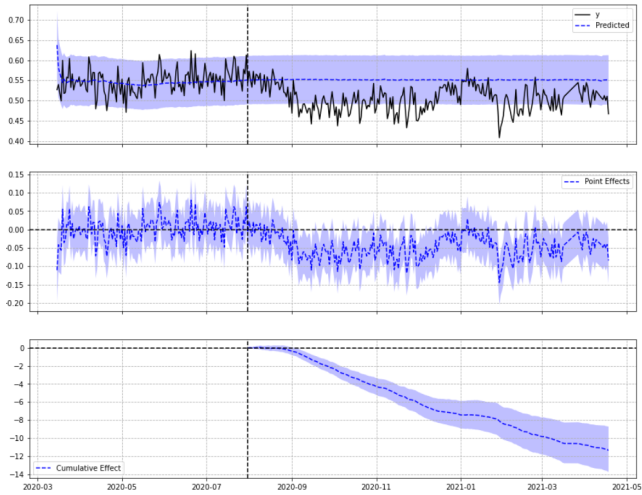


(c) Stress

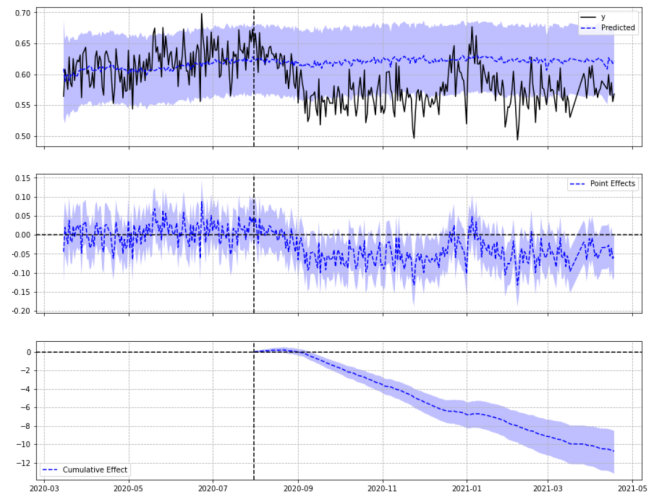


(d) Suicidal Ideation

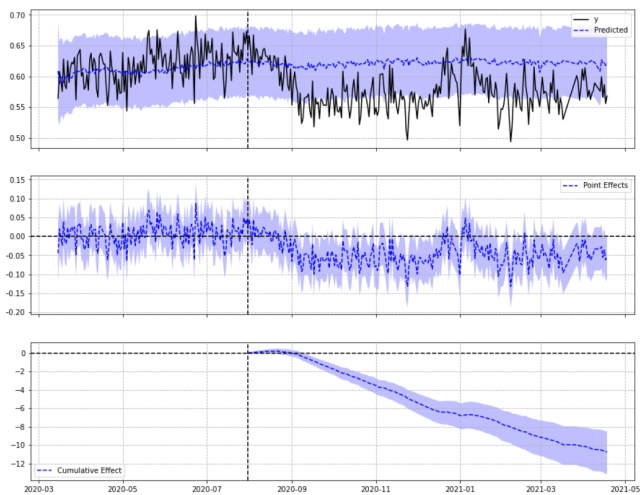
Figure A4: Causal impact estimations for the symptomatic mental health expressions for *Isolation* phase.



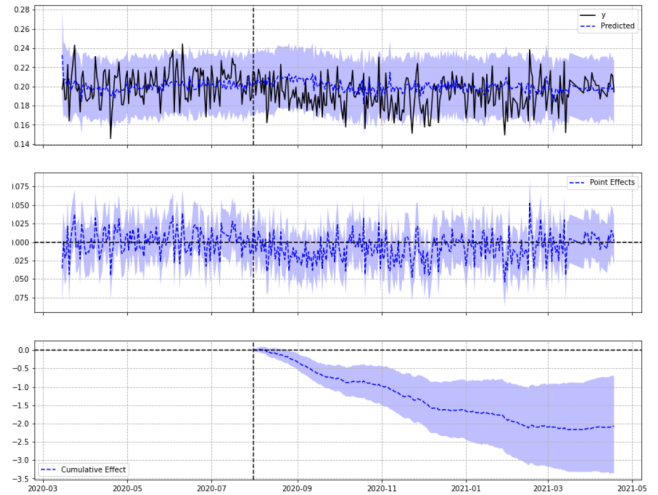
(a) Depression



(b) Anxiety

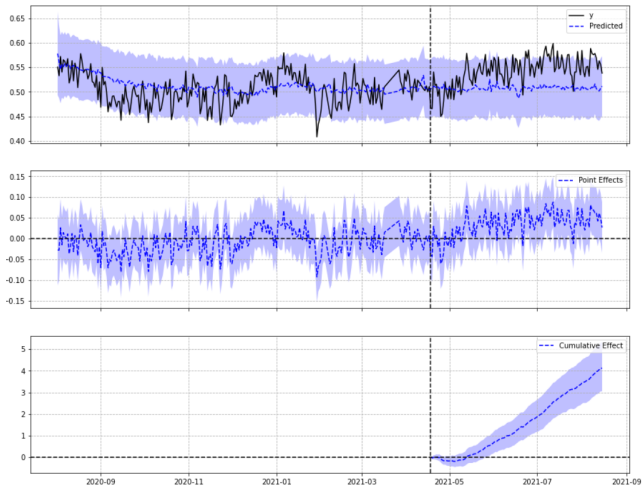


(c) Stress

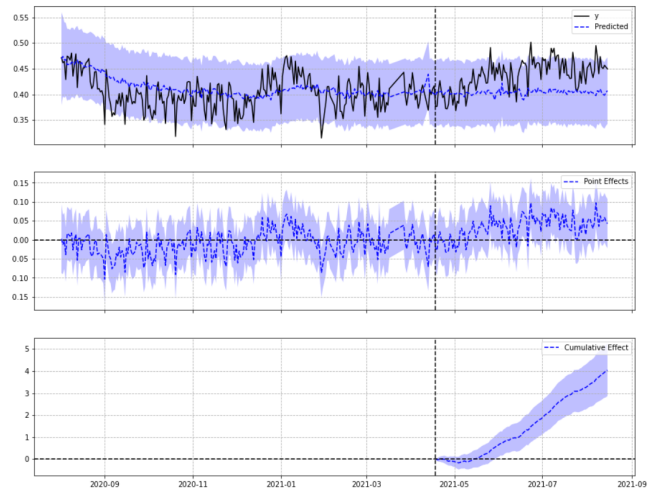


(d) Suicidal Ideation

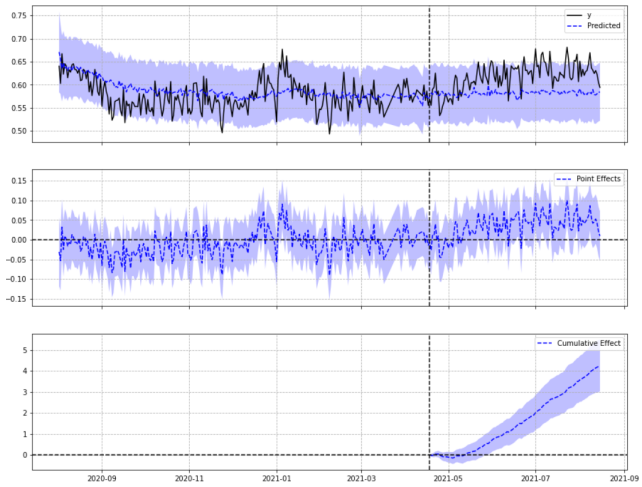
Figure A5: Causal impact estimations for the symptomatic mental health expressions for *Normalization* phase.



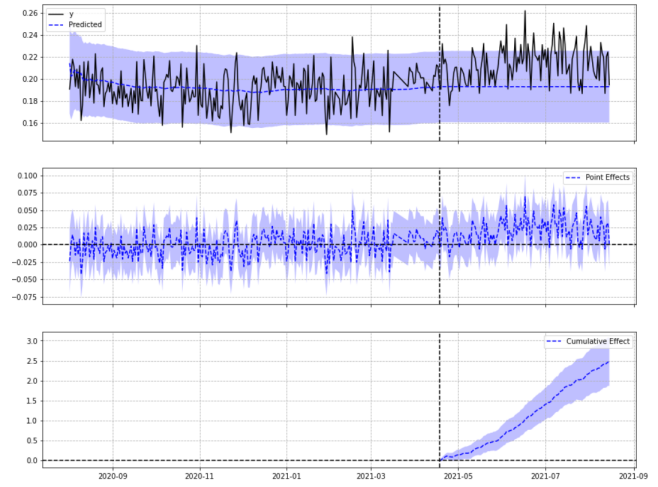
(a) Depression



(b) Anxiety



(c) Stress



(d) Suicidal Ideation

Figure A6: Causal impact estimations for the symptomatic mental health expressions for *Vaccination* phase.

List of all the topics identified in our dataset. The topics with “NA” did not have coherent themes and therefore were not labeled with a theme.

Theme	Topic Keywords
Parking	parking, park, pass, permits, garage, ticket, passes, parked, cars, garages
Outdoor Events	tickets, ticket, selling, sale, stadium, tailgate, tours, tailgating, bought, sold
Physical Readings	textbooks, textbook, books, pdf, adobe, storage, bookstore, copy, isbn, pdfs
Economics	econ, microeconomics, credit, eco, econometrics, major, 10a, majors, 100a, macroeconomics
Exams	calc, exam, mcat, exams, rsos, mcdb, msw, mae, credit, test
Meeting Friends	friends, meet, lonely, dating, friend, social, homeless, group, clubs, meeting
Classroom Response System	iclicker, python, cengage, zot, ecology, noninteger, evolution, fail, zybooks, contradiction
Housing	roommate, roommates, rent, room, apartment, house, bedroom, lease, rooms, beds
Career opportunities	internship, interview, internships, interviews, job, applying, intern, apply, resume, applied
College Mascots	mascot, capes, wolf, hero, wolfie, avengers, warhammer, cape, ridge, war
NA	ochem, ants, tht, 2021, vidos, happy, fuck, domino, efan
Chemistry	chem, chegg, chemistry, che, lab, exam, labs, exams, 1c, cus
Fire alarm	fire, alarm, smell, loud, fireworks, noise, sirens, alarms, smells, emergency
Beverages	coffee, bottle, alcohol, drinking, drink, fountain, marijuana, beer, fountains, cannabis
Course waitlists	waitlist, waitlisted, waitlists, position, wait, spot, waitlisting, seats, spots, waiting
NA	heerf, horns, clusters, stickers, doooooownhorns, transcendentals, handydandy, cluster, runiversityofhouston, coogs
Physical Mail	mail, delivered, mailroom, email, mailing, mailbox, emails, sent, shipped, send
Subleasing	sublease, sublet, subleasing, bedroom, rent, room, furnished, subletting, bath, roommates
Online classes	online, asynchronous, synchronous, webreg, courses, webassign, async, schedule, instruction, sync
NA	epstein, professor, library, transferring, hokie, alumni, professors, teacher, aggies, teachers
Humanities	writing, humanities, journalism, writers, reading, poetry, writer, write, read, journal
Biology	bio, biol, biology, lab, microbi, microbiology, biochem, textbook, bios, biomg
2019	2019, 2019sept, 2019july, thread, proctorio, ucertify, 2019aug, 2019june, 2019may, 2019oct
Biking	bike, bikes, scooter, scooters, biking, bicycle, cycling, bikers, cyclists, rides
Aerospace	aerospace, engineering, drone, engineer, aviation, suspension, electrical, helicopter, pilot, engineers
Pet Animals	dogs, pets, squirrelsdogs, pets, squirrel, pet, squirrels, geese, animal, cats, huskies, animals
College Admits	decision, decisions, appeal, admissions, acceptance, applied, admission, received, accepted, selection
Outdoor sports	soccer, volleyball, hockey, team, sports, games, game, athletics, lacrosse, players
Physics	physics, quantum, 5a, professor, exams, 5c, exam, 7a, calc, study
NA	ta, pokemon, pokmon, raids, tas, pods, panda, coyote, raid, taing
Music	music, piano, headphones, pianos, synth, listen, listening, earbuds, audio, violin
Computer Science	cs, cse, csci, csc, csn, unlv, courses, workload, programming, database
On-Campus Housing	dorms, dorm, tutoring, freshman, classrooms, bunk, tutors, beds, classroom, freshmen
NA	caf826rt1, email, cmsc, chem, lab, isom, runiversityofhouston, exam, mydaq, ece
Languages	linguistics, speaking, arabic, languages, speaker, hindi, english, speak, proficiency, speech
Home Equipments	laundry, fridge, microwave, microwaves, shower, kitchen, wash, dryer, kitchens, washing
NA	curated, outstudying, hmmm, discuss, server, mock, world, event, internships, interesting
Cold Weather	winter, cold, weather, heat, rain, tornado, temperature, freezing, heater, snow
California Universities	uci, cc, ucsc, credits, uc, ucb, uconn, ucsb, ucf, ucsc
Academics	gpa, extracurriculars, sorority, junior, aps, score, rush, honors, varsity, stats
Holidays and Breaks	thanksgiving, break, christmas, sunset, holiday, picnic, holidays, disneyland, disney, calendar
NA	bimm, bild, esa, bicd, bipn, bibc, iowa, admissions, ames, exam
Outdoor dining	dining, food, restaurants, tassel, lunch, taco, restaurant, cafe, tacos, mcdonald
NA	airpods, airpod, swim, fishing, fish, seas, marine, scuba, oceanography, diving
Purdue	purdue, mgmt, purdues, boiler, iu, boilermakers, university, boilerkey, boilermaker, polytechnic
Fee payments	fee, fees, charger, charged, charge, charging, paying, charges, payments, billing
Dropping of Courses	drop, dropping, dropped, adddrop, deadline, credits, add, credit, hours, zoomers
Public transport	bus, buses, shuttle, busses, shuttles, stops, transit, metro, transportation, subway
Exam Prep Websites	chegg, ge, bot, registration, ges, change, leetcode, step, changing, concentrations
NA	horrible, halloween, homeless, costume, haunted, nonprofit, leaders, organization, hate, communities
Standardized Tests	pell, gre, grant, frisbee, test, tests, iss, std, testing, pe
SAT Prep.	megathread, sat, thread, scores, lecture, questions, admissions, prospective, megathreads, threads
Professor Reviews	professor, professors, rate, ratemyprofessor, prof, ratemyprofessors, ratemydormcom, recommendations, rating, ratecollegeclubscom
Degree	diploma, residency, diplomas, dartmouth, degree, graduated, alumni, regents, graduate, bachelors
NA	ra, bme, sas, ras, utsa, rasu, gs, lmao, gsi, utma
Physical Health	yoga, workout, fitness, meditation, exercises, pilates, stretch, stretches, massage, exercise
Laptop Software	laptop, macbook, mac, windows, pc, laptops, computer, monitor, dell, computers
Crime/Cheating	probation, assault, criminal, harassment, crime, archery, gunshots, rape, shooting, assaulted
Positive Results	luck, mood, happy, congrats, emotions, congratulations, moodle, satisfaction, emotion, wishes
Online Gaming	minecraft, server, discord, legends, genshin, java, join, link, lost, cyberpunk
Racism	racism, racist, black, diversity, racial, hispanic, ethnic, african, latino, diverse
Mathematics	math, equations, algebra, calculus, mathematics, calc, credits, credit, courses, course
Refunds	refund, refunds, deposit, stole, refunded, stoles, tuition, received, fee, fees
NA	googleapple, syllabuddycom, calendar, dates, calendars, syllabus, week, schedule, catalog, spring
NA	fake, secret, confessions, reminder, sheldon, funny, favorite, bang, fakes, show
NA	registration, register, grocery, shopping, groceries, bookstore, ordering, stores, store, walmart
International Travel	abroad, french, international, italian, visa, immigrants, immigrant, immigration, europe, foreign

Digital Games	minigames, game, robotics, play, mcthequaddev, minecraft, games, vr, players, dragons
NA	permission, declare, senate, ethics, government, consent, declaring, senator, chancellor, department
Scholarships	scholarship, scholarships, merit, award, received, awards, applied, awarded, expos, nomination
Cheating in Exam	cheating, cheat, integrity, plagiarism, scam, dishonesty, cheated, misconduct, cheaters, phishing
Finance	finance, entrepreneurship, bank, portfolio, investing, investment, business, banking, banks, entrepreneurs
Medicine	premed, med, pharmacy, pre, bsn, opioid, overdose, premeds, prehealth, medicine
LGBTQ	lgbt, gender, lgbtq, gay, sex, queer, sexual, sexuality, transgender, lgbtqia
Vaccination	vaccine, immunization, vaccination, vaccines, vaccinated, covid, quarantine, covid19, vaccinations, immunizations
Grading	grades, grade, transcript, professors, professor, nr, submitted, graded, gpa, grading
Wifi	wifi, ethernet, vpn, connection, router, network, connecting, connected, wireless, networks
COVID tests	testing, covid, test, tested, results, tests, appointment, covid19, weekly, symptoms
Photography	photos, photographer, pictures, camera, photographers, graduation, photo, pics, picture, pic
Examinations	midterm, midterms, fpf, studying, fws, exam, exams, questions, lecture, readings
Indoor Games	chess, poker, games, nintendo, playlist, gaming, game, jeopardy, bingo, chesscom
Leadership	survey, oa, uiuc, uic, responses, leadership, success, questions, nsls, preassessment
Housing	coop, coops, lease, co, housing, communities, gainesville, charter, residents, community
College Members	swipes, swipe, professors, freshman, freshmen, undergraduate, emailing, professor, email, assistant
Statistics	stat, stats, statistics, distribution, distributions, statistical, courses, stat400, regression, stat
NA	deans, wharton, rutgers, abortion, parents, mumps, dad, gpa, college, transcript
NA	nightmare, runescapeampreg, bot, yup, tribot, philosophical, readers, scholarly, fuck, definitive
Student clubs	clubs, join, group, joining, gang, members, meetings, partners, groups, gangs
Walking	commute, traffic, walk, walking, crosswalk, crossing, street, cars, urban, intersection
Degree	honors, college, evaluations, apply, benefits, applying, evaluation, gpa, programs, university
Discussions	discussion, participants, discussions, participate, social, session, project, participating, compensated, participation
Financial	fafsa, aid, tax, taxes, tuition, filing, grant, loans, pay, fax
University Names	northwestern, northeastern, pros, grades, schools, return, college, betatesting, returning, professors
Professors	professors, professor, motivation, procrastination, motivated, lectures, studying, focus, lecture, procrastinating
Delay in Exams	cancelled, canceled, cancel, closed, exam, exams, shut, postponed, cancels, shutdown
Eyewear	glasses, tattoo, laude, honors, goggles, sunglasses, piercing, necklace, piercings, tattoos
Gardening	organic, plants, tree, plant, trees, flowers, arboretum, garden, cactus, blossoms
College Advisors	advisor, advising, advisors, appointment, consulting, appointments, email, emails, adviser, emailed
Orientation	orientation, register, freshman, date, orientations, dates, registration, freshmen, session, registering
Asian languages	chinese, japanese, korean, hong, china, mandarin, japan, asian, hk, thai
East Coast Colleges	cornell, stony, ithaca, cornells, cornellians, rochester, binghamton, ivy, triton, university
Room mates	roommate, roommates, housing, find, finding, selection, dorm, rooms, search, apartment
Masks	masks, mask, face, vaccinated, coverings, masking, vaccination, nose, vaccine, covering
Astronomy	astronomy, astro, conference, seminar, seminars, astrophysics, solar, fellowship, scholars, moon
martial arts	boxing, martial, mma, karate, ufc, taekwondo, fight, fights, judo, kickboxing
Graduation	gown, capstone, cap, hat, graduation, dress, gowns, suit, attire, hats
Psychology	psych, psychology, social, cognitive, behavioral, psy, psyc, lab, psychological, tested
Algebra	algebra, linear, statics, pathways, routines, discrete, pathway, tracking, find, tracing
Rescinded	rescind, deferred, deflation, grade, grades, credits, update, gpa
Fund Raising	donate, volunteer, volunteering, boost, donation, donations, donating, boosters, donors, donor
Grade Rounding	curve, curved, curves, grade, grades, curving, rounding, round, grading, exams
Mobile Technologies	ipad, pencil, phone, iphone, tablet, ipads, chalk, pen, verizon, notebook
University Names	harvard, stanford, tufts, oxford, fordham, yale, university, nyu, essay, alma
Elections	vote, voting, ballot, election, voter, elections, polling, ballots, voted, voters
Wall art	art, paint, painting, wall, wallpaper, walls, graffiti, paintball, decorate, marksmanship
Health Insurance	insurance, waiver, waive, ucship, waived, uc, fee, waiving, medicaid, tuition
Housing/Room	swap, housing, swapping, transfer, transfers, swaps, rooms, tacoma, apartment, roommate
Sports and Games	tennis, esports, play, smash, tournament, tryouts, badminton, spikeball, tournaments, melee
Grades	pnf, gpa, grade, grading, pf, crnc, grades, pnc, transcript, npb
Computer Science courses	cis, cs, majors, major, cse, courses, covid, covid19, cise, majoring
Moving	move, movein, enrollment, date, utc, dates, enroll, appointment, freshmen, freshman
Music Band	band, music, jam, drummer, musicians, jazz, bands, drums, bassist, drum
Funding	grant, certificate, funding, aid, grants, certificates, funds, awarded, loans, scholarshi
Religion	temple, knights, forgiveness, jewish, israel, lord, temples, jews, palestine, peace
Roommates	sleep, roommate, roommates, sleeping, asleep, dream, woke, tired, dorm, dreams
NA	ripplematch, comm, segfault, echols, tmp, covid, companies, resume, mfos, echo
School	nyu, nyggers, nyus, nyc, york, aid, schools, school, nyush, ny, '
Fast Food	pizza, cookies, cake, cream, cookie, donuts, chocolate, krispy, pizzas, pie
Skating	skate, skating, skateboard, skateboarding, longboard, skaters, skates, skateboards, longboarding, collegeboard
Research Studies	survey, study, coping, compensation, restriction, participants, community, experiences, restrictions, challenge
Clothing	shirts, apparel, shoes, sweatshirt, hoodies, store, clothes, sweatshirts, clothing, sweaters
Grades	calculator, gpa, calculators, ti84, ti, titan, calculate, titanium, calculated, count
Religious	byu, church, religion, religious, churches, lds, catholic, mormon, byui, preachers
Logging In	password, clicker, browser, lockdown, login, log, chrome, logging, clickers, account
Grooming	hair, haircut, barber, salon, haircuts, barbers, salons, barbershop, barbershops, dentists
Professors	professor, yang, professors, chen, prof, zhang, teaching, chinese, chungus, kanye
Tutoring	threads, thread, weekly, chem, financial, task, tutor
Jobs	jobs, job, hiring, salary, work, employment, employees, fired, hire, occupational
Discord Chats	discord, join, channels, channel, invite, chat, discords, talk, discuss, friends

Syllabus	syllabus, syllabi, copy, chem, syllabuses, share, access, chapters, professor, prof
Films	film, movie, video, cinema, films, filming, filmmaking, filmmakers, filmed, cinematic
Wallets and Cards	wallet, gift, lost, cards, card, win, raffle, chipotle, leather, finds
Concert and Events	blacksburg, climbing, virginia, boulder, rock, admission, summit, climb, christiansburg, concert
NA	geog, gis, gi, geol, geology, gies, geo, geogator, recommendations, courses
Housing	apartments, estate, living, apartment, housing, village, homes, place, courtyards, townhomes
Bathroom	bathrooms, bathroom, toilet, restroom, restrooms, toilets, urop, shower, pee, urinal
Comedy	comedy, joke, funny, eyebrows, laugh, comedians, comedian, eyebrow, prank, chuckle
Politics	political, politics, poli, conservative, polisci, pol, liberal, pols, conservatives, republicans
University Names	duke, baylor, bears, auburn, lsu, dukes, predictions, prediction, baylors, durham
NA	ross, sja, sjacs, banana, chavez, campanile, chaveztubman, bois, bananas, canes
Group Me	groupme, group, chat, chats, groupchat, groupmes, join, groupchats, groups, invite
Cybersecurity	cybersecurity, cyber, spam, hackathon, hacking, hacks, hackathons, hack, hackers, hackru
NA	uva, coronavirus, corona, virus, uvm, uv, uvas, deng, asbestos, charlottesville
NA	placement, calc, test, situs, exam, grading, chem, boi, professors, grade
Buildings	architecture, bridge, newark, building, rutgers, architectural, clocktower, tower, hartford, spire
Memes	meme, memes, bored, fun, minecraft, suggestions, recommendations, easiest, courses, graduate
Driving	car, driving, busch, drivers, cars, gt, driver, racing, f1, gri
Locations	barnard, hopkins, gracies, armstrong, sanders, dallas, bro, houston, cruz, arlington
Graduation	graduation, commencement, tickets, ceremony, homecoming, yearbook, celebration, graduating, ceremonies, graduates
Mental Health	counseling, therapy, therapist, therapists, depression, suicide, counselor, counselors, psychiatrist, depressed
Performing Arts	dance, theatre, theater, audition, orchestra, auditions, ballet, choir, dancers, symphony
Designing	design, graphic, animation, software, developer, app, engineering, developers, planner, visualizing
Long Distance travel	airport, uber, driving, uberlyft, bus, train, station, amtrak, transportation, drivers
COVID Symptoms	flu, coronavirus, cough, sick, coughing, virus, respiratory, infections, viral, fever
Cryptocurrencies	crypto, authentication, blockchain, cryptography, atms, atm, bitcoin, cryptocurrency, currency, coin
Schedule	schedule, happeningsfree, discuss, courseprofessor, discussion, huskies, weekly, topics, posts, projections
Data Analytics	data, analytics, ai, learning, informatics, thread, intelligence, registration, information, msda
Printing	print, printing, printer, printers, printed, laser, paper, library, prints, libraries
Gymming	gym, gyms, locker, lockers, fitness, crowded, workout, weights, packed, room
University Names	notre, dame, louisville, ucla, basketball, ncaacom, spartans, rutgers, score, ncaa
Podcasts	podcast, hbo, podcasts, streaming, channel, chromecast, torrenting, channels, webcast, torrent
Jobs	strike, workstudy, work, unemployment, jobs, job, labor, workers, strikers, paid
Seats	seats, furniture, chair, chairs, seat, seating, ikea, desks, reservations, sofa
Marketing	poster, advertising, flags, posters, balloons, banner, signs, balloon, banners, ads
Essays	essay, essays, write, supplement, writing, applying, submit, feedback, wrote, questionnaire
Grief	cry, sad, death, depressed, died, crying, die, depression, weeps, funeral
Reddit	reddit, subreddit, creditno, credits, courses, subreddits, 1credit, onecredit, reddits, 3credit
Legal	llc, courts, lles, court, lawsuit, legal, lawyer, sue, hoop, attorney
Loans	aid, financial, loans, loan, disbursement, subsidized, disbursed, federal, pay, offered
Libraries	gap, places, spaces, library, studying, libraries, rooms, room, areas, space
Food and Eating	meal, dining, meals, food, breakfast, eat, waffle, egg, dinner, eating
Courses	neuroscience, eduroam, anatomy, brain, kin, kinesiology, neuro, kine, kines, biology
Thesis Writing	digikey, tutor, freeessayssamplescom, dissertations, willing, attention, georgia, assist, gem, hiring
Electrical Engineering and Computer Science	ece, eeecs, ecs, ee, eemb, workload, courses, schedule, ee16b, ec, '
Calculus	calculus, quad, exam, calc, prerequisite, prereq, precalculus, precalc, pre, quads
Microsoft Tools	library, microsoft, libraries, windows, office, excel, excelsior, 365, software, hours,
History	anthropology, history, statue, athens, athena, historical, archaeology, statues, oldest, ancient, '
Book rentals	rental, books, renters, bookstore, book, rentals, renting, textbooks, leasing, textbook
Snow	snow, sledding, snowball, snowman, snowing, snowboard, winter, snowboarding, snowfall, snowy, '
Dissent	petition, protest, protests, sign, petitions, petitioning, protesting, protesters, petitioned, protestors
Mentorship	mentor, facebook, mentors, startup, group, groups, join, mentoring, signing, mentorship, '
Weight loss	vegan, weight, fat, vegetarian, fitbit, diet, lbs, weigh, pounds, calories
Cancelations and Refunds	cancel, cancellation, cancelling, canceling, cancelled, dismissal, canceled, cancelation, terminate, refund, '
Engineering Tools	matlab, mat, iolab, degreeworks, engineering, kit, degree, mechatronics, masters, cmu, '
Psychology Courses	psyc, psy, ps4, pstat, psc, psa, ps5, psci, ps3, psych, '
Course Instruments	canvas, sakai, added, page, enrolled, courses, syllabus, posted, professor, swapped, '
Accounting	accounting, audit, acct, auditing, inspections, exam, checks, inspection, exams, cpa, '
Course Withdrawal	withdraw, withdrawal, outage, withdrawing, drop, dropping, withdrawn, withdrew, dropped, withdrawals