ABSTRACT

YADAV, SHRIKANTH MURALIDHAR. Mixture Models to Understand Help-Seeking Behavior in U.S. Youth Before and During COVID-19. (Under the direction of Julie Ivy).

Objective: This study characterized the unobserved patterns in crisis response among youth in the United States from March to December 2020 and determined the characteristics of vulnerable subgroups who were at increased risk for suicide due to the pandemic.

Methods: Latent class analysis of crisis support-seeking from a national text-based crisis platform, (n= 179,497, ages 24 or younger) for 11 crisis concerns (e.g., depression, anxiety/stress, suicidal thoughts, isolation, abuse, bereavement, relationships) for three study periods: 1) January 2017 to December 2020; 2) Pre-pandemic: 1 Jan 2017 to 12 March 2020; 3) Pandemic: 13 March to 20 December 2020. Demographic characteristics (age, race/ethnicity, sexual orientation, and gender identity) were used as predictors for class membership using the three-step method.

Results: Four latent classes were identified: (1) Depression/Isolation/Self-Harm (D/I/S) (18,694 texters, 10.4%); (2) Interpersonal Stress/Mood-Anxiety (I/M) (32,640 texters, 18.2%); (3) Suicidal Thoughts/Depressed (S/D); and (4) Adjustment/Stress (A/S) (94,096 texters, 52.4%). During the pandemic, an increase in suicidal thoughts and active rescues occurred in the D/I/S and S/D higher risk subclasses. Characteristics of vulnerable groups in higher risk classes since the pandemic included children, LGBTQ, Native American, White, Black, Asian, female, and gender-nonconforming youth.

Conclusions: Results identified a strong association with class membership in more severe risk classes during the pandemic and an increase in suicidal behaviors, particularly among children and LGBTQ youth. Low-cost and targeted crisis text-based platforms for support-seeking in youth may be one potential safety net strategy to address the effects of the COVID pandemic on mental health in youth.

Mixture Models to Understand Help-Seeking Behavior in U.S. Youth Before and During COVID-19

by
Shrikanth Muralidhar Yadav

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

Industrial Engineering

Raleigh, North Carolina
2021

APPROVED BY:

Jennifer Runkle (External Member)	Semra King
Negash Medhin	Julie Ivy Chair of Advisory Committee

ACKNOWLEDGEMENTS

I would like to thank my advisor Dr. Julie S. Ivy for her tremendous patience and support. I would also like to thank Dr Jennifer Runkle for generously offering her time and expertise on numerous occasions, my committee members, Dr Medhin Negash and Dr Semra King for their inputs, and Dr Maggie Sugg and Dr Kurt Michael from Appalachian State University for their time and guidance. I also owe a great deal of thanks to Dr Aravind Saibaba for his support.

Thanks also to Dr Yahya Fathi and ISE department at NC State University for giving me the opportunity to grow.

Finally, I would like to thank my family and friends for their unconditional love and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER

1

INTRODUCTION

1.1 Motivation

Since the outbreak of the coronavirus disease in 2019 (COVID-19), public health policy has mostly focused on preventing healthcare systems from being overloaded (Chowdhury et al. (2020)). This has led to many countries imposing lockdown restrictions, limiting travel and social interactions. Several studies suggest external stressors like limited contact, financial/economic uncertainty, and loss of kin can contribute to poor mental health in vulnerable populations and exacerbate pre-existing mental health conditions such as depression and anxiety (Pfefferbaum and North (2020); Brooks et al. (2020); Fiorillo and Gorwood (2020); Albuquerque and Santos (2021)).

Emotional reactions to stress can arise due to the pandemic on a personal level (for instance, emotional isolation, poor nutrition, substance use and lack of physical activity (Pfefferbaum and North (2020))) and on a social level (for instance, loss of kin, loss of social support due to public health measures, and loss of conventional coping mechanisms (Albuquerque and Santos (2021))). Emphasis is currently being placed on restructuring mental healthcare delivery systems (Moreno et al. (2020)). One example of this is telehealth.

Novel digital applications such as text-based crisis platforms are uniquely poised to

provide mental and behavioral support to youth in need (Ye (2020)). These emerging technologies can be used to fill a critical gap in national public health surveillance efforts on the psychological impact of the pandemic on children and adolescents (de Figueiredo et al. (2021)). This work uses data from Crisis Text Line (CTL), a national not-for-profit organization in the U.S. that provides a free and confidential crisis texting service that is available 24/7 for support and counselling.

This works uses Latent Class Analysis (LCA): a class of mixture models where the estimated hidden parameter is a categorical variable identified from a set of measurements. This study aims to characterize latent classes from a cross-sectional cohort of children and young adults (≤ 24 years or younger) who reach out to Crisis Text Line, free national text-based crisis counselling platform for crisis intervention and counselling. The analysis period ranges from January, 2017 to December, 2020. The study seeks to understand the impact of the COVID-19 pandemic on help-seeking behavior using standard LCA and Multiple-Group LCA. Further, this work also presents a methodological framework to study the impacts of external events from cross-sectional data using Multiple-Group LCA.

1.2 Literature Review

Several studies indicate lifestyle changes due to the pandemic resulting in economic uncertainty, social isolation and bereavement can be detrimental to mental health (Sher (2020); Verdery et al. (2020); Cheung et al. (2008)). Additionally, studies also indicate that children and young adults from affected families are likely to be psychologically impacted due to loss of social support groups due to risk of infection, and poor communication with parents (Albuquerque and Santos (2021); Larsen et al. (2021)).

The following section highlights studies that focus on understanding the impact of pandemics on mental health (particularly in children). It also looks at how latent class analysis has been previously used to study mental health and high-risk behaviour. The following subsections also highlight the gaps in literature pertaining to understanding high-risk behavior in adolescents and young adults due to the COVID-19 pandemic.

1.2.1 Pandemics and Mental Health

Sher (2020) highlights studies that focus on the relationship between pandemics and poor mental health. For instance, Cheung et al. (2008) studied the relationship between suicides among elderly and the 2003 SARS outbreak in Hong Kong. In this study, 321 suicide deaths were classified as “SARS related” and “non-SARS related” based on the victims’ suicide

notes and witnesses. The motives identified in SARS-related deaths were associated with isolation, stress and the disease burden from long term illnesses.

For the COVID-19 pandemic, Larsen et al. (2021) highlight the impact of school closures on children using the Norwegian Family Dynamics Study (Norwegian Institute for Public Health (2021)). The longitudinal cohort study of 442 families assessed children's reactions to home schooling and isolation. The measurement variables included questions capturing three outcome reactions (emotional, cognitive, and worry), home schooling experience, parent stress levels, arguments within the family, and screen time usage. Children's psychological vulnerability for depression and anxiety were also measured. Hierarchical linear regression models were developed with the three reactions as outcome variables. COVID-19 related predictors (except for screen time use) showed significant association with all outcomes. Family stress and instability were found to have the strongest effect. Older children were found to have higher levels of reported reactions. However, this study was limited to interviews with parents and children who are at low risk of self-harm relative to most (roughly 60%) texters appearing in this study. In contrast, our work covers approximately 180,000 texters from diverse racial groups throughout the United States from 2017 to 2020 thus providing a significantly larger baseline population.

In Czeisler et al. (2020), the authors examined suicidal ideation and substance abuse in adults from April to June, 2020. They use a multivariable Poisson regression model with prevalence of behavioural health conditions and trauma- and stressor- related disorder (TSRD) as target variables. They report high suicide ideation among youth (between 18 and 24 years of age), and greater prevalence of suicidal ideation in Hispanic, non-Hispanic Black and essential workers. In Ettman et al. (2020), the prevalence of depression related symptoms in US adults before and during the pandemic from a nationally representative sample was studied. The authors examined depression and suicidal thoughts. Both studies do not consider children (age less than 18 years) in their analysis.

Hill et al. (2021) is the study that comes closest to our study with a focus on Suicide Ideation and Attempts in a Pediatric Emergency Department among youth in the United States. They report an increase in the prevalence of suicide ideation between March and July of 2020. They also use logistic regression models to study the effect of and demographic variables on positive suicide-risk. The interaction between race (and gender) with year were not significant. The next subsection describes Latent Class Analyses used in the context of mental health.

1.2.2 Latent Class Models for Mental Health

In Szlyk et al. (2020), the authors use CTL data to analyze latent classes with measurements related to psycho-social issues such as psychological health, relationships, abuse and environmental stressors. They classify texters into three groups based on the number of times they contact CTL: (1) one-time users, (2) two-time users and (3) three or more. They identify a solution with three latent classes, with the largest class corresponding to the group with the least distress across all three groups. Similar observations were made in our study and is discussed in Chapter 4. Our study was different from this work in several aspects. The focus of our work is on the impact of demographic characteristics during the COVID-19 pandemic on latent class identification. As CTL makes it optional for texters to disclose demographic information, the study by Szlyk et al. (2020) considers all texters reaching out to CTL. Our study, however, is also focused on learning the associations between demographic variables and the identified latent classes. We also delve further by using Multiple-Group LCA to understand how these clusters changed during the COVID-19 pandemic.

Choi et al. (2017) and Kwan et al. (2016) are two examples of studies examining mental health in young adults. They identify separate latent classes for varying levels of self-harm risk and depression. Both studies provide prevention strategies and policy recommendations targeting vulnerable subgroups.

Choi et al. (2017) investigated homogeneous subgroups related to perpetration and victimization in teen-dating related violence. They identified predictors of membership in these subgroups, and the mental health consequences associated with the membership. Their data was roughly uniformly distributed across adolescents from White, Hispanic, African American, and other racial and ethnic groups corresponding to 30%, 32%, 29% and 9% of the sample, respectively. They use the three-step method (proposed in Asparouhov and Muthén (2014)) to estimate the latent class model with covariates. The identified latent classes were as follows: 'nonviolence', 'emotional/verbal abuse', 'forced sexual contact', 'psychological and physical violence', and 'psychological abuse'. The study used several covariates like gender, ethnicity, parental education, acceptance of female-to-male violence, and acceptance of male-to-female violence. Relative to the nonviolent class, African-Americans were more likely to be members of the emotional/verbal abuse and psychological + physical violence classes. The proportion of females were significantly greater across the three violent classes. Finally, the more accepting youth were of female-to-male violence, the more likely they were to be members of forced sexual contact, psychological + physical violence, or psychological abuse classes compared to the nonviolence class.

In Kwan et al. (2016), health risk-behavior in university students is examined. Latent Class Analysis was used on a survey consisting of 300 items to assess mental, health behavior and status of 837 undergraduates with mean age of 20. The respondents were primarily female (64%) and Caucasian (65%). The authors identify a latent class model with three classes. These classes were labeled as 'typical', 'high-risk', and 'moderately healthy'. The measurement variables included survey questions pertaining to marijuana use, other illicit drug use, risky sex, smoking, binge drinking, poor diet, physical inactivity, and insufficient sleep. It was further reported that high risk students reported higher levels of stress relative to the low-risk group.

In summary, the majority of the studies seeking to understand the relationship between the COVID-19 pandemic and mental health do not report on children (age $\leq$ 18 years). These studies, however, suggest a greater prevalence of adverse depressive mental health symptoms relative to pre-pandemic levels. Among studies that do consider adolescents, they consider those respondents that are hospitalized or those that visit the Emergency Department. This work addresses key gaps in literature that study help-seeking behaviour in the context of the COVID-19 pandemic. This work also characterizes the changes in suicidal and self-harm related tendencies during the pandemic. Table 1.1 summarizes relevant literature for a quick glance.

1.3 Research Objectives

The objective of our study was to perform a latent class analysis (LCA) to a) characterize crisis concern profiles in young people using a national anonymized crisis texting platform, b) determine if the COVID pandemic was associated with distinct crisis profiles, c) examine socio-demographic characteristics for the distinct crisis profiles, and d) assess the likelihood of being flagged as high risk for suicide during the pandemic compared to the pre-pandemic period. LCA is a class of mixture models where the latent class is a categorical variable. In other words, these latent classes indicate the presence of homogeneous classes within a much larger and complex population and set of outcome(s).

This manuscript is structured as follows. Chapter 2 describes the data source, measurements pertaining to high risk behavior, statistical models used in the analysis which includes methods for selecting the number of classes, inclusion of covariates, and inclusion of grouping variables. Chapter 3 highlights the results for each of the three research questions. Chapter 4 compares findings from Chapter 3 with other studies and highlight how this work contributes to existing literature. Chapter 5 summarizes the findings, con-

tributions, strengths, limitations, and future work.

Table 1.1: Summary of Relevant Literature.

Authors	Objective	Country	Approach Measurements outcome of interest	Model	Parameters of interest	Results
Larsen et al. 2021	Impact of school closure and isolation on children	Norway	Home school experience, health related stress and instability, screen time, missing friends	Hierarchical linear models	Emotional, somatic and cognitive reactions	Significant association between COVID-19 related predictors (except screen time use) and the three outcomes.
Albuquerque et al. 2021	Opinion paper studying grieving children during the pandemic period					Grief exacerbated by loss of conventional coping mechanisms and lack of support. Identifying children at greater risk is challenging due to lack of social support, depletion of emotional resources.
Czeisler et al. 2020	Mental Health and Suicidal ideation during the COVID-19 pandemic	United States	Symptoms of behavioral health conditions, trauma- and stressor-related disorder.	Multivariable Poisson Regression	Sex, age, sexual orientation, race/ethnicity, census region	High suicidal ideation among Youth (18-24), Hispanic, non-Hispanic Black and essential workers.
Czeisler et al. 2021	Follow-up survey in September 2020	United States	Symptoms of behavioral health conditions, trauma- and stressor-related disorder.	Multivariable Poisson Regression	Sex, age, sexual orientation, race/ethnicity, census region	Acute increase in prevalence of poor mental health symptoms during the early phase of the pandemic.
Ettman et al. 2020	Prevalence of depression-related symptoms in US adults before and during the pandemic from a nationally representative sample.	United States	Depression Symptoms and COVID-19 stressor score	Multivariate Logistic Regression	Demographic characteristics, education, marital status, household income etc.	Increased prevalence in depressive symptoms during pandemic across all demographic groups; Asians with 18% higher prevalence of depression during the pandemic; Married had a lower rate of depression symptoms; individuals in the highest and the lowest income categories
Szlyk et al. 2020	Latent Class Analysis on Suicidal texters engaging with Crisis Text Line	United States	Clusters characterized by self reported psychological and relational issues.	Latent Class Model	Mental health indicators similar to those used in this study	Three latent classes were identified across data from 2013 to 2017.
Carr et al. 2021	Effects of COVID-19 on primary-care recorded mental illness and self harm episodes	United Kingdom	monthly incidence rates of psychotropic prescriptions and self harm	Negative Binomial Regression	Demographic variables and socioeconomic status	Reduction in incidents of depression in primary care, anxiety disorders and self harm in adults.
Gomez-Ramiro et al. 2021	Trends in psychiatric emergency service admissions during the COVID-19 pandemic	Spain	Hospital Admission rates	Multivariate Poisson Regression		Significant increase in the percentage of psychiatric hospitalizations and substance use related disorders. However, 37.9% decrease in Emergency department-related admissions.
Hill et al. 2021	Suicidal Ideation in Pediatric Emergency Departments during the COVID-19 pandemic	United States	Rates of suicide-risk screening results in a pediatric emergency department	χ^2 difference test and logistic regression	Prevalence of suicide risk and demographic variables	Higher rates of suicidal ideation in some months but not high all the time.

CHAPTER

2

METHODS

2.1 Data and Measurements

Data on crisis response were derived from a free, national text-based crisis counselling platform, Crisis Text Line (CTL). The texting service is a non-profit organization in the U.S. that provides confidential support and counselling that is available 24/7. When at-risk texters reach out to the service, the platform connects crisis counselors with individual texters and assist them in resource sharing and collective problem-solving. Crisis counselors are volunteers who are trained with supervision to interact and document calls with at-risk texters. Each conversation lasts for roughly 45 minutes and there is no limit on the number of conversations a texter can have. To date, approximately six million text messages have been exchanged and CTL currently has the largest repository of near real-time mental health data in the world (Crisis Text Line (2021)). After each conversation, a counselor answers a questionnaire concerning a texter's psychosocial health. Information regarding the measurements are given in Subsection 2.1.1 and demographic variables are given in Section 2.1.3.

Three cross-sectional cohorts were constructed (elaborated in Subsection 2.1.2): 1) Full: texters in contact with CTL from January 1, 2017 to December 2, 2020; 2) pre-pandemic:

texters from January 1, 2017 to March 12, 2020; and 3) Pandemic: texters from March 12, 2020 to December 2, 2020. We use March 13, 2020 to identify texters as belonging to the pandemic cohort as this was the day the White House announced stay-at-home orders in the United States (Proclamation No. 9994 March 13th (2020)).

Texters reaching out to the service have the option to remain anonymous and a majority of the texters choose to opt out of disclosing demographic information. Consequently, the original dataset contained 354,814 texters and only 179,497 texters had a valid response for age, race/ethnicity, and gender identity. The distribution of missingness is available in Table 2.1. Each row of the first three columns refer to the combination of 'Unusuable' and 'Useful' inputs. The corresponding number of texters and conversations are shown in fourth and fifth columns, respectively. For instance, there were 32,814 texters (116,855 conversations) under the age of 24 years who responded to the question pertaining to the age category but declined to identify their race and gender. Consequently, the cohort of interest in this analysis would have 179,497 texters. The LCA study by Szlyk et al. (2020) that considered all texters reaching out to CTL (specifically, not filtering based on the usability of demographic information) revealed similar classes.

Table 2.1: Summary of Missing Data

Age Category	Race Category	Gender Category	Texters	Conversations
Age<=24	Unusable	Unusable	32814	116855
Age<=24	Unusable	Useful	12048	48379
Age<=24	Useful	Unusable	95	402
Age<=24	Useful	Useful	179497	634077
Age>24	Unusable	Unusable	16199	47885
Age>24	Unusable	Useful	5189	17294
Age>24	Useful	Unusable	31	127
Age>24	Useful	Useful	67717	195729
Unusable	Unusable	Unusable	2406	9808
Unusable	Unusable	Useful	5781	29860
Unusable	Useful	Unusable	120	603
Unusable	Useful	Useful	32917	173105

In the context of the demographic variables, as race and ethnicity disclosures were made by text messages, regular expressions were used to assign texters to specific racial and gender groups. Table 2.2 highlight the regular expressions used in the analysis. The expressions were chosen based on the most frequently occurring race and gender responses

in the dataset. For the purpose of the analysis, 'Mixed' and 'Other' races were combined into the 'Other' race category. Similarly, 'transgender' and 'non-binary' gender categories were combined into the 'non-conforming' gender category.

Table 2.2: Regular Expressions for Race and Gender Categories.

Race Category	Regular Experession
African American	"black", "african american"
American Indian / Alaska Native	"native american", "american indian", "alaska native"
Asian	"asian", "indian", "pakistani"
Hispanic	"hispanic", "mexican"
Mixed Race	"mix", "multi", "bi-racial", "biracial", "bi racial"
White	"white"
Other	Race descriptions that do not fall in the above categories

Gender Category	Regular Expression
Female	"female"
Male	"male"
Transgender	"trans"
non-Binary	"non" & "binary"
Other	All other gender descriptions

2.1.1 Crisis Response Topics

Several studies suggest a wide array of emotional and psychological reactions for children, adolescents, and young adults in response to the pandemic. These typically include elevated stress, anxiety, depression, suicidal thoughts, loneliness, grief, eating disorders, abuse, and substance abuse (Murata et al. (2021); Hill et al. (2021); Fernández-Aranda et al. (2020); Thompson et al. (2021); Duan and Zhu (2020)). These can arise from changes in coping mechanisms and poor communication with support groups (Albuquerque and Santos (2021)). For the identification of latent classes, we had access to only those crisis topics pertaining to psychological health. The following crisis topics were treated as binary responses in our analysis: depressed, suicidal thoughts, self-harm, stress and anxiety, relationship, substance, bereavement, bully, eating (eating/body image issues), isolated, and abuse (emotional, physical, or sexual). Crisis counselors labeled text conversations with mental health topics (e.g., suicidal thoughts, self-harm, depression). Each of the item responses was flagged as '1' by the crisis counselor if a texter met the corresponding description. As texters can be flagged with multiple tags, the Jaccard similarity coefficient (Gower and Legendre (1986)) was used to visualize the co-occurrence of crisis tags. The Jaccard Index is defined as

$$S3_{XY} = \frac{a}{a+b+c}$$

where a, b, c and, d are elements of the contingency table corresponding to binary data; a denotes the condition where both crisis topics X and Y are present, b represents conditions where X is present and Y is absent, c represents conditions where Y is present and X is absent, and b represents conditions where both were absent.

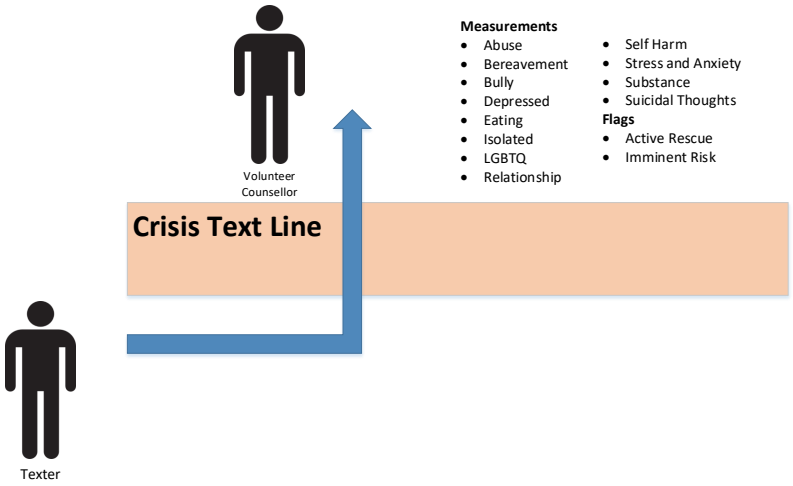

Figure 2.1: Description of the Crisis Text Line Service and the Measurement Variables.

2.1.2 Construction of Cross-Sectional Data

This section highlights the construction of the cross-sectional dataset. As each texter can reach out to CTL multiple times, we use a 'worst-case' criteria to identify a 11×1 dimensional item response vector for each texter. The elements of the vector constitute the crisis topics mentioned in Subsection 2.1.1. Table 2.3 shows an example of a dataset with three item responses (S_1, S_2, S_3) from three texters (with three, four and four conversations respectively). Note that the item response $S_1 = 1$ if the crisis response S_1 appears in at least one conversation. The final dataset would have the same number of rows as the number of texters. Finally, texters with conversations in both pandemic and the pre-pandemic periods will be flagged as being in the pandemic period. The next subsection highlights the covariates used in this study.

2.1.3 Covariates

Mental health response in young people varies by age, gender identity, sexual orientation, and across racial/ethnic groups. The covariates used in the analysis include: age (children '13 or younger', adolescents and young adults '14 to 24 years'), gender identity (female, male, nonconforming, other), and race/ethnicity (American Indian/American Native, Asian, Hispanic, Black or African American, White, Other). A COVID-19 indicator variable was created and assigned '1' to a texter who had any contact with CTL after the 13th of March, 2020 (start of White House stay-at-home order (Proclamation No. 9994 March 13th

Table 2.3: The table shows an example where longitudinal data is transformed into a cross-sectional dataset.

Longitudinal Data					**Crosssectional Data**		
Texter ID	**Conversation ID**	S_1	S_2	S_3	S_1	S_2	S_3
1	1	0	1	0			
1	2	1	1	0			
1	3	1	1	1	1	1	1
2	1	1	0	0			
2	2	1	0	0			
2	3	1	0	1			
2	4	1	0	1	1	0	1
3	1	0	0	0			
3	2	0	0	1			
3	3	0	1	1			
3	4	0	1	1	0	1	1

(2020))) and '0' otherwise.

A texter's sexual orientation was flagged as LGBTQ (yes=1, 0=no). To understand the most severe crisis concerns, we included texters at high risk for suicide (e.g., Imminent Risk or Active Rescue). A texter presenting with suicidal ideation, a plan to end their life, means, and an imminent time frame (typically within 48 hours) were considered to be at Imminent Risk (IR). If the counselor was not able to de-escalate and help the texter separate from the means of harm and work towards a safety plan, an Active Rescue (AR) was initiated by CTL . AR and IR were flagged as '1' if a texter met the criteria. About 1.3% of the texters were flagged in need of an active rescue.

Descriptive statistics were calculated to examine demographic characteristics and crisis topics. Overall proportions of texters belonging to each demographic group and the proportions of texters with each crisis response is compared with a χ^2-test for equal proportions with the appropriate degrees of freedom.

2.2 Latent Class Analysis

Clustering is an unsupervised approach for identifying homogeneous subgroups in a set of observations based their attributes. In epidemiology and health sciences, Latent Class Analysis (LCA) is a model-based clustering method that aims to identify clusters (or hidden

classes) from a set of observations (typically arising from a survey).

To further illustrate the idea of latent classes, consider a dataset in a clinical setting where each patient is recorded to have four symptoms $a-d$ (corresponding to the item responses in this study). Assume the range of patients can be diagnosed with diseases X, Y, and Z, and that disease X is associated with the presence of symptoms a, b, and c, disease Y with symptoms b, c, and d, and disease Z with symptoms a, c and d. LCA will attempt to detect the presence of latent classes (the disease entities) by identifying patterns of association in the symptoms. The following section elaborates on the specifications of the mathematical model.

2.2.1 Notation

Table 2.4: Summary of Notation

Summary of Notation			
Symbol	**General Notation**	**Parameters in this study**	**Example Problem**
J	Number of measurements	11	4
r_j	Response categories	2 (yes and no)	2 (yes and no)
j	A measurement $j \in \{1,2,\ldots,J\}$	Crisis topics (depression, ..., abuse)	Symptoms (a, b, c, and d)
y	Vector of observations with J elements	Vector of observations with the 11 crisis topics	Vector of recorded symptoms (a,b,c,d)
c	Latent Class $c \in [1,2,\ldots,C]$	Homogeneous subgroups that have a unique combination of crisis responses	Disease entities X, Y and Z
γ_c	Prevalence of latent class c	The proportion of texters belonging to class c	Prevalence of disease entities X, Y and Z
$\rho_{j,r_j\|c}$	Probabilty for respones r_j in observation j conditioned on class c	Probability being flagged with response r_j for crisis topic j conditioned on latent class c	Probability being flagged with response r_j for symptom j conditioned on disease c
θ	Set of all parameters		
x	The complete dataset		
C	The number of latent classes		

Table 2.4 provides a summary of the notation along with the connection to this study. The column 'General Notation' lists the general model and the third and the fourth columns elaborate on the parameters used in the study and the example problem respectively. The column 'Parameters in this study' provides an overview of how the model is related to both the clinical example and the general notation. The 'Example Problem' is provided

for the sake of completeness. Consider a dataset with J observed variables each with $r_j, \forall j \in \{1,\ldots,J\}$ response categories. The response categories can be answers to questions in a survey or other measurements in a cross-sectional dataset. For the current analysis, the observed variables are the 11 crisis topics mentioned in Subsection 2.1.1 and the response categories are binary ($r_j = 2 \ \forall j \in \{1,\ldots,J\}$) corresponding to responses 'yes' and 'no'. Consequently, the contingency table will have 2^J cells. If each cell corresponds to a unique combination y ($y \in \mathbf{R}^{J\times 1}$) of the responses, we can define Y as the discrete random variable used to refer to the response pattern $y_i, \forall i \in \{1,2,\ldots,2^J\}$. The sum of probabilities of observing all response patterns y is equal to one such that

$$\sum_{i=1}^{2^J} P(Y = y_i) = 1. \tag{2.1}$$

In terms of the model used in this work, y_i refers to a 11 (J) dimensional vector that contains the categorical responses (depression, stress and anxiety, ..., abuse) to each of the crisis responses. In terms of the clinical example, y_i is a four dimensional vector containing information on whether a patient was diagnosed with each of the four symptoms. Eq 2.1 corresponds to the the sum over the probabilities of observing all possible combinations of responses from the texter. For instance, in this study, $P(Y = y_1)$ corresponds to the probability of a response where no crisis topic was flagged and $P(Y = y_{2048})$ corresponds to the probability of a response where all crisis topics were flagged.

Further, each latent class model has two types of parameters: class prevalence (denoted by γ) and item-response probabilities (denoted by ρ). The parameter γ_c denotes the prevalence of the latent class c such that $\sum_{c=1}^{C} \gamma_c = 1$. This holds based on the definition of γ where the latent classes are mutually exclusive and exhaustive.

Let C denote the number of latent classes and L the random variable denoting the latent class. If $c = 1,\ldots,C$, the probability of observing latent class c is given by

$$P(L = c) = \gamma_c.$$

In this study, γ_c corresponds to the faction of texters belonging to class c.

Let ρ denote the item response probability such that $\rho_{j,r_j|c}$ represents the probability of observing response r_j ($r_j \in \{0,1\}$ as all measurements are binary) in measurement $j \in \{1,2,\ldots,J\}$ conditioned on latent class c. Consequently, we have

$$\sum_{r_j=\{0,1\}} \rho_{j,r_j|c} = 1$$

and the fundamental expression for observing response y is given by

$$P(Y=y)=\sum_{c=1}^{C}\gamma_c P(Y=y|L=c) \tag{2.2}$$

where the conditional probability can be defined as

$$P(Y=y|L=c)=\prod_{j=1}^{J}\prod_{r_j=\{0,1\}}\rho_{j,r_j|c}\mathbb{1}_{y_j=r_j}. \tag{2.3}$$

Note that there are $C-1$ terms for each class prevalence γ_c and $C\times\sum_{j=1}^{J}(R_j-1)$ terms for the class specific item response probabilities $\rho_{j,r_j|c}$.

In this study, Equations 2.2 and 2.3 correspond to the probability of observing response $y=$ (depression, stress and anxiety, ..., abuse) being parameterized by the class prevalence and the item response probabilities (depression, stress and anxiety, ..., abuse) that are conditioned on the latent class c. Table 2.4 summarizes all notation used above.

2.2.2 Model Estimation

The Model is estimated using Maximum Likelihood Estimation (usually the EM algorithm (Collins and Lanza (2009))). Specifically, given the dataset s, for the set of all parameters θ (all $\gamma_c, \rho_{j,r_j|c}$), the learning problem is defined by

$$\hat{\theta}=\arg\max_\theta log(p(s|\theta))=\arg\max_\theta \sum_c \gamma_c log(p(s,c|\theta))$$

As the iterative algorithm begins by assigning arbitrary starting values to all parameters, a model is considered 'identifiable' if at least 50% of the random starts lead to the same model (Collins and Lanza (2009)). Identifiable models converge to the same solution (i.e., same item response probabilities, class prevalence, and negative log-likelihood) regardless of the starting values whereas unidentifiable models converge to slightly different solutions and are sensitive to the starting value. An unidentifiable model suggests that the number of parameters are too high (Collins and Lanza (2009)) and consequently, the number of latent classes. This is analogous to overfitting in machine learning.

2.2.3 Model Selection

After a model is estimated and successfully identified, there are several methods for selecting the number of latent classes. These methods can be classified into Information Criterion ((e.g., Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC))) and likelihood-based tests (Lo et al. (2001); McLachlan et al. (2019)). Monte Carlo simulation studies from synthetic datasets in Nylund et al. (2007) recommend using BIC and the Bootstrap Likelihood Ratio Test (BLRT). In Nylund et al. (2007), the authors test each of the selection methods (information criteria and likelihood-based tests) and quantify the methods by comparing the number of times the correct number of hidden classes were successfully identified; BLRT and BIC were the best performing methods. Note that the information regarding the correct number of latent classes were available as the study used a synthetically generated dataset.

The BLRT provides an objective selection criteria where the improvements from increasing the number of parameters can be accurately tested. The null hypothesis is that the $k-1$ class model is the true model and the alternate hypothesis is that the k class model is the true data generating model. An outline of the test between two latent class models is outlined below:

1. Two models with $k-1$ and k classes are estimated using Maximum Likelihood Estimation from the real data and $-2\Delta L$ is calculated. Note that

$$\Delta L = L_k - L_{k-1}$$

   where L_k and L_{k-1} refer to the maximum log-likelihoods for the latent class models with k and $k-1$ classes respectively.

2. Synthetic data is generated from the $k-1$ class model.

3. Two sets of latent class models are re-estimated with k and $k-1$ classes to identify the null distribution of $-2\Delta\hat{L}$ for data corresponding to $k-1$ latent classes where

$$\Delta\hat{L}_i = \hat{L}_{k,i} - \hat{L}_{k-1,i}$$

   where $\hat{L}_{k,i}$ and $\hat{L}_{k-1,i}$ refer to the maximum log-likelihoods for the latent class and the subscript $i \in \{1,2,\dots B\}$ refers to the bootstrap sample. Note that B refers to the total number of bootstrap samples.

4. The bootstrap p-value is estimated by comparing $-2\Delta L$ from the real dataset and

the distribution of $-2\Delta\hat{L}_i$ which was estimated from the bootstrap samples in step 3. (McLachlan et al. (2019), Chapter 6). Specifically,

$$p-\text{value} = \sum_{i=1}^{B} \frac{\mathbf{1}_{-2\Delta L > -2\Delta\hat{L}_i}}{B},$$

where $\mathbf{1}$ is the indicator function. Note that if k is in fact the number of latent classes represented by the data, the improvement in log-likelihoods between the k and $k-1$ class model will be larger. Consequently, $-2\Delta L$ will be larger than the majority of the samples in $-2\Delta\hat{L}_i$.

The MPlus uses a more advanced (and efficient) variant of the BLRT method described above and is elaborated in Asparouhov and Muthén (2014). Once the number of latent classes are determined, the item response probabilities are interpreted and labelled based on defining characteristic (Collins and Lanza (2009)). However, additional inferences can be made by the inclusion of covariates to predict latent class membership. This is described in the following subsection.

2.2.4 Inclusion of Covariates

Covariates can be incorporated into LCA through a multinomial regression framework to predict latent class membership (Collins and Lanza (2009)). In the original latent class model, the probability of observing response $P(Y = y)$ is given by equations 2.2 and 2.3. In the latent class model with a single covariate x, the class prevalence γ_c is treated as an output of a logistic regression model:

$$\gamma_c = P(L = c|X = x) = \frac{e^{\beta_c^T x}}{1 + \sum_{a=1}^{C-1} e^{\beta_a^T x}}$$

where $X = [0, x]^T$ and the model parameters are $\beta_c = [\beta_{0,c}, \beta_{1,c}]^T$.

For multiple covariates, each variable is dummy-coded and the above equation generalizes to vector equivalents of x and the parameters β_c. As in the case of multinomial logistic regression, a latent class is chosen as the reference class and the odds ratios are reported for inference. In this analysis, the demographic variables are used as covariates to predict class memberships.

The simplest approach of directly including covariates in a Latent Class Analysis and performing the inference step. This is referred to as the 'one-step' method (Bolck et al. (2004); Vermunt (2010)). However, the method of including covariates increases model

building complexity: (1) The number of covariates in practical settings are large. It is possible that the measurement model itself can change when new covariates are introduced; (2) Including covariates could effect class selection. It is not clear if the class selection procedure should consider the model with covariates; Additional disadvantages are listed in Vermunt (2010). Consequently, the 3-Step method by Bolck et al. (2004) for LCA is used in MPlus for including the covariates. The steps are outlined below:

1. In the first step, latent classes are identified using indicator variables and no covariates. In other words, the latent classes are estimated without the covariates influencing the class assignments on the texters.

2. In the second step, classes $1:C$ are assigned to each response y_i to create a 'most-likely class' variable $W \in \mathbf{R}^{N\times K}$ where N and K is the number of observations and latent classes, respectively. The most common assignment rules are modal or proportional assignment. Modal assignment is equivalent to hard partitioning where the class with the greatest probability is assigned to response y_i (Vermunt (2010); Asparouhov and Muthén (2014)). For instance, if the posterior probabilities ($P(C|y_i)$) for response y_i with a three class model are $P(C=1|y_i)=0.8$, $P(C=2|y_i)=0.1$ and $P(C=3|y_i)=0.1$, modal assignment for $w_{i,k}$ would be $w_{i,1}=1$, $w_{i,2}=0$ and $w_{i,3}=0$. These weights are used to identify the classification error matrix

$$P(C=t|X=s)=\frac{\sum_{i=1}^{N}P(X=s|Y_i)w_{i,k}}{P(X=s)} \tag{2.4}$$

   where $s,t\in\{1,2,\ldots,C\}$ and $w_{i,k}$ are elements of the weight matrix W. Specifically, classification error $P(C=t|X=s)$ shows the probability of the estimated value conditional on the true value. The derivations for Equation 2.4 can be found in Vermunt (2010).

3. The weights (which are assigned to each texter/respondent through y_i) from step 2 account for classification uncertainty. In the last step, the regression coefficients are estimated for predicting latent class membership from the covariates using the weighted multinomical logistic regression (Vermunt (2010); Bolck et al. (2004)).

In summary, latent class models with covariates essentially have two components: the measurement model (or the latent classes) and the structural model (or the multinomial regression). The three-step method separates the estimation of the measurement model (i.e. in step 1) from the causal model and uses these weights to estimate the causal model

(i.e. step 2 and step 3). Consequently, this step addressed the model building complexity described above and in Vermunt (2010). In the next section, the Multiple Group Latent Class analysis is described in detail.

2.2.5 Inclusion of Grouping Variables

In the original latent class model, we have

$$P(Y=y)=\sum_{c=1}^{C}\gamma_c\prod_{j=1}^{J}\prod_{r_j=\{0,1\}}\rho_{j,r_j|c}\,\mathbb{1}_{y_j=r_j}.$$

Grouping variables can be included to study differences in item response probabilities. For example, in our study, while the crisis topics can be used as the item response variables, gender can be used as a grouping variable to compare latent classes that appear from male and female populations. Consequently, we have the model

$$P(Y=y|V=q)=\sum_{c=1}^{C}\gamma_{c|q}\prod_{j=1}^{J}\prod_{r_j}^{R_j}\rho_{i,r_j|c,q}\,\mathbb{1}_{y_j=r_j}$$

where V is the discrete random variable indicating the grouping variable and q is the group of interest. Assuming there are no parameter restrictions, the number of parameters estimated are $Q(C-1)$ for each γ and $QC(\sum_{j=1}^{J}(R_j-1))$ for each ρ. This is because each group will have a set of item response probabilities (ρ) associated with them. In terms of this study, the grouping variable corresponds to the two cohorts (pre-pandemic and the pandemic). Specifically, V is a binary random variable for the cohort and $q \in$ (pre-pandemic, pandemic). The grouping variables can be interpreted as studying 'known classes'; for example, this can involve studying the differences in the identified latent classes between different gender groups.

Consequently, differences between the latent classes identified under each grouping variable can manifest in one of several ways (Collins and Lanza (2009)):

1. Type 1: Both the number of classes and their defining characteristics are the same. This can occur when introducing the grouping variable brings no new information.

2. Type 2: The number of classes are the same but the item response probabilities defining the classes that appear in the two groups are different. Specifically, this implies the emergence of a new latent class.

3. Type 3: Both the number of classes and the item responses are different. This implies

the model would be estimated by explicitly defining the model with additional parameters.

These differences can be verified by inspecting individual populations using separate latent class models for each grouping level. For example, in the gender example, two separate latent class models (one for male and one for female) can be estimated to learn about the type of fit expected in a Multiple-Group analysis (Collins and Lanza (2009)).

The statistical software MPlus (Muthén and Muthén (2017)) also allows the definition of additional parameters using the `MODEL CONSTRAINT` command. This allows users to test significance of transformations of existing parameters. Our work uses said functionality to test differences between two item responses from the pandemic and the pre-pandemic periods. Generally, assuming $\rho_{i,r_j|c_1,q_1}$ and $\rho_{i,r_j|c_2,q_2}$ are item responses from comparable classes c_1 and c_2 between the grouping variables q_1 and q_2, the new variable δ_j can be defined by

$$\delta_j = \rho_{i,r_j|c_1,q_1} - \rho_{i,r_j|c_2,q_2}$$

for observing response r_j of item response j. In the case of the gender example for Multiple-Group LCA, the idea can be illustrated as follows. If the crisis topic of interest is *depression*, gender specific item response probabilities can be denoted as $\rho_{dep,1|c,female}$ and $\rho_{dep,1|c,male}$ for female and male cohorts, respectively, where '1' corresponds to response level $r_j = 1$. Then, we have the transformed variable

$$\delta_{dep} = \rho_{dep,1|c,female} - \rho_{dep,1|c,male}.$$

The `MODEL CONSTRAINT` command in MPlus computes standard errors for additional (transformed) parameters using the delta method (Raykov and Marcoulides (2004)) and statistical significance is tested using a Z-test (Muthén and Muthén (2017)).

2.2.6 Summary

In this study, the 3-step LCA was used to identify vulnerable subclasses of youth with similar crisis response patterns. Four models corresponding to distinct research questions were developed: (1) Model 1: Latent Class Model with covariates, (2) Model 2: Multiple-Group Latent Class Model; and (3) Model 2a: Latent Class Model for the pre-pandemic; and Model 2b: Latent Class Model for the pandemic period. The specifications of the model are described in the Section 2.3. For all models, a preliminary model with one to six latent classes were estimated and each estimation was replicated 100 times (Collins and Lanza

(2009)) to ensure the models were identifiable. Models that did not converge to the same solution in at least 50% (Collins and Lanza (2009)) of the replications were considered 'unidentifiable' (see Subsection 2.2.2 for details).

The number of latent classes were finalized by inspecting Bayesian Information Criterion for Models 1, 2, 2a and 2b. The Bootstrap Likelihood Ratio Test was used to test for differences between k and k-1 class models (McLachlan et al. (2019)) for Models 1, 2a and 2b. The 3-Step BCH method was used to simplify the covariate selection and to account for the classification uncertainty of latent classes (Asparouhov and Muthén (2012)). The following section introduces the research questions.

2.3 Research Questions

Is class membership for texters from the pandemic period different from those in the pre-pandemic period? (Model 1)

Specifically, the *Full* cohort was examined to determine if the pandemic was associated with class membership. The latent class model (denoted as Model 1) used dummy coded demographic variables (age, race/ethnicity, gender identity, sexual orientation), risk flags (Active Rescue (AR) and Imminent Risk (IR)), and 'COVID-19' as covariates. The demographic variables were included to understand the baseline characteristics of the latent classes. In Model 1, a texter was flagged as $COVID-19 = 1$ if they had at least one conversation with CTL after March 13, 2020 (conceptualized start date for lock down restrictions).

Does crisis response among texters engaging with CTL in the pre-pandemic and pandemic periods differ? (Model 2)

A multi-group LCA analysis was performed (referred to as Model 2) to determine whether texters engaging with CTL before and during the pandemic should be considered distinct user populations. In this model, we compared the pre-pandemic cohort to the pandemic cohort. As different groups can have similar characteristics, a wide range of constraints could be added before the model is estimated to reduce the number of parameters. However, significant differences were observed in most of the item response probabilities between the pandemic and the pre-pandemic period,so no parameter constraints were added (Collins and Lanza (2009)).

What are the characteristics of vulnerable subgroups and how do they change during the pandemic?

As sociodemographic factors can be external stressors, two models are estimated

(Model 2a and Model 2b) with each model corresponding to texters from the pre-pandemic and pandemic period, respectively. Differences between identified clusters in terms of the covariates can point to changes that can potentially guide intervention and outreach programs for public health professionals. Other covariates used in the model like Active Rescue and Imminent Risk could help with the characterization of suicidal thoughts in youth due to the pandemic.

All research questions were chosen based on gaps identified in literature pertaining to help-seeking behavior in young adults due to the COVID-19 pandemic. The following section highlights the software and the statistical tools used in the analysis.

2.4 Software

LCA was conducted using Mplus version 8.5 (Muthén and Muthén (2017)) to characterize the unobserved classes of crisis response in young people. The R software (R Core Team (2020)) was used for processing the data. In addition, several R packages were used: 'MPlusAutomation' (Hallquist and Wiley (2018)) was used to automate MPlus; 'Lubridate' (Grolemund and Wickham (2011)) was used to handle dates; and the 'tidyverse' (Wickham et al. (2019)) suite of packages was used for data wrangling.

2.5 Summary

This chapter highlights the epidemiological framework that is developed using LCA and Multiple-Group LCA to assess the impact of lockdown restrictions in young adults and children during the COVID-19 pandemic. Research questions were developed based on gaps in existing literature studying the impacts of COVID-19 lockdown restrictions on young adults. The next chapter summarizes the fit statistics and inferences.

CHAPTER

3

RESULTS

3.1 Descriptive Statistics

Tables 3.1 and 3.2 highlight the demographic attributes and prevalence of crisis responses of the texters, and the number of texters flagged with each mental health outcome for the pre-pandemic, and the pandemic cohorts. In general, a larger proportion of children (13 years or younger), gender non-conforming youth, and texters identifying as Asian, American Indian/Alaska Native, or Hispanic engaged with the service during the pandemic. A notable increase in the prevalence of all crisis topics was observed during the pandemic.

Tables 3.1 and 3.2 also report the test statistic and the p-value from the χ^2 test for equal proportions before and during the pandemic. Table 3.3 shows the Jaccard similarity matrix for the 13 crisis concerns. The greatest overlap (shaded in green) is observed between Depression and Relationship (S3 score = 0.427), Relationship and Stress and Anxiety (S3 score = 0.402), Depression and Stress and Anxiety (S3 score = 0.433), and Depression and Suicide (S3 score = 0.422).

Table 3.1: Summary of the demographic variables from both the pre-pandemic and the pandemic periods.

				dof:	1
Groups		Pandemic	Pre-Pandemic	Chi-sq test statistic	p-value
Age	**13 or younger**	8369 (15.59)	15467 (12.29)		
	14-24	45302 (84.41)	110359 (87.71)	355.95	p-value < 0.001
Gender	**Female**	42656 (79.48)	98276 (78.1)		
	Male	5411 (10.08)	16542 (13.15)		
	non-Conforming	2534 (4.72)	4647 (3.69)		
	Other	3070 (5.72)	6361 (5.06)	429.04	p-value < 0.001
Race	**African American**	6795 (12.66)	14636 (11.63)		
	American Indian / Alaska Native	2226 (4.15)	6001 (4.77)		
	Asian	3875 (7.22)	7863 (6.25)		
	Hispanic	10549 (19.65)	24160 (19.2)		
	Mixed Race	1576 (2.94)	3509 (2.79)		
	Other	433 (0.81)	1150 (0.91)		
	White	28217 (52.57)	68507 (54.45)	155.56	p-value < 0.001

Table 3.2: Summary of Measurement Variables

			dof:	1
Crisis Response	Pandemic	Pre-Pandemic	Chi-sq test statistic	
Abuse	8826 (16.44)	13551 (10.77)	971.94	p-value < 0.001
Active Rescue	935 (1.74)	1464 (1.16)	94.24	
Bereavement	5519 (10.28)	7466 (5.93)	983.86	
Bully	4360 (8.12)	8660 (6.88)	79.89	
Depressed	34244 (63.8)	65228 (51.84)	971.72	
Eating	6259 (11.66)	8118 (6.45)	1275.03	
Imminent Risk	3576 (6.66)	4487 (3.57)	803.22	
Isolated	24290 (45.26)	36978 (29.39)	2775.72	
LGBTQ	3687 (6.87)	5706 (4.53)	391.93	
Relationship	31224 (58.18)	56413 (44.83)	1371.83	
Self-Harm	16079 (29.96)	24785 (19.7)	1739.87	
Stress and Anxiety	32581 (60.71)	56708 (45.07)	1849.22	
Substance	2345 (4.37)	3483 (2.77)	297.05	
Suicidal Thoughts	23063 (42.97)	44541 (35.4)	572.77	

Table 3.3: Jaccard Index to visualize the extent of overlap between crisis topics. The darker share of green indicates higher co-occurrence between the responses.

	Abuse	Active Rescue	Bereavement	Bully	Depressed	Eating	Imminent Risk	Isolated	LGBTQ	Relationship	Self-Harm	Stress and Anxiety	Substance
Active Rescue	0.034	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Bereavement	0.081	0.027	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Bully	0.113	0.036	0.067	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Depressed	0.145	0.021	0.092	0.088	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Eating	0.092	0.031	0.059	0.094	0.096	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Imminent Risk	0.084	0.118	0.064	0.082	0.067	0.068	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Isolated	0.152	0.026	0.101	0.106	0.398	0.106	0.080	0.000	0.000	0.000	0.000	0.000	0.000
LGBTQ	0.072	0.027	0.045	0.082	0.062	0.069	0.064	0.073	0.000	0.000	0.000	0.000	0.000
Relationship	0.163	0.020	0.087	0.091	0.427	0.092	0.065	0.377	0.070	0.000	0.000	0.000	0.000
Self-Harm	0.152	0.038	0.085	0.107	0.258	0.121	0.104	0.233	0.084	0.228	0.000	0.000	0.000
Stress and Anxiety	0.147	0.019	0.080	0.085	0.433	0.100	0.059	0.335	0.066	0.402	0.236	0.000	0.000
Substance	0.063	0.054	0.051	0.036	0.044	0.050	0.069	0.050	0.038	0.043	0.059	0.043	0.000
Suicide	0.161	0.035	0.092	0.105	0.422	0.102	0.118	0.330	0.075	0.339	0.301	0.326	0.053

3.2 Research Question 1

This question corresponds to the inferences from Model 1 where covariates are included to study class membership. Figure 3.1 shows the bar-plot of the number of times the Maximum Likelihood Estimator arrived at the same solution with different starting values. The x-axis shows the negative log-likelihood and y axis denotes the frequency of models with the corresponding likelihood values. As described in Subsection 2.2.2, a latent class model is considered as 'not identified' if the same model is not estimated in at least 50% random starts. As a result, Table 3.4 shows the fit statistics for those models that were identified. It is observed that while both BIC and AIC decrease, models with five and six classes were poorly identified (Figure 3.1). Consequently, a four-class model best represented our data and the corresponding item response probabilities are shown in Figure 3.3. The four classes were:

(1) Depression/Isolation/Self-Harm (D/I/S) (18,694 texters, 10.4%) with the highest probabilities for all crisis topics. Most prominent crisis topics (and their probabilities) were Depression (0.97), Isolation (0.81), Relationship issues (0.91), Self-Harm (0.7), Stress and Anxiety (0.90), and Suicidal Thoughts (0.90).

(2) Interpersonal Stress/Mood-Anxiety(I/M) (32,640 texters, 18.2%) reporting high probabilities of Depression (0.85), Isolation (0.71), Relationship issues (0.83), and Stress and Anxiety (0.66).

(3) Suicidal Thoughts/Depressed (S/D) (34067 texters, 19%) characterized by a high probability of Depression (0.74) and Suicidal Thoughts (0.81).

(4) Adjustment/Stress (A/S) (94,096 texters, 52.4%) is characterized by a low probability of being flagged with a majority of the crisis tags. Unlike the other classes, it is noted that the Stress and Anxiety (0.44) and Relationship issues (0.42) have the highest probability of occurrence.

Figure 3.2 shows the distribution of the number of crisis response topics. It is observed that lower risk classes like *A/S* are associated with fewer topics. However, it is also observed that the *D/I/S* are associated with significantly higher number of crisis topics. Table 3.5 shows covariates associated with class membership in Model 1.

Relative to the A/S class (the reference group), a texter engaging during the pandemic was 3.7 times more likely to be in the D/I/S class (95% CI: 3.60, 3.97) and 2.4 times more likely to be in the I/M class (95% CI: 2.25, 2.46). Compared to A/S, D/I/S texters were more likely to be children (13 years or younger), female, gender non-conforming, White, American Indian/Alaska Native, or LGBTQ. Compared to A/S texters, S/D texters were more likely to be African American, gender non-conforming, or American Indian/Alaska

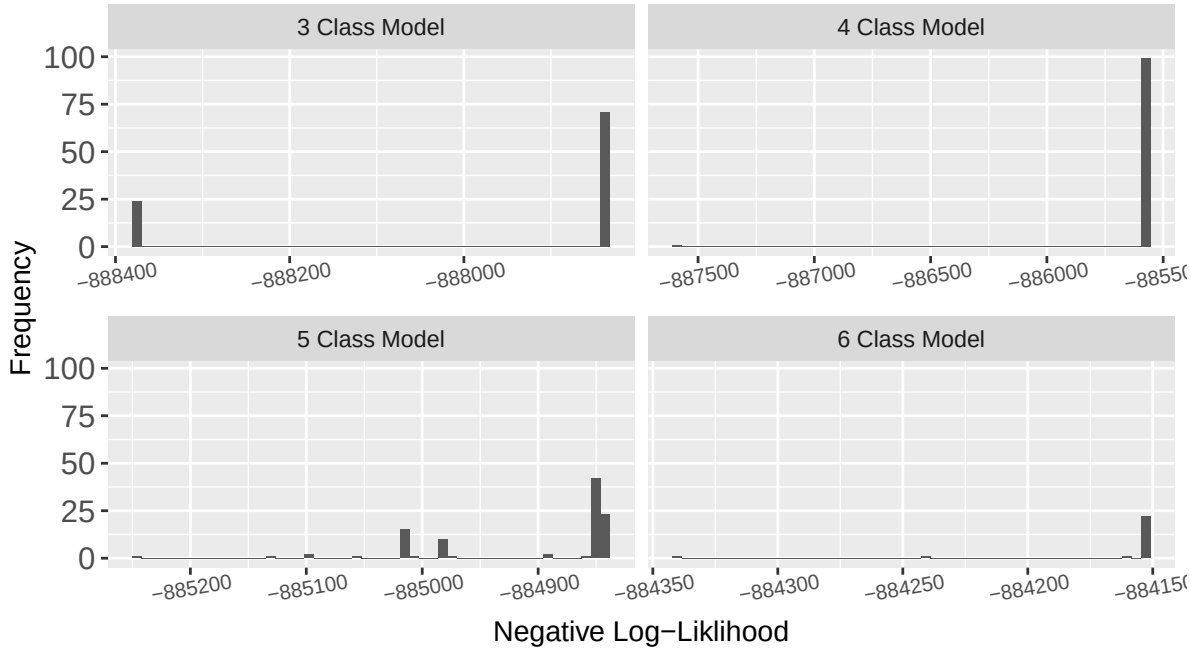

Figure 3.1: Distribution of the negative log-likelihood. Each model was estimated 100 times with different starting values.

Native; and I/M texters were more likely to be gender-nonconforming or Asian. As shown in Table 3.11, the large odds ratios for Active Rescue (AR) and Imminent Risk (IR) suggest high probabilities of membership in D/I/S, I/M, S/D classes relative to the A/S class.

In summary, the model with four latent classes best represented our data. Each of these latent classes were labelled based on their item response probabilities. Texters in the pandemic period were more likely to be associated with the higher risk classes. Demographic covariates indicated American Indian/Alaska Native and White texters were more likely to be in the highest risk groups. While this research question aimed to study the relationship between class membership and the pandemic period, the next research question considers texters in the pandemic period as belonging to a separate grouping variable.

Table 3.4: Model fit statistics for Model 1. The four-class model has the least AIC, BIC. Models with more than five classes are not identified.

Negative Log-Likelihood	Number of Classes	AIC	BIC	Entropy	Model Identification (%n)	BLRT (p-value<0.001)
-890857	2	1781761	1781993	0.734	100%	96596.933
-887838	3	1775746	1776100	0.627	73%	6038.288
-885577	4	1771248	1771723	0.576	100%	4522.01
-884841	5	1769801	1770397	0.548	Not Identified	1471.404
-884153	6	1768447	1769164	0.489	Not Identified	1377.588

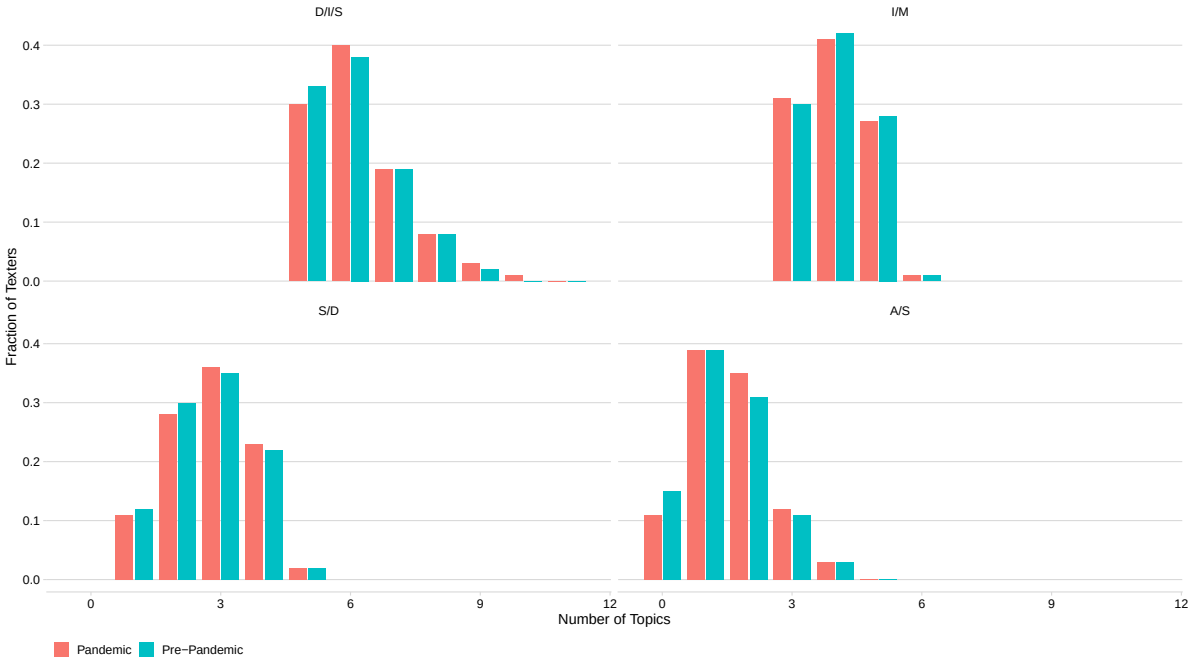

Figure 3.2: Prevalence of the number of crisis topics before and during the pandemic.

3.3 Research Question 2

Multi-group LCA analysis views texters from different groups as belonging to separate populations. Consequently, introducing texters belonging to the pandemic period as different group allows us to compare latent classes and item response probabilities that appear in both populations. Multi-group LCA analysis identified four latent classes (Table 3.7) across both groups. This was further verified by two LCA models estimated over each population (with the pandemic and pre-pandemic cohort) as described in Table 3.10. Specifically, both models indicate the presence of four latent classes.

Inspecting the latent classes and the item response probabilities suggest that although the two cohorts are similar in terms of their latent classes, the probability of many item responses being flagged were much higher for the pandemic cohort. The extent of difference between grouping variables lies in-between the Type 1 and Type 2 differences listed in Subsection 2.2.5. Specifically, while the item response probabilities during the pandemic were not different to warrant relabeling of the identified latent class, the item response probabilities appeared higher.

The classes were labeled based on the most prevalent crisis topics similar to the naming criteria used for Model 1: A/S, I/M, S/D, and D/I/S (Figure 3.4). As the defining item response probabilities in each class were similar to those identified in Figure 3.3, the same

Figure 3.3: Class-wise item response probabilities.

Table 3.5: Odds Ratios and 95% confidence intervals of the covariates with A/S as the reference class in Model 1.

Covariate		**The Odds Ratio of being in either D/I/S, I/M or S/D relative to A/S (95% CI)**		
		D/I/S - A/S	I/M - A/S	S/D - A/S
COVID		3.767 (3.579-3.965)***	2.354 (2.251-2.461)***	0.782 (0.742-0.825)***
LGBTQ		6.398 (5.896-6.944)***	2.162 (1.934-2.417)***	1.117 (0.998-1.249)***
Age	13 or younger	1	1	1
	14-24	0.578 (0.542-0.615)***	1.025 (0.953-1.102)	0.573 (0.542-0.607)***
Gender	Male	1	1	1
	Female	1.917 (1.743-2.107)***	1.036 (0.97-1.106)	0.859 (0.806-0.914)***
	non-Conforming	2.282 (1.988-2.619)***	0.769 (0.654-0.904)**	2.053 (1.843-2.287)***
	Other	2.778 (2.449-3.151)***	0.848 (0.741-0.97)*	1.652 (1.497-1.823)***
Race	White	1	1	1
	African American	0.628 (0.577-0.685)***	1.008 (0.941-1.079)	1.166 (1.092-1.245)***
	Hispanic	0.57 (0.53-0.613)***	0.886 (0.837-0.938)***	0.942 (0.891-0.996)*
	Asian	0.703 (0.63-0.785)***	1.148 (1.057-1.246)**	0.988 (0.904-1.079)
	Other	0.935 (0.828-1.057)	0.95 (0.842-1.073)	1.101 (0.989-1.226)
	American Indian / Alaska Native	1.123 (1.006-1.254)*	1.011 (0.901-1.134)	1.407 (1.283-1.543)***

*p-value<0.05, **p-value<0.01, ***p-value<0.001

Table 3.6: Odds Ratios and 95% confidence intervals of the covariates with I/M and S/D as the reference class in Model 1.

Covariate		The Odds Ratio of being in either D/I/S and S/D relative to I/M or S/D (95% CI)		
		D/I/S- I/M	D/I/S- S/D	S/D - I/M
COVID		1.601 (1.492-1.717)***	4.815 (4.517-5.134)***	0.332 (0.311-0.355)***
LGBTQ		2.96 (2.611-3.354)***	5.73 (5.129-6.401)***	0.517 (0.448-0.595)***
Age	13 or younger	1	1	1
	14-24	0.564 (0.511-0.622)***	1.008 (0.938-1.083)	0.559 (0.513-0.610)***
Gender	Male	1	1	1
	Female	1.85 (1.639-2.089)***	2.232 (2.016-2.471)***	0.829 (0.759-0.906)***
	non-Conforming	2.968 (2.394-3.681)***	1.111 (0.957-1.29)	2.67 (2.221-3.211)***
	Other	3.277 (2.713-3.959)***	1.682 (1.466-1.93)***	1.949 (1.66-2.288)***
Race	White	1	1	1
	African American	0.623 (0.556-0.699)***	0.539 (0.49-0.593)***	1.157 (1.056-1.268)**
	Hispanic	0.643 (0.584-0.709)***	0.605 (0.557-0.657)***	1.063 (0.983-1.149)
	Asian	0.613 (0.531-0.707)***	0.712 (0.627-0.809)***	0.861 (0.766-0.967)*
	Other	0.984 (0.823-1.177)	0.849 (0.737-0.979)*	1.159 (0.991-1.355)
	American Indian / Alaska Native	1.111 (0.943-1.309)	0.798 (0.706-0.902)***	1.392 (1.21-1.602)***

*p-value<0.05, **p-value<0.01, ***p-value<0.001

naming criterion was used to label the classes in the Multiple-Group model. However, a majority of the item responses were significantly higher for the pandemic period suggesting that the crisis topics were more prevalent in each class. This is also evidenced by Figure 3.2 where a greater proportion of texters in the pandemic period had a higher number crisis topics than those in the pre-pandemic period.

The `MODEL CONSTRAINT` command was used to test for significant differences between the item responses for corresponding classes. The differences between the item responses are summarized in detail below and in Tables 3.8 and 3.9. In both tables, differences that were not significant were left as blanks. The rest of the reported group-specific item response probabilities were statistically different (p-value < 0.0005).

A/S - Adjustment/Stress: For the pre-pandemic cohort, the crisis tags with the highest probabilities were stress and anxiety (0.38) and relationship issues (0.35). The probability of the texter being depressed was 0.30. All other item responses were less than 0.13. For the pandemic cohort, response probabilities for stress and anxiety (0.47), and feelings of depression (0.4) were significantly higher. While the probability of being flagged with suicidal thoughts was low in the pre-pandemic cohort, this outcome tripled in 2020 (0.13 compared to 0.04).

I/M - Interpersonal Stress/Mood-Anxiety: For the pre-pandemic cohort, depression (0.82), isolation (0.67), relationship issues (0.70), and stress and anxiety (0.59) were the most prominent crisis tags; followed by suicidal thoughts (0.43). Responses for the pandemic cohort increased significantly for the following: isolation (0.78), relationship issues (0.89), and stress and anxiety (0.77). While no a tag is as common as the other tags in this cohort, the probability of abuse doubled since the pandemic (0.1 vs. 0.2).

S/D - Suicidal Thoughts/Depressed: Compared to the pre-pandemic cohort, a 41% increase in suicidal thoughts (0.68 vs 0.96), and 43% increase in depression (0.54 vs 0.77) since the pandemic. A significant increase was observed for isolation (0.17 vs 0.36), relationship issues (0.30 vs 0.47), and stress and anxiety (0.32 vs 0.50).

D/I/S - Depression/Isolation/Self-Harm: Compared to the pre-pandemic cohort, a 9% increase in suicidal thoughts (0.86 vs 0.94), a 3% increase in depression (0.95 vs 0.98), a 14% increase in stress and anxiety (0.83 vs 0.95), a 16% increase in relationship issues (0.81 vs 0.94), a 22% increase in feelings of isolation (0.71 vs 0.87), and a 20% increase in self-harm (0.65 vs 0.78). Bereavement (0.15 vs. 0.24) and abuse (0.35 vs. 0.47) were also significantly higher for the pandemic compared to pre-pandemic.

In general, an increase in the item response probabilities since the pandemic was observed in all latent classes for the most prevalent crisis topics (depression, isolation, relationship, stress and anxiety). An increase in the probability of having suicidal thoughts was observed in A/S, S/D, and D/I/S.

Table 3.7: Model fit statistics for the Multiple-Group Latent Class Analysis

Model	Negative Log-Likelihood	Number of Classes	AIC	BIC	Entropy	Model Identification (%n)
Model 2	-994734.146	2	1989562.292	1990036.894	0.856	100%
	-991945.188	3	1984032.375	1984749.327	0.728	50%
	-989698.922	4	1979587.845	1980547.146	0.706	92%
	-988964.895	5	1978167.791	1979369.442	0.69	Not identified

3.4 Research Question 3

In the third research question, the characteristics of vulnerable subgroups in the pandemic and the pre-pandemic cohorts (Model 2a and 2b) are examined. Figure 3.5 shows the barplot of the multiple repetitions for both models. Table 3.10 shows model fit statistics

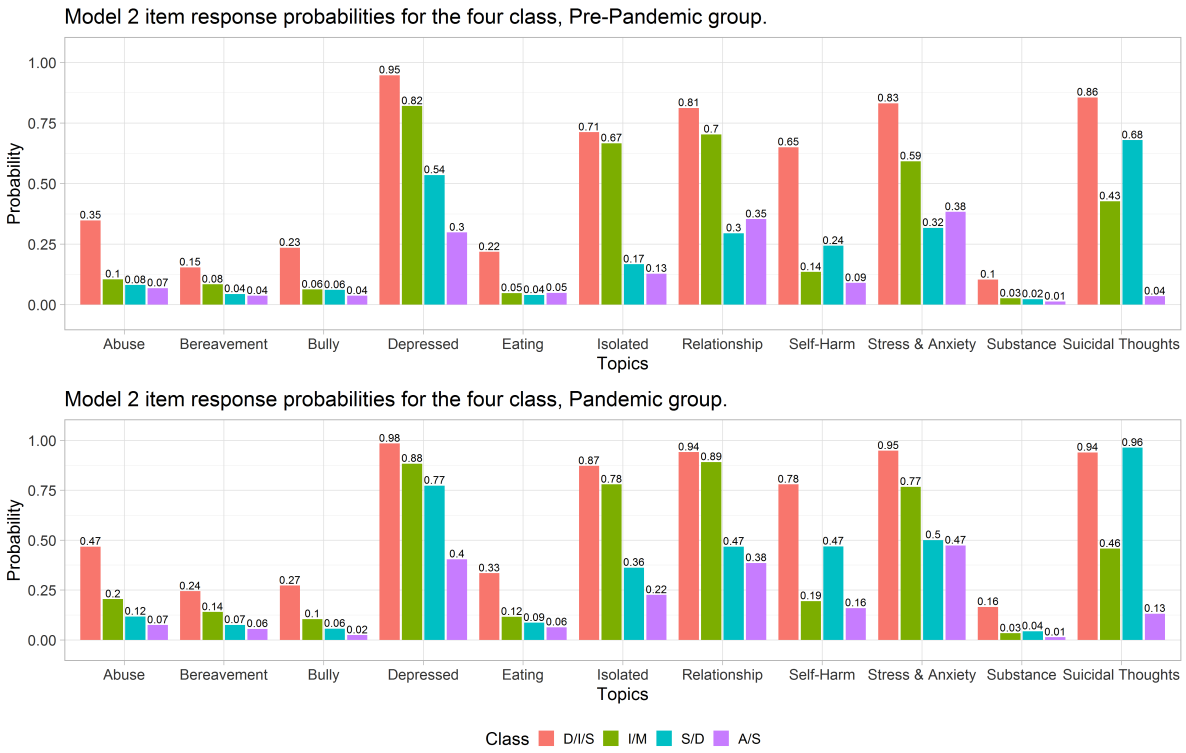

Figure 3.4: Item response probabilities of the Multiple-Group Latent Class Model.

Table 3.8: Differences in group-specific item response probabilities that were statistically significant (p-value<0.0005).*

Indicator	High Risk		Depressed Relationship	
	Pre-Pandemic	Pandemic	Pre-Pandemic	Pandemic
Abuse	0.35	0.47	0.10	0.20
Bereavement	0.15	0.24		
Bully				
Depressed	0.95	0.98	0.82	0.88
Eating				
Isolated	0.71	0.87	0.67	0.78
Relationship	0.81	0.94	0.70	0.89
Self Harm	0.65	0.78		
Stress and Anxiety	0.83	0.95	0.59	0.77
Substance				
Suicide	0.86	0.95	0.43	0.46

*Those item responses that were not statistically different before and during the pandemic are not reported.

Table 3.9: Differences in group-specific item response probabilities that were statistically significant (p-value<0.0005).*

Indicator	Depressed Suicidal		Normal Risk	
	Pre-Pandemic	Pandemic	Pre-Pandemic	Pandemic
Abuse				
Bereavement				
Bully				
Depressed	0.54	0.77	0.30	0.40
Eating				
Isolated	0.17	0.36		
Relationship	0.30	0.47	0.35	0.38
Self Harm				
Stress and Anxiety	0.32	0.50	0.38	0.47
Substance				
Suicide	0.68	0.96	0.04	0.13

*Those item responses that were not statistically different before and during the pandemic are not reported.

and it is observed that both BLRT and BIC suggested the four class model best represented both cohorts.

Generally, texters identifying as LGBTQ during the pandemic were more likely to appear in all three higher risk groups relative to the A/S group (D/I/S: 5.887 (95% CI: 5.127-6.759) compared to 4.922 (95% CI: 4.437-5.459), I/M: 2.308 (95% CI: 1.798-2.962) compared to 1.705 (95% CI: 1.437-2.023), S/D: 1.312 (95% CI:1.059-1.626) compared to 0.987 (95% CI: 0.861-1.131) (Table 3.12). However, it is also observed that the odds of being in D/I/S relative to A/S was lower in the pandemic period for children 0.471 (95% CI: 0.438-0.507) compared to 0.583 (95% CI: 0.531-0.639). In terms of relationship-related stress, both Model 1 (baseline) and Model 2b (pandemic cohort) indicated that the 14-24 group and children had comparable odds ratios of being in I/M relative to A/S (for the baseline: 1.03 (95% CI: 0.95-1.10) and for the pandemic period: 0.92 (95% CI: 0.804-1.053)). This was in contrast to Model 2a (pre-pandemic cohort) which showed 14-24 year old texters having greater odds ratios of being in the I/M class (1.33 (95% CI: 1.176-1.503)).

3.4.1 Association Between Suicide Risk and Active Rescue Events

The odds ratios for AR and IR flags were large; therefore, we report the more symmetric natural log of the odds ratios in Table 3.11 to better compare the odds of AR and IR for individual classes. For the pre-pandemic group, the classes in terms of the highest to lowest risk of AR and IR can be ordered as $D/I/S > S/D > I/M > A/S$ based on the log-odds ratios.

For the pandemic group, the AR and IR were ordered by class as follows: $D/I/S > S/D >$

Table 3.10: Model fit statistics for Model 2a and 2b.

Model	Negative Log-Likelihood	Number of Classes	AIC	BIC	Entropy	Model Identification (%n)	BLRT(p-value<0.0005)
Model 2a	-604973.321	2	1209992.641	1210216.722	0.705	100%	49171.041
	-603417.412	3	1206904.824	1207245.817	0.524	62%	3111.817
	-602013.122	4	1204120.243	1204578.148	0.528	96%	2808.581
	-601604.327	5	1203326.655	1203901.472	0.557	Not identified	817.588
Model 2b	-268892.957	2	537831.914	538036.398	0.709	100%	27303.426
	-268060.165	3	536190.33	536501.502	0.613	75%	1665.583
	-267269.696	4	534633.393	535051.252	0.593	99%	1580.938
	-266979.814	5	534077.628	534602.175	0.518	Not identified	579.765

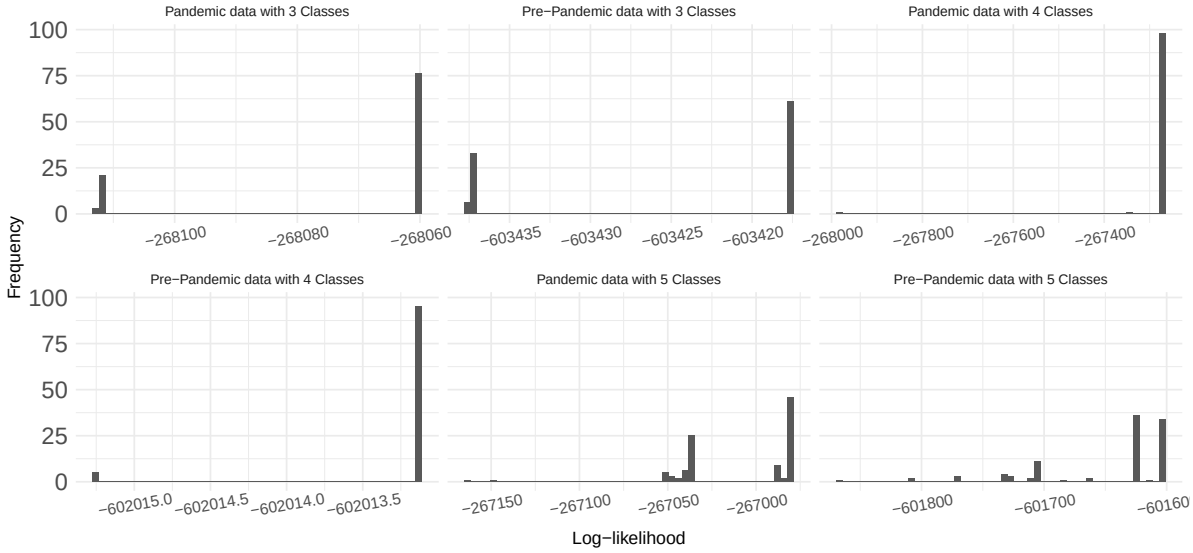

Figure 3.5: Model Identification for LCA models estimated with pre-pandemic and pandemic cohorts separately.

$I/M = A/S$ (high to low). In the pandemic period, the I/M class changed and was no longer associated with IR and AR as supported by: (1) The log-odds were large and negative for both AR and IR for I/M class relative to the A/S class in Model 2b. (2) The log-odds of AR corresponding to class membership in D/I/S relative to I/M were much larger for the pandemic compared to the pre-pandemic cohort (6792.3 (SE < 0.0005, p-value < 0.0005) vs. 55.82 (SE < 0.0005, p-value < 0.0005)). Similarly, IR was a statistically significant predictor of class membership for D/I/S relative to I/M in the pandemic cohort. This was not the case for the pre-pandemic cohort (185.2 (SE < 0.0005, p-value < 0.0005) vs. 6.11 (SE: 3.466, p-value = 0.08)).

In summary, models estimated from the two cohorts separately indicated that the odds of LGBTQ during the pandemic were more likely to appear in higher risk groups. Some changes were observed in the age categories where the texters 13 or younger had comparable likelihoods of being in the I/M class relative to the 14-24 year old during the pandemic period. In addition, the suicide risk indicators were characterized by the two models. During the pandemic, the I/M and A/S class had comparable association to suicide risk. This was in contrast to the pre-pandemic model where the I/M had suicide risk in between S/D and the A/S class.

3.5 Summary

The model with four latent classes best represented our data. Each of these latent classes were labeled based on their item response probabilities. Texters in the pandemic period were more likely to be associated with the higher risk classes. Demographic covariates indicated American Indian/Alaska Native and White texters were more likely to be in the highest risk classes.

When considering the pandemic cohort as a distinct population, an increase in the item response probabilities since the pandemic was observed in all latent classes for the most prevalent crisis topics (depression, isolation, relationship, stress and anxiety). An increase in the probability of having suicidal thoughts was observed in A/S, S/D, and D/I/S.

The models estimated from the two cohorts separately indicated that the odds of the LGBTQ population during the pandemic were more likely to appear in higher risk groups. Some changes were observed in the age categories where the texters 13 or younger had comparable likelihoods of being in the I/M class relative to the 14-24 year old group during the pandemic period. During the pandemic, the I/M and A/S class had comparable association to suicide risk. This was in contrast to the pre-pandemic model where the

association with suicidal thoughts was greater in the I/M class than in the A/S class.

Table 3.11: Covariates for Active Rescue and Imminent Risk from Models 1, 2a, and 2b.

	Log-odds (SE) with p-value<0.0005 unless specified otherwise					
	Active Rescue			**Imminent Risk**		
Class / **Model**	**D/I/S - A/S**	**I/M - A/S**	**S/D- A/S**	**D/I/S - A/S**	**I/M - A/S**	**S/D- A/S**
Model 1	32.37 (0.07)	26.39 (4.14)	30.96 (SE<0.0005)	31.85 (0.04)	28.29 (0.2)	30.52 (SE<0.0005)
Model 2a	80.81 (SE < 0.0005)	24.99(SE<0.0005)	79.17 (0.09)	34.66 (3.46)	28.55 (SE<0.0005)	33.09 (3.46)
Model 2b	24.18 (SE < 0.0005)	-6768 (SE < 0.0005)	23.42 (0.12)	25.09 (SE< 0.0005)	-160.7 (SE<0.0005)	24.49 (0.06)
Class / **Model**	**D/I/S - I/M**	**D/I/S - S/M**	**S/D- I/M**	**D/I/S - I/M**	**D/I/S - S/M**	**S/D- I/M**
Model 1	5.98 (4.136) (p-value = 0.149)	1.4 (0.071)	4.58 (4.139) (p-value = 0.269)	3.56 (0.203)	1.33 (0.04)	2.229 (0.205)
Model 2a	55.82 (SE < 0.0005)	1.64 (0.088)	54.17 (0.088)	6.11 (3.466) (p-value = 0.08)	1.57 (0.048)	4.54 (3.463) (p-value = 0.19)
Model 2b	6792.3 (SE <0.0005)	0.76 (0.118)	6791.5 (0.118)	185.2 (SE<0.0005)	0.605 (0.61)	185.19 (0.061)

Table 3.12: Covariates from the Pre-pandemic and the pandemic Models

Covariates		Pre-pandemic Cohort (Model 2a)			Pandemic Cohort (Model 2b)		
		D/I/S - A/S	*I/M - A/S*	*S/D - A/S*	D/I/S - A/S	I/M - A/S	S/D - A/S
LGBTQ		4.922 (4.437-5.459)***	1.705 (1.437-2.023)***	0.987 (0.861-1.131)	5.887 (5.127-6.759)***	2.308 (1.798-2.962)***	1.312 (1.059-1.626)*
Age	13 or younger	1	1	1	1	1	1
	14-24	0.471 (0.438-0.507)***	1.33 (1.176-1.503)***	0.613 (0.574-0.655)***	0.583 (0.531-0.639)***	0.92 (0.804-1.053)	0.504 (0.459-0.554)***
Gender	Male	1	1	1	1	1	1
	Female	1.78 (1.6-1.982)***	0.862 (0.794-0.935)***	0.816 (0.763-0.872)***	1.644 (1.42-1.904)***	1.038 (0.906-1.189)	1.032 (0.901-1.182)
	non-Conforming	2.399 (2.028-2.837)***	0.734 (0.578-0.932)*	2.016 (1.769-2.299)***	1.809 (1.475-2.218)***	0.42 (0.286-0.618)***	1.924 (1.586-2.333)***
	Other	2.538 (2.186-2.948)***	0.776 (0.646-0.933)**	1.502 (1.34-1.683)***	2.172 (1.79-2.634)***	0.653 (0.489-0.871)**	1.863 (1.549-2.24)***
Race	White	1	1	1	1	1	1
	African American	0.717 (0.649-0.793)***	1.169 (1.067-1.281)**	1.165 (1.083-1.254)***	0.73 (0.645-0.825)***	0.988 (0.862-1.133)	1.201 (1.071-1.347)**
	Hispanic	0.629 (0.578-0.685)***	0.942 (0.872-1.019)	0.946 (0.891-1.005)	0.688 (0.62-0.765)***	0.979 (0.874-1.097)	0.982 (0.888-1.087)
	Asian	0.776 (0.682-0.884)***	1.126 (1-1.268)*	1.012 (0.918-1.114)	0.808 (0.693-0.942)**	1.18 (1.005-1.384)*	0.918 (0.781-1.078)
	Other	0.839 (0.72-0.978)*	1.055 (0.896-1.244)	1.148 (1.019-1.295)*	1.156 (0.971-1.377)	0.985 (0.775-1.254)	0.868 (0.7-1.076)
	American Indian / Alaska Native	1.228 (1.084-1.391)**	1.027 (0.876-1.204)	1.419 (1.278-1.575)***	1.26 (1.064-1.492)**	1.093 (0.866-1.38)	1.255 (1.043-1.509)*

*p-value<0.05, **p-value<0.01, ***p-value<0.001

CHAPTER

4

DISCUSSION

The objective of the study was to address gaps in literature pertaining to (1) understanding help seeking behaviour in young people, (2) determining if the COVID pandemic was associated with specific crisis profiles, and (3) examining socio-demographic characteristics for the said crisis profiles. This study identifies four homogeneous classes within data from Crisis Text Line using Latent Class Analysis. This work considers 11 crisis topics from children and young adults seeking counselling in the United States. This is also the first study to examine help-seeking behavior in children through December 2020.

Further, demographic information was used to assess the likelihood of being in a high-risk class (D/I/S, I/M, and S/D) relative to the least risk (A/S) class. Demographic covariates to predict class membership suggested that White and Native Americans, children under 13 years and members belonging to the LGBTQ community were more likely to be in higher risk groups than others. In addition, texters during the pandemic were associated with higher probability of being in a high risk class.

This work contributes to a growing body of literature seeking to understand help-seeking behavior. Most notably, the three identified latent classes from the previous study in Szlyk et al. (2020) with CTL data from 2013 to 2017 are similar to D/I/S, I/M and A/S. This implies the S/D class was a new cluster that appeared in the population after 2017 that

was uniquely characterized by suicidal thoughts and depression but without relationship issues. Other studies with focus on understanding the COVID-19 pandemic and mental health (such as Czeisler et al. (2020, 2021); Ettman et al. (2020); Carr et al. (2021); Killgore et al. (2020a,b)) did not report on children (age $\leq$ 18 years). In Czeisler et al. (2020) and Ettman et al. (2020), the prevalence of adverse depressive mental health symptoms was high relative to the pre-pandemic levels. In Czeisler et al. (2020), a majority of the young adults (18-24 years) reported having more suicidal thoughts relative to the older population. In this work, the identified latent classes attempt to deconstruct the underlying process. A 20% increase in crisis conversations for self-harm and a 9% increase in texts for suicidal thoughts were observed during COVID-19 (e.g., D/I/S). While Active Rescue (AR) and Imminent Risk (IR) were associated with fewer latent classes during the pandemic (relative to the pre-pandemic period), this suggests a decrease in the number of texters that CTL would consider high-risk. Further, it is important to note that texters in the high-risk classes (D/I/S or S/D) were more likely to report depression, stress and anxiety, isolation and relationship issues during the pandemic period.

Regarding the general population, several studies also report greater association with symptoms like depression and anxiety (Lai et al. (2020); Xiong et al. (2020)) during the pandemic period. In Lai et al. (2020), it is reported that female (OR, 1.94; 95% CI, 1.26-2.98), and professional doctors working on diagnosis and treatment (OR, 1.82; 95% CI, 1.38-2.39) were at greater risk of being associated with depressive symptoms. González-Sanguino et al. (2020) is a study from Spain using cross-sectional data from an online survey during the early periods of the COVID-19 pandemic. They also report female respondents and students were associated with greater risk of depression, anxiety and PTSD related symptoms.

In Carr et al. (2021), it was reported that adults experienced a decrease in rates of psychotropic prescriptions and self-harm related injuries in the pandemic period in the United Kingdom in April 2020 and increased to normal levels by September 2020. In Gómez-Ramiro et al. (2021), data from a single hospital in Spain suggested that while there was a 37.9% decrease in Emergency Department admissions, a significant increase in diagnosis of substance use disorders (95% CI: 0.126-0.181, $p < 0.001$) was observed. Several other studies report a reduction in hospitalization due to psychiatric emergencies in Italy (Capuzzi et al. (2020); Clerici et al. (2020)) and France (Pignon et al. (2020)). All studies reported are limited to hospitalizations among adults.

Among studies in the Unites States on psychiatric symptoms, our study supports the findings of Killgore et al. (2020a), Killgore et al. (2020b), and Thompson et al. (2021). Killgore et al. (2020a,b) report on trends in suicidal ideation and loneliness, respectively, among

adults during the first three months of the COVID-19 pandemic. The authors found an increase in both suicidal ideation and loneliness among respondents under lockdown. Those respondents not under lockdown were found to have near constant levels of suicidal ideation. In Thompson et al. (2021), the authors examine psychiatric hospitalizations among adolescents during the COVID-19 pandemic. They study self-reported ratings of suicidal ideation and suicide attempts in a hospital in Rhode Island and compare trends to a similar analysis window in the previous year. The authors found higher ratings of suicidal attempts (8% for past-week and 11% for past-month).

In the context of pediatric admissions, Hill et al. (2021) study suicide ideation before and during the COVID-19 pandemic. Specifically, they examined suicide-risk screening data from patients aged between 11 and 21. Their analysis period was from January to July from both years (2019 and 2020). The odds of suicide ideation were 1.60 times higher in March and 1.45 times higher in July. The odds of suicide attempts were 1.58, 2.34, 1.75, and 1.77 times higher in the months between February to April, respectively.

Findings from this study demonstrated that LGBTQ youth were significantly more likely to be in the highest risk groups since the pandemic; whereby the likelihood of these youth being a member in the I/M or D/I/S group increased by 130 and 490 percent, respectively, during the pandemic period. LGBTQ youth are at a much higher risk of physical and sexual abuse, depression, suicide ideation and self-harm; estimated rates of attempted suicide are two to ten-fold higher than their peers (Cohen and Bosk (2020)). Texters reporting 'other' and 'non-Conforming' genders were associated with the highest risk classes (D/I/S and S/D) for both the pre-pandemic and pandemic periods.

Among vulnerable youth, e.g., those belonging to the LGBTQ community, homeless, and those with history of substance abuse, COVID-19-related disruptions to important social networks, mental health access/services, and school-related resources have resulted in loneliness, loss of in-person support, and a range of mental health needs and difficulties (Cohen and Bosk (2020)). Online or text-based support tools were important to this group before COVID-19 and recent research has found texting platforms are a continued support line for LGBTQ youth since the pandemic (Sampson et al. (2021)).

In summary, while several studies indicate increasing prevalence of psychiatric symptoms during the COVID-19 pandemic for young adults, our study uses the largest available dataset to identify and characterize latent classes (clusters) based on reported symptoms. Our study also identifies how the clusters changed before and during the pandemic and the associated changes in the crisis topics. Comparing the latent classes identified before and during the pandemic identified vulnerable populations. From a methodological perspective, this work also demonstrates how Multiple-Group Latent Class analysis can be

leveraged to study impacts of external events using MPlus' `MODEL CONSTRAINT` functionality.

CHAPTER

5

CONCLUSION

5.1 Summary of Results

This study identified four types of crisis profiles in U.S. youth: (1) Depression/Isolation/Self-Harm (D/I/S) (18,694 texters, 10.4%); (2) Interpersonal Stress/Mood-Anxiety (I/M) (32,640 texters, 18.2%); (3) Suicidal Thoughts/Depressed (S/D); and (4) Adjustment/Stress (A/S) (94,096 texters, 52.4%). Membership in more severe risk classes was strongly associated with the pandemic and an increase in suicidal behaviors, particularly among children and LGBTQ youth.

During the pandemic, an increase in suicidal thoughts and active rescues occurred in the higher risk (D/I/S and S/D) sub-classes. The identified clusters and their associations suggest that demographic differences in crisis response profiles and suicidal risks may require tailored intervention and prevention programs. The identified classes in the pandemic period pointed to a reduced number of texters seeking help with suicidal thoughts. However, the item responses of the high-risk classes displayed the highest likelihoods of depression, isolation, relationship issues, stress and anxiety, and suicidal thoughts. This suggests that during the pandemic period, the identified patterns in help-seeking behavior were more severe in terms of the crisis topics that come up during conversations relative

to the baseline.

5.2 Contributions

The study addressed several gaps in the existing literature on help-seeking behavior among children and young adults during the COVID-19 pandemic. While the results of our study aligned with a previous study from 2017 in terms of the latent class models identified from the same data source, it also identified the emergence of a new latent class characterized by feelings of depression and suicidal thoughts. The affiliation of latent classes to the socio-demographic variables also aided the efforts to identify vulnerable subgroups with respect to race/ethnicity, gender and age. In addition, this study population included members from racial groups like the Asian and American Indian/Pacific Islanders not included in previous studies. These findings support the development of better intervention strategies for high risk and vulnerable populations. For example, the clusters (and their evolution with time) can be useful to policy makers when planning interventions.

5.3 Strengths and Limitations

Table 5.1: Race and ethnicity in the study relative to the country.

Race	CTL data (%n)	U.S. Census Bureau 2019 (%n)
African American	11.94	14
American Indian / Alaska Native	4.58	1
Asian	6.54	5
Hispanic	19.34	26
Other	3.71	4
White	53.89	50

Key strengths of this study included the use of CTL's national platform to capture the near real-time responses of help-seeking youth throughout the pandemic with data on a baseline comparative population of users. We used a person-centered approach (e.g., LCA) to examine a wide range of crisis-response behaviors for at-risk racial/ethnic groups and minority gender/sexual orientation groups. As shown in Table 5.1, this study also addresses gaps in the literature studying high-risk adolescents from underrepresented racial

and ethnic groups including the Asian and American Indian/Alaska Native populations. However, the African American and Hispanic populations were underrepresented in the cohort relative to the true proportions.

There are a few limitations that must be considered in interpreting results. For this exploratory analysis, we used a cross-sectional cohort to explore the potential changes in crisis response for the pre-pandemic and pandemic periods. Future research should employ a latent transition model to examine longitudinal data for each CTL user. This would help understand movement between classes due to the pandemic. Further, the data consists of measurements (crisis responses) based on volunteer counselor judgements, and thus cannot be validated using ICD-10 diagnosis codes. Finally, data was also collected from a select group of individuals willing to seek support when in crisis and disclose demographic information. However, given the anonymity and free 24/7 access to CTL's digital platform, help-seekers may not be impeded by stigma or commonly encountered financial/access barriers to mental health services for this population.

5.4 Future Work

Regarding the study design with measurements similar to the current dataset, the time component can be introduced to account for movement between latent classes. Specifically, Latent Transition Analysis (LTA) and the more recent Random Intercept LTA (RI-LTA) (Muthén and Asparouhov (2020)) can be used to study the dynamic component of latent classes from longitudinal datasets. Coupled with the aggregation method described in this work, the impact of natural disasters and political movements on mental health can be studied using two sample RI-LTA and LTA models. Inclusion of covariates as main or interaction effects can be used to incorporate socio-demographic characteristics to identify transitions between latent classes over time. Specifically, transition-probability matrices can themselves be made functions of external covariates (similar to how covariates are used in Latent Class Analysis). This framework would also enable the inclusion of control populations.

While our study identifies an increase in suicide ideation during the pandemic, it is not clear if this necessarily translates into higher incidents of self-harm related hospitalizations. Specifically, as studies report reduction in hospitalization due to psychiatric emergencies in Spain (Gómez-Ramiro et al. (2021)), Italy (Capuzzi et al. (2020); Clerici et al. (2020)) and France (Pignon et al. (2020)), a complete characterization of high-risk behavior is still not achieved for the United States. Further, all studies above are limited to adults.

Future studies and surveillance strategies could identify contributing factors like access to mental healthcare services, dietary behavior, and physical activity (such as in the Youth Risk Behavior Surveillance System (Centers for Disease Control and Prevention, US and others (2021))) These are factors that can inform policy makers and ultimately result in better guidelines for the public. Further, given that access to CTL is far greater than conventional mental healthcare deliver methods, incorporating ICD-10 diagnosis codes into the service can contribute to clinically rigorous studies on large scale mental health programs.

5.5 Conclusion

The objective of this study was to perform a latent class analysis to characterize crisis concern profiles in young people using a national anonymized crisis texting platform; determine if the COVID pandemic was associated with distinct crisis profiles; examine socio-demographic characteristics for the distinct crisis profiles; and assess the risk of suicide during the pandemic compared to the pre-pandemic period.

Chapter 3 highlighted the results for each of our three research questions. The model with four latent classes best represented the baseline dataset. Each of these latent classes was labelled based on their item response probabilities. Texters in the pandemic period were more likely to be associated with the higher risk classes. Demographic covariates indicated Native American and White texters were in the highest risk groups. When considering the pandemic cohort as a grouping variable, an increase in the item response probabilities since the pandemic was observed in all latent classes for the most prevalent crisis topics.

The models estimated from the two cohorts separately indicated that the odds of appearing in higher risk groups was greater for respondents identifying as LGBTQ during the pandemic. Some changes were observed in the age categories where the texters who were 13 or younger had comparable likelihoods of being in the I/M class relative to the 14-24 year old during the pandemic period. During the pandemic, the I/M and A/S classes had comparable association to suicide risk. This was in contrast the pre-pandemic model where the association with suicidal thoughts was greater in the I/M class than in the A/S class.

Our study also identifies how the clusters changed before and during the pandemic and the associated changes in the latent classes. In addition, comparing the latent classes identified before and during the pandemic identified vulnerable populations. Consequently,

these findings address gaps in literature aimed at studying help-seeking behaviors among children and youth in the United States due to the COVID-19 pandemic.

REFERENCES

Albuquerque, S. and Santos, A. R. (2021). "in the same storm, but not on the same boat": Children grief during the covid-19 pandemic. *Frontiers in Psychiatry*, 12:23.

Asparouhov, T. and Muthén, B. (2012). Using mplus tech11 and tech14 to test the number of latent classes. *Mplus web notes*, 14(22):1–17.

Asparouhov, T. and Muthén, B. (2014). Auxiliary variables in mixture modeling: Using the bch method in mplus to estimate a distal outcome model and an arbitrary secondary model. *Mplus web notes*, 21(2):1–22.

Bolck, A., Croon, M., and Hagenaars, J. (2004). Estimating latent structure models with categorical variables: One-step versus three-step estimators. *Political analysis*, 12(1):3–27.

Brooks, S. K., Webster, R. K., Smith, L. E., Woodland, L., Wessely, S., Greenberg, N., and Rubin, G. J. (2020). The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *The lancet*, 395(10227):912–920.

Capuzzi, E., Di Brita, C., Caldiroli, A., Colmegna, F., Nava, R., Buoli, M., and Clerici, M. (2020). Psychiatric emergency care during coronavirus 2019 (covid 19) pandemic lockdown: results from a department of mental health and addiction of northern italy. *Psychiatry research*, 293:113463.

Carr, M. J., Steeg, S., Webb, R. T., Kapur, N., Chew-Graham, C. A., Abel, K. M., Hope, H., Pierce, M., and Ashcroft, D. M. (2021). Effects of the covid-19 pandemic on primary care-recorded mental illness and self-harm episodes in the uk: a population-based cohort study. *The Lancet Public Health*, 6(2):e124–e135.

Centers for Disease Control and Prevention, US and others (2021). Youth risk behavior surveillance system (yrbss). *Atlanta, GA: Centers for Disease Control and Prevention*.

Cheung, Y., Chau, P. H., and Yip, P. S. (2008). A revisit on older adults suicides and severe acute respiratory syndrome (sars) epidemic in hong kong. *International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences*, 23(12):1231–1238.

Choi, H. J., Weston, R., and Temple, J. R. (2017). A three-step latent class analysis to identify how different patterns of teen dating violence and psychosocial factors influence mental health. *Journal of youth and adolescence*, 46(4):854–866.

Chowdhury, R., Luhar, S., Khan, N., Choudhury, S. R., Matin, I., and Franco, O. H. (2020). Long-term strategies to control covid-19 in low and middle-income countries: an options overview of community-based, non-pharmacological interventions. *European journal of epidemiology*, 35(8):743–748.

Clerici, M., Durbano, F., Spinogatti, F., Vita, A., de Girolamo, G., and Micciolo, R. (2020). Psychiatric hospitalization rates in italy before and during covid-19: did they change? an analysis of register data. *Irish journal of psychological medicine*, 37(4):283–290.

Cohen, R. I. S. and Bosk, E. A. (2020). Vulnerable youth and the covid-19 pandemic. *Pediatrics*, 146(1).

Collins, L. M. and Lanza, S. T. (2009). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*, volume 718. John Wiley & Sons.

Crisis Text Line (2021). Crisis trends.

Czeisler, M. É., Lane, R. I., Petrosky, E., Wiley, J. F., Christensen, A., Njai, R., Weaver, M. D., Robbins, R., Facer-Childs, E. R., Barger, L. K., et al. (2020). Mental health, substance use, and suicidal ideation during the covid-19 pandemic—united states, june 24–30, 2020. *Morbidity and Mortality Weekly Report*, 69(32):1049.

Czeisler, M. É., Lane, R. I., Wiley, J. F., Czeisler, C. A., Howard, M. E., and Rajaratnam, S. M. (2021). Follow-up survey of us adult reports of mental health, substance use, and suicidal ideation during the covid-19 pandemic, september 2020. *JAMA network open*, 4(2):e2037665–e2037665.

de Figueiredo, C. S., Sandre, P. C., Portugal, L. C. L., Mázala-de Oliveira, T., da Silva Chagas, L., Raony, Í., Ferreira, E. S., Giestal-de Araujo, E., Dos Santos, A. A., and Bomfim, P. O.-S. (2021). Covid-19 pandemic impact on children and adolescents' mental health: biological, environmental, and social factors. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 106:110171.

Duan, L. and Zhu, G. (2020). Psychological interventions for people affected by the covid-19 epidemic. *The lancet psychiatry*, 7(4):300–302.

Ettman, C. K., Abdalla, S. M., Cohen, G. H., Sampson, L., Vivier, P. M., and Galea, S. (2020). Prevalence of depression symptoms in us adults before and during the covid-19 pandemic. *JAMA network open*, 3(9):e2019686–e2019686.

Fernández-Aranda, F., Casas, M., Claes, L., Bryan, D. C., Favaro, A., Granero, R., Gudiol, C., Jiménez-Murcia, S., Karwautz, A., Le Grange, D., et al. (2020). Covid-19 and implications for eating disorders. *European Eating Disorders Review*, 28(3):239.

Fiorillo, A. and Gorwood, P. (2020). The consequences of the covid-19 pandemic on mental health and implications for clinical practice. *European Psychiatry*, 63(1).

Gómez-Ramiro, M., Fico, G., Anmella, G., Vázquez, M., Sagué-Vilavella, M., Hidalgo-Mazzei, D., Pacchiarotti, I., Garriga, M., Murru, A., Parellada, E., et al. (2021). Changing trends in psychiatric emergency service admissions during the covid-19 outbreak: Report from a worldwide epicentre. *Journal of Affective Disorders*, 282:26–32.

González-Sanguino, C., Ausín, B., Castellanos, M. Á., Saiz, J., López-Gómez, A., Ugidos, C., and Muñoz, M. (2020). Mental health consequences during the initial stage of the 2020 coronavirus pandemic (covid-19) in spain. *Brain, behavior, and immunity*, 87:172–176.

Gower, J. C. and Legendre, P. (1986). Metric and euclidean properties of dissimilarity coefficients. *Journal of classification*, 3(1):5–48.

Grolemund, G. and Wickham, H. (2011). Dates and times made easy with lubridate. *Journal of Statistical Software*, 40(3):1–25.

Hallquist, M. N. and Wiley, J. F. (2018). MplusAutomation: An R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling*, pages 1–18.

Hill, R. M., Rufino, K., Kurian, S., Saxena, J., Saxena, K., and Williams, L. (2021). Suicide ideation and attempts in a pediatric emergency department before and during covid-19. *Pediatrics*, 147(3).

Killgore, W. D., Cloonan, S. A., Taylor, E. C., Allbright, M. C., and Dailey, N. S. (2020a). Trends in suicidal ideation over the first three months of covid-19 lockdowns. *Psychiatry Research*, 293:113390.

Killgore, W. D., Cloonan, S. A., Taylor, E. C., and Dailey, N. S. (2020b). Loneliness: A signature mental health concern in the era of covid-19. *Psychiatry research*, 290:113117.

Kwan, M., Arbour-Nicitopoulos, K., Duku, E., and Faulkner, G. (2016). Patterns of multiple health risk–behaviours in university students and their association with mental health: application of latent class analysis. *Health promotion and chronic disease prevention in Canada: research, policy and practice*, 36(8):163.

Lai, J., Ma, S., Wang, Y., Cai, Z., Hu, J., Wei, N., Wu, J., Du, H., Chen, T., Li, R., et al. (2020). Factors associated with mental health outcomes among health care workers exposed to coronavirus disease 2019. *JAMA network open*, 3(3):e203976–e203976.

Larsen, L., Helland, M. S., and Holt, T. (2021). The impact of school closure and social isolation on children in vulnerable families during covid-19: a focus on children's reactions. *European child & adolescent psychiatry*, pages 1–11.

Lo, Y., Mendell, N. R., and Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88(3):767–778.

McLachlan, G. J., Lee, S. X., and Rathnayake, S. I. (2019). Finite mixture models. *Annual review of statistics and its application*, 6:355–378.

Moreno, C., Wykes, T., Galderisi, S., Nordentoft, M., Crossley, N., Jones, N., Cannon, M., Correll, C. U., Byrne, L., Carr, S., et al. (2020). How mental health care should change as a consequence of the covid-19 pandemic. *The Lancet Psychiatry*.

Murata, S., Rezeppa, T., Thoma, B., Marengo, L., Krancevich, K., Chiyka, E., Hayes, B., Goodfriend, E., Deal, M., Zhong, Y., et al. (2021). The psychiatric sequelae of the covid-19 pandemic in adolescents, adults, and health care workers. *Depression and Anxiety*, 38(2):233–246.

Muthén, B. and Asparouhov, T. (2020). Latent transition analysis with random intercepts (ri-lta). *Psychological Methods*.

Muthén, B. and Muthén, L. (2017). *Mplus*. Chapman and Hall/CRC.

Norwegian Institute for Public Health (2021). The dynamics of family conflict study.

Nylund, K. L., Asparouhov, T., and Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A monte carlo simulation study. *Structural equation modeling: A multidisciplinary Journal*, 14(4):535–569.

Pfefferbaum, B. and North, C. S. (2020). Mental health and the covid-19 pandemic. *New England Journal of Medicine*, 383(6):510–512.

Pignon, B., Gourevitch, R., Tebeka, S., Dubertret, C., Cardot, H., Dauriac-Le Masson, V., Trebalag, A.-K., Barruel, D., Yon, L., Hemery, F., et al. (2020). Dramatic reduction of psychiatric emergency consultations during lockdown linked to covid-19 in paris and suburbs. *medRxiv*.

Proclamation No. 9994 March 13th (2020). United states : Proclamation on declaring a national emergency concerning the novel coronavirus disease (covid-19) outbreak. *Office of the Federal Register, National Archives and Records Service, General Services Administration*.

R Core Team (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raykov, T. and Marcoulides, G. A. (2004). Using the delta method for approximate interval estimation of parameter functions in sem. *Structural Equation Modeling*, 11(4):621–637.

Sampson, L., Ettman, C. K., Abdalla, S. M., Colyer, E., Dukes, K., Lane, K. J., and Galea, S. (2021). Financial hardship and health risk behavior during covid-19 in a large us national sample of women. *SSM-population health*, 13:100734.

Sher, L. (2020). The impact of the covid-19 pandemic on suicide rates. *QJM: An International Journal of Medicine*, 113(10):707–712.

Szlyk, H. S., Roth, K., and Garcia-Perdomo, V. (2020). Subgroups of suicidal texters engaging with crisis text line. *Psychiatric services (Washington, DC)*, 71(4):319.

Thompson, E. C., Thomas, S. A., Burke, T. A., Nesi, J., MacPherson, H. A., Bettis, A. H., Kudinova, A. Y., Affleck, K., Hunt, J., and Wolff, J. C. (2021). Suicidal thoughts and behaviors in psychiatrically hospitalized adolescents pre-and post-covid-19: a historical

chart review and examination of contextual correlates. *Journal of affective disorders reports*, 4:100100.

Verdery, A. M., Smith-Greenaway, E., Margolis, R., and Daw, J. (2020). Tracking the reach of covid-19 kin loss with a bereavement multiplier applied to the united states. *Proceedings of the National Academy of Sciences*, 117(30):17695–17701.

Vermunt, J. K. (2010). Latent class modeling with covariates: Two improved three-step approaches. *Political analysis*, 18(4):450–469.

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., and Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686.

Xiong, J., Lipsitz, O., Nasri, F., Lui, L. M., Gill, H., Phan, L., Chen-Li, D., Iacobucci, M., Ho, R., Majeed, A., et al. (2020). Impact of covid-19 pandemic on mental health in the general population: A systematic review. *Journal of affective disorders*.

Ye, J. (2020). Pediatric mental and behavioral health in the period of quarantine and social distancing with covid-19. *JMIR pediatrics and parenting*, 3(2):e19867.