Patterns of Mental Health Problems among Children: Multilevel Joint Latent Class Analysis

By

Mst Farzana Akter

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

MASTER OF SCIENCE

Department of Community Health Sciences

University of Manitoba

Winnipeg

Abstract

Background:

Child and Adolescent mental health disorders (MHD) are a global issue that pose a huge burden for individuals, families, society, and the economy, impeding the transition from childhood to adulthood. About half of all mental disorders in adulthood start showing symptoms from adolescence, and symptoms of mental illness in childhood increase the risk of mental health problems later in life. Identifying specific mental health patterns within children helps to plan targeted interventions and improves their mental wellbeing. Frequently used methods to identify subgroups within population with similar pattern of behavior include latent class models and clustering algorithm. However, these methods may not be effective when analyzing the behaviours reported by multiple sources and may lead to inaccurate findings. The application of multilevel extension of latent class models has not yet been explored in clustering schools, where children are nested and show similar pattern of behaviour within school. The purpose of the study was to identify mental health subgroups among children in schools and clusters of schools using assessments by multiple informants, considering the variability of these subgroups across different school environments.

Objectives:

The objectives of this thesis are: i) to analyze the mental health patterns in children by using assessments from multiple informants and comparing these patterns with those identified by a single informant; ii) to capture the heterogeneity of mental health patterns of children across individual school and clusters of schools, based on the patterns of mental health within schools; and iii) to assess the effect of school-level and individual-level factors on clusters of schools and the mental health patterns of children, respectively.

Methods:

Healthy Child Manitoba Office of the Government of Manitoba, Canada, conducted Manitoba Grade 5 Menta Health Survey across all schools in Manitoba in 2015/2016. The data collected includes information on students' behaviour over the last six months, or the current school year, as reported by teachers and students themselves. This thesis project proposes using joint latent class analysis to classify students into the mental health patterns based on these dual assessments. The latent patterns represent the underlying mental health classes across these assessments. The optimal number of mental health patterns and classes were decided using fit indices such as information criteria and likelihood ration tests (LRT). The study also extended this analysis to a multilevel joint latent class analysis to accommodate the nested structure of the data, specifying categorical latent variable to cluster schools. This approach assessed whether mental health patterns vary across different schools, clustering them based on the prevalence of mental health patterns among students. Finally, we investigated the effects of school-level and children-level factors on clusters of schools and mental health patterns using multinomial logistic regression.

Results:

We identified six mental health classes (high, moderately high, medium, mild, mild internalizing and low risk) for each informant, as well as three mental health patterns: high-risk, low-risk and self-reported risk, integrating assessments from both teachers and students. The high-risk pattern (19.3%) consisted of students identified as high mental health risk by both self-report and teacher assessment, while the low-risk pattern (38.3%) included students deemed low risk on both reports. A third pattern (42.4%) emerged, labeled self-reported risk, comprising students reporting moderate to high risk themselves but assessed as low to mild risk by teachers. Furthermore, three clusters of schools were identified: high-risk (30 schools, 10.6%), low-risk (76 schools, 26.8%), and student-reported risk (178 schools, 62.6%), based on variations of mental health patterns among schools. Male students, students who born in Canada, and those engaged in bullying activities reported by teachers had higher odds of being in the high-risk and self-reported risk patterns compared to their counterparts. Moreover, a higher prevalence of bullying in school settings was associated with increased odds of being in the high-risk and student-reported risk school clusters.

Conclusions and significance:

This study provides an approach for classification to effectively manage information provided by multiple informants by considering the possible correlation among assessments and within higher-level groups. This approach will serve as a valuable tool for researchers working with data from multiple sources within a hierarchical framework. This approach will enable the evaluation of the impact of group-based intervention plans using information from multiple sources. The mental health patterns identified in this study offer insights for policymakers to develop training program for teachers and implement targeted interventions to prevent the long-term effects of mental health disorders. This research provides valuable information about factors that can inform the development of targeted interventions tailored to the specific needs of students. This enables educators and policymakers to plan interventions more effectively, ensuring that resources are allocated efficiently, and interventions are tailored to address the root causes of mental health issues within school settings.

Keywords: Mental health problems, Children, Latent class analysis, Multilevel data, Multiple informants, Multilevel joint latent class analysis.

Acknowledgment

I am deeply grateful to my thesis supervisor, Dr. Depeng Jiang, for his unwavering support, expert guidance, and invaluable mentorship throughout the research process. Dr. Jiang's encouragement and insightful advice have played a pivotal role in shaping this thesis and motivating me to pursue excellence in my work. I am truly fortunate to have had the opportunity to learn from his expertise and to benefit from his dedication to my academic and professional development.

I would like to extend my sincere appreciation to the funding bodies that have made my academic journey possible. Firstly, I am grateful for the financial support and training provided by the VADA program, which has enabled me to pursue my master's degree. Additionally, I am thankful for the generous support received from the Jenny Fraser Scholarship and the RFHS Graduate Studentship.

I am immensely grateful to my fellow lab mates from Dr. Jiang's lab, who have been instrumental in shaping my academic journey and enriching my research experience. Their willingness to share insights on various research topics and generously impart their knowledge and expertise have been invaluable assets throughout the development of my thesis. Moreover, their experiences and perspectives gained from working on their own theses have provided me with invaluable guidance and inspiration, motivating me to overcome challenges.

Furthermore, I extend my heartfelt appreciation to my parents, siblings and best friend, Sabikunnaher Mitu for their unwavering encouragement, understanding, and unconditional support during the ups and downs of this journey. Their belief in my abilities, patience, and words of encouragement have been a constant source of strength and motivation, inspiring me to persevere through challenges and pursue my academic aspirations.

I am deeply grateful to everyone who has played a part, big or small, in making this thesis a reality. Your support and encouragement have been invaluable, and I am profoundly thankful for the opportunity to undertake this research endeavor.

Contents

List of Tables

List of Figures

List of Tables in Appendix

1. Introduction

Mental health refers to the condition of the mind that allows a person to acquire life skills, use those skills to be productive, handle the stress of everyday life, and contribute to society [1]. It is important to pay extra attention to nurturing mental health during childhood and adolescence because symptoms of mental health problems often begin to appear during these stages and can continue into adulthood. When children are mentally healthy, it means they have emotional and social skills, can cope with various situations, and behave appropriately for their age [2]. Mentally healthy children have a good perspective toward their life, perform well at school and at home, and have good relationships with their peers. Some characteristics that are considered normal during childhood development may actually indicate more serious mental health conditions in children [3]. Those characteristics can impact a child's development and hinders a smooth transition into adulthood [3,4]. Adults with mental health disorders often reflect on the impact of their mental health conditions during childhood and express regret for not receiving help during that time.

Common mental health disorders among children include anxiety, depression, obsessive-compulsive, attention-deficit/hyperactivity disorder (ADHD), aggression, and oppositional defiant behaviour. 50% of mental health problems of course of life occur during childhood and adolescence and about 10-20% of children worldwide experience mental health problems [5–8]. Although mental health illness is common among children, most of them remain unknown [5,9]. Identification of mental health problems at early stage has important role to reduce the negative impact on children's development [5]. It is important to note that the specific types of mental health disorder experienced by children can vary depending on the context. For example, conduct problems and ADHD [10–12] are often more noticeable at home, while externalizing problems and behavioral disorders are more prevalent in school setting [13–15]. Children usually report their internalizing problems [16,17]. Collecting information from multiple sources can improve the accuracy of identifying mental disorders in children [18,19].

In the field of mental health, researchers are currently identifying patterns of mental health disorders among children and categorizing subgroups with similar symptoms or behaviour within a population. By studying these subgroups, practitioners can better understand the prevalence of different patterns of mental health conditions, factors that influence these disorders, and plan the most effective interventions for each specific mental health pattern [20]. These efforts aim to improve the overall condition of children with mental health issues.

Mental health profiles are created based on variations in various aspects of behaviour/symptoms. Different profiles show different outcomes in the future. Some profiles show better physical health, improved peer relationships, and good academic outcomes in the long run, despite having some symptoms of mental disorders [21]. It is also important to distinguish between children who already have mental disorders and those who are at risk of developing them, as separate interventions are necessary for each group to achieve successful outcomes.

In studies that aim to identify mental health disorders among children, a common approach is to use a cut-off method. This method categorizes children into two groups: mentally healthy or mentally unhealthy, based on their total score on a checklist of symptoms [21,22]. However, this method may introduce large variability within each group since one group can have a wide range of scores. This approach also tends to force individuals to fit into common mental health patterns within the population [23]. There have been debates regarding labeling of certain mental health patterns as just good or bad. Some argue that patterns that exhibit a mix of characteristics should be separated from conventional mental disorder groups [21,24]. It is important to recognize that there are several subtypes of mental health profiles within the population, and researchers should employ methods that allow for exploration of those subtypes in their studies [23,25].

To identify mental health subgroups, several alternative statistical techniques have been applied. These techniques are chosen based on specific requirement and type of information about mental health indicators. Popular methods to discover subgroups within the population belong to a finite mixture modelling subset [26]. Techniques aim to capture the existing heterogeneity for grouping most similar individuals and creating enough distinctness among the groups like latent class models, clustering algorithm [27–30]. Cluster algorithms classify individuals within a population into clusters based on set of variables, calculating similarity and dissimilarity using distance among individuals [31–33]. Unlike probability-based models, cluster algorithms does not support statistical tests to check the clusters [34].

Latent class models also group individuals into latent classes based on the response for a set of variables where individuals within class show similar pattern of response and between classes have different patterns [20,35]. Latent class model assigns probability of belonging to a class to each individual and assumes that variables are independent of each other within a latent class [36]. It has been used in various fields to identify subgroups within population e.g., classifying internet users based on the purpose of using internet [37], adolescents based on their interest toward tobacco products [38]. In the field of mental health, latent class model also identified mental health conditions among children based on specific symptom, disorder, combination of disorders, symptoms reported by multiple sources or domains of mental health problems [20]. Mental health subgroups identified through LCA include high/severe, moderate, slight/mild or low level of disorder group [39–42], multimorbid/high risk class, internalizing problem, externalizing problem group [43–45], intermediate conduct problems with high ADHD and oppositional defiant disorder (ODD), intermediate ADHD and ODD but low/no conduct problems group [46,47], problematic behavior at home group, at school, cross-context (both school and home), problems reported by mother, reported by teacher and generalized problem (reported by mother and teacher) group [30,48].

Multilevel latent class analysis, extended version of LCA, considers the nested structure of data (lower-level units belong to higher-level units, e.g., students from several schools) and model the dependency among individuals occurred due to the nested structure [36]. Parametric and nonparametric version of multilevel latent class model incorporates this dependency in different ways where parametric version includes random effects for higher-level units and non-parametric version not only classify individuals into latent classes but also higher-level units into latent clusters based on the heterogeneity of lower-level classes [49]. Studies to discover mental health patterns modeled the dependency among observations due to nested structure of the data using parametric version and classification of higher-level units were not conducted [50,51].

In most studies, information used for latent class analysis was gathered from a single informant. In studies that did consider symptoms reported by multiple informants, all symptoms were included in the same latent class model, regardless of the informant [30,48]. However, this approach may violate the assumption of the latent class model because symptoms reported by different informants for the same children may be correlated. Unfortunately, this correlation was not considered in those studies.

In this study, we proposed a joint Latent Class Analysis (LCA) to identify mental health patterns among students using data from multiple informants, such as reports from both teachers and students themselves. We compared these mental health patterns using multiple informant data with those obtained from a single informant, such as Student's self report or Teacher's report. We also proposed multilevel joint LCA by extending the joint LCA to consider the nested structure of the data where students were nested within schools. This allowed us to identify clusters of schools based on difference in the prevalence of student's mental health patterns. Additionally, we explored the effects of student-level and school-level factors on mental health patterns and clusters of schools, respectively.

2. Literature Review

The literature review consists of five sections. The first section describes the mental health of children and the importance of identifying different subgroups within population. The second section reviews the statistical methods used to identify these subgroups based on observed information. The third section reports on the mental health subgroups identified in various studies. The fourth section provides the overview of the Strength and Difficulties Questionnaire (SDQ), and fifth section reviews individual and contextual factors that influence the mental health of children. The final section outlines the objectives of this thesis study, which use the data from a mental health survey conducted on Grade-5 students across Manitoba Province.

2.1 Child Mental Health

Approximately 14% of children aged 10-19 worldwide experience mental health problems [52], but most of those children go undiagnosed and untreated [8]. In the US, this percentage rises to nearly 20% for children aged 3-17 [53]. In Canada, about 1.2 million children and youth are affected by mental issues, with this number increasing to over 7 million among those in their mid-20s [54]. Unfortunately, only about one in five children and youth with mental health problems receive the appropriate care they need. In 2020, about 23% of hospitalization for children and youth (age 5 to 24) in Canada were due to mental health problems [55]. A 2016 report from the Manitoba Center for Health Policy (MCHP) revealed that 14% of children in Manitoba have mental illness while 8.5% have behavioural disorders [56]. Being mentally healthy is crucial for children as it allows them to develop appropriately, acquire social skills, maintain emotional well-being, and contribute positively to their family and society [1,2]. Mental health problems can disrupt normal childhood growth and have long-lasting effects into adulthood when the proper support is lacking. Early intervention for mental health problems in childhood has been shown to have long-lasting effect in preventing disorders in adulthood, but unfortunately, many children do not receive effective support for early symptoms of mental problems [57]. As a result, these symptoms often progress into severe mental disorders, leading to a continued prevalence of mental disorders in the population [58].

Psychological problems can broadly be classified into two main types of symptoms: internalizing problems and externalizing problems [59]. Internalizing problems encompass emotional symptoms and peer relationship problems, while externalizing problems includes conduct problems and hyperactivity/inattention [60]. Internalizing problem refers to attitudes and behaviors that are directed inwardly [59] and often goes unnoticed by others. These behaviors are influenced by negative emotions and can manifest as anxiety, depression/sadness, fear, social withdrawal, physical sickness, low confidence, and difficulty in verbal communication [45,60–64]. Emotional symptoms such as somatic, depressed, unhappiness, nervousness, low confidence, fear can arise from internal factors or external environments [65]. Despite their prevalence, these symptoms often go undetected [66]. Children who experience peer relationship problems, such as social withdrawal, exclusion, difficulty making friends, and being bullied, face long-term challenges in social and emotional adjustment [67]. Children with internalizing problems have difficulty to make friends at school and seem to avoid conversation with peers and teachers purposefully [62,68], show poor academic achievements, irregularity in school, poor communication and social skill [65,69–72] and have an increased risk of developing mental disorder such as antisocial behaviour, anxiety disorder, and depression [73].

Externalizing problems, in contrast to internalizing problems, are directed toward others and easily noticeable by people [64]. These problems include hyperactivity, losing temper, fidgeting, bullying others, lying, distracted, impulsive, stealing, and inattention. These problems can lead to behavioural disorder, delinquency, substance dependence, and attention-deficit/hyperactivity disorder (ADHD) [45,63,74,75]. Conduct problems, such as losing temper, fighting, lying, stealing, bullying others, and disrespecting adults, often results in detention at school, skipping classes, dropping-out from school, having poor relationship with peers and less friends [76,77]. About 3% school-going children experience conduct disorder [78], and this disorder in childhood increases the risk of developing anti-social personality in the long run [79]. Hyperactivity/Inattention, such as restlessness, fidgeting, being distracted easily, being impulsive, and difficulty in focusing or being attentive, can develop into attention-deficit/hyperactivity disorder (ADHD) and impulsive disorder [80]. Children with ADHD tend to have bad peer relationship and unstable friendships [77], suffer academically, may not develop mentally according to their age [81]. Externalizing problems have been found to be associated with academic difficulties [82], school dropout, and underachievement in school [83].

Pro-social behaviour, which reflects positive development, is an additional domain of mental health. Pro-social behaviours are actions directed toward others with the intention of helping them, such as being considerate of someone's feeling, sharing thing, helping people in need, being kind to others and offering help [84]. People show pro-social behaviours with the urge to ease some's suffering/pain and sympathy/empathy helps one person to understand other's pain, circumstances, problems [85]. Pro-social behaviour is often seen as the opposite of anti-social behaviour [86] and discourage aggressive behaviour toward others, fostering positive peer relationships [87].

The classification of individuals based on their mental health problems has long been a topic of concern in the field of mental health research and practice. This issue continues to be debated due to the nature of classification itself [88–90]. The accurate classification of individuals is important for practitioners and researchers to improve the care for mental problems and develop targeted intervention plan [20]. Especially accurate differentiation among children with symptoms can have a significant impact on their long-term outcome as social and behavioural development largely depends on the condition during childhood [91,92]. It is crucial to investigate the factors influencing various mental health problems, the growth of children with and without problems and the severity in occurrence of different forms of mental health problems within a population [20,90]. Current approaches, such as Diagnostic and Statistical Manual of Mental Disorders (DSM-5-TR) [93] and cut-off point method, provide pre-determined groups of mental health, but may not accurately represent the common mental health profiles in population [20,21,90]. The literature on mental health profiles of children argued in the favor of the existence of heterogeneity within pre-specified groups, subtypes of common mental health issues, as well as varying levels of difficulties [23,25,90,94,95].

However, the issue of classification becomes complicated due to the co-existence of multiple problems within research samples [96,97]. Comorbidity, refers to the co-existence of multiple mental disorders in one individual, is commonly observed in both clinical and community studies [98]. Displaying the symptoms of comorbid condition of mental disorder is common among children [99,100] and children with externalizing problems found to have internalizing problems during adulthood [45,60]. The high correlation between internalizing and externalizing symptoms suggests the presence of underlying factors that influence mental problem and co-occurrence of disorders [45,63]. Using strategy to screen one disorder by excluding other disorder made the interpretation of the mental disorder results problematic and the high rate of comorbidity among disorders implies that “pure” disorder instance (only depression or only anxiety) would not be a representative of the population [101]. To fulfill the purpose of identifying the symptoms or issues within children effectively, collecting information from different perspectives is recommended as some symptoms might not be noticeable to one perspective, but to other [18,19]. Perspectives of parents and teachers on the behaviour of children bring more information to the table along with the self-reported assessment from children [17,102–104]. Mental disorders among children have been found to be situation-specific, highlighting the importance of including multiple perspectives on children’s symptoms and behaviour [19,48,105]. Parents were more prone to identify conduct problems, ADH [10–12]; self-report mostly validated the presence of internalizing problems [16,17], and conduct problems, externalizing problems and behavioural disorders were more noticeable by teachers compared with other mental problems [13–15].

2.2 Statistical Methods for Identifying Subgroups

This section discusses common statistical methods used to identify subgroups within a population by analyzing their response pattern for a specific topic.

2.2.1 Latent Class Analysis (LCA)

Latent Class analysis is a person-centered statistical approach which clusters the respondents into subgroups based on their response for a set of observed variables [106,107]. The resulted subgroups are identified as Latent Classes and the variables whose responses are used to label the classes are called Indicators. Human behaviors and choices are believed to be guided by unobserved (or latent) characteristic and that characteristic or variable worked as explanatory variable to explain the relationship of the observed variables (which are the indicators here) [108,109]. Latent Classes are the categories of unobserved characteristic or latent variable. LCA identifies classes with individuals providing similar responses for indicators and individuals from different classes are different from each other based on the responses to indicators. LCA has several advantages over other traditional methods to identify mental health subgroups like cut-off point method. Traditional methods leads to cluster almost similar individuals in different subgroups and degree of heterogeneity can be high within a subgroup [20]. LCA applies maximum likelihood estimation method to separate individuals into different classes which consists of homogenous individuals and are heterogenous externally [27]. LCA is a model-based method which provides fit statistics based on which researchers choose final model and enables the comparison among models [110,111].

Moreover, LCA reports probability for each individual belonging to a specific class and classification error resultant from fitting the model in hand.

LCA can cluster individuals into classes based on the responses for a set of categorical or continuous variables (indicators) [112]. Usually LCA indicates the analysis to cluster individuals into classes based on categorical indicators whereas the term "Latent Profile Analysis" is used to refer to the same techniques for continuous indicators [112]. We preferred to use LCA to refer to this analysis. We used "Classes" and "Profiles" interchangeably for the rest of the study and indicated the type of indicator separately. For LCA using continuous indicators, all indicators are assumed to follow normal distribution and the means and covariance structures of the indicators are specific to each class. According to the research interests and nature of the responses of indicators, the covariance structure can take different forms. The general form of LCA with continuous indicators, wherewith class specific means and variances of indicators are used, can lead to overparameterization and less reliable results [113,114]. However, to reduce the bias and preserve the structure of the raw data, it is recommended to have class-specific variances of indicators [115–118].

LCA has been utilized in various studies to classify individuals into distinct groups based on complex characteristics. One study employed LCA to classify the internet users into different classes, such as traditional use, knowledgeable use, and interactive use, based on their frequency of internet use and reasons for using the internet [37]. Another study employed LCA to identify groups of individuals based on their complex characteristics considering bariatric surgery into three groups based on their concerns: cost, benefit, and procedure [119]. Similarly, LCA was employed to classify pediatricians into two groups: poorly trained with poor satisfaction and trained with high satisfaction based on their performance, perception, satisfaction toward managing sleeping problems among children [120]. Study on population-based cohort of adults in Tehran identified four groups of Metabolic Syndrome profiles: low risk group, metabolic syndrome with diabetes medicine, without diabetes medicine group and another group with low high-density lipoprotein (HDL) and high triglyceride based on their socio-demographic and behavioural characteristics using latent class model [121]. Another study discovered three groups among younger adults from Southern Connecticut based on their curiosity, interest to try and experiment tobacco products: influenced by all tobacco products (e.g., cigarettes, e-cigarettes, hookah, cigars, smokeless tobacco), influenced by e-cigarettes, hookah, and blunts, and mostly not influenced by any products [38]. LCA reported five classes of families based on several measures of socio-economic status using student health and well-being survey data of UK and the classes were nonworking families, highly affluent families, deprived families, lower affluent, and affluent families with deprived schools [122]. Using an econometric approach of latent class in a field experiment, four groups of individuals were identified based on their food purchase behaviour and preference: prestige lovers, ambitious shoppers, shoppers concerned about product functionality and shoppers with high income and preference for expensive brands [123].

2.2.2 Multilevel Latent Class Analysis (MLCA)

A multilevel extension of the latent class model is employed to analyze data that has a nested structure. This is referred to as multilevel latent class model [49,124]. The nested structure implies that there are multiple observations within each group, such as multiple measurements from individuals, and individuals sampled from higher-level entities. Responses collected from students (lower level) belonged to several schools (higher level) comprise nested data structure where observations of students from same school tend to be similar in nature and different than students from other schools. This produces the dependency among responses of students from same schools. The multilevel modeling takes this dependency into account to avoid the violation of independency assumption in the latent class model and to prevent bias in parameter estimation [49,124]. If researchers ignore the nested structure of the data, it can lead to poor classification of lower-level units [125] and the selection of incorrectly specified models [126].

Two approaches are used in the multilevel latent class modelling to account for the dependency caused by the nested structure. The first approach is the parametric approach, which allows for random-effects of lower-level latent class across higher-level units. The second approach is the non-parametric approach, which classifies not only lower-level units but also higher-level units into latent classes [49]. As the number of lower-level latent classes increases, the parametric approach becomes compute-intensive [127]. The random-effects of higher-level units are assumed to follow normal distribution for the parametric approach, while the non-parametric approach considers a multinomial distribution [128]. This makes the non-parametric approach computationally lighter and does not depend on the assumption of normality [35]. Another advantage of the non-parametric approach is that it provides meaningful insights and explanation of the dependency due to the higher-level units. It can show severe characteristics in some higher-level units and the absence of the same characteristics in others [129]. This clustering of higher-level units allows for the discovery of whether the presence of a specific lower-level class is higher in some higher-level units and absent in others.

MLCA has been widely applied in various fields for different purposes. For example, it has been used to cluster cities in Taiwan based on the internet use segments. In one study, Taiwan was clustered into three class: Metropolitan, Southern Northern Taiwan, based on the response on internet usage from consumers from 25 cities [37]. MCLA has also been used to cluster higher-level units such as universities based on their student's research practices [130]. Another study used MCLA to cluster residential areas based on residents alcohol dependency [131]. Additionally, MCLA has been applied to cluster countries based on their ownership of financial products [132].

The MLCA model has also been applied to large-scale data from the health-care industry in the Netherlands. The final model provided two classes at employee level and two classes at team level [36]. The "Diverse class" at the employee level consisted of employees with high level of work variation, diversity, and creative work. On the other hand, the "Structured class" had a low level of work variation and a high level of repetitive work. At the team level, the "Diverse group" included a higher proportion of teams with diverse employees, while most of the teams in the "Uniform group" consisted of structured employees.

The parametric version of MLCA was applied on smoking behaviour of Community Drug and Alcohol Survey (CDAS) data among 9th grade students from several communities in the United States [35]. The model reported three student-level classes: heavy-smokers, moderate smokers, and non -smokers. The chance of belonging to any student level class was allowed to vary across the communities. This model supported the fact that, the chance of girl belonging to heavy smoker group might be higher for some communities and very low for some other communities. The non-parametric MLCA model discovered two community-level clusters: high use communities and low use communities, along with the three student-level class from the same response data. However, the parametric model was found to be the best fit for the data.

2.2.3 Cluster Analysis

Cluster analysis is a method to classify individuals based on their similarities and dissimilarities in some criteria [31]. The goal is to group individuals into clusters where individuals within a cluster are similar to each other, while individuals from different clusters are different from each other based on those criteria [133]. The similarity and dissimilarity among individuals are decided through classification procedure which relied on density of population or closeness/distance between individuals [32,33]. There are several clustering techniques that are used as per the situation of the research like hierarchical techniques (subtypes: agglomerative or divisive), optimizing techniques, density or mode-seeking techniques, clumping techniques. To compute the similarity or dissimilarity between individuals, measurements like Manhattan distance, correlation, and mutual information are commonly used [33,134]. These measurements help quantify the degree of similarity or dissimilarity between individuals in the clustering process.

The resulting clusters in the cluster analysis depend on the researcher's choice for techniques and measurements. Cluster analysis reports clear-cut membership of individuals in clusters [134,135]. Cluster analysis is not based on probabilistic model like latent class analysis and it assumes association among variables within a cluster to show similarity between individuals [135,136]. Examples of using cluster analysis in studies include clustering college students based on their performance on perfectionism [137], clustering countries based on the poor living condition and work condition for people with disabilities (such as high risk of poverty cluster) [138], and clustering authors from social science and humanities based on their publication patterns [139].

However, there are several disadvantages of cluster analysis compared to latent class analysis. Firstly, the cluster analysis does not use probabilistic model like LCA, which allows for the inclusion of different form of the model with different restrictions and supports statistical tests to validate the findings. Secondly, cluster analysis does not provide different formal fit indices to compare models, as LCA does. Thirdly, the choice of clusters in cluster analysis is more arbitrary compared to the classes of LCA. Fourthly, cluster analysis does not allow for the analysis of mixed types of variable. Lastly, cluster analysis does not report any error due to misclassification, as it does clear-cut assignment of individuals in clusters [34,110].

2.3 Mental Health Subgroups within Population

In of the field of mental health, researchers aim to classify participants into different subgroups within a population in order to plan targeted interventions for these subgroups. These studies may focus on specific mental health disorder, symptom, or a broad range of mental health conditions. By investigating subgroups of mental health, researchers have discovered several intermediary levels different from conventional “all-or-nothing” classes. This suggests that mental health conditions can be better understood and addressed by considering the various subgroups and their specific needs.

LCA is recommended for analyzing complex concept structure like mental health and developmental psychopathology. It helps to capture the heterogeneity within a population and study individual psychological aspects [20,140–142]. LCA is an inductive approach that investigates subgroups and identifies non-conventional subgroups. A cross-sectional study on students of United Kingdom aged 8-9 years explored anxiety profiles by applying latent class model on continuous scores for general, test and math anxiety. Four ordinal anxiety profiles: high anxiety, moderate anxiety, slight anxiety and low anxiety profiles were discovered with slight anxiety profile having most students [39]. Similar age group in another study on social and separation anxiety depicted two classes after analysing through latent class model which are high/moderate problem and low problem class [40]. Two samples: one consists of children who were referred for mental health services and another of nonreferred children from USA were analyzed to explore comorbidity of anxiety and depression. Latent class analysis reported three classes among children for both samples separately and those classes represented the combination of anxiety and depression. The classes were different from each other in terms of severity of anxiety and depression together and supported the comorbid status [143].

In a study conducted in Western China, a Mental Health Test (MHT) was used to assess anxiety among children. LCA was applied, and three types of children were identified: low risk, mild risk and high risk [144]. Another study conducted in Netherlands focused on identifying subtypes of ADHD among children aged 10. The study used six response samples: 1) conditions of boys reported by mother, 2) boys reported by father, 3) girls reported by father, 4) girls reported by mother, 5) boys reported by teachers, and 6) girls reported by teachers, separately. Investigator reported subtypes of ADHD from latent class analysis like no/mild symptoms of ADHD type, mild inattentive, severe inattentive, hyperactive-impulsive, severe ADHD type for separate samples [145]. A study in the US focused on attention problem among children. LCA was used to analyze non-referred children and clinically referred children separately. For non-referred children, three latent classes were identified: low/none, mild and moderate attention problem classes. For clinically referred children, three classes were identified: mild, moderate and severe attention problem classes [41].

A study on Dutch twins for ODD reported four groups of children based on symptoms reported by mothers through multilevel LCA [51]. The multilevel LCA model incorporated the dependency among twins through estimating robust standard error and the identified classes were no/low symptom class, defiance problem class, irritation problem class, and high problematic class. Similar studies focusing on specific symptom/disorder or combination of closely related mental issues discovered ordinal mental health classes like high/severe, moderate, slight/mild or low problem classes through latent class technique [146,147].

Classes being different from each other based on type and risk were discovered on the studies investigating combination of symptoms or disorders. In a school-based study on elementary school-going children in the Netherlands, indicators of conduct problems, oppositional defiant disorder (ODD) problems, attention-deficit/Hyperactivity (ADH) problems were analyzed together using latent class analysis for categorical indicators. Three classes: intermediate conduct problems and high ADH and ODD problems, intermediate ADHD and ODD but low conduct problems, and low for all problem class were found among those children [46,47]. Another study on US children aged 6-18 years and children (aged 7, 10 and 12 years) from the Netherlands explored the obsessive-compulsive behavior subgroups among children through the application of latent class analysis. Authors reported four subgroups: no symptom group, worries and has to be perfect group, thought problems group and obsessive compulsive disorder group [148]. Only conduct problem among children from New Zealand was assessed through latent class technique but the condition of children was reported by mother and teachers. Latent class model was applied on two dichotomous variables: one representing child having conduct problems as per mother and another representing teacher's perspective. Study discovered four subgroups of children: no conduct problem, mother-reported, teacher reported, and generalized conduct problem group (children with conduct problem as per both mother and teacher) [48]. Similar subgroups were discovered for conduct problems among children from US where total 10 ordinal variables (five for parent's and 5 for teacher's perspective) representing conduct issues of a child were analyzed through latent class analysis [30]. Studies considering the multiple perspective to identify mental health subgroups included all variables representing different perspectives in the same latent class model and reported the subgroups based on that latent class model with the assumption of variables being independent of each other in each subgroup. Symptoms of mental problems of children reported by multiple informants using same set of questions might introduce correlation among the responses of same questions from informants.

In a cross-sectional study on children with disabilities in Sweden, cluster analysis techniques were employed to group children into different clusters based on their mental health symptoms. Fiver clusters were identified among children [149]. These clusters were named conduct problem cluster, emotional problem, prosocial difficulties, all difficulties, and no symptom cluster. The clusters were determined based on the mean score for mental difficulties scale of children in each cluster. Another study was conducted in Sweden using data from 2017-18 Health Behaviour in School-aged Children (HBSC) survey on children aged 11 years. Cluster analysis was applied to measures such as headache, backache, stomach problems, depressed, anger issue, anxious, sleeping problems, dizziness and self-reported overall health [150]. Four clusters were identified. One cluster included participants with good health status, another included participants with poor health status, third included individuals with high symptoms of mental issues, and last one included individuals with average symptoms and overall health. A study conducted in Canada analyzed survey responses regarding the change in mental health condition of autistic children after COVID-19 pandemic using cluster analysis technique [151]. The study identified two clusters, namely deteriorated clusters and unchanged clusters. The clustering technique was applied to six indicators representing two domains of mental health problems which were internalizing problems (indicators: depressed, anxious and OCD) and externalizing problems (indicators: inattentive, hyperactive and irritation). In another study involving participants aged 9-17 from Italy who had a history of non-suicidal self-injury, cluster analysis was used to identify mental health subgroups [152]. The analysis was based on indicators representing social competence, social problems, and affection problem. The researchers reported four clusters among those participants as 1) moderate affection problems with good social functioning, 2) socio-affective difficulties, 3) good with socio-affection and 4) low affection problem with low social competency.

2.4 Strength and Difficulties Questionnaire (SDQ)

Screening instruments are used in studies to target specific symptoms or disorders. One such instrument is the Strength and Difficulties Questionnaire, which assesses mental health conditions in children. The SDQ covers internalizing (emotional symptoms and peer relationship problems) and externalizing difficulties (conduct problems and hyperactivity/inattention) and pro-social behaviour through 25 statements [153]. Informants, such as parents, teachers, or children themselves, respond to these statements based on their observations of the child being screened. The SDQ can serve as an initial assessment tool for clinical diagnosis and measuring treatment outcomes [154]. Previous studies have used the SDQ to determine the prevalence of common mental health disorders and their risk factors among children. Conventional cut-off methods for difficulties scores were used to define these disorders [64,66,155–157]. Some studies have also used responses to SDQ statements related to mental health to discover subtypes and combination of common mental health problems. For example, a study in Spain involved children aged 6-8 reported four mental health groups using the SDQ: high difficulties group, internalizing difficulties, externalizing difficulties and well-adjusted (high score for pro-social behaviour and low for all difficulties) group [63]. Similar subgroups were discovered among elementary school children in Pennsylvania using the SDQ, with teachers reporting on the children's externalizing and internalizing problems [60].Another study on adolescents from the Netherlands using adolescent and parent versions of SDQ reported six subgroups: no problems group, borderline hyperactivity problems, borderline conduct and social problems, emotional problems, emotional and social problems and overall problems group [19]. Adolescents were unable to distinguish the borderline conduct and social problem group and emotional problem group from no problem group whereas parents reported at least one problem for each group. Five subgroups were reported based on the responses provided by parents through SDQ among children with disabilities in Sweden [149]. The subgroups were no problem group, conduct problem, emotional problem, all problems and prosocial problem group.

2.5 Factors Associated with Mental Health Patterns

The development of children usually depends on conditions and environment in which they grow up. Various factors within those conditions or environment can have a negative impact on the mental and physical development of children. Researchers are investigating various factors, such as personal, environmental, or contextual, that can influence the mental conditions of children.

The risk, symptoms and prevalence of mental disorders have consistently shown gender difference, which has led researchers in the field of mental health to focus on gender as a significant factor. The national survey of Great Britain of 2004 reported that among primary school children, internalizing problems are higher in girls and externalizing problems are higher in boys [158]. A study on sample aged 12-15 years from central Norway found that depression symptoms were more common among girls than boys [159]. Similarly a study involving high school students in western China revealed that females had a higher risk of anxiety compared to males [144]. According to a meta-analytic review of 166 studies, sadness and anxiety were significantly higher among girls than boys while boys showed higher level of anger [160]. Studies on children with and without disabilities reported that boys had higher risk of hyperactivity issue compared with girls [64,160–163]. Studies on depression have identified factors that increase the risk of depression, and these factors were more commonly found among women, ultimately leading to a higher prevalence of depression among them [164]. The studies mentioned above have highlighted the gender difference in mental health problems. Additionally, when identifying subgroups within the population based on symptoms of mental health issues, the gender of an individual could possibly influence the placement of individual in any given subgroups. Therefore, gender is a crucial factor that needs to be studied in mental health field. Thus, understanding is important in order to design targeted interventions for specific mental health patterns, taking into consideration the characteristics of gender that is more prevalent in that group.

The condition and characteristics of school are important to analyze in order to protect children's mental development, as they spend most of their time there. This has been a major focus in this field [165–167]. Negative experiences with peers and mistreatment by older students can influence children's development, causing them to withdraw from social interactions, peer relationships, verbal communication with others and show increased level of depression, anxiety [168].

A cohort of Dutch school students were examined to understand the effect of bullying on their behaviour and attitude in school. The authors reported their findings for four groups of students: bully, victims of bully, bully-victim (who bullied others and got bullied) and uninvolved [169]. The bully and bull-victim groups showed higher level of aggressive behaviour compared to the uninvolved and victim groups, with the uninvolved groups displaying the lowest level of aggressions. Victims and bully victims experienced the most social isolation, and bully-victims were the least liked by others. A cross-sectional study conducted in the US on school students aged 6 to 10 years found that both victims of bullying and those who bully others showed externalizing symptoms of mental disorders, such as violent attitudes [170]. Children who are bully, victim or bully-victim at school are evidently showing higher risk of having mental health problems. Schools where bullying occurs to some extent may have a greater number of students with various mental issues. The bullying environment in school can be an influential factor for mental health problems, providing valuable insights for planning school-based interventions.

Difference between immigrants and non-immigrants in mental health problems has found to be mixed in nature from previous studies. According to community health survey in Canada, the prevalence of mental disorder among immigrant population was lower compared to Canadian-born individuals, and those who recently immigrated had the lowest level of depression and alcohol problems [171]. Studies conducted in British Colombia and Ontario on children also found that the prevalence of mental disorders were lower for immigrant children compared to the non-immigrant children, which was reflected in the number of children seeking mental health services [172–174]. Also, the difference in using mental health services among immigrants and non-immigrants might have happened due to barriers in seeking help for immigrants such as language, unaware of services, stigma of mental health problems. However, immigrant children were found to have a higher risk of depression and anxiety disorders, which could be attributed to the challenging experiences they faced after migration [174]. Mental disorders were less prevalent among immigrants compared to the non-immigrants based on nationally representative survey in U.S. but the prevalence increased to be at same level over time [175]. The difference among immigrants and non-immigrants in mental health conditions may be related with the cultural settings, sociodemographic factors and other factors related to their home country. These differences may also impact the chance of having different mental health problems with different severity. Including immigrant status as a factor in the research can provide insights into the influence of immigration status on mental health patterns.

2.6 Objectives

In this thesis project, we aimed to propose multilevel joint latent class analysis using data from multiple informants/sources (teacher's reports and youth self reports) to identify the patterns of mental health problems among children in Manitoba. We also examined whether these patterns vary across schools and explored individual and contextual predictors associated with these patterns and clusters respectively.

To the best of our knowledge, existing literature has not considered the possible correlation among responses from multiple perspectives and the nested structure of the data when identifying mental health patterns among children. Previous studies have typically identified subgroups based on the symptoms reported by either parents/guardians, teachers, or children themselves. Some studies have included multiple perspectives but either investigated them separately or together without considering possible correlation. A few studies have considered the multilevel/nested structure of the data, where lower-level units (e.g., students) are nested within higher-level units (e.g., schools). However, these studies have often mitigated the dependency among lower-level units by incorporating random effects for heterogeneity across higher-level units, without considering the classification of higher-level units.

Therefore, our thesis project aimed to address these gaps by proposing a multilevel joint latent class analysis that considers the correlation among responses from multiple informants and the nested structure of the data.

Firstly, we proposed the joint latent class analyses for items reported by teachers and students simultaneously. We introduced a latent pattern variable to represent the connection between the two informants. We compared the patterns of mental health identified by the joint LCA approach with that by separate LCA for items teacher-reported and students reported items.

Secondly, we proposed multilevel joint LCA to explore whether there was the heterogeneity in these mental health patterns across schools. By incorporating the multilevel structure, we examined if there were variations in school clusters, which could be explained by an underlying latent variable.

Thirdly, we investigated how individual and contextual factors were associated with these mental health subgroups identified by multilevel joint LCA. In the selected multilevel joint LCA model structure, we included student-level factors such as sex, Canadian-born and bullying as student-level factors to investigate the effects of these factors on latent mental health patterns of students.

We also added the school level factors such as students bullying at school to examine their impact on the latent clusters of schools.

3. Methodology

3.1 Grade-5 Mental Health Survey

The Grade 5 Mental Health Survey was conducted in Manitoba during the 2015/2016 school year by the Healthy Child Manitoba Office (HCMO), in collaboration with the Manitoba Association of School Superintendents, Manitoba School Boards Association, Manitoba Teachers Society [176]. The survey aims to assess the mental health condition of Grade 5 students using the Strength and Difficulty Questionnaire, which was completed by both students and their teachers. This questionnaire consisted of 25 items that assess attributes related to children's behaviour, including both positive and negative attributes. These items covered five domains of behaviour: 1) Emotional Symptoms, 2) Conduct Problems, 3) Hyperactivity/Inattention, 4) Peer relationship Problems and 5) Prosocial Behaviour. Informants responded to each item by choosing one option from a 3-Likert scale ("Not True", "Somewhat True", "Certainly True") to indicate the extent to which the item applied to the student.

The survey consisted of three questionnaires: i) SDQ for teachers, ii) SDQ for students and iii) "More Information about Yourself" for students. Students completed the SDQ based on their feeling and experience over the last six months, while teachers provided their perspective on student's behaviour over the same time period, or sometimes over the past year. The questionnaire completed by Teachers to assess the behaviour of students contains exact same items [177], but the wording was slightly changed in the version of questionnaire for Student's self-reporting [178]. The teacher version included additional questions to assess whether the student has any difficulties in areas such as emotions, concentration, behaviour or social interactions. It also asked about the duration of these difficulties, the level of suffering experienced by the student, and whether these problems affected the entire class or the teacher [179]. Personal information collected included student ethnicity, birth-country, self-reported health and mental health condition, age and sex.

3.1.1 Ethics Approval

This thesis used the existing Manitoba Grade 5 mental Health survey dataset to investigate the mental health subgroups among grade 5 students during 2015/2016 school year by applying proposed joint latent class models. The ethical approval to conduct investigation of the data had previously approved by Health Research Ethics Board of University of Manitoba (H2017:233). Since we are using this data for this graduate thesis project, we also obtained ethics approval from Health Research Ethics Board of University of Manitoba (H2024:004).

3.2 Data

The original dataset had information of 11406 grade-5 students from 412 schools in Manitoba. Both teachers and students themselves reported the same set of 25 SDQ items.

3.2.1 SDQ for Students

The SDQ for students provided 25 variables representing 25 items which covered the five domains of student's self-reported behaviour. All 25 items were measured in 3-Likert scale ("Not True=0", "Somewhat True=1", "Certainly True=2") and yielded 25 ordinal variables in data.

3.2.2 SDQ for Teachers

The SDQ for teachers provided 25 variables representing 25 items which covered the five domains of student's behaviour identified by teachers. Same as students' part, all 25 items were measured in 3-point Likert scale ("Not True=0", "Somewhat True=1", "Certainly True=2") and yielded 25 ordinal variables in data.

3.2.3 Student's Personal Information

As part of student's personal information, one variable presented the sex of students and another variable represented the self-rated mental well-being of student through 5-point Likert scale ("Excellent=0", "Very good=1", "Good=2", "Fair=3", and "Poor=4"). One dichotomous variable measured whether student was born in Canada ("Yes=1", "No=2").

3.3 Data Analysis

Data analysis was conducted using R software version 4.1.2 and LatentGold software version 6.0. Data were processed before conducting latent class analyses. Five items in prosocial domain (Q1, Q4, Q9, Q17, and Q20) were reverse coded as "Not True=2", "Somewhat True=1", "Certainly True=0" for teacher's and student's variables. Self-reported mental well-being of students was recoded as "Excellent=4", "Very good=3", "Good=2", "Fair=1", and "Poor=0". For Student's and Teacher's questionnaire individually, the scores for emotional symptoms, conduct problems, hyperactivity-inattention, peer problems and prosocial behavior domains were calculated from 25 items [180]. The scores were calculated by summing values of valid items of each domain, then dividing the scores with the number of valid items, then multiplying the result with 5. In this way, even if one item was missing, the score was not missing for a domain and reduce the amount of information loss. We obtained in total 10 scores (five from student's perspective and five from teacher's perspective). High scores on a domain indicate high level of difficulties experienced by students in that area. The observations with missing information for any of these scores were removed from the dataset. We also removed observations with missing for covariates such as sex, Canadian-born etc. The demographics and distribution of domain scores for students, both before and after removing observations, showed no significant differences, indicating that the missing data had little impact on the overall results. To reduce biasness, accurately select model and by considering the complexity of the proposed models, the number of students per group is restricted to be at least 10. So, we removed observations of students who belonged to those schools from which less than 10 student's information was added in this data. Finally, information of 8167 students from 284 schools and 31 division was included in the analysis. Two variables were created based on the item "Often fights with other youth or bullies them" from the Teacher's questionnaire: one variable to indicate if the student is a bully at an individual level, and another variable to measure the prevalence of bullying at school. This item reported by teachers was converted (recoded "Not True=0" and "Somewhat True=1", "Certainly True=1") into a dichotomous variable (Yes/No) to determine if the student is a bully or not. The percentage of students who were bullies was used as a measure of the prevalence of bullying at school. A British cut-offs [181,182] for categorizing the scores of SDQ domains into three categories: "normal", "borderline" and "abnormal" was used to identify students with difficulties score in abnormal range according to self-report and teacher's report. The cut-offs are different based on whether it was reported by teachers or students and the slightly adjusted cut-offs for continuous scores of SDQ domains were presented in Table A1 of Appendix.

3.3.1 Identifying Mental Health Classes among Students

3.3.1.1 Latent Class Analysis (LCA)

We applied the latent class model with five domain scores reported by teachers as indicators and separately to five domain scores reported by students themselves. Five domain scores were treated as continuous indicators, which were used to group students into different latent classes.

Consider $Y_i = (Y_{i1}, Y_{i2}, \dots \dots \dots, Y_{iK})$ where Y_{ik} represents the response of individual i for continuous indicator k with $i = 1,2, \dots \dots \dots, n$ and $1 \leq k \leq K$. The simple Latent Class model for K continuous indicators is:

$$f(Y_i) = \sum_{t=1}^{T} P(C = t) f(Y_i | C = t)$$

Where C represents the latent variable containing T latent classes as categories with $1 \leq t \leq T$. Different form of LP model can be obtained by specifying different form for $f(Y_i|C = t)$. Assuming all continuous indicators follows class -specific multivariate normal distributions, the form becomes:

$$f(Y_i|C = t) = (2\pi)^{-T/2} \, |\Sigma_t|^{-1/2} \exp\{-\frac{1}{2}(Y_i - \mu_t)' \, \Sigma_t^{-1} \, (Y_i - \mu_t)\}$$

In this model, each latent class has its own mean μ_t and variance-covariance matrix Σ_t. Another model which assumes that all covariances are zero is known as the most restrictive model. This model holds the local independence assumption. This model takes form as follows: $f(Y_i) = \sum_{t=1}^{T} P(C = t) \prod_{k=1}^{K} f(Y_{ik}|C = t)$ with $P(C = t) = \frac{\exp(\gamma_t)}{\sum_{t=1}^{T} \exp(\gamma_t)}$ and $f(Y_{ik}|C = t) = \frac{1}{\sqrt{2\pi\sigma_{kt}^2}} \, exp\left\{-\frac{(Y_{ik} - \mu_{kt})^2}{\sigma_{kt}^2}\right\}$.

The variance-covariance matrix Σ_t can specified in different form. It is possible to specify some off-diagonal elements to be zero and others not based on the pair of indicators. Also, we can specify the model with class-independent variances by replacing the σ_{kt}^2 by σ_k^2 or Σ_t by Σ. This yields a more restrictive version of latent class model where classes are different from each other based on the mean values of the indicators and variances for indicators are same across all classes. The violation of local independence assumptions can lead to a poorly fitted latent class model. Usually, the number of latent classes are increased to obtain the better fitted model and fulfill the assumption. This might provide complex model as better fit. Another way is to relax this assumption for the model by adding the direct relationship between indicators which have large value for bivariate residuals indicating strong relationship. The parameters of the LCA for continuous indicators is obtained by maximizing the log-likelihood as follows,

$$\log L = \sum_{i=1}^{n} \log f(Y_i)$$

which is known as maximum likelihood (ML) estimation technique through expectation-maximization (EM) and Newton-Raphson (NR) algorithm [183]. If the indicators are non-normally distributed, robust ML estimation techniques is applied [184]. The log- likelihood might not reach its global solution and might be stuck at local [185]. To solve this problem, there was a set of random start values, and the model was iterated several times. This procedure increases the chance of getting a global solution with large number of random start values and iterations [26,27,183]. Fit indices of several models of different values for T was compared to select the optimal number of classes T. After selecting the number of classes T, students were assigned to mental health classes based on posterior class membership probabilities.

Model Specification: For latent class analysis models, we used four different specifications for variance-covariance matrix $\boldsymbol{\Sigma_t}$: a) Class-dependent diagonal $\boldsymbol{\Sigma_t}$ – all covariances are assumed to be zero and variances are dependent on classes ; b) Class-independent diagonal $\boldsymbol{\Sigma_t}$ – all covariances are assumed to be zero and variances are independent of classes; c) Class-dependent unrestricted - variances and covariances are dependent on classes; and d) Class-independent unrestricted $\boldsymbol{\Sigma_t}$ – variances and covariances are independent of classes.

Table 3.1: Specifications for variance-covariance matrix of the latent class models

Five domain scores reported by Teachers	**Five domain scores reported by Students**
a) Class-dependent diagonal $\boldsymbol{\Sigma_t}$	a) Class-dependent diagonal $\boldsymbol{\Sigma_t}$
b) Class-independent diagonal $\boldsymbol{\Sigma_t}$	b) Class-independent diagonal $\boldsymbol{\Sigma_t}$
c) Class-dependent unrestricted $\boldsymbol{\Sigma_t}$	c) Class-dependent unrestricted $\boldsymbol{\Sigma_t}$
d) Class-independent unrestricted $\boldsymbol{\Sigma_t}$	d) Class-independent unrestricted $\boldsymbol{\Sigma_t}$

We started by fitting models with diagonal $\boldsymbol{\Sigma_t}$ to maintain the local independency assumption of indicators. The models with diagonal variance-covariance matrix $\boldsymbol{\Sigma_t}$ gave bivariate residuals for each combinations of indicators which were larger than 2. This indicates that there is local dependency among indicators. Because of this, we fitted models with unrestricted $\boldsymbol{\Sigma_t}$ to relax the local independency assumptions for each combination of indicators, starting from the combination with highest bivariate residuals. We were informed that there was local dependency among every combination of indicators. So, we concluded to run models with unrestricted $\boldsymbol{\Sigma_t}$. We compared the Class-dependent unrestricted $\boldsymbol{\Sigma_t}$ models and Class-independent unrestricted $\boldsymbol{\Sigma_t}$ models and concluded that class -dependent unrestricted $\boldsymbol{\Sigma_t}$ models performed better in terms of information criteria values [Table A2 of Appendix]. LCA provided one set of mental health classes among students based on the indicators reported by teachers and another set of mental health classes based on the indicators reported by students themselves.

3.3.1.2 Joint Latent Class Analysis (Joint LCA)

In previous section, we conducted separate LCA on indicators reported by teachers and students themselves. However, our objective is to explore groups of students based on their mental health condition reported by both teachers and students simultaneously. The data contains responses for the same students on same indicators from two sources (informants), which may introduce correlation among same indicators from different sources. This might violate the local independence assumption of LCA for indicators. To address this, we proposed introducing another latent variable (latent pattern) to capture the correlation among same indicators from different sources. This approach provided mental health patterns which is the combination of mental health classes from sources.

Suppose we have responses for n individuals on K continuous indicators from S sources and these indicators are representing a common idea. The response of one individual from source s is denoted as $Y_i^{\,s} = ({Y_{i1}}^s, {Y_{i2}}^s, \dots \dots \dots \dots, {Y_{iK}}^s)$ where ${Y_{ik}}^s$ represents the response for k indicator of i individual from source s with $1 \leq k \leq K$. Joint LCA assumes that n individuals can be classified into T classes of an unobserved latent variable C_s for each source s which directs the responses of individuals for these K indicators. And latent classes from sources are assumed to be associated with each other as responses are for same indicators from different sources and this association is measured using another latent variable G (Latent Pattern). This yields joint LCA model with T classes and L patterns from responses for K continuous indicators from S sources.

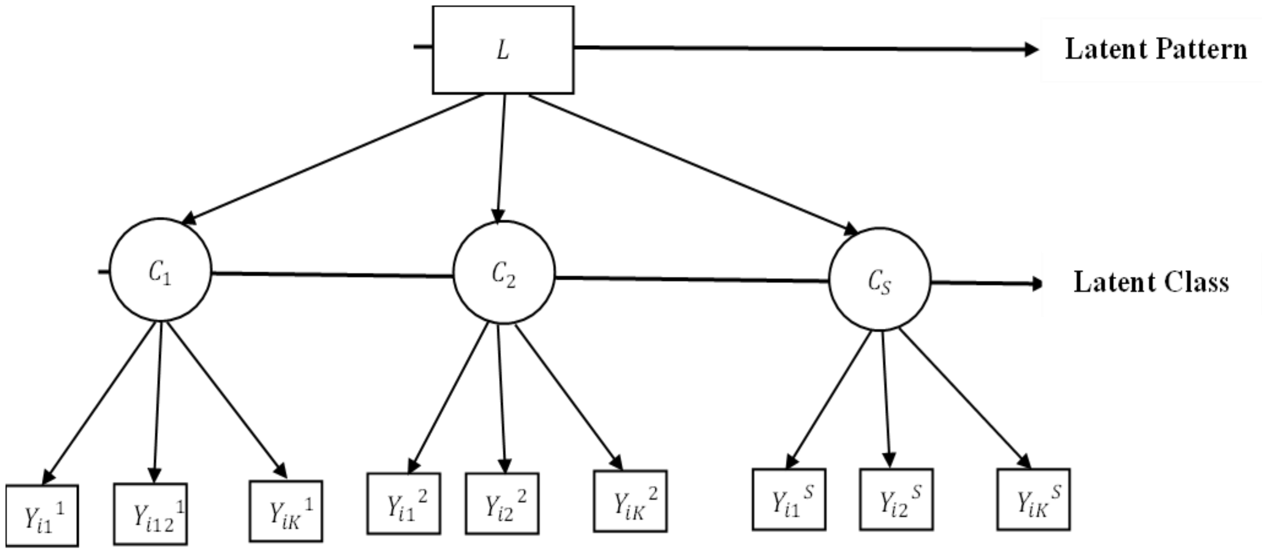

Figure 3.1: Graphical example of a joint LCA

The model or probability of a responses pattern Y_i (response of one individual) can be written as

$$P(Y_i) = \sum_{l=1}^{L} P(G = l) \prod_{s=1}^{S} \left\{ \sum_{t=1}^{T} P(C_s = t \mid G = l)\, f({Y_i}^s | C_s = t) \right\}$$

Where, a) $f({Y_i}^s | C_s = t)$ represents the multivariate normal density function,

b) $P(C_s = t \mid G = l) = \eta_{t|l}^{s}$ represents the probability of belonging to class t for source s for a given pattern l,

c) $P(G = l) = \gamma_l$ represents the probability of belonging to pattern l.

This model assumes that each latent class has its own set of mean μ_t and variance-covariance matrix Σ_t and these parameters of classes are equal for sources to maintain the same classes for sources which ensures invariant interpretation of classes across sources. This model holds the following local independence assumptions:

a) The continuous indicators for source s are conditionally independent of each other given a latent class for source $s; s = 1, \dots \dots, S$.

b) The latent classes of all sources are conditionally independent of each other given a latent pattern.

This model takes form as follows:

$$P(Y_i) = \sum_{l=1}^{L} P(G = l) \prod_{s=1}^{S} \{\sum_{t=1}^{T} P(C_s = t \mid G = l) \prod_{k=1}^{K} f(Y_{ik}{}^{s} | C_s = t)\}$$

$$= \sum_{l=1}^{L} \gamma_l \prod_{s=1}^{S} \left\{\sum_{t=1}^{T} \eta_{t|l}^{s} \prod_{k=1}^{K} f(Y_{ik}{}^{s} | C_s = t)\right\}$$

with $f(Y_{ik}{}^{s} | C = t) = \frac{1}{\sqrt{2\pi\sigma_{kt}^2}} exp\left\{-\frac{(Y_{ik} - \mu_{kt})^2}{\sigma_{kt}^2}\right\}$.

As mentioned in LCA, the violation of LCA is mitigated by increasing the number of categories of latent variables (here, latent classes and latent patterns). Also, if needed, this assumption is relaxed by allowing the possible dependence in the model. The parameter of the model with different number of latent classes and latent patterns are usually estimated by maximum likelihood method. The log-likelihood function which is maximized by this method is:

$$\log L = \sum_{i=1}^{n} \log P(Y_i)$$

Parameter estimation of joint LCA follows the techniques like the LCA. In addition, the expectation step of EM algorithm applies forward-backward variant [186] and NR algorithm applies forward recursion scheme [187] to incorporate the structure of large of individuals per sources [183]. Then the optimal number of classes T and number of patterns L are selected by comparing the several models of different values for T and M based on fit indices.

Model specification: Simple LCA suggested that there was local dependency among the pairs of indicators and supported class-dependent unrestricted $\boldsymbol{\Sigma_t}$ models. Based on that information, we fitted joint LCA models with class-dependent unrestricted $\boldsymbol{\Sigma_t}$. The Joint LCA reported mental health classes and mental health patterns among students based on both sets of indicators reported by teachers and students themselves.

3.3.1.3 Statistical Criteria for Model Selection

Final class model was selected based on statistical criteria and interpretability of the classes and patterns and several criteria were considered to evaluate models. For selecting final model for both LCA and joint LCA, following selecting criteria were considered:

- Information Criteria (IC): Lower value for ICs indicates better model. In practical, it is common to decrease value of ICs as the number of classes increase and not reach any minimum point. In such instances, the values of ICs are plotted to visually detect the class after which only small decrease happen with each additional class, known as "elbow plot".

Table 3.2: Information Criteria

Information Criteria	**Formulae**
Akaike Information Criteria (AIC)[188]	$-2LL + 2P$
AIC3[189]	$-2LL + 3P$
Consistent Akaike Information Criteria (CAIC)[190]	$-2LL + (1 + \log(n))P$
Bayesian Information Criteria (BIC)[191]	$-2LL + \log(n)\, P$
Sample-size Adjusted BIC[192]	$-2LL + \log\left(\frac{n+2}{24}\right) P$

- Voung-Lo-Mendell-Rubin adjusted likelihood ratio test (VLMRLRT): This test compares a model of T class with another model of $T-1$ class and reported p-value. If the p-value is insignificant, model of $T-1$ class is better model than model of T class.
- Bootstrap likelihood ratio test (BLRT): This test works as same as VLMRLRT.
- Average Latent Class Posterior Probability: This diagnostic statistic represents the average probability of classifying individuals correctly through the model in hand. Researchers suggested having the average probability for each class of the model greater than 0.90 but if other statistical criteria supported as model, the value between 0.80 to 0.90 is acceptable.
- Entropy R^2: Another diagnostic statistic indicates how the model correctly defines the classes, and the value of entropy should be close to 1.
- Class size: Researcher suggests that classes should not contain fewer than 50 cases or 5% of sample size but if the statistical criteria supported the model with small classes and conceptually makes sense, the model with small classes can be selected.

First, Information criteria were used to compare the models. BIC and sample-size adjusted BIC are commonly reported in studies and these incorporate penalties for complexity and number of parameters. The model with the lowest values for ICs is the better fitted model. Second, likelihood ratio test was utilized to compare two neighbouring models. The p-value indicates the better model. Average latent class posterior probability, entropy $\boldsymbol{R^2}$, and class size are diagnostic criteria, and the final class was not selected solely based on these. But these classification diagnostics were considered while selecting final class as addition to support the final model. Along with statistical criteria, theoretical knowledge and the interpretability of the classes was considered while selecting the final model. It is to avoid selecting a solution with too complex or too simple model and high statistical value which does not make sense practically.

3.3.2 Identifying Mental Health Clusters among Schools

To achieve our second objectives, we applied a non-parametric version of multilevel LCA on the five domain scores from both teachers and students. This approach takes into account the nested structure of the data, allowing us to classify not only students into latent classes but also schools into latent clusters based on mental health classes/patterns of students. Furthermore, we employed a proposed multilevel joint LCA to analyze the combined five domains score reported by teachers and students. This analysis provided insight into the mental health patterns among students as well as the mental health clusters among schools. Later, we compared the mental health clusters among schools obtained from this analysis with the clusters derived from separate multilevel LCAs based on the reports from teachers and students individually.

3.3.2.1 Multilevel Latent Class Analysis (MLCA)

Multilevel Latent Class Analysis (LCA) is the extended version of latent class analysis for continuous indicators which is modified to consider the nested structure of the data and explore the heterogeneity in mental health classes across school. Suppose there are n number of observations from J groups for K continuous indicators. $Y_{ij} = (Y_{ij1}, Y_{ij2}, \dots \dots, Y_{ijK})$ indicates the set of response of individual i of group j where Y_{ijk} denotes the response of individual i of group j for continuous indicator k with each group having n_j observations, $i = 1,2, \dots \dots \dots, n, 1 \leq j \leq J$ and $1 \leq k \leq K$.The non-parametric version of Multilevel Latent class analysis model for K continuous indicators can be written as

$$f(Y_{ij}) = \sum_{m=1}^{M} P(W = m)\left(\sum_{t=1}^{T} P(C = t \mid W = m) f(Y_{ij} \mid C = t, W = m)\right)$$

Where C and W represent the latent variables of level-1 class and level-2 cluster respectively, $P(W = m) = \frac{\exp(\delta_m)}{\sum_{m=1}^{M} \exp(\delta_m)}$, $P(C = t \mid W = m) = \frac{\exp(\gamma_{tm})}{\sum_{t=1}^{T} \exp(\gamma_{tm})}$ and $f(Y_{ij} \mid C = t, W = m) = (2\pi)^{-T/2} \, |\Sigma_{tm}|^{-1/2} \exp\{-\frac{1}{2}(Y_i - \mu_{tm})' \, \Sigma_{tm}^{-1} \, (Y_i - \mu_{tm})\}$. By replacing the $f(Y_{ij} \mid C = t, W = m)$ with $f(Y_{ij} \mid C = t)$, we get the restricted version of non-parametric multilevel latent class model which let the items-conditional probability density to be independent of group-level units where $f(Y_{ij} \mid C = t) = (2\pi)^{-T/2} \, |\Sigma_t|^{-1/2} \exp\{-\frac{1}{2}(Y_i - \mu_t)' \, \Sigma_t^{-1} \, (Y_i - \mu_t)\}$. The non-parametric multilevel model can use different forms of variance-covariance matrix and mean of continuous indicators as per the behavior of the data by means of specifications of $f(Y_{ij} \mid C = t)$. As mentioned in LCA, the $f(Y_{ij} \mid C = t)$ takes another form under local independence assumptions and the violation of this assumption is treated using the usual ways.

The parameters of the multilevel LCA for continuous indicators is obtained by maximizing the log-likelihood as follows,

$$\log L = \sum_{j=1}^{J} \sum_{i=1}^{n_j} \log f(Y_{ij})$$

The estimation procedure of multilevel LCA is similar to LCA and the upward-downward scheme [49] is used in expectation step of EM algorithm with forward recursion scheme [187] for NR algorithm to accommodate the nested structure [183]. The final number of patterns at student-level T and number of clusters at school-level M are selected by comparing the fit indices of the models with different values of M and T. We had two sets of mental health clusters among schools from the application of MLCA on indicators reported by teachers and self-reported indicators separately.

3.3.2.2 Multilevel Joint Latent Class Analysis (Multilevel Joint LCA)

Multilevel joint LCA is the extension of joint LCA to incorporate the nested structure of the data and reported mental health clusters among schools from the distribution of mental health patterns among students. In MLCPA model, there are three latent variables: latent class and latent pattern variable (student-level) and latent cluster variable (school-level). Consider ${Y_{ij}}^s = ({Y_{ij1}}^s, {Y_{ij2}}^s, \dots\dots, {Y_{ijK}}^s)$ indicates the set of response of one individual from source s where ${Y_{ijk}}^s$ denotes the response of individual i of group j for continuous indicator k from source s with each group having n_j observations, $i = 1,2, \dots\dots\dots, n$, $1 \le j \le J$ and $1 \le k \le K$. The multilevel joint LCA model (non-parametric version) is constructed with M clusters, L patterns and T classes from the responses for K indicators from S sources.

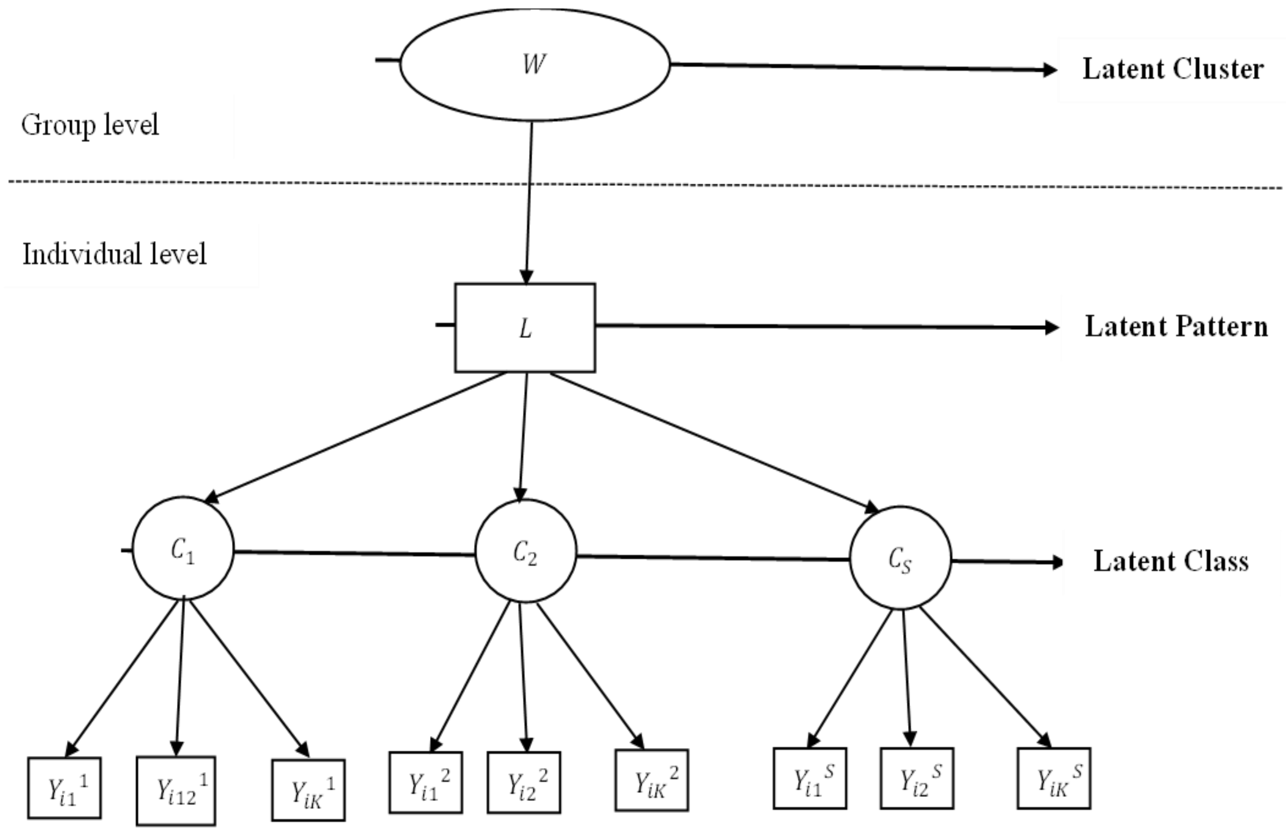

Figure 3.2: Graphical example of a multilevel joint LCA

The model for Y_{ij} can be written as,

$$P(Y_{ij}) = \sum_{m=1}^{M} P(W = m) \left(\sum_{l=1}^{L} P(G = l \mid W = m) \prod_{s=1}^{S} \left\{ \sum_{t=1}^{T} P(C_s = t \mid G = l)\, f({Y_{ij}}^s \mid C_s = t) \right\} \right)$$

Where, a) groups are classified into M clusters of latent cluster variable W, G represents latent pattern variable and C_s represents latent class variable,

b) $f({Y_{ij}}^s|C_s = t)$ represents the multivariate normal density function,

c) $P(C_s = t \mid G = l)$ represents the probability of belonging to class t for source s for a given pattern l,

d) $P((G = l\,|W = m)$ represents the probability of belonging to pattern l for a given cluster m,

e) $P(W = m)$ represents the probability of belonging to cluster m.

This model assumes that each latent class has its own set of mean μ_t and variance-covariance matrix Σ_t and these parameters of classes are equal for sources to maintain the same classes for sources. Also, this conditional density is not dependent on group and this model holds the following local independence assumptions:

a) The continuous indicators for source s are conditionally independent of each other given a latent class for source $s; s = 1, \dots \dots, S$.

b) The latent classes of all sources are conditionally independent of each other given a latent pattern.

c) The latent patterns of all individuals of group j are conditionally independent of each other given a latent cluster of group j.

This model takes form as follows:

$$P(Y_{ij}) = \sum_{m=1}^{M} P(W = m) \left(\sum_{l=1}^{L} P(G = l\,|W = m) \prod_{s=1}^{S}\left\{\sum_{t=1}^{T} P(C_s = t \mid G = l) \prod_{k=1}^{K} f({Y_{ijk}}^s|C_s = t)\right\}\right)$$

$$= \sum_{m=1}^{M} \delta_m \left(\sum_{l=1}^{L} \gamma_{l|m} \prod_{s=1}^{S}\left\{\sum_{t=1}^{T} \eta_{t|l}^{s} \prod_{k=1}^{K} f({Y_{ijk}}^s|C_s = t)\right\}\right)$$

with $f({Y_{ik}}^s|C = t) = \frac{1}{\sqrt{2\pi\sigma_{kt}^2}}\, exp\left\{-\frac{(Y_{ik}-\mu_{kt})^2}{\sigma_{kt}^2}\right\}$. The violation of local independence assumption is solved through previously mentioned ways. The log-likelihood function which is maximized to estimate the parameters of this model is:

$$\log L = \sum_{j=1}^{J} \sum_{i=1}^{n_j} \log P(Y_{ij})$$

ML estimates of the multilevel joint LCA are obtained using EM and NR algorithm where the expectation step of EM algorithm uses both forward-backward [186] and upward-downward schemes [49] and NR algorithm uses forward recursion scheme [187]. The expectation step is modified due to the structure of data where individuals are nested within groups and large number of individuals per sources [183]. The local solution situation is avoided through iterating the model with several random start values as mentioned in LCA. Comparing the fit indices of the model with different values of M, T and L, the final number of classes, patterns and clusters are selected. The application of multilevel joint LCA provided health cluster among schools by using the responses provided by teachers and students themselves for indicators of mental health.

3.3.2.3 Statistical Criteria for Multilevel Model selection

Information Criteria were used to select the final class as instructed in latent class analysis part. It is computationally intensive to run several combinations of latent classes and clusters or latent classes, patterns, and clusters simultaneously. We first selected the final number of latent classes through LCA and then choose optimal clusters through MLCA. Similarly, for proposed joint analysis, we ran joint LCA to select optimal number of latent classes and patterns and then extended the model to multilevel joint LCA to select final number of latent clusters for schools. The final number of latent classes and patterns were selected using the same formulae of Information Criteria as mentioned in latent class part and to select the optimal number of latent clusters for schools, the below modified formulas of ICs were used where the number of groups (schools) was used as sample size.

Table 3.3: Information Criteria for Multilevel model

Criterion	**Formulae (based on total number of groups as sample size)** [193]
Consistent Akaike Information Criteria (CAIC)	$-2LL + (1 + \log(J))P$
Bayesian Information Criteria (BIC)	$-2LL + \log(J)\, P$
Sample-size Adjusted BIC	$-2LL + \log\left(\frac{J+2}{24}\right) P$

3.3.3 Effects of student-level and school-level covariates

We investigated the effect of student-level covariates, such as sex, self-reported mental well-being, bullying and Canadian born status, on the mental health patterns among students derived from the multilevel joint LCA. Further, we explored the influence of school-level covariates, including bullying at school, on the mental health clusters among schools.

The extension of multilevel joint LCA to include the student and school-level covariates is to understand the effects of student-level covariates on the latent pattern memberships and school-level covariates on the latent cluster memberships. The extension is allowing individual-level covariates to influence only pattern memberships, not latent classes because covariates might affect the local independence of latent classes within a given pattern [194]. Assume we have two covariates, X_{1ij}: student-level covariate and X_{2j}: school-level covariate, then,

$$P(G = l \mid X_{1ij}, W = m) = \frac{\exp\left(\beta_{0l|m} + \beta_{1l|m} X_{1ij}\right)}{\sum_{l=1}^{L} \exp\left(\beta_{0l|m} + \beta_{1l|m} X_{1ij}\right)}$$

Where $\beta's$ measures the change in log-odds of belonging to pattern l compared to the reference pattern L for one-unit change in covariate X_{1ij} and $\beta's$ can be assumed to be same for all clusters (e.g., $\beta_{1|m} = \beta_{1|1} = \cdots = \beta_{1|M}$).

And

$$P(W = m \mid X_{2j}) = \frac{\exp\left(\lambda_{0l} + \lambda_{1m} X_{2j}\right)}{\sum_{l=1}^{L} \exp\left(\lambda_{0l} + \lambda_{1m} X_{2j}\right)}$$

Where, $\lambda's$ denote the change in the log-odds of belonging to cluster m compared to reference cluster M for one-unit change in covariate X_{2j}.

The extended multilevel joint LCA model can be written as,

$$P\left(Y_{ij}|X_{1ij}, X_{2j}\right)$$

$$= \sum_{m=1}^{M} P\left(W = m|X_{2j}\right)\left(\sum_{l=1}^{L} P(G = l\,|X_{1ij}, W\right.$$

$$\left.= m)\prod_{s=1}^{S}\left\{\sum_{t=1}^{T} P(C_s = t \mid G = l)\prod_{k=1}^{K} f({Y_{ijk}}^{s}|C_s = t)\right\}\right)$$

4. Results

4.1 Sample Characteristics

Table 4.1 presents the final sample of 8167 students from 284 schools across 31 divisions that were used for the analysis. The sample consisted of 4054 (49.60%) girls and 4113 (50.40%) boys, with 82.80% of students being Canadian born. Among these students, 3101 (38%) rated their mental health as excellent, while 182 (2.20%) rated as poor. According to teacher's report, 77.7% of 8167 students did not engage in fighting or bullying their peers. The average percentage of students reported by teachers to be involved in bullying behaviour at school was 22.33 (±12.14).

Table 4.1: Characteristics of Students (n = 8167) from Manitoba Grade-5 Mental Health Survey

Variables	N (%)
Sex	
Female	4054 (49.60)
Male	4113 (50.40)
Teacher-reported bullying incidents	
No	6343 (77.70)
Yes	1824 (22.30)
Self-rated Mental Health	
Poor	182 (2.20)
Fair	535 (6.60)
Good	1712 (21.00)
Very Good	2637 (32.30)
Excellent	3101 (38.00)
Canadian born	
Yes	6762 (82.80)
No	1405 (17.20)

Table 4.2 shows the means and standard deviations for five SDQ domains based on reports from teachers and students. On average, students reported experiencing higher level of difficulties compared to what their teachers reported. Alternatively, teachers reported higher levels of prosocial difficulties compared to what students reported by themselves.

Table 4.2: Outcome of five domains from two perspectives, mean (±std. dev)

Domain Scores	Self report	Teacher's report

Emotional difficulties	**3.14 (±2.37)**	1.92 (±2.30)
Peer difficulties	**2.38 (±1.94)**	1.45 (±1.93)
Conduct difficulties	**1.72** (±1.74)	1.17 (±**1.93**)
Hyperactivity difficulties	**3.53** (±2.28)	3.08 (±**3.14**)
Prosocial difficulties*	2.10 (±1.72)	**2.27 (±2.48)**

*Prosocial difficulty is the opposite of prosocial behaviour. The higher the score, the greater the challenges or problems in this area.

The cut-offs suggested by Goodman were used to classify the students into three categories: "normal", "borderline" and "abnormal"; this categorization was created based on data from a community sample in England and various epidemiological studies [177,178]. Different cut-off values were recommended depending on whether it was self-reported by students or reported by teachers. These cut-offs were slightly adjusted to better align with the continuous SDQ sub-scale scores, which are shown in Appendix. Table 4.3 presents the percentage of students falling into the "abnormal" range on each SDQ sub-scale, as reported by teachers and students separately.

Table 4.3: Percentage of students with scores in the abnormal range on SDQ sub-scales reported by teachers and students

Domains	Self report	Teacher's report
Emotional difficulties	**9.8%**	9.1%
Peer difficulties	7.3%	**9.4%**
Conduct difficulties	7.7%	**12.9%**
Hyperactivity difficulties	10.3%	**17.0%**
Pro-social difficulties	3.7%	**11.7%**

Teacher reported higher percentage of students in the abnormal category compared to students' report across various domains. However, students reported a higher percentage of themselves experiencing abnormal emotional difficulties than what their teachers reported. The disparity between the average difficulties and the percentage of student in abnormal category stemmed from the distribution of scores reported by teachers and students. The low average difficulties score in teachers' report reflected a mix of many students with low scores and a notable proportion with high scores. In contrast, the high average difficulties score in students' report resulted from fewer students with low scores compared to teachers' report and a fair of amount of students with high scores.

4.2 Identifying Mental Health Classes/Patterns among Students

4.2.1 Latent Class Analysis

4.2.1.1 Informant: Student

LCA models, utilizing the scores of five SDQ domains reported by students as indicators, were used to identify the patterns of mental health among students. LCA was conducted for one to nine classes models and the final model was selected by comparing the results across different models. Table 4.4 presents the goodness of fit statistics of the LCA models using self-reported scores for strengths and difficulties domains.

Table 4.4: Goodness of fit statistics of LCA models [Informant: Student]

Model	LL	AIC	BIC	SABIC	VLMRLRT p-value
1 Class	-54083.1894	108206.3787	108346.5359	108282.9797	-
2 Class	-46099.7871	92281.5743	92568.8964	92438.6062	0.0000
3 Class	-40294.6937	80713.3874	81147.8746	80950.8504	0.0046
4 Class	-35655.8552	71477.7104	72059.3625	71795.6044	0.0000
5 Class	-32522.7962	65253.5923	65982.4094	65651.9173	0.0000
6 Class	**-30477.3976**	**61204.7952**	**62080.7773**	**61683.5512**	**0.0000**
7 Class	-29667.6408	59627.2816	60650.4288	60186.4686	0.0000
8 Class	-27651.2383	55636.4766	56806.7887	56276.0946	0.0000
9 Class	-27129.8664	54635.7327	55953.2098	55355.7818	0.0000

The results of information criteria showed that the BIC, AIC, and SABIC values continued to decrease as the number of classes increased. But after six-class models, the further increase in number of classes did not show equivalent reduction in these values. Notably, the more parsimonious model with five-class provided classes that closely resembled those of six-class model, with the latter revealing two unique classes not identified in the former. The classes in the six-class and seven-class models were largely similar and additional class in seven-class model did not provide extra information compared to six-class model. Thus, we selected the six-class model as the final model based on fit indices, parsimony and interpretability of classes. The entropy R^2 values for six-class model was 0.986, and the lowest value of average latent class posterior probability was 0.979, both falling within an acceptable range. The profile plot for six-class model is shown in Figure 4.1, where the lines represent the classes, x-axis represents five SDQ indicators, and y-axis shows the mean score of the indicators.

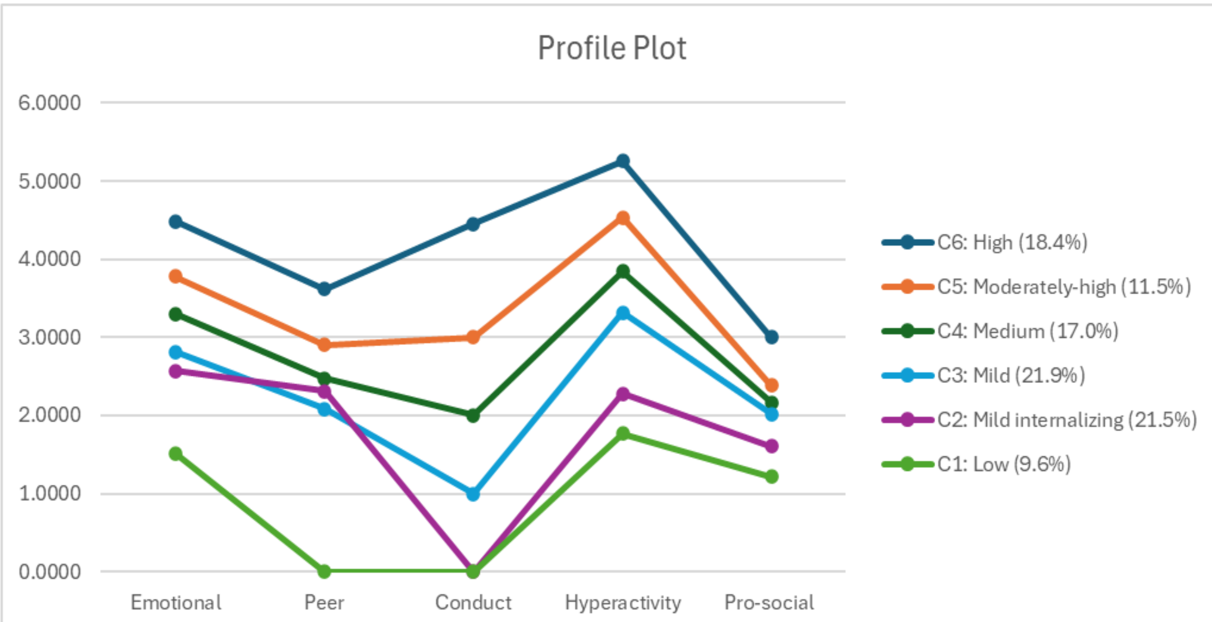

Figure 4.1: Profile plot of Six-class LCA model [Informant: Student]

In the profile plot of six-class model [Figure 4.1], Class 1, comprising 9.6% of students, represented individuals with an overall low risk, characterized by the absence of peer-relationships and conduct difficulties. Class 2 had mild internalizing risk without conduct difficulties, while Class 3 consisted of students with a mild risk profile, making it the largest class at 21.9%, followed closely by Class 2 at 21.5%. Class 4, accounting for 17.0% of students, displayed an overall medium risk level, whereas students in Class 5 (11.5%) showed overall a moderately high risk. The highest overall high risk was observed in Class 6, encompassing 18.4% of students, who exhibited elevated externalizing difficulties compared to other classes.

Prevalence of Abnormal Scores Across Mental Health Classes: The percentage of students with scores in the abnormal range according to the British three-fold classification in each domain by mental health classes from self-report is presented in Table 4.5. The students classified in the high-risk class consistently exhibited higher rates of abnormal scorers while those in the low-risk class consistently displayed lower rates across all domains compared to other classes. Among students in the mild-internalizing risk class, 4.7% had abnormal scores for emotional difficulties and 3.6% for peer-difficulties, with relatively lower percentages in other domains. As the level of mental health classes advanced from low to high based on the mean scores of sub-scales, there was a corresponding increase in the percentage of students with score in abnormal range across all domains.

Table 4.5: Percentage of Students with Scores in the Abnormal Range on SDQ Sub-scales by mental health classes [Informant: Student]

Classes	Emotional	Peer	Conduct	Hyperactivity	Pro-social
Low	0.9%	0.0%	0.0%	0.8%	0.6%
Mild internalizing	4.7%	3.6%	0.0%	1.8%	1.1%
Mild	6.7%	3.8%	0.0%	4.5%	2.2%
Medium	9.2%	5.7%	0.0%	9.9%	3.2%
Moderately high	13.5%	10.3%	0.0%	16.8%	3.9%
High	22.6%	19.4%	42.0%	28.6%	10.5%

4.2.1.2 Informant: Teacher

Similar analysis procedures were employed to teacher-reported mental health scores as we did for student-reported scores. Table 4.6 show the statistical fit measures for the LCA models utilizing five sub-scale scores reported by teachers.

Table 4.6: Fit statistics of LCA models [Informant: Teacher]

Model	LL	AIC	BIC	SABIC	VLMRLRT p-value
1 Class	-51125.9576	102291.9153	102432.0724	102368.5162	-
2 Class	-26893.6164	53869.2329	54156.555	54026.2648	0.000
3 Class	-17133.1291	34390.2583	34824.7454	34627.7213	0.000
4 Class	-12064.0434	24294.0868	24875.7389	24611.9807	0.000
5 Class	-9245.1207	18698.2413	19427.0585	19096.5663	0.000
6 Class	**-4268.2669**	**8786.5338**	**9662.5159**	**9265.2898**	**0.000**
7 Class	-2164.262	4620.524	5643.6711	5179.711	0.000

8 Class	-397.7105	1129.421	2299.7331	1769.039	0.000
9 Class	1180.29	-1984.5799	-667.1028	-1264.5309	0.000

The AIC values decreased with the addition of classes in the model, showing a consistent pattern with the BIC and SABIC values. We focused on the BIC values, which depicted a diminishing rate of decrease after the 6-class model. After comparing the classes of six-class model and other models, we found that there were variations among the classes from six-class solution and those of five-class solution. Seven and eight-class models depicted classes which were largely like six-class model, only distributed closely related students in different classes. Finally, we selected six-class model as the final model considering fit, interpretability and parsimony and profile plot of this model is depicted in Figure 4.2. The smallest class within six-class model consisted of 8% of the students. The average latent class posterior probabilities and the entropy were greater than 0.90 for the six-class solution.

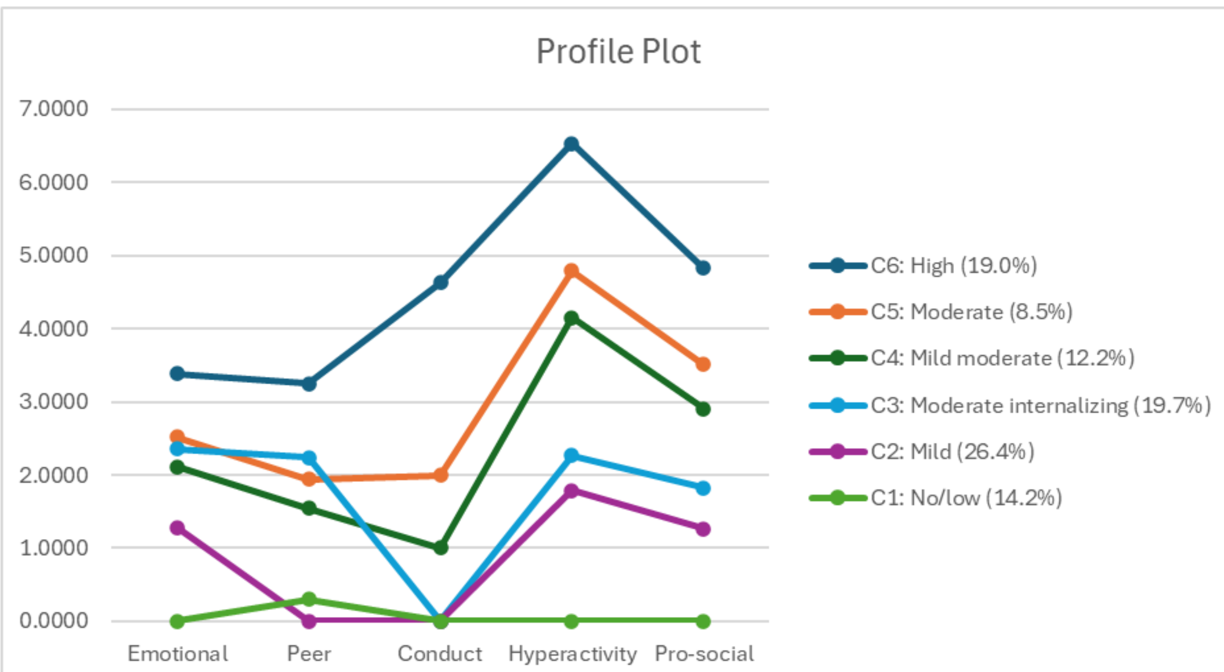

Figure 4.2: Profile plot of Six-class LCA model [Informant: Teacher]

In the profile plot of the six-class model [Figure 4.2], Class 1, comprising 14.2% of students, represented individuals with no or low risk, and Class 2, the largest class at 26.2%, exhibited mild risk with no peer and conduct difficulties and mild other difficulties. Class 3, consisting 19.7% of students, showed no conduct difficulties but moderate internalizing difficulties. Class 4

(12.2%) included students with mild to moderate risk, while Class 5 (8.4%) comprised students with moderate risk levels. Class 6 was the overall high-risk class, with 19.2% students exhibiting more externalizing and pro-social difficulties compared to other classes.

Prevalence of Abnormal Scores Across Mental Health Classes: Table 4.7 illustrates the percentage of students exhibiting scores in abnormal range based on British cut-offs in each domain for each mental health classes from teachers' report. Among students classified with high-risk class, the percentages of students with scores in abnormal range were higher compared to students in other classes. In the moderate internalizing risk class, 10.3% of students had abnormal scores for emotional difficulties, a rate similar to the 10.4% observed in the mild moderate risk class. However, 9.4% of students had scores in abnormal range for peer difficulties within moderate internalizing risk class, higher than that of mild moderate risk class. Notably, among students belonging to no/low risk classes, no student had abnormal scores in any domain. As the classes progressed from low to high based on the mean values of domain scores, the prevalence of students with scores in abnormal range increased across all domains.

Table 4.7: Percentage of Students with Scores in the Abnormal Range on SDQ Sub-scales by mental health classes [Informant: Teacher]

Classes	**Emotional**	**Peer**	**Conduct**	**Hyperactivity**	**Pro-social**
No/low	0.0%	0.0%	0.0%	0.0%	0.0%
Mild	2.3%	0.0%	0.0%	4.2%	1.8%
Moderate internalizing	10.3%	9.4%	0.0%	6.2%	4.5%
Mild moderate	10.4%	8.4%	0.0%	20.1%	13.0%
Moderate	13.2%	11.5%	0.0%	28.1%	18.3%
High	21.3%	29.3%	67.7%	51.9%	38.0%

4.2.2 Joint Latent Class Analysis (Joint LCA)

We fitted joint LCA models where the parameters for classes were assumed to be equal across two informants to maintain measurement invariance of classes. The joint LCA models provided mental health classes for each informant and mental health patterns representing combinations of mental health classes from both informants. We started with fitting a series of joint LCA models with various combination of latent classes and latent patterns. Table 4.8 reports the fit statistics of the joint LCA models that converged.

Table 4.8: Fit statistics for a series of joint latent class model with different number of classes and patterns

Number of Classes	Number of Patterns	LL	BIC	AIC	SABIC
2	1	-140967	282312	282017.7	282178.5
	2	-140637	281678.6	281363.2	281535.6
3	1	-131527	263631.3	263182.8	263427.9
	2	-131084	262790.3	262306.8	262571.1
4	1	-124833	250441.5	249838.8	250168.2
	2	-124160	249157.4	248505.6	248861.8
5	1	-111003	222979.4	222222.6	222636.2
	2	-110457	221968	221148.1	221596.2
6	1	-108648	218467.1	217556.1	218054
	2	-108063	217395.8	216407.7	216947.7
	3	**-107589**	**216547.5**	**215482.3**	**216064.5**
7	1	-103978	209324.4	208259.2	208841.3
	2	-104223	209931.8	208775.5	209407.5

The joint LCA model with seven mental health classes and one mental health pattern had lowest values across all information criteria [Table 4.8], followed by the model with six-class and three-pattern. The classes from six-class and three-pattern model were considerably similar to the classes from seven-class and one-pattern model. Students which could be identified in the same class were distributed in the neighbouring classes in the model with seven-class. Considering interpretability of classes and patterns, we selected the six-class and three-pattern joint LCA model as the final model (BIC=216547.5; AIC=215482.3).

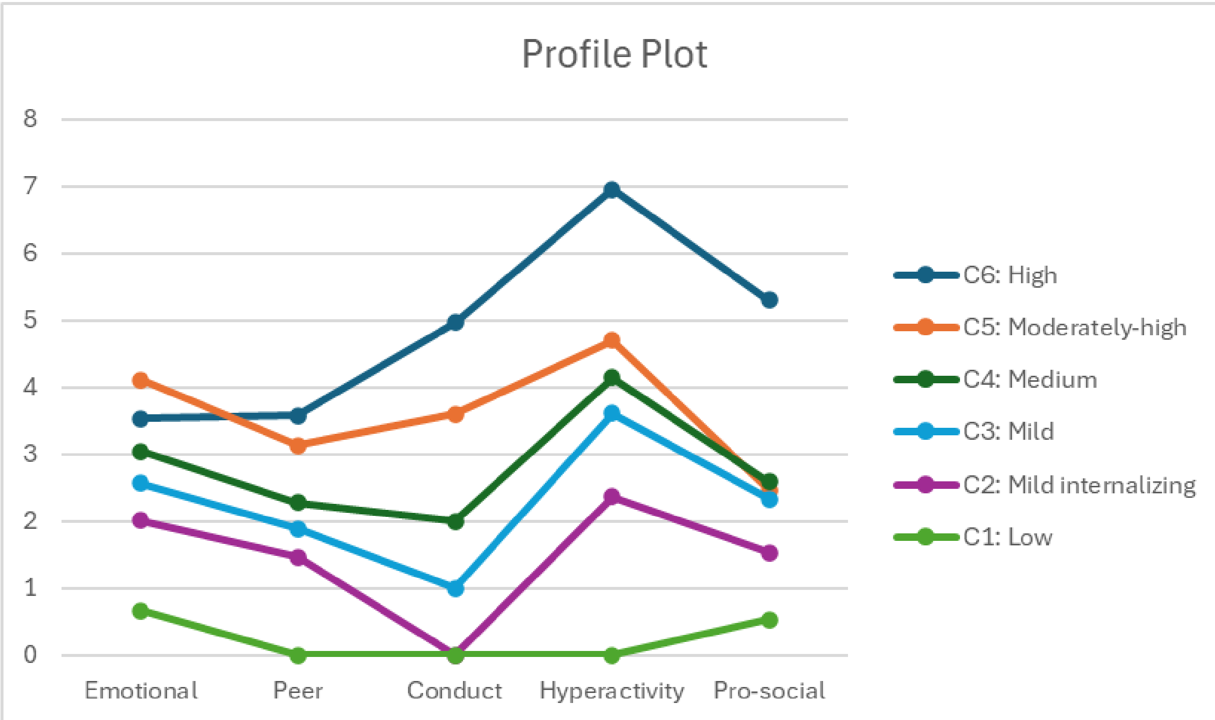

Figure 4.3: Profile plot of six-class/three pattern joint LCA model

The profile plot of six-class and three-pattern model can be seen in Figure 4.3. Low risk class had low mean over all domains and high-risk class had high mean value across all domains. Table 4.9 presents the estimated mean parameters of the domains for each class under the final model.

Table 4.9: Estimated means of the domains for each class under six-class and three-pattern joint LCA model

Classes	Emotional	Peer	Conduct	Hyperactivity	Pro-social
Low	0.664	0.000	0.000	0.000	0.531
Mild internalizing	2.008	1.469	0.000	2.368	1.532
Mild	2.565	1.887	1.000	3.608	2.329
Medium	3.040	2.281	2.00	4.141	2.587
Moderately high	4.106	3.135	3.600	4.694	2.464
High	3.530	3.579	4.970	6.949	5.302

While we maintained equal class parameters across informants, the prevalence percentages of students in each class were allowed to vary across informants. This means that although the rules for categorizing students remained the same, the distribution of students across classes differed depending on the informants. Table 4.10 presents the percentages of students belonging to a specific class based on the self-report and teacher's report separately. The prevalence of mild internalizing risk was higher compared to other classes for both informants. Teachers reported 21.9% of students belonging to low risk, whereas students reported only 3.0% to be in this category. On the other hand, 26.7% of students were categorized as moderately high risk based on students' self-reports, while only 4.0% of students were in the moderately high-risk class based on teacher's report. 15.2% of students were placed in the high-risk class by teachers, whereas 3.2% of students belonged to that class according to self-report.

Table 4.10: Class prevalence percentages of students by informants under six-class and three-pattern joint LCA model

Classes	Self-report	Teacher's report
Low	3.0%	21.9%
Mild internalizing	**28.1%**	**38.2%**
Mild	21.9%	12.2%
Medium	17.0%	8.5%
Moderately high	26.7%	4.0%
High	3.2%	15.2%

Table 4.11 shows the estimated conditional probabilities of class membership for each informant for a mental health pattern based on the reports by both teachers and students. For students in the high-risk pattern, there was a higher likelihood of being classified in the high-risk or moderately high class by both informants, with probability of 0.6351 being in the high-risk class in the teacher's report and the probability of 0.4633 being in the moderately high-risk class in student's report. In the low-risk pattern, 47.5% of students in the teacher's report were classified as the mild internalizing risk class and 41.8% as low-risk class, while 48.7% of students in the self-report were placed in the mild internalizing risk class. In the self-reported risk pattern, 31.7% of students from self-report belonged to the moderately high class and 21.4% to the medium class, while in the teacher's report, 44.6% of students belonged to the mild internalizing class and 8% to low-risk class. The most common pattern was self-reported risk (42.4%), followed by low risk (38.3%) and high risk (19.3%).

Table 4.11: Estimated probabilities of belonging to a particular class based on a certain report for each pattern and pattern prevalence percentages of students under six-class and three-pattern joint LCA model

Patterns	Classes	Self-report	Teacher's report
High risk (19.3%)	Low	0.0037	0.0161
	Mild internalizing	0.0483	0.0684
	Mild	0.1001	0.0802
	Medium	0.1857	0.1198
	Moderately high	**0.4633**	0.0805
	High	0.199	**0.6351**
Self-reported risk (42.4%)	Low	0.006	**0.0822**
	Mild internalizing	0.1551	**0.4462**
	Mild	0.2993	0.2455
	Medium	**0.2139**	0.1163
	Moderately high	**0.3177**	0.0636
	High	0.008	0.0462
Low risk (38.3%)	Low	0.0607	**0.4187**
	Mild internalizing	**0.4868**	**0.4752**
	Mild	0.21	0.0433
	Medium	0.1235	0.0402
	Moderately high	0.1162	0.0105
	High	0.0029	0.0121

Table 4.12: Prevalence of student's self-rated mental health by three latent patterns under the six-class and three-pattern joint LCA model

Self-rated Mental Health	High-risk pattern	Self-reported risk pattern	Low-risk pattern
Poor	88 (5.6%)	71 (2.1%)	23 (0.7%)
Fair	193 (12.2%)	271 (7.8%)	71 (2.3%)
Good	432 (27.4%)	814 (23.5%)	466 (14.9%)
Very good	426 (27.0%)	1184 (34.2%)	1027 (32.8%)
Excellent	439 (27.8%)	1120 (32.4%)	1542 (49.3%)

Table 4.12 presents the prevalence of students reporting their mental health on a scale from poor to excellent, based on latent patterns from the joint model. We found that percentage of students who rated their mental health as poor, fair and good were higher in high-risk and self-reported risk pattern compared to the low-risk pattern. About half of students from low-risk pattern rated their mental health as excellent, which was higher than those in high-risk and self-reported risk

patterns. Among students in the low-risk and self-reported risk patterns, around 32% and 34% of them rated their mental health as very good, while about 28% of students from high-risk pattern rated their mental health as very good.

Comparison of patterns from joint LCA with classes from LCA: When comparing the patterns identified by the joint LCA model with the classes derived from a simple LCA based on self-report in Figure 4.4, we found that about half of the students classified in the high-risk pattern by the joint LCA were also categorized in the high-risk class, with about 20% falling into the moderately high-risk class according to the LCA model based on self-report only. For students in the low-risk pattern identified by the joint model, more than 70 % of them were assigned to the low or mild internalizing risk class according to the LCA model based on student's report only. Additionally, more than half of the students classified in the self-reported risk pattern by the joint LCA were classified into higher levels of risk classes (high, moderate and medium risk classes) according to the LCA model based on the self-report only.

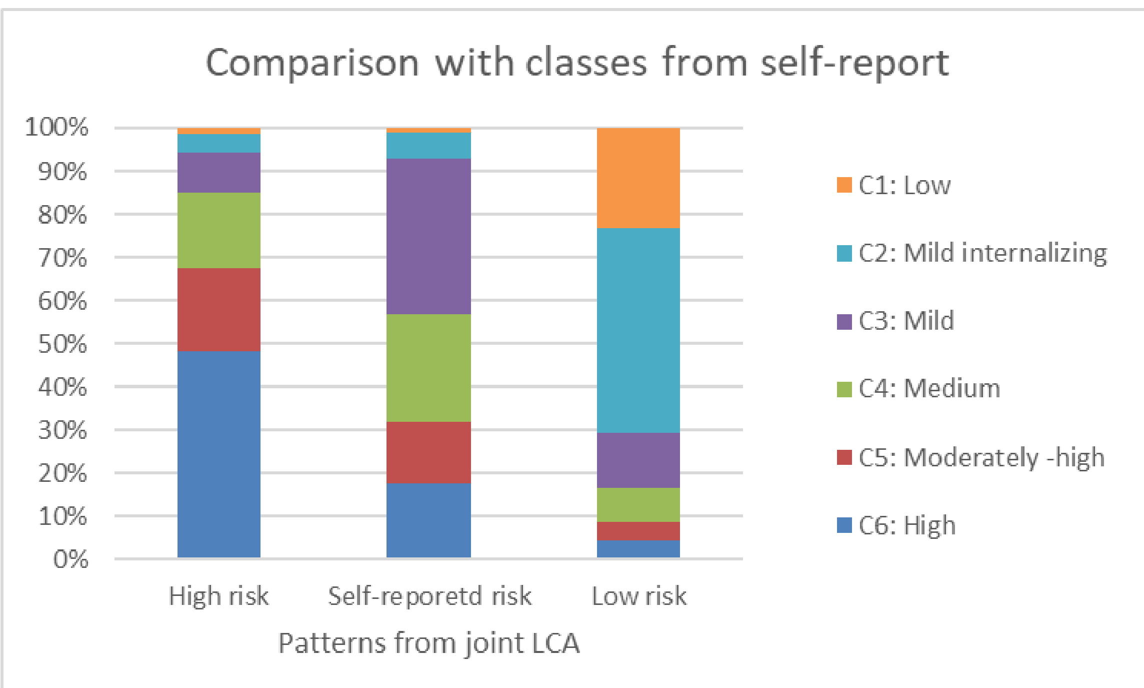

Figure 4.4: Comparison of patterns from joint LCA model with classes from simple LCA model for self-report

When comparing the latent patterns derived from the joint LCA model with the latent classes identified by a simple LCA model based on teacher's report only in Figure 4.5, we found that about 90% of students classified in high-risk pattern by joint LCA were classified in the high-risk class according to LCA model based on the teacher's report only. More than 70 % of students classified in the low-risk pattern by joint LCA model were classified into low or mild risk classes according to LCA model based on the teacher's report only. On the other hand, for students in the self-reported risk pattern identified by joint LCA model, about 30% of them were classified into lower level of risk classes (low/no and mild risk classes) and 28.5% into moderate internalizing risk class according to the simple LCA model based on the teacher's report only.

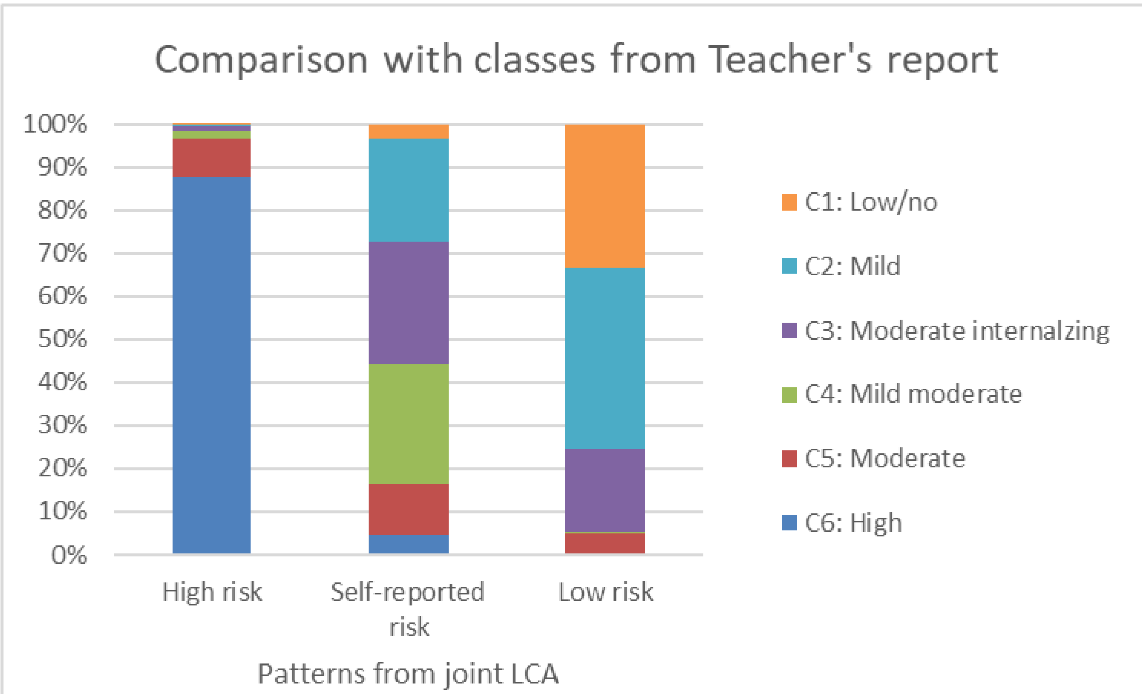

Figure 4.5: Comparison of patterns from joint LCA model with classes from simple LCA model for teacher's report

The joint LCA model effectively served as a bridge between the latent classes identified by self-reports and teacher's report separately. The latent patterns generated by the joint LCA model were the composite of the latent classes identified based on both teachers' and students' reports.

4.3 Identifying Mental Health Clusters among Schools

4.3.1 Multilevel Latent Class Analysis (MLCA)

4.3.1.1 Informant: Student

The conventional latent class analysis of students' self-reports identified six latent classes representing different mental health profiles. To account for the hierarchical structures of the data, where students are nested within schools, we expanded the six-class solution. By comparing one to seven clusters for schools, we aimed to identify school clusters based on the distribution of six student-level classes. The fit statistics of these multilevel LCA models are presented in Table 4.13.

Tabel 4.13: Fit Statistics of multilevel LCA models considering six-class among students [Informant: Student]

Model	LL	BIC	CAIC	SABIC
1-Cluster 6-Class	-14295.3456	28618.9361	28623.9361	28603.0809
2-Cluster 6-Class	-14231.1599	**28524.4586**	**28535.4586**	28489.5772
3-Cluster 6-Class	-14214.654	28525.3405	28542.3405	**28471.4329**
4-Cluster 6-Class	-14209.0635	28548.0534	28571.0534	28475.1196
5-Cluster 6-Class	-14202.8355	28569.4912	28598.4912	28477.5311
6-Cluster 6-Class	-14198.0801	28593.8744	28628.8744	28482.8881
7-Cluster 6-Class	-14193.6766	28618.9612	28659.9612	28488.9487

When comparing the information criteria values across models with one to seven school clusters, both BIC and CAIC values favored the two-cluster model, but SABIC favored three-Cluster models. Upon further exploration of two-cluster and three-cluster models, we ultimately chose the three-Cluster model as our final model.

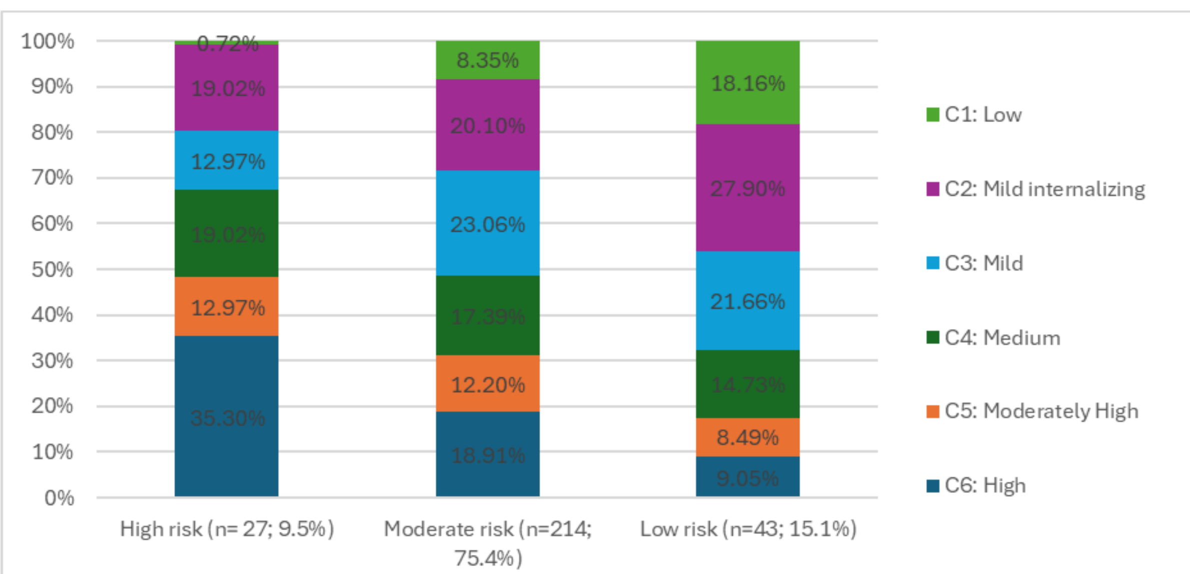

Figure 4.6: Distribution of classes within Clusters for Six-class and Three-cluster MLCA model [Informant: Student]

Figure 4.6 shows the distribution of student-level classes within school-level clusters. Figure 4.6 revealed that among 284 schools, Cluster 1 (High-risk schools) comprised 27 schools, Cluster 2 (Moderate-risk schools) included 214 schools, and Cluster 3 (Low-risk schools) encompassed 43 schools.

High-risk schools exhibited the highest proportion of students in Class 6 (High risk) and Class 4 (Medium risk), with a notably low representation of students in Class 1 (Low risk). Conversely, low-risk schools had the highest proportions of students in Class 2 (Mild internalizing risk) and Class 1 (Low risk). Moderate risk schools predominantly consisted of students in Class 3 (Mild risk). In addition, there was a slight variation in the proportion of students in Class 5 (Moderately high-risk) for the Moderate risk cluster (12.20%) and the High-risk cluster (12.97%). The proportions of students in Class 4 (Medium risk) were 19.02% and 17.39% in High-risk and Moderate-risk clusters, respectively.

4.3.1.2 Informant: Teacher

The conventional latent class analysis of teacher's report identified six latent classes representing different mental health profiles. To accommodate the hierarchical structure of the data, where students are nested within schools, we extended the six class LCA model. By comparing one to nine clusters for schools, we aimed to identify school clusters based on the distribution of six student-level classes. Table 4.14 represents the fit statistics for these multilevel LCA models.

Table 4.14: Fit Statistics of multilevel LCA models considering six-class among students [Informant: Teacher]

Model	LL	BIC	CAIC	SABIC
1-Cluster 6-Class	-14136.369	28300.9829	28305.9829	28285.1277
2-Cluster 6-Class	-13966.0234	27994.1855	28005.1855	27959.3041
3-Cluster 6-Class	-13898.1952	**27892.423**	**27909.423**	**27838.5154**
4-Cluster 6-Class	-13877.2554	27884.4372	27907.4372	27811.5034
5-Cluster 6-Class	-13857.4574	27878.7351	27907.7351	27786.775
6-Cluster 6-Class	-13842.8112	27883.3364	27918.3364	27772.3502

7-Cluster 6-Class	-13832.5768	27896.7615	27937.7615	27766.749
8-Cluster 6-Class	-13824.5911	27914.684	27961.684	27765.6453
9-Cluster 6-Class	-13817.7916	27934.9788	27987.9788	27766.9139

Among models with one to nine school clusters, BIC and CAIC value favoured the five clusters and four clusters model respectively. SABIC value continued to decrease until the eight-cluster model, after which it slightly increased for the nine-cluster model. After examining the plot of these values in Figure 4.7, we discovered that the elbow point of the plot occurred at the three-cluster model. Additionally, increasing the number of clusters beyond three did not result in a noticeable decrease in these values. Considering fit and parsimony, we decided to select three-cluster and six-class model as the final model.

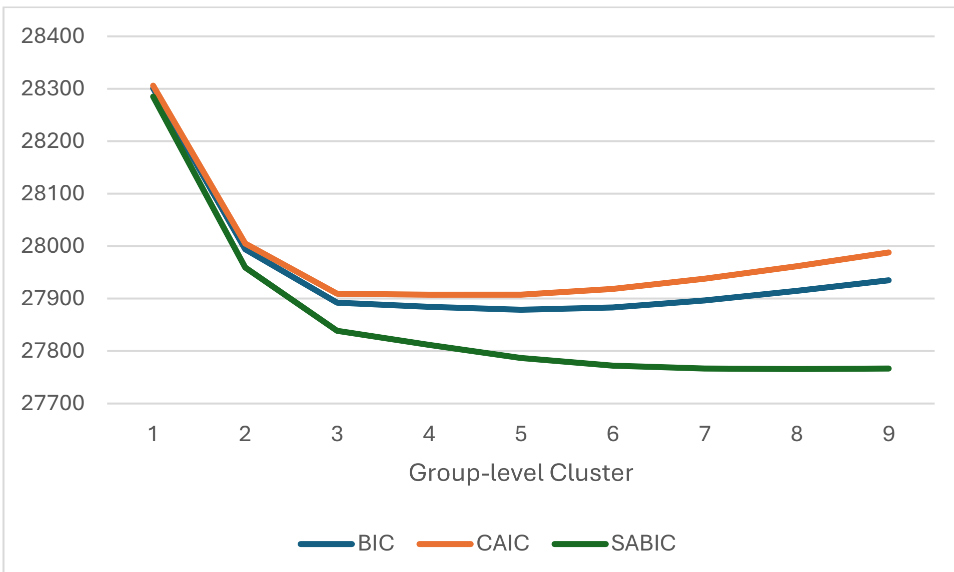

Figure 4.7: Plot of Information Criteria of multilevel LCA [Informant: Teacher]

Figure 4.8 shows the distribution of school-level classes in school-level clusters for the final model. Figure 4.8 revealed that Cluster 1 (Moderate-high risk schools) included 117 schools, Cluster 2 (Mild-risk schools) comprised 105 schools and Cluster 3 (Low-risk schools) encompassed 62 schools.

Low risk schools exhibited the highest proportions of students in Class 1 (No/low risk) and Class 2 (Mild risk). Mild risk schools had highest proportion of students in Class 2 (Mild risk) and Class 3 (Moderate internalizing risk). Moreover, the Mild risk schools and Low risk schools displayed nearly identical proportions of students across other classes. Moderate-High risk schools predominantly consisted of students in Class 6 (High risk) and Class 3 (Moderate internalizing risk).

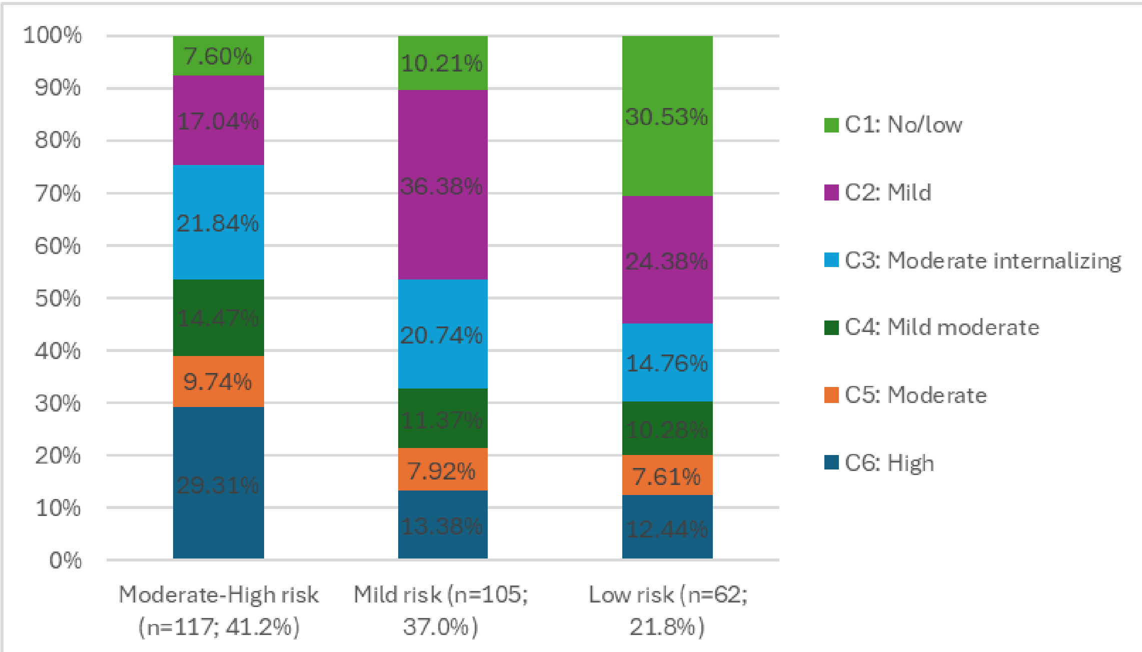

Figure 4.8: Distribution of classes within Clusters for Six-class and Three-cluster MLCA model [Informant: Teacher]

4.3.2. Multilevel Joint Latent Class Analysis (Multilevel Joint LCA)

Using joint LCA, we identified six classes and three patterns that reflect various mental health patterns derived from both students' self-report and teachers' report. To address the hierarchical nature of the data, with students nested within schools, we extended the six-class and three-pattern solution. Our goal was to determine school clusters by comparing one to six clusters based on the distribution of mental health patterns among students.

Table 4.15 presents the fit statistics for the multilevel joint LCA models. BIC value was lowest for the four-cluster model, CAIC values was lowest for the four-cluster model and SABIC supported the five-cluster model. The clusters from four-cluster, five-cluster and three-cluster models were compared with each other. After considering the parsimony and interpretability of the model, we selected the three-cluster model for the school-level as the final model (BIC=16897.9303; SABIC=16872.562).

Table 4.15: Fit statistics for a series of multilevel joint LCA models with six-class and three pattern

Clusters	LL	BIC	CAIC	SABIC
1	-8567.6307	17146.5594	17148.5594	17140.2173
2	-8457.13	16942.5049	16947.5049	16926.6498
3	**-8426.3692**	**16897.9303**	**16905.9303**	**16872.562**
4	-8416.1324	16894.4036	16905.4036	16859.5222
5	-8408.7242	16896.534	16910.534	16852.1395
6	-8405.8933	16907.8192	16924.8192	16853.9116

Table 4.16 provides the conditional probability of mental health patterns given a particular cluster of schools and the classification of schools based on the distribution of these patterns. The high-risk cluster consisted of schools where about 54.8% of students exhibited high-risk patterns, while the low-risk cluster included schools with a higher probability of students displaying low-risk pattern. The third cluster consisted of schools where the students were more likely to exhibit a self-reported risk pattern, indicating that students fell into the medium to high-risk class based on self-reports and the mild risk class based on teachers' reports. This student-reported risk cluster encompassed 62.6% of schools.

Table 4.16: Estimated conditional probabilities of belonging to a specific pattern given a particular cluster of schools under the final multilevel joint LCA model

Patterns	Clusters		
	Student-reported risk (n=178; 62.6%)	**Low risk (n=76; 26.8%)**	**High risk (n=30; 10.6%)**
High risk	0.2078	0.0938	**0.5484**
Self-reported risk	**0.4273**	0.2413	0.2669
Low risk	0.3649	**0.6649**	0.1846

Comparison of clusters from joint model with clusters from separate models: Figures 4.9 and 4.10 compare clusters obtained from multilevel joint LCA with those from separate multilevel LCAs. The multilevel LCA based solely on students' self-report identified three school clusters: high risk, low risk, and moderate risk clusters. Conversely, the multilevel LCA based only on teachers' report categorized schools into three clusters: low risk, mild risk and moderate high-risk clusters. In contrast, the joint multilevel LCA models integrated both self-report and teachers' report revealing three clusters: high risk, low risk and student-reported risk clusters. Schools classified under the high-risk cluster in the joint model corresponded to the moderate high-risk cluster in the teacher-report-only model and were placed in either the high or moderate risk cluster based solely on self-reports.

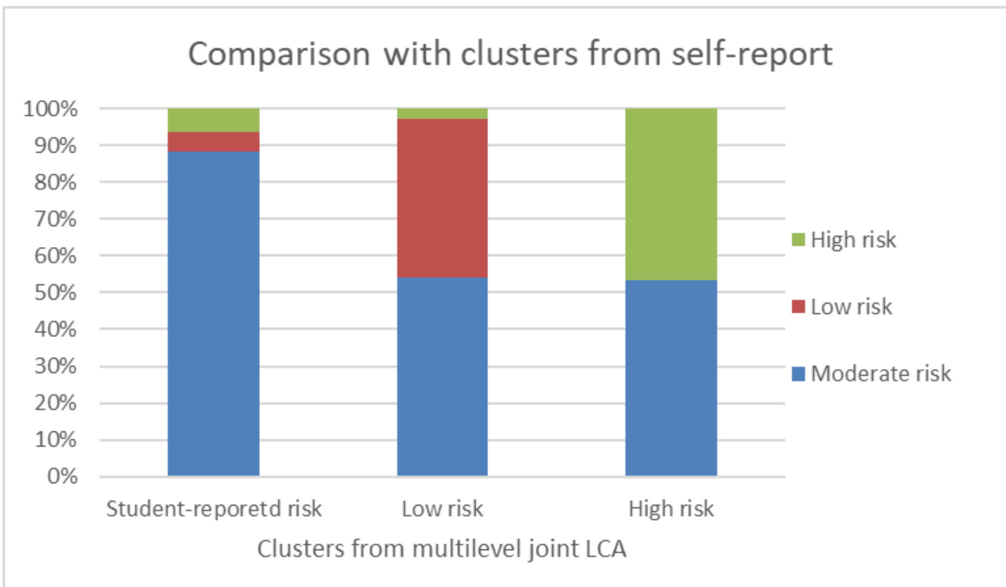

Figure 4.9: Comparison of clusters from multilevel joint LCA model with simple LCA model for self-report

The student-reported risk cluster in the multilevel joint LCA model mainly comprised of schools that were mostly placed in moderate risk cluster based solely on self-report. Conversely, 50% of schools in the student-reported risk cluster from joint model were categorized into either the low or mild risk clusters based on only teachers' report. Additionally, the majority of schools in the low-risk cluster of the joint model were classified into either low or mild risk cluster in the separate model based on teacher's report. And most of the schools in low-risk cluster of joint model were clustered into either low or moderate risk cluster in the separate model based on self-report.

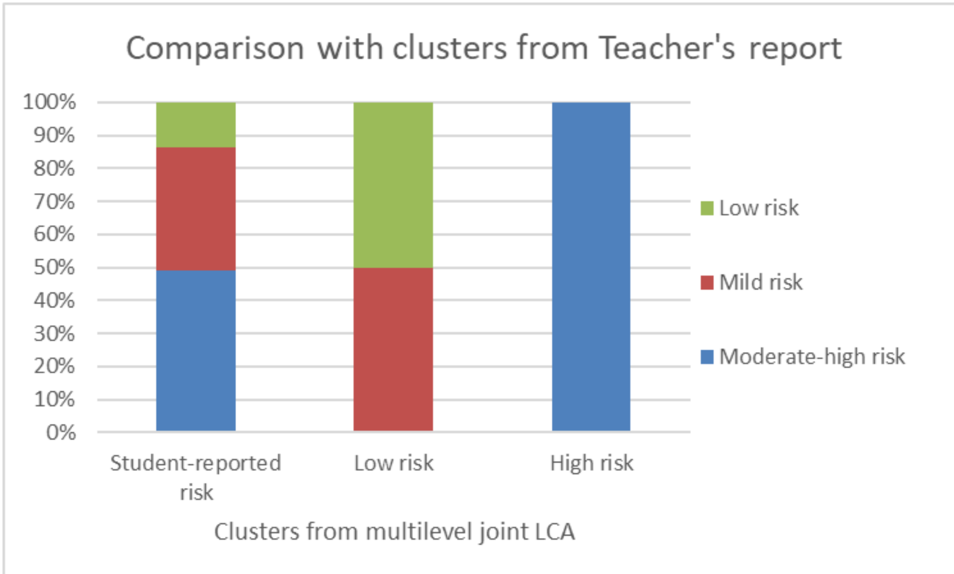

Figure 4.10: Comparison of clusters from multilevel joint LCA model with simple LCA model for teacher's report

Therefore, the joint model generated clusters that were the combination of clusters of schools discovered through analyzing self-report and teacher's report separately. Contrary to clusters defined by different risk levels obtained through separate models, the joint model revealed a distinct cluster of schools where the majority of students self-reported being in the moderate risk, while teachers reported a mixture of risk classes.

4.4. Effects of student-level and school-level covariates

The mental health patterns and clusters obtained from final joint LCA and multilevel join LCA models were used to assess the influence of various factors. We examined the effects of student-level factors on the mental health patterns among students and school-level factors on mental health clusters of schools. Table 4.17 shows the effects of student-level factors- such as sex, bullying at school, and Canadian-born status - on mental health patterns. Female students and students not bullying their peers at school were less likely to be associated with high risk and self-reported risk patterns compared to the low-risk pattern. Canadian-born students were more likely than immigrant students to be linked with the high risk on both reports rather than the low-risk pattern.

Table 4.17: Multinomial logistic regression results for student-level factors – estimated odds ratios and 95% confidence intervals

Factors	Patterns*		d.f.	Chi-square	p-value
	High risk	Self-reported risk			
Sex [Ref=male]					
Female	0.212 [0.178, 0.252]	0.491 [0.444, 0.543]	2	195.3683	<0.001
Teacher-reported bullying incidents [Ref=Yes]					
No	0.007 [0.005, 0.009]	0.136 [0.106, 0.174]	2	29.1562	<0.001
Canadian-born [Ref=No]					
Yes	1.570 [1.246, 1.977]	1.089 [0.957, 1.238]	2	11.8778	0.0026

*Reference category was low risk pattern

Table 4.18 presents the effects of school-level factor - specifically, the percentage of students engaging in bullying behavior at school - on school clusters. As each point increased in the percentage of students bullying at school, the odds of belonging to the high-risk cluster increased by 42% and the odds increased by 17% for student-reported risk cluster compared to the low-risk cluster.

Table 4.18: Multinomial logistic regression results for school-level factors – estimated odds ratios and 95% confidence intervals

Factor	Clusters*		d.f.	Chi-square	p-value
	Student-reported risk	High risk			
Percentage of students engaging in bullying behaviour at school	1.178 [1.126, 1.233]	1.425 [1.317, 1.542]	2	27.6737	<0.001

*Reference category was low risk cluster

5. Discussions and Conclusions

5.1 Discussion

This study proposed a joint latent class model to identify mental health classes and patterns by integrating the perspectives from both teachers and students. By extending the joint latent class model to a multilevel joint latent class model, we could also assess whether the distribution of these mental health patterns varied across schools. Our joint latent class model introduced an additional latent variable, latent pattern, to capture dependencies among same indicators from different sources. In addition, our multilevel extension of joint LCA incorporated another latent variable, latent cluster, to address dependencies among students within a higher-level group (school). Consequently, the joint model not only reflected individual perspective but also revealed interconnected patterns, thus enhancing our comprehension of student's mental health.

In our study, students reported experiencing more internalizing and externalizing difficulties on average compared to their teachers' reports. Previous literature using SDQ has shown that self-report usually underestimates the difficulties compared to teachers' reports [178]. However, one study focusing on a clinical sample of adolescents aged 11-17 found that student reported higher emotional and conduct difficulties than what their teachers reported [11]. The discrepancies in findings across studies emphasizes the importance of study characteristics and the involvement of other factors contributing to these divergent outcomes.

Upon examining the prevalence of students whose difficulty scores fell within the abnormal range for SDQ sub-scales, it was found that more students reported abnormally high scores in emotional difficulties compared to teachers' reports. Emotional symptoms like unhappy, depressed, anxious, self-conscious etc. are not easily observable or visible to others except for the individual experiencing them, leading to teachers potentially under reporting internalizing issues [65,195,196]. Conversely, teachers' report showed higher proportions of students being in abnormal range on other difficulties (peer, externalizing and pro-social difficulties) compared to students' report. Teachers form their perceptions of students by observing their behavior during classroom activities and interactions with peers, focusing on issues that may hinder teaching [15]. Some researchers also discussed that teachers might find behavioral issues to be more burdensome compared to emotional issues due to the effect of disruptive behaviour in classroom learning environment [13,14,197]. The discrepancies in outcomes between the two informants in our study highlight the importance of considering multiple perspectives to enhance the identification of mental health risks in children.

Through proposed joint LCA analysis, this study discovered six mental health classes (high, moderately high, medium, mild, mild internalizing and low risk) for each informant. By integrating these classes from both self-report and teachers' assessment, three mental health patterns: high risk, low risk and self-reported risk patterns among students were identified. The high-risk pattern (19.3%) consisted of students identified as high mental health risk by both self-report and teacher assessment, while low risk pattern (38.3%) included students deemed low risk on both reports. Additionally, a third pattern (42.4%) emerged, labeled self-reported risk, comprising students reporting moderate to high risk themselves but assessed as low to mild risk by teachers. This group of students represented the discrepancy among perspectives and the importance of considering perspectives from multiple informants, as these provided both subjective insights and objective observations on mental health issues. Moreover, this indicated towards barriers like gaps in awareness, understanding, and communication between teachers and students concerning mental health [198–200].

Tailed education and training programs for teachers can enhance their awareness and understanding of mental health issues. Programs targeting the collaboration among teachers, students, and mental health professionals are valuable for bridging perspectives. This collaboration enables teachers to identify and report any mental health concerns they observe in students. Students, in turn, feel empowered to share their challenges with teachers, fostering a supportive environment. and professionals integrate these perspectives to offer personalized interventions. The comparison of mental health patterns derived from joint LCA model with those obtained from separate LCAs offered notable advantages in integrating the perspectives of both teachers and students while effectively handling the dependencies inherent in the data structure. The proposed joint model discovered mental health patterns among students, capturing the combined insights from both sources, aligning with the vision of the proposed model.

Our analysis of exploring the effects of sex on mental health patterns revealed that female students were less likely to belong to high risk or self-reported risk pattern, indicating females having low risk compared to males. This finding contradicted with the conclusions of previous studies. In the study on a community sample of students aged 14-19 from Spain [201], gender differences were reported in terms of SDQ domains reported by students themselves. Female students had higher prevalence of having high scores for total difficulties. Also, emotional difficulties and behavioural difficulties were more prevalent among female and male students respectively. Another study on secondary school students from Austria reported that females perceived to have more overall problems compared to males [202]. Girls were found to have more mental health problems than boys, indicating girls at higher risk based on adolescents from population-based study of Finland [203].

A cross-national study also reported similar findings and mentioned that cultural variation and gender equality condition in respective countries can influence the gender gap and the direction (boys might have higher mental health issues in some countries) [204]. Depression and anxiety symptoms were more common among girls compared to boys in the data of Canadian Health Behaviour in School-aged Children (HBSC) study, but they had not considered any other symptoms like conduct, hyperactivity or pro-social problems [205]. Although previous research has reported consistent findings, our study presents contrasting results. This suggests that there may be more complexities to consider in understanding the topic, emphasizing the importance of continued exploration and refinement of our understanding.

Compared to immigrants, students who born in Canada were more likely to be classified as high risk in both teacher's report and self-report in our study. This finding is corroborated by previous studies. A study on grade 7 and 8 students in Montreal reported that Canadian-born students had higher level of emotional and behavioural problems compared to the immigrants counterparts (Cambodian and Central American) [206]. Another study in Montreal reported Caribbean and Filipino adolescents having less behavioural issues compared to Canadian-born counterpart [207]. This result is related to the assumptions obtained from population studies on migrated children that immigrants typically exhibit superior health outcomes relative to the native-born population of the host country or the populace they relocate to [206,208]. This hypothesis derives from empirical observations indicating that immigrants frequently present with more favorable health statuses upon arrival in their new country than the prevailing demographic and may sustain this advantage for a period following immigration [207].

The proposed multilevel joint LCA provided three clusters of schools: student-reported risk (178 schools, 62.6%), low risk (76 schools, 26.8%) and high-risk (30 schools, 10.6%) clusters based on the variations of mental health patterns among schools. Our study uncovered a distinct cluster that deviates from conventional categories as either low risk or high risk. Student-reported risk being the largest cluster of schools suggested that in most of the schools, students reported themselves to have moderate to high risk of mental problems while teachers reported students having low to mild risk. Previous literature mentioned that teachers did not feel confident to identify mental health problems of students [15] and even if they reported some students, most of them were being reported for their behavioural problems [14]. Furthermore, Teachers expressed a lack of sufficient training in identifying mental health issues of students and there is a tendency among teachers to overlook internalizing problems [209]. The student-reported cluster of schools comprised those institutions where teachers might not have understanding or received adequate training to identify students' symptoms of mental problems but students themselves reported their behaviour when asked through in-depth psychological evaluation. Policymakers should implement training programs for teachers to recognize children's mental health issues effectively. They should also prioritize creating supportive school environments that encourage open communication between students and teachers. By investing in teacher training and promoting supportive school environments, policymakers can significantly increase the chance of early identification of mental issues.

Through separate LCAs based on data from a single informant, clusters were identified that varied in terms of risk levels (e.g., high, low, mild, etc.). However, through a joint analysis that incorporated perspectives from both informants, we discovered clusters that integrated viewpoints from both sources.

Our analysis revealed that students not engaged in bullying activities reported by teachers were less likely to have high or self-reported risk. Our findings are quite similar to what other studies have discovered. The study using Youth and Mental Health Study in Norway found that those involved in bullying in adolescence tended to have higher risk of mental health problem in adulthood [210]. Adolescents engaging in bullying or physical fights are usually more prone to have overall high difficulties across all domains especially in emotional, peer and conduct difficulties [211]. Students who bully others experienced heightened levels of behavioral and emotional challenges, as well as difficulties at school [212]. Consistent findings across multiple studies indicate the significant impact of bullying on mental health, revealing its association with mental problems among students.

This study also explored the influence of the prevalence of bullying in schools on clusters of schools and discovered that as the prevalence of bullying increased within schools, the likelihood of a school being in high risk or student-reported risk clusters also increased. The prevalence of bullying increased within a school setting, indicating that both the students who engaged in bullying behavior and those who were victims of bullying increased. Thus, those students were reported having difficulties through the questionnaire by teachers and/or students themselves. Consequently, these students were categorized into the high-risk or self-reported risk group. As the number of students in the high-risk or self-reported risk group rose in a school, it ultimately led to the classification of the school itself as a high-risk school or student-reported risk school. This finding suggests adopting policies and intervention plans focusing on preventing bullying to reduce its occurrence in school setting and providing counselling or support services to both bullies and victims.

In conclusion, based on the application of proposed methods on grade 5 mental health survey, we suggest the proposed methods for analyzing the data from multiple informants simultaneously and identify the higher-level groups based on the variation of individual-level classes. Additionally, the study outcomes recommended the use of multiple informants to integrate diverse perspectives on mental health of students. This study suggests policymakers to organize training for teachers, intervention programs targeting teachers, students and mental health professionals to foster a supportive environment and to reduce the incidence of bullying at school.

5.2 Limitations and Future research

This study has some limitations and offers some suggestions for future research. There is a possibility of response and recall biases due to the self-reported data. Additionally, the dataset lacked more information on school-level factors, restricting a comprehensive understanding of variations among schools. Future research with access to more comprehensive school-level data would be invaluable in enhancing our understanding of the multifaceted dynamics that contribute to classification of schools based on student's mental health. More research is needed to understand what other factors make certain schools to be high risk or low risk in terms of mental health problems. This knowledge is important for designing mental disorder prevention programs that can be effective enough by targeting those factors which will change the situation. Moreover, the results from latent class analysis are very sample-sensitive, and we did not conduct cross-validation due to the potential underrepresentation of small class observations in both training and test datasets. Therefore, the findings may be limited to the current dataset. Analyzing different datasets could potentially reveal different latent classes.

In our proposed joint LCA, we constrained the definition of classes to be invariant over informants. Further research can be conducted by incorporating the variation in the definition and interpretation of classes by measuring item-response parameters to be different over informants/sources. In addition, we assigned equal weights to information from both sources in the joint model. Future research could explore assigning varying weights to these sources for assessing whether one source can provide more insights into student's mental health condition than the other. Furthermore, the relationship among domains of psychological problems can be explored through casual inference and other machine learning techniques like random forest model can be utilized to identify subgroups within the population in future studies.

5.3 Significance

Early identification of symptoms of mental health conditions is crucial in preventing the onset of severe mental disorders later in life. The primary objective of this thesis research was to identify patterns of mental health problems. Identification of groups with various mental health patterns will help policymakers and health practitioners to develop the targeted intervention and treatment plans that are tailored to the specific needs of these groups.

When assessing psychological status, it is common practice to gather reports from multiple informants (such as patient themselves, clinicians, and guardians) to ensure a more accurate judgment since the reports varies across informants. The proposed joint latent class offered a valuable tool for analyzing the information from these informants. This approach integrated several informants' view and provided more valid assessment of mental health conditions. In addition, the multilevel joint latent class model addressed the dependency among observations due to the nested structure of the data. In cluster-based intervention studies, it is crucial to consider the non-independence of observations within clusters when analyzing the data. The proposed method not only combined the perspectives of multiple informants but also mitigated the correlation among informant reports and the dependency of observations within groups or clusters.

To gain a comprehensive understanding of an individual's mental health problems, it is crucial to examine the contextual and perspective differences in mental health conditions. Another objective of this research was to explore the symptoms of students' mental health conditions from both the students' and teachers' perspectives simultaneously. This study obtained mental health patterns based on multiple perspectives and compared them with the mental health classes derived from each perspective individually. This comparative analysis provided the insights into the contribution of multiple informants in detecting mental health problems among children.

Considering that students spend a significant amount of time at school and share the same environment, it is likely that they may exhibit similar mental health conditions. Therefore, this study also investigated the heterogeneity among schools in terms of mental health patterns and reported that certain patterns were more prevalent in some schools compared to others. This study provides valuable insight to policymakers for school-oriented intervention programs targeting teachers, students and mental health professionals together.

The final objective of this study was to examine the effect of student-level and school-level factors on mental health patterns and clusters. The findings informed females, not engaging in bullying activities, being immigrants and having good subjective mental health influenced the students to have lower risk of mental health problems. This will guide to create intervention plans that address the unique needs of individuals and the overall population.

By understanding the mental health patterns among students and the effect of factors on different conditions, the study will serve as a guide for tailoring intervention plan for children as per individual and population needs. This will involve implementing intervention plan, understanding the potential implications, and anticipating the expected outcome. The ultimate goal is to promote successful development from childhood to adulthood and reduce the risk of mental disorders in adulthood.

Reference

1. García-Carrión R, Villarejo-Carballido B, Villardón-Gallego L. Children and Adolescents Mental Health: A Systematic Review of Interaction-Based Interventions in Schools and Communities. *Front Psychol* 2019; 10: 918.

2. What Is Children's Mental Health? *Centers for Disease Control and Prevention*, https://www.cdc.gov/childrensmentalhealth/basics.html (2023, accessed 13 July 2023).

3. Children and Mental Health: Is This Just a Stage? *National Institute of Mental Health (NIMH)*, https://www.nimh.nih.gov/health/publications/children-and-mental-health (accessed 13 July 2023).

4. Whitney DG, Peterson MD. US National and State-Level Prevalence of Mental Health Disorders and Disparities of Mental Health Care Use in Children. *JAMA Pediatrics* 2019; 173: 389–391.

5. Koning NR, Büchner FL, Vermeiren RRJM, et al. Identification of children at risk for mental health problems in primary care—Development of a prediction model with routine health care data. *eClinicalMedicine* 2019; 15: 89–97.

6. Reiss F, Meyrose A-K, Otto C, et al. Socioeconomic status, stressful life situations and mental health problems in children and adolescents: Results of the German BELLA cohort-study. *PLOS ONE* 2019; 14: e0213700.

7. Belfer ML. Child and adolescent mental disorders: the magnitude of the problem across the globe. *J Child Psychol Psychiatry* 2008; 49: 226–236.

8. Polanczyk GV, Salum GA, Sugaya LS, et al. Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *J Child Psychol Psychiatry* 2015; 56: 345–365.

9. Kieling C, Baker-Henningham H, Belfer M, et al. Child and adolescent mental health worldwide: evidence for action. *The Lancet* 2011; 378: 1515–1525.

10. He J-P, Burstein M, Schmitz A, et al. The Strengths and Difficulties Questionnaire (SDQ): the factor structure and scale validation in U.S. adolescents. *J Abnorm Child Psychol* 2013; 41: 583–595.

11. Becker A, Hagenberg N, Roessner V, et al. Evaluation of the self-reported SDQ in a clinical setting: do self-reports tell us more than ratings by adult informants? *Eur Child Adolesc Psychiatry* 2004; 13 Suppl 2: II17-24.

12. Vugteveen J, De Bildt A, Hartman CA, et al. Using the Dutch multi-informant Strengths and Difficulties Questionnaire (SDQ) to predict adolescent psychiatric diagnoses. *Eur Child Adolesc Psychiatry* 2018; 27: 1347–1359.

13. Loades ME, Mastroyannopoulou K. Teachers' Recognition of Children's Mental Health Problems. *Child and Adolescent Mental Health* 2010; 15: 150–156.

14. Meltzer H, Bebbington P, Brugha T, et al. The reluctance to seek treatment for neurotic disorders. *Int Rev Psychiatry* 2003; 15: 123–128.

15. Rothì DM, Leavey G, Best R. On the front-line: Teachers as active observers of pupils' mental health. *Teaching and Teacher Education* 2008; 24: 1217–1231.

16. Cantwell DP, Lewinsohn PM, Rohde P, et al. Correspondence between adolescent report and parent report of psychiatric diagnostic data. *J Am Acad Child Adolesc Psychiatry* 1997; 36: 610–619.

17. Vazire S. Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology* 2010; 98: 281–300.

18. Guze SB. Diagnostic and Statistical Manual of Mental Disorders, 4th ed. (DSM-IV). *AJP* 1995; 152: 1228–1228.

19. Vugteveen J, de Bildt A, Hartman CA, et al. The combined self- and parent-rated SDQ score profile predicts care use and psychiatric diagnoses. *Eur Child Adolesc Psychiatry* 2021; 30: 1983–1994.

20. Petersen KJ, Qualter P, Humphrey N. The Application of Latent Class Analysis for Investigating Population Child Mental Health: A Systematic Review. *Front Psychol* 2019; 10: 1214.

21. Petersen KJ, Humphrey N, Qualter P. Latent Class Analysis of Mental Health in Middle Childhood: Evidence for the Dual-Factor Model. *School Mental Health* 2020; 12: 786–800.

22. Parkes J, White-Koning M, Dickinson HO, et al. Psychological problems in children with cerebral palsy: a cross-sectional European study. *Journal of Child Psychology and Psychiatry* 2008; 49: 405–413.

23. Kim EK, Dowdy E, Furlong MJ, et al. Mental health profiles and quality of life among Korean adolescents. *School Psychology International* 2017; 38: 98–116.

24. Suldo SM, Shaffer EJ. Looking Beyond Psychopathology: The Dual-Factor Model of Mental Health in Youth. *School Psychology Review* 2008; 37: 52–68.

25. Dowdy E, Furlong M, Raines TC, et al. Enhancing School-Based Mental Health Services With a Preventive and Promotive Approach to Universal Screening for Complete Mental Health. *Journal of Educational and Psychological Consultation* 2015; 25: 178–197.

26. Masyn KE. Latent Class Analysis and Finite Mixture Modeling. In: *The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2*. Oxford University Press, 2013. Epub ahead of print 2013. DOI: 10.1093/oxfordhb/9780199934898.013.0025.

27. Berlin KS, Williams NA, Parra GR. An Introduction to Latent Variable Mixture Modeling (Part 1): Overview and Cross-Sectional Latent Class and Latent Profile Analyses. *Journal of Pediatric Psychology* 2014; 39: 174–187.

28. Gustafsson J, Lyyra N, Jasinskaja-Lahti I, et al. Mental health profiles of Finnish adolescents before and after the peak of the COVID-19 pandemic. *Child and Adolescent Psychiatry and Mental Health* 2023; 17: 54.

29. Eriksson C, Stattin H. Secular trends in mental health profiles among 15-year-olds in Sweden between 2002 and 2018. *Frontiers in Public Health*; 11, https://www.frontiersin.org/articles/10.3389/fpubh.2023.1015509 (2023, accessed 27 August 2023).

30. Sulik MJ, Blair C, Greenberg M. Child conduct problems across home and school contexts: A person-centered approach. *Journal of Psychopathology and Behavioral Assessment* 2017; 39: 46–57.

31. Bergman LR, Magnusson D. Person-centered Research. In: Smelser NJ, Baltes PB (eds) *International Encyclopedia of the Social & Behavioral Sciences*. Oxford: Pergamon, pp. 11333–11339.

32. Gad SC, Rousseaux CG. 15 - Use and Misuse of Statistics in the Design and Interpretation of Studies. In: Haschek WM, Rousseaux CG, Wallig MA (eds) *Handbook of Toxicologic Pathology (Second Edition)*. San Diego: Academic Press, pp. 327–418.

33. Rousseaux CG, Gad SC. Chapter 30 - Statistical Assessment of Toxicologic Pathology Studies. In: Haschek WM, Rousseaux CG, Wallig MA (eds) *Haschek and Rousseaux's Handbook of Toxicologic Pathology (Third Edition)*. Boston: Academic Press, pp. 893–988.

34. Karnowski V. Latent Class Analysis. In: *The International Encyclopedia of Communication Research Methods*. John Wiley & Sons, Ltd, pp. 1–10.

35. Henry KL, Muthén B. Multilevel Latent Class Analysis: An Application of Adolescent Smoking Typologies with Individual and Contextual Predictors. *Struct Equ Modeling* 2010; 17: 193–215.

36. Bakk Z, Di Mari R, Oser J, et al. Two-stage Multilevel Latent Class Analysis with Covariates in the Presence of Direct Effects. *Structural Equation Modeling: A Multidisciplinary Journal* 2022; 29: 267–277.

37. Hsieh T-C, Yang C. Multi-Level Latent Class Analysis of Internet Use Pattern in Taiwan. In: Yonazi JJ, Sedoyeka E, Ariwa E, et al. (eds) *e-Technologies and Networks for Development*. Berlin, Heidelberg: Springer, 2011, pp. 197–208.

38. Bold KW, Buta E, Simon P, et al. Using Latent Class Analysis to Examine Susceptibility to Various Tobacco Products Among Adolescents. *Nicotine Tob Res* 2019; 22: 2059–2065.

39. Carey E, Devine A, Hill F, et al. Differentiating anxiety forms and their role in academic performance from primary to secondary school. *PLOS ONE* 2017; 12: e0174418.

40. Ferdinand RF, Bongers IL, van der Ende J, et al. Distinctions between separation anxiety and social anxiety in children and adolescents. *Behav Res Ther* 2006; 44: 1523–1535.

41. Hudziak JJ, Wadsworth ME, Heath AC, et al. Latent class analysis of Child Behavior Checklist attention problems. *J Am Acad Child Adolesc Psychiatry* 1999; 38: 985–991.

42. Romano E, Baillargeon R, Tremblay RE. *Prevalence of hyperactivity-impulsivity and inattention among Canadian children: findings from the first data collection cycle (1994-1995) of the National Longitudinal Survey of Children and Youth: final report*. Hull, Quebec: Human Resources Development Canada, 2002.

43. Basten MMGJ, Althoff RR, Tiemeier H, et al. The dysregulation profile in young children: empirically defined classes in the Generation R study. *J Am Acad Child Adolesc Psychiatry* 2013; 52: 841-850.e2.

44. Basten M, Tiemeier H, Althoff RR, et al. The Stability of Problem Behavior Across the Preschool Years: An Empirical Approach in the General Population. *J Abnorm Child Psychol* 2016; 44: 393–404.

45. McElroy E, Shevlin M, Murphy J. Internalizing and externalizing disorders in childhood and adolescence: A latent transition analysis using ALSPAC data. *Comprehensive Psychiatry* 2017; 75: 75–84.

46. Van Lier PAC, Verhulst FC, Crijnen AAM. Screening for disruptive behavior syndromes in children: The application of latent class analyses and implications for prevention programs. *Journal of Consulting and Clinical Psychology* 2003; 71: 353–363.

47. van Lier PAC, Verhulst FC, van der Ende J, et al. Classes of disruptive behaviour in a sample of young elementary school children. *Journal of Child Psychology and Psychiatry* 2003; 44: 377–387.

48. Fergusson DM, Boden JM, Horwood LJ. Situational and generalised conduct problems and later life outcomes: evidence from a New Zealand birth cohort. *J Child Psychol Psychiatry* 2009; 50: 1084–1092.

49. Vermunt JK. Multilevel Latent Class Models. *Sociological Methodology* 2003; 33: 213–239.

50. Jiang D, Santos R, Josephson W, et al. A Comparison of Variable- and Person-Oriented Approaches in Evaluating a Universal Preventive Intervention. *Prev Sci* 2018; 19: 738–747.

51. Kuny AV, Althoff RR, Copeland W, et al. Separating the domains of oppositional behavior: Comparing latent models of the Conners' Oppositional Subscale. *Journal of the American Academy of Child & Adolescent Psychiatry* 2013; 52: 172–183.

52. Mental health of adolescents, https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health (accessed 21 August 2023).

53. CHILD AND ADOLESCENT MENTAL HEALTH. In: *2022 National Healthcare Quality and Disparities Report [Internet]*. Agency for Healthcare Research and Quality (US), https://www.ncbi.nlm.nih.gov/books/NBK587174/ (2022, accessed 30 August 2023).

54. Children and Youth. *Mental Health Commission of Canada*, https://mentalhealthcommission.ca/what-we-do/children-and-youth/ (accessed 30 August 2023).

55. Mental health of children and youth in Canada | CIHI, https://www.cihi.ca/en/mental-health-of-children-and-youth-in-canada (accessed 30 August 2023).

56. Chartier M, Brownell M, MacWilliam L, et al. *The Mental Health of Manitoba's Children*. Winnipeg, MB.: Manitoba Centre for Health Policy, http://mchp-appserv.cpe.umanitoba.ca/deliverable.php?referencePaperID=77480 (Fall 2016, accessed 30 August 2023).

57. Weisz JR, Kazdin AE. *Evidence-Based Psychotherapies for Children and Adolescents, Second Edition*. Guilford Press, 2010.

58. Tully LA, Hawes DJ, Doyle FL, et al. A national child mental health literacy initiative is needed to reduce childhood mental health disorders. *Aust N Z J Psychiatry* 2019; 53: 286–290.

59. Achenbach TM, Ivanova MY, Rescorla LA, et al. Internalizing/Externalizing Problems: Review and Recommendations for Clinical and Research Applications. *Journal of the American Academy of Child & Adolescent Psychiatry* 2016; 55: 647–656.

60. Willner CJ, Gatzke-Kopp LM, Bray BC. The dynamics of internalizing and externalizing comorbidity across the early school years. *Development and Psychopathology* 2016; 28: 1033–1052.

61. Achenbach TM, Edelbrock CS. The classification of child psychopathology: A review and analysis of empirical efforts. *Psychological Bulletin* 1978; 85: 1275–1301.

62. Gage NA. Characteristics of Students with Emotional Disturbance Manifesting Internalizing Behaviors: A Latent Class Analysis. *Education and Treatment of Children* 2013; 36: 127–145.

63. Morales A, Rodríguez-Menchón M, Tomczyk S, et al. Internalizing and externalizing symptoms in Spanish children aged 6-8: Results of a latent profile analysis. *Journal of Affective Disorders* 2021; 279: 617–623.

64. O'Rourke RH, Brown D, Martin Ginis KA, et al. An Examination of the Mental Health Status of Canadian Children and Youth with Disabilities. *Canadian Journal of Community Mental Health* 2022; 41: 75–103.

65. Pedersen ML, Holen S, Lydersen S, et al. School functioning and internalizing problems in young schoolchildren. *BMC Psychology* 2019; 7: 88.

66. Heiervang E, Stormark KM, Lundervold AJ, et al. Psychiatric Disorders in Norwegian 8- to 10-Year-Olds: An Epidemiological Survey of Prevalence, Risk Factors, and Service Use. *Journal of the American Academy of Child & Adolescent Psychiatry* 2007; 46: 438–447.

67. Asher SR, Parker JG. Significance of Peer Relationship Problems in Childhood. In: Schneider BH, Attili G, Nadel J, et al. (eds) *Social Competence in Developmental Perspective*. Dordrecht: Springer Netherlands, pp. 5–23.

68. Rutherford RB, Quinn MM, Mathur SR. *Handbook of Research in Emotional and Behavioral Disorders*. Guilford Press, 2007.

69. Grover RL, Ginsburg GS, Ialongo N. Psychosocial outcomes of anxious first graders: a seven-year follow-up. *Depress Anxiety* 2007; 24: 410–420.

70. Weidman AC, Augustine AA, Murayama K, et al. Internalizing symptomatology and academic achievement: Bi-directional prospective relations in adolescence. *Journal of Research in Personality* 2015; 58: 106–114.

71. Riglin L, Petrides K v., Frederickson N, et al. The relationship between emotional problems and subsequent school attainment: A meta-analysis. *Journal of Adolescence* 2014; 37: 335–346.

72. Wagner M, Kutash K, Duchnowski AJ, et al. The Special Education Elementary Longitudinal Study and the National Longitudinal Transition Study: Study Designs and Implications for Children and Youth With Emotional Disturbance. *Journal of Emotional and Behavioral Disorders* 2005; 13: 25–41.

73. Woodward LJ, Fergusson DM. Childhood Peer Relationship Problems and Psychosocial Adjustment in Late Adolescence. *J Abnorm Child Psychol* 1999; 27: 87–104.

74. Liu J. Childhood Externalizing Behavior: Theory and Implications. *J Child Adolesc Psychiatr Nurs* 2004; 17: 93–103.

75. Colder CR, Scalco M, Trucco EM, et al. Prospective Associations of Internalizing and Externalizing Problems and Their Co-Occurrence with Early Adolescent Substance Use. *J Abnorm Child Psychol* 2013; 41: 667–677.

76. Lau MA, Temcheff CE, Poirier M, et al. School dropout: The role of childhood conduct problems and depressive symptoms. *Psychology in the Schools* 2021; 58: 2313–2327.

77. Mikami AY, Lorenzi J. Gender and Conduct Problems Predict Peer Functioning among Children with Attention-Deficit/Hyperactivity Disorder. *J Clin Child Adolesc Psychol* 2011; 40: 777–786.

78. Fairchild G, Hawes DJ, Frick PJ, et al. Conduct disorder. *Nat Rev Dis Primers* 2019; 5: 1–25.

79. Krueger RF, Hicks BM, Patrick CJ, et al. Etiologic connections among substance dependence, antisocial behavior, and personality: modeling the externalizing spectrum. *J Abnorm Psychol* 2002; 111: 411–424.

80. McDougal E, Tai C, Stewart TM, et al. Understanding and Supporting Attention Deficit Hyperactivity Disorder (ADHD) in the Primary School Classroom: Perspectives of Children with ADHD and their Teachers. *J Autism Dev Disord*. Epub ahead of print 1 July 2022. DOI: 10.1007/s10803-022-05639-3.

81. Attention-Deficit/Hyperactivity Disorder. *National Institute of Mental Health (NIMH)*, https://www.nimh.nih.gov/health/topics/attention-deficit-hyperactivity-disorder-adhd (accessed 24 July 2023).

82. Palmu IR, Närhi VM, Savolainen HK. Externalizing behaviour and academic performance – the cross-lagged relationship during school transition. *Emotional and Behavioural Difficulties* 2018; 23: 111–126.

83. Olivier E, Morin AJS, Langlois J, et al. Internalizing and Externalizing Behavior Problems and Student Engagement in Elementary and Secondary School Students. *J Youth Adolesc* 2020; 49: 2327–2346.

84. Hammond SI, Waugh W, Satlof-Bedrick E, et al. Prosocial Behavior during Childhood and Cultural Variations. In: Wright JD (ed) *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*. Oxford: Elsevier, pp. 228–232.

85. Eisenberg N, Fabes RA, Spinrad TL. Prosocial Development. In: *Handbook of child psychology: Social, emotional, and personality development, Vol. 3, 6th ed*. Hoboken, NJ, US: John Wiley & Sons, Inc., 2006, pp. 646–718.

86. Wispé LG. Positive Forms of Social Behavior: An Overview. *Journal of Social Issues* 1972; 28: 1–19.

87. Obsuth I, Eisner MP, Malti T, et al. The developmental relation between aggressive behaviour and prosocial behaviour: A 5-year longitudinal study. *BMC Psychology* 2015; 3: 16.

88. Achenbach TM. Diagnosis, Assessment, Taxonomy, and Case Formulations. In: Ollendick TH, Hersen M (eds) *Handbook of Child Psychopathology*. Boston, MA: Springer US, pp. 63–87.

89. Achenbach TM. Challenges and Benefits of Assessment, Diagnosis, and Taxonomy for Clinical Practice and Research. *Aust N Z J Psychiatry* 2001; 35: 263–271.

90. Dowdy E, Kamphaus RW. A Comparison of Classification Methods for Use in Predicting School-Based Outcomes. *Contemp School Psychol* 2007; 12: 121–132.

91. Jimerson SR, Coffino B, Sroufe LA. Building school-based interventions on attachment theory and research. *Journal of Early Childhood and Infant Psychology* 2007; 3: 79–94.

92. Sroufe LA, Egeland B, Carlson EA, et al. *The Development of the Person: The Minnesota Study of Risk and Adaptation from Birth to Adulthood*. Guilford Press, 2009.

93. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, Text Revision*. 5th edition. Washington, DC: American Psychiatric Association Publishing, 2022.

94. Jensen PS, Watanabe HK, Richters JE, et al. Scales, diagnoses, and child psychopathology: II. Comparing the CBCL and the DISC against external validators. *J Abnorm Child Psychol* 1996; 24: 151–168.

95. Cantwell DP. Classification of child and adolescent psychopathology. *Child Psychology & Psychiatry & Allied Disciplines* 1996; 37: 3–12.

96. Clark LA, Watson D, Reynolds S. Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology* 1995; 46: 121–153.

97. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Arch Gen Psychiatry* 1994; 51: 8–19.

98. Krueger RF. The Structure of Common Mental Disorders. *Archives of General Psychiatry* 1999; 56: 921–926.

99. Wilens TE, Biederman J, Brown S, et al. Patterns of Psychopathology and Dysfunction in Clinically Referred Preschoolers. *Journal of Developmental & Behavioral Pediatrics* 2002; 23: S31.

100. Briggs-Gowan MJ, Carter AS, Bosson-Heenan J, et al. Are Infant-Toddler Social-Emotional and Behavioral Problems Transient? *Journal of the American Academy of Child & Adolescent Psychiatry* 2006; 45: 849–858.

101. Krueger RF, Caspi A, Moffitt TE, et al. The structure and stability of common mental disorders (DSM-III-R): a longitudinal-epidemiological study. *J Abnorm Psychol* 1998; 107: 216–227.

102. De Los Reyes A, Kazdin AE. Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychol Bull* 2005; 131: 483–509.

103. Achenbach TM, McConaughy SH, Howell CT. Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin* 1987; 101: 213–232.

104. Goodman R, Meltzer H, Bailey V. The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *Int Rev Psychiatry* 2003; 15: 173–177.

105. Fergusson DM, Horwood LJ. Estimation of method and trait variance in ratings of conduct disorder. *J Child Psychol Psychiatry* 1989; 30: 365–378.

106. Goodman LA. Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* 1974; 61: 215–231.

107. Lazarsfeld PF, Henry NW. *Latent structure analysis*. New York: Houghton, Mifflin, http://www.gbv.de/dms/hbz/toc/ht000685628.pdf (1968, accessed 22 July 2023).

108. Hagenaars JA, McCutcheon AL (eds). *Applied Latent Class Analysis*. Cambridge: Cambridge University Press. Epub ahead of print 2002. DOI: 10.1017/CBO9780511499531.

109. Eshima N. *An Introduction to Latent Class Analysis: Methods and Applications*. Singapore: Springer. Epub ahead of print 2022. DOI: 10.1007/978-981-19-0972-6.

110. Vermunt JK, Magidson J. Latent Class Cluster Analysis. In: McCutcheon AL, Hagenaars JA (eds) *Applied Latent Class Analysis*. Cambridge: Cambridge University Press, pp. 89–106.

111. Miettunen J, Nordström T, Kaakinen M, et al. Latent variable mixture modeling in psychiatric research--a review and application. *Psychol Med* 2016; 46: 457–467.

112. Oberski D. Mixture Models: Latent Profile and Latent Class Analysis. In: Robertson J, Kaptein M (eds) *Modern Statistical Methods for HCI*. Cham: Springer International Publishing, pp. 275–287.

113. Bauer D, Curran P. Distributional Assumptions of Growth Mixture Models: Implications for Overextraction of Latent Trajectory Classes. *Psychological methods* 2003; 8: 338–63.

114. Chen F, Bollen KA, Paxton P, et al. Improper Solutions in Structural Equation Models: Causes, Consequences, and Strategies. *Sociological Methods & Research* 2001; 29: 468–508.

115. Diallo TMO, Morin AJS, Lu H. Impact of Misspecifications of the Latent Variance–Covariance and Residual Matrices on the Class Enumeration Accuracy of Growth Mixture Models. *Structural Equation Modeling: A Multidisciplinary Journal* 2016; 23: 507–531.

116. Enders CK, Tofighi D. The Impact of Misspecifying Class-Specific Residual Variances in Growth Mixture Models. *Structural Equation Modeling: A Multidisciplinary Journal* 2008; 15: 75–95.

117. Kim M, Lamont AE, Jaki T, et al. Impact of an equality constraint on the class-specific residual variances in regression mixtures: A Monte Carlo simulation study. *Behav Res* 2016; 48: 813–826.

118. Peugh J, Fan X. Modeling Unobserved Heterogeneity Using Latent Profile Analysis: A Monte Carlo Simulation. *Structural Equation Modeling: A Multidisciplinary Journal* 2013; 20: 616–639.

119. Rozier MD, Ghaferi AA, Rose A, et al. Patient Preferences for Bariatric Surgery: Findings From a Survey Using Discrete Choice Experiment Methodology. *JAMA Surg* 2019; 154: e184375.

120. Nosetti L, Paglietti MG, Brunetti L, et al. Application of latent class analysis in assessing the awareness, attitude, practice and satisfaction of paediatricians on sleep disorder management in children in Italy. *PLOS ONE* 2020; 15: e0228377.

121. Ahanchi NS, Hadaegh F, Alipour A, et al. Application of Latent Class Analysis to Identify Metabolic Syndrome Components Patterns in adults: Tehran Lipid and Glucose study. *Sci Rep* 2019; 9: 1572.

122. Lowthian E, Page N, Melendez-Torres GJ, et al. Using Latent Class Analysis to Explore Complex Associations Between Socioeconomic Status and Adolescent Health and Well-Being. *Journal of Adolescent Health* 2021; 69: 774–781.

123. Palma MA, Ness ML, Anderson DP. Fashionable food: a latent class analysis of social status in food purchases. *Applied Economics* 2017; 49: 238–250.

124. Vermunt JK. An EM algorithm for the estimation of parametric and nonparametric hierarchical nonlinear models. *Statistica Neerlandica* 2004; 58: 220–233.

125. Chen Q, Kwok O-M, Luo W, et al. The Impact of Ignoring a Level of Nesting Structure in Multilevel Growth Mixture Models: A Monte Carlo Study. *Structural Equation Modeling: A Multidisciplinary Journal* 2010; 17: 570–589.

126. Kaplan D, Keller B. A Note on Cluster Effects in Latent Class Analysis. *Structural Equation Modeling: A Multidisciplinary Journal* 2011; 18: 525–536.

127. Van Horn ML, Fagan AA, Jaki T, et al. Using Multilevel Mixtures to Evaluate Intervention Effects in Group Randomized Trials. *Multivariate Behav Res* 2008; 43: 289–326.

128. Vermunt JK. Latent class and finite mixture models for multilevel data sets. *Stat Methods Med Res* 2008; 17: 33–51.

129. Park J, Yu H-T. The Impact of Ignoring the Level of Nesting Structure in Nonparametric Multilevel Latent Class Models. *Educ Psychol Meas* 2016; 76: 824–847.

130. Mutz R, Daniel H-D. University and student segmentation: multilevel latent-class analysis of students’ attitudes towards research methods and statistics. *Br J Educ Psychol* 2013; 83: 280–304.

131. Rindskopf D. Heavy Alcohol use in the “Fighting Back” Survey Sample: Separating Individual and Community Level Influences using Multilevel Latent Class Analysis. *Journal of Drug Issues* 2006; 36: 441–462.

132. Bijmolt THA, Paas LJ, Vermunt JK. Country and consumer segmentation: multi-level latent class analysis of financial product ownership. *International Journal of Research in Marketing* 2004; 21: 323–340.

133. Nikita E. Chapter 9 - Statistical Methods in Human Osteology. In: Nikita E (ed) *Osteoarchaeology*. Academic Press, pp. 355–442.

134. Boccard J, Rudaz S. Chapter 27 - Mass Spectrometry Metabolomic Data Handling for Biomarker Discovery. In: Issaq HJ, Veenstra TD (eds) *Proteomic and Metabolomic Approaches to Biomarker Discovery*. Boston: Academic Press, pp. 425–445.

135. Weller BE, Bowen NK, Faubert SJ. Latent Class Analysis: A Guide to Best Practice. *Journal of Black Psychology* 2020; 46: 287–311.

136. Bunge JA, Judson DH. Data Mining. In: Kempf-Leonard K (ed) *Encyclopedia of Social Measurement*. New York: Elsevier, pp. 617–624.

137. Bolin JH, Edwards JM, Finch WH, et al. Applications of cluster analysis to the creation of perfectionism profiles: a comparison of two clustering approaches. *Front Psychol* 2014; 5: 343.

138. Frączek B. The Use of Cluster Analysis to Assess the Threats of Poverty or Social Exclusion in EU Countries: The Case of People with Disabilities Compared to People without Disabilities. *Sustainability* 2022; 14: 14223.

139. Verleysen FT, Weeren A. Clustering by publication patterns of senior authors in the social sciences and humanities. *Journal of Informetrics* 2016; 10: 254–272.

140. von Eye A, Bergman LR. Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology* 2003; 15: 553–580.

141. Rosato NS, Baer JC. Latent Class Analysis: A Method for Capturing Heterogeneity. *Social Work Research* 2012; 36: 61–69.

142. Lanza ST, Cooper BR. Latent Class Analysis for Developmental Research. *Child Development Perspectives* 2016; 10: 59–64.

143. Wadsworth ME, Hudziak JJ, Heath AC, et al. Latent class analysis of child behavior checklist anxiety/depression in children and adolescents. *J Am Acad Child Adolesc Psychiatry* 2001; 40: 106–114.

144. Liu F, Yang D, Liu Y, et al. Use of latent profile analysis and k-means clustering to identify student anxiety profiles. *BMC Psychiatry* 2022; 22: 12.

145. Althoff RR, Copeland WE, Stanger C, et al. The latent class structure of ADHD is stable across informants. *Twin Res Hum Genet* 2006; 9: 507–522.

146. Baillargeon RH, Tremblay RE, Willms D. The prevalence of physical aggression in Canadian children: A multi-group latent class analysis of data from the first collection cycle (1994-1995) of the NLSCY, http://ruor.uottawa.ca/handle/10393/12870 (1999, accessed 25 July 2023).

147. Lee K-H, Baillargeon RH, Vermunt JK, et al. Age differences in the prevalence of physical aggression among 5–11-year-old Canadian boys and girls. *Aggressive Behavior* 2007; 33: 26–37.

148. Althoff RR, Rettew DC, Boomsma DI, et al. Latent class analysis of the Child Behavior Checklist Obsessive-Compulsive Scale. *Compr Psychiatry* 2009; 50: 584–592.

149. Täljedal T, Granlund M, Almqvist L, et al. Patterns of mental health problems and well-being in children with disabilities in Sweden: A cross-sectional survey and cluster analysis. *PLOS ONE* 2023; 18: e0288815.

150. Eriksson C, Stattin H. Mental-health profiling with person-centred analysis: A study of adolescents in Sweden. *Scand J Public Health* 2023; 51: 628–635.

151. Charalampopoulou M, Choi EJ, Korczak DJ, et al. Mental health profiles of autistic children and youth during the COVID-19 pandemic. *Paediatrics & Child Health* 2022; 27: S59–S65.

152. Raffagnato A, Iannattone S, Fasolato R, et al. Psycho-Behavioral Profiles of Pediatric Inpatients with Past and Recent Onset of Nonsuicidal Self-Injury: A Cluster Analysis Approach. *Journal of Clinical Medicine* 2022; 11: 4602.

153. Goodman A, Lamping DL, Ploubidis GB. When to use broader internalising and externalising subscales instead of the hypothesised five subscales on the Strengths and Difficulties Questionnaire (SDQ): data from British parents, teachers and children. *J Abnorm Child Psychol* 2010; 38: 1179–1191.

154. Goodman R. Psychometric Properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child & Adolescent Psychiatry* 2001; 40: 1337–1345.

155. Goodman R, Slobodskaya H, Knyazev G. Russian child mental health--a cross-sectional study of prevalence and risk factors. *Eur Child Adolesc Psychiatry* 2005; 14: 28–33.

156. Lynch F, Mills C, Daly I, et al. Challenging times: prevalence of psychiatric disorders and suicidal behaviours in Irish adolescents. *J Adolesc* 2006; 29: 555–573.

157. Mullick MSI, Goodman R. The prevalence of psychiatric disorders among 5-10 year olds in rural, urban and slum areas in Bangladesh: an exploratory study. *Soc Psychiatry Psychiatr Epidemiol* 2005; 40: 663–671.

158. Green H, McGinnity A, Meltzer H, et al. *Mental Health of Children and Young People in Great Britain, 2004*. PALGRAVE MACMILLAN. Epub ahead of print 2005. DOI: 10.1037/e557702010-001.

159. Undheim AM, Sund AM. School factors and the emergence of depressive symptoms among young Norwegian adolescents. *EuropChild & Adolescent Psych* 2005; 14: 446–453.

160. Chaplin TM, Aldao A. Gender differences in emotion expression in children: a meta-analytic review. *Psychol Bull* 2013; 139: 735–765.

161. Bourdon KH, Goodman R, Rae DS, et al. The Strengths and Difficulties Questionnaire: U.S. normative data and psychometric properties. *J Am Acad Child Adolesc Psychiatry* 2005; 44: 557–564.

162. Matsuishi T, Nagano M, Araki Y, et al. Scale properties of the Japanese version of the Strengths and Difficulties Questionnaire (SDQ): a study of infant and school children in community samples. *Brain Dev* 2008; 30: 410–415.

163. Zahn-Waxler C, Shirtcliff EA, Marceau K. Disorders of childhood and adolescence: gender and psychopathology. *Annu Rev Clin Psychol* 2008; 4: 275–303.

164. Kuehner C. Why is depression more common among women than among men? *Lancet Psychiatry* 2017; 4: 146–158.

165. Källmén H, Hallgren M. Bullying at school and mental health problems among adolescents: a repeated cross-sectional study. *Child and Adolescent Psychiatry and Mental Health* 2021; 15: 74.

166. Högberg B. Educational stressors and secular trends in school stress and mental health problems in adolescents. *Soc Sci Med* 2021; 270: 113616.

167. Kidger J, Araya R, Donovan J, et al. The effect of the school environment on the emotional health of adolescents: a systematic review. *Pediatrics* 2012; 129: 925–949.

168. Arseneault L, Bowes L, Shakoor S. Bullying victimization in youths and mental health problems: 'Much ado about nothing'? *Psychological Medicine* 2010; 40: 717–729.

169. Veenstra R, Lindenberg S, Oldehinkel AJ, et al. Bullying and Victimization in Elementary Schools: A Comparison of Bullies, Victims, Bully/Victims, and Uninvolved Preadolescents. *Developmental Psychology* 2005; 41: 672–682.

170. Nansel TR, Overpeck MD, Haynie DL, et al. Relationships Between Bullying and Violence Among US Youth. *Archives of Pediatrics & Adolescent Medicine* 2003; 157: 348–353.

171. Kirmayer LJ, Narasiah L, Munoz M, et al. Common mental health problems in immigrants and refugees: general approach in primary care. *CMAJ* 2011; 183: E959–E967.

172. Georgiades K, Duncan L, Wang L, et al. Six-Month Prevalence of Mental Disorders and Service Contacts among Children and Youth in Ontario: Evidence from the 2014 Ontario Child Health Study. *Can J Psychiatry* 2019; 64: 246–255.

173. Chiu M, Amartey A, Wang X, et al. Ethnic Differences in Mental Health Status and Service Utilization: A Population-Based Study in Ontario, Canada. *Can J Psychiatry* 2018; 63: 481–491.

174. Gadermann AM, Gagné Petteni M, Janus M, et al. Prevalence of Mental Health Disorders Among Immigrant, Refugee, and Nonimmigrant Children and Youth in British Columbia, Canada. *JAMA Netw Open* 2022; 5: e2144934.

175. Breslau J, Aguilar-Gaxiola S, Borges G, et al. Risk for Psychiatric Disorder among Immigrants and Their US-Born Descendants: Evidence from the National Comorbidity Survey-Replication. *J Nerv Ment Dis* 2007; 195: 189–195.

176. Government of Manitoba. 2015-2016 Annual Report | Annual Reports | Publications | Child and Youth Programs | Province of Manitoba, https://www.gov.mb.ca/healthychild/about/annual.html (accessed 14 July 2023).

177. Goodman R. The Strengths and Difficulties Questionnaire: a research note. *J Child Psychol Psychiatry* 1997; 38: 581–586.

178. Goodman R, Meltzer H, Bailey V. The Strengths and Difficulties Questionnaire: a pilot study on the validity of the self-report version. *Eur Child Adolesc Psychiatry* 1998; 7: 125–130.

179. Goodman R. The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *J Child Psychol Psychiatry* 1999; 40: 791–799.

180. Scoring the SDQ, https://www.sdqinfo.org/py/sdqinfo/c0.py (accessed 14 July 2023).

181. SDQscore, https://sdqscore.org/ (accessed 1 June 2024).

182. SDQplus, https://sdqscore.org/sdqplus/ (accessed 1 June 2024).

183. Vermunt JK, Magidson J. Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax.

184. He J, Fan X. Latent Profile Analysis. In: Zeigler-Hill V, Shackelford TK (eds) *Encyclopedia of Personality and Individual Differences*. Cham: Springer International Publishing, pp. 2577–2580.

185. McLachlan GJ, Peel D. *Finite Mixture Models*. John Wiley & Sons, 2004.

186. Baum LE, Petrie T, Soules G, et al. A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. *The Annals of Mathematical Statistics* 1970; 41: 164–171.

187. Lystig TC, Hughes JP. Exact Computation of the Observed Information Matrix for Hidden Markov Models. *Journal of Computational and Graphical Statistics* 2002; 11: 678–689.

188. Akaike H. Information Theory and an Extension of the Maximum Likelihood Principle. In: Parzen E, Tanabe K, Kitagawa G (eds) *Selected Papers of Hirotugu Akaike*. New York, NY: Springer, pp. 199–213.

189. Bozdogan H. Choosing the Number of Component Clusters in the Mixture-Model Using a New Informational Complexity Criterion of the Inverse-Fisher Information Matrix. In: Opitz O, Lausen B, Klar R (eds) *Information and Classification*. Berlin, Heidelberg: Springer, 1993, pp. 40–54.

190. Bozdogan H. Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions. *Psychometrika* 1987; 52: 345–370.

191. Schwarz G. Estimating the Dimension of a Model. *The Annals of Statistics* 1978; 6: 461–464.

192. Sclove SL. Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika* 1987; 52: 333–343.

193. Yu H-T, Park J. Simultaneous Decision on the Number of Latent Clusters and Classes for Multilevel Latent Class Models. *Multivariate Behav Res* 2014; 49: 232–244.

194. Chung H, Anthony JC, Schafer JL. Latent class profile analysis: an application to stage-sequential process in early-onset drinking behaviours. *J R Stat Soc Ser A Stat Soc* 2011; 174: 689–712.

195. Kolko DJ, Kazdin AE. Emotional/behavioral problems in clinic and nonclinic children: correspondence among child, parent and teacher reports. *J Child Psychol Psychiatry* 1993; 34: 991–1006.

196. van der Ende J, Verhulst FC, Tiemeier H. Agreement of informants on emotional and behavioral problems from childhood to adulthood. *Psychol Assess* 2012; 24: 293–300.

197. Rothì DM, Leavey G. Mental Health Help-Seeking and Young People: A Review. *Pastoral Care in Education* 2006; 24: 4–13.

198. O'Farrell P, Wilson C, Shiel G. Teachers' perceptions of the barriers to assessment of mental health in schools with implications for educational policy: A systematic review. *Br J Educ Psychol* 2023; 93: 262–282.

199. Duby Z, Maruping K, Jonas K, et al. "We can't share things with our teachers": Narratives of mistrust and disconnect between South African female learners and their teachers. *Front Educ (Lausanne)* 2022; 7: 10.3389/feduc.2022.882959.

200. Cortes SA. INCREASING TEACHER AWARENESS OF MENTAL HEALTH IN CHILDREN. *Electronic Theses, Projects, and Dissertations*; 1606.

201. Fonseca-Pedrero E, Ortuño-Sierra J, Pérez-Albéniz A. Emotional and behavioural difficulties and prosocial behaviour in adolescents: A latent profile analysis. *Revista de Psiquiatría y Salud Mental (English Edition)* 2020; 13: 202–212.

202. Philipp J, Zeiler M, Waldherr K, et al. Prevalence of emotional and behavioral problems and subthreshold psychiatric disorders in Austrian adolescents and the need for prevention. *Soc Psychiatry Psychiatr Epidemiol* 2018; 53: 1325–1337.

203. Carlén K, Suominen S, Augustine L, et al. Teenagers' mental health problems predict probable mental diagnosis 3 years later among girls, but what about the boys? *Child and Adolescent Psychiatry and Mental Health* 2022; 16: 41.

204. Campbell OLK, Bann D, Patalay P. The gender gap in adolescent mental health: A cross-national investigation of 566,829 adolescents across 73 countries. *SSM - Population Health* 2021; 13: 100742.

205. Gariépy G, Elgar FJ. Trends in Psychological Symptoms among Canadian Adolescents from 2002 to 2014: Gender and Socioeconomic Differences. *Can J Psychiatry* 2016; 61: 797–802.

206. Rousseau C, Drapeau A, Platt R. Living conditions and emotional profiles of Cambodian, Central American, and Québécois youth. *Can J Psychiatry* 2000; 45: 905–911.

207. Rousseau C, Hassan G, Measham T, et al. Prevalence and correlates of conduct disorder and problem behavior in Caribbean and Filipino immigrant adolescents. *Eur Child Adolesc Psychiatry* 2008; 17: 264–273.

208. Beiser M, Hou F, Hyman I, et al. Poverty, family process, and the mental health of immigrant children in Canada. *Am J Public Health* 2002; 92: 220–227.

209. *Bright Futures: Promoting Children and Young People's Mental Health*. Mental Health Foundation, 1999.

210. Sigurdson JF, Undheim AM, Wallander JL, et al. The long-term effects of being bullied or a bully in adolescence on externalizing and internalizing mental health problems in adulthood. *Child and Adolescent Psychiatry and Mental Health* 2015; 9: 42.

211. Nansel TR, Overpeck M, Pilla RS, et al. Bullying Behaviors Among US Youth: Prevalence and Association With Psychosocial Adjustment. *JAMA* 2001; 285: 2094–2100.

212. Ganesan K, Shakoor S, Wertz J, et al. Bullying behaviours and other conduct problems: longitudinal investigation of their independent associations with risk factors and later outcomes. *Soc Psychiatry Psychiatr Epidemiol* 2021; 56: 2041–2052.

Appendix

Table A1: Categorization of continuous scores of SDQ domains

	Three-bound categorization		
	Normal	**Borderline**	**Abnormal**
Self-report			
Emotional difficulties	[0, 5.50)	[5.50, 6.50)	[6.50, 10]
Peer Difficulties	[0, 3.50)	[3.50, 5.50)	[5.50, 10]
Conduct difficulties	[0, 3.50)	[3.50, 4.50)	[4.50, 10]
Hyperactivity difficulties	[0, 5.50)	[5.50, 6.50)	[6.50, 10]
Pro-social difficulties	[0, 4.50)	[4.50, 5.50)	[5.50, 10]
Teacher's report			
Emotional difficulties	[0, 4.50)	[4.50, 5.50)	[5.50, 10]
Peer Difficulties	[0, 3.50)	[3.50, 4.50)	[4.50, 10]
Conduct difficulties	[0, 2.50)	[2.50, 3.50)	[3.50, 10]
Hyperactivity difficulties	[0, 5.50)	[5.50, 6.50)	[6.50, 10]
Pro-social difficulties	[0, 4.50)	[4.50, 5.50)	[5.50, 10]

Table A2: BIC values of LCA models for self and teacher's report

	Number of Classes								
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Self-report									
Profile-independent unrestricted $\boldsymbol{\Sigma_t}$	108346	106963	106191	105977	105502	105591	105452	104981	105073
Profile-dependent unrestricted $\boldsymbol{\Sigma_t}$	108346	92568	81147	72059	65982	62080	60650	56806	55953
Teacher's report									
Profile independent unrestricted $\boldsymbol{\Sigma_t}$	102432	98030	96441	94246	93399	92320	91311	85352	85673
Profile dependent unrestricted $\boldsymbol{\Sigma_t}$	102432	54156	34824	24875	19427	9662	5643	2299	-667