Data-Driven Analytics to Explore Associations between Risk and Protective Factors and School Absenteeism for Secondary School Students

A DISSERTATION

SUBMITTED TO THE FACULTY OF THE FACULTY OF THE

UNIVERSITY OF MINNESOTA

BY

Gunwoo (Knoo) Lee

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Connie W. Delaney, PhD, RN, FAAN, FACMI, FNAP

July 2021

Acknowledgements

I express my deepest gratitude to wonderful people here at the University of Minnesota including faculties, staffs and my colleagues. To Connie W. Delaney, thank you for putting your faith in me and showing support with the guidance along the course of the Ph.D. program. Past three years working with you, learning how to be a nursing informatics researcher was a privilege and a pure joy. To Barbara J. McMorris, thank you for sharing your thoughts and helping me with meticulous directions throughout the overall progress of the program. It was a joy to meet and work with you, I have learned not only to be a good researcher but also to be a provider who knows how to give back and serve those who are in need. To Wendy S. Looman, since the day I nervously asked you to be my committee member, you've been nothing but the most supportive person that I can ask of things that are not just about the research but also how to keep myself on track and focus. To Chih-Lin Chi, thank you for your advice and help along the way of me going through the dissertation. It was with your support and expertise which made this study as how I've wanted it to be. And Erich Kummerfeld, before getting into your one-day course of causal discovery on one cold winter day, I'd never thought I would get into it as how I've actually accomplished it. Love your energy and your support with all the emails we've had, it was quite a ride but also an enjoyable one. Thank you so much.

Dedications

I'd like to dedicate this work to people who've helped me get through this adventurous agony of lonely road called becoming a Ph.D. Kevin, I was a lucky guy to meet you when I first got here. Still can't get my head out of the time we've shared together walking around the ancient Greece with Papa Johns and Costco pizza. Becca and Mary, thank you so much for your help and everything that you've done to make me feel like I belong here. You ladies truly helped me above and beyond and it really meant a lot and still does. Loved the office, loved the virtual band even more. To the three engineering stooges (Jiyong, Ryan, Moonki), thanks for putting up with me throughout the COVID and everything. It's sad that we couldn't enjoy all the sunny days and lakes as much as we'd wanted but I'm sure you guys will make it and get to places where you want to be. Good luck and it was a pleasure. To Garim, I wanted to let you know that without your energy and kindness, meeting those stooges would have had been a bit too much of a jungle ride than a pleasant evening. You know you've balanced things out and I do appreciate it and also you putting things up with us. I'm genuinely happy to see how things are going between you and Moonki. Send me invitation, I'll swim the Pacific if I have to. To all my cohorts (Tipper, Grey, Pamela, Donna, Becka), can't believe how the time just flicks in front of your face and vanishes. Two years of my life I'm not getting back but it was a joy since I've spent it with you. To my friends in Korea (Ray, Hoony, JW), my Ph.D. went with your endless midnight messages about politics, real estates, bitcoin, and all the groundbreaking news that had happened for the past three years. Guys, don't worry and be happy sometimes and thank you for keeping me up at night. That helped, in a way. Last, to my parents and my sister with her family, I'm sorry I couldn’t be there for the past two years. I do come off as a prodigal son (and brother) sometimes but please know that I love you and you mean the world to me. I’ll see you soon.

Abstract

Study Purpose

Chronic absenteeism (CA) is an administrative term defining extreme failure for students to be present at school. CA is recognized as a national problem in the U.S. that has devastating long-term impacts on students. However, in consideration of what counts as students missing school, the partial-day absence (PDA) is inconsistently used across the U.S. as opposed to the full-day absence (FDA). This is because the impact of PDA on student outcomes is less studied due to diverse policies at the local school district level. Applying causal discovery analysis techniques to student-level data, this study analyzed the interconnectivity of partial-day absence and full-day absence by comparing risk and protective factors operationalized by specific student-reported factors were included in the analysis based on Bronfenbrenner's bioecological model of development.

Methods

Using machine learning techniques (i.e. feature selection, prediction model performance comparison) on de-identified student-level data (*n* = 121,005) from the Minnesota Student Survey 2016, factors associated with school absences were identified as the Aim 1. For Aim 2, which was conducting a mixed-methods approach, a focus-group interview with licensed school nurses (LSNs) in Minnesota helped to identify factors associated with CA and how it's different between PDA and FDA in a qualitative perspective. Then a mixed-methods approach utilizing a casual discovery method was conducted using the Minnesota Student Survey 2019 (*n* = 125,375). In the mixed-methods approach, identified factors and knowledge gained from both the quantitative (feature selection and prediction model performance comparison) and qualitative (LSNs

focus-group interview) approaches were used separately and also combined to compare and validate the results during the causal discovery analysis process.

Results

For the Aim 1, a total of 18 risk and protective factors (out of 113) associated with school absences were identified which were within either micro- or mesosystem in the bioecological multisystem. With the results of Aim 1 and LSN focus-group interview, causal discovery analyses were conducted. Findings indicated a) PDA directly affecting FDA, b) PDA shown to be the main linkage between FDA and other school absences surrounding factors (e.g. school engagement, student-teacher relationships), and c) an implication of PDA covering school absence related factors within micro-, meso-, and macrosystem which is wider than that of FDA (i.e. only directly affected by factors within micro- and mesosystem).

Implications

Results suggest PDA's fundamental differences with FDA which calls for recognition of PDA in the field of school absences. This dissertation study also revealed the current impact the LSNs have on students who are missing schools (i.e. assessing the student-in-risk for CA, providing breakfast or space for support) from the focus-group interview with current limitations they have such as low student to school nurse ratio which was also reflected in the data used in quantitative approaches. From these results, future researchers would benefit from differentiating school absences into PDA and FDA as it enables those studies to point out which aspect of school absences they are focusing on. Also, attention to validating what's identified in this study is needed, i.e., Utilizing data from different time periods to replicate the results as the study only served its purpose as an exploratory study of PDA. Locating the data with a) a sufficient amount of LSN features, b) a balanced ratio of factors throughout the hierarchical multisystem (i.e. factors from micro-, meso-, exo-, macrosystem), and c) a definition of CA used which are unexcused and excused absences combined will help to better understand the interconnection of school absences surrounding factors and LSNs.

Table of Contents

List of Tables

List of Figures

Chapter I: The Research Problem

Background and Significance

In the United States, over five million school-age children are chronically absent from school each year (Robert Wood Johnson Foundation, 2016). Chronic absenteeism (CA) is an administrative term defining extreme failure of students to be present in school including both excused and unexcused occasions (Rafa, 2017). The U.S Department of Education (2016a) defines CA as missing at least 15 days or more of a given academic year which is approximately 10% of the school year (Attendance Works and Everyone Graduates Center, 2018). CA is a national problem that can be devastating to a child's education (U.S Department of Education, 2016b). For example, frequent absences have been associated with students engaging in daytime juvenile crime (Henry et al., 2012). CA may also result in long-term impacts to young people's physical and mental health and well-being, as well as substance abuse later in life (Dahlgren & Whitehead, 1993; Gottfried, 2014b; Henry et al., 2012).

Despite the significant impact of school absences on students' long-term life, a preliminary review of research focused on absenteeism showed a lack of consistency of measures and definitions utilizing the term 'absenteeism' which characterizes the field (Cicutto et al., 2013; Rogers & Feller, 2018; Spirito et al., 2018). Further, a review of guidelines on reporting school absences and CA from multiple states resulted in mixed findings regarding what is reported as absenteeism (State of New Jersey Department of Education, 2020; State of Pennsylvania Department of Education, n.d.; South Carolina Department of Education, n.d.). Furthermore, very few studies acknowledge the potential differences between partial and full day absences and the risk and protective factors that might be associated differently when comparing two different types of absences (Whitney & Liu, 2017). According to Whitney and Liu (2017), partial-day absences (e.g. arriving late due to transportation, or arrive late in order to get the lunch) are as prevalent as full-day absences. However, only a few studies recognize this, which lead to studies excluding partial-day absence in the study without valid examination or reference. Therefore, evaluating the context in which absences occur using a wide variety of risk and protective factors (measured as individual, developmental, parental, family, socioeconomic, and community influences) is necessary to thoroughly comprehend the topic of school absences (Teasley, 2004). This is especially important given that the factors leading to school engagement for students are known to promote positive youth development and health outcomes through the school years and into adulthood (Archambault et al., 2019).

Regarding the presence of school nurses in terms of their relations with school absenteeism, school absences are one of the school-setting factors school nurses are closely connected with. The presence of school nurses in school settings demonstrated an impact on school absenteeism, whether it is due to chronic health issues or identifying at-risk students (Maughan, 2003; Rodriguez et al., 2013). In addition, students' behavior of partial-day absence could be associated with school nurses as students need to visit the school nurse office to be excused. However, studies focusing on the dynamics of how the student would be chronically absent in the context of school nurses' role with regards to partial and full-day absence have been scarce.

Previous studies on school absenteeism also suggest that either administrative data from school records or student/parent reports were used to measure absenteeism

(Fornander & Kearney, 2019; Gottfried, 2014b, 2019; Richardson et al., 2018). While studies based on these types of data are acceptable, using an aggregated version of data (i.e. school or district level) could lack the results representing useful information on individual entities (i.e. student-level data derived results). There have been studies utilizing student-level data with large sample sizes. For example, Stempel (2017) conducted a secondary analysis of data from a national survey (58,765 students) examining associations between adverse childhood experiences and school absenteeism. The data were gathered from parents or caregivers via a telephone survey. The study did not examine any additional relations associated with absenteeism other than adversarial childhood experiences. While it is considered as a norm to conduct a study with only a few variables of interest based on empirical evidence or a specific theory, the field of school absenteeism may benefit from research that includes a variety of factors which are known to be associated with absenteeism, especially given a wide variety of risk and protective factors.

To expand our knowledge on school absenteeism in order to address such challenges, conducting a study with a data-driven approach utilizing ‘big data’ may be a solution. Big data deals with a broad range of phenomena focused on the analysis of large data sets. This methodology is being used in areas such as intelligence analytics, behavior and preference modeling, sustainability studies, online and offline commerce, biomedical research and healthcare, and various other forms of scientific and social research (Mittelstadt & Floridi, 2016). Broadly speaking, big data can refer to the process of analyzing big data sets, and the datasets themselves. The term ‘big’ can be defined variably in terms of the sizes or quantities of entries, individuals, or events that are represented by the data. Milton (2017) stated “data are quantitative and used to track and profile behaviors, preferences, and other characteristics of individuals for the purpose of predicting future behavior, and potentially future healthcare decision-making” (p. 300).

Based on current knowledge, no studies to date used data-driven approaches to investigate associations and causal relationships between self-reported measures of both risk and protective factors and school absences in a large student-level dataset collected for state-wide surveillance of the health of young people. This study addresses the gaps mentioned above (i.e. inconsistent standards to school absenteeism, data with limited size and/or dimensions) by utilizing two hypothesis-free data-driven machine learning approaches which are predictive models comparison and causal inference methodology. Using a data-driven approach, this dissertation research will, 1) identify school absences associated factors and their interconnectivity, and 2) establish a model to help predict or support decision making around school absences (e.g. prediction model of frequent absences by schools, an algorithm that helps to evaluate a school’s absenteeism rate from its surrounding factors which helps the decision-making of school to reduce such rate). A prior research has found that an algorithm based on machine learning from the data of 1,962 young people presenting to youth mental service had helped to predict their experience of self-harm in the six months after the first assessment which shows that if the model was trained using the right data, it will help to detect/predict the outcome of interest (Iorfino et al., 2020).

An acquisition of data that represents a variety of risk and protective factors with both partial- and full-day absences is integral to this study. The Minnesota Student Survey (MSS) fulfills such a necessity by implementing student self-reported survey to all types of schools in the state of Minnesota. The survey also asks a series of questions that pertains to the risk and protective factors that associates with school absences (e.g. questions related to school climate, bullying, out-of-school activities, emotional and mental health, relationships, substance usage).

Statement of Purpose

The purpose of this dissertation study is to examine the interconnectivity between risk and protective factors that are linked to school absences among secondary school students.

This dissertation research addresses two specific aims and related research questions.

Aim 1. Identify secondary student-reported factors that are associated with school absence in 2016 using data-driven approaches.

Research question for Aim 1:

1. Which risk and protective factors are associated with school absences of the last 30 days (defined as skipping school/cutting class) among secondary school students in 2016?

Aim 2. Identify factors that distinguish between students' reports of past month partial-day absences and full-day absences in 2019, using a mixed-method approach.

Research questions for Aim 2:

1. What are licensed school nurse's perceptions of chronic absenteeism and the differences between partial and full-day absence?

2. Based on knowledge gained from Aim 1 (quantitative analysis), how are the risk and protective factors associated with partial and full-day absences?
3. Informed by perceptions of licensed school nurses (qualitative analysis), how are the risk and protective factors associated with partial-day absence, different or similar to, full-day absences?
4. How are the risk and protective factors associated with partial and full-day absence based on knowledge gained from Aim 1 (quantitative analysis) with perceptions of licensed school nurses (qualitative analysis) combined and how are they distinguished compared to the Aim 2 - research questions 2 and 3?

This dissertation study addresses the gaps in research by discovering how different types of school absences are associated with risk and protective factors in three ways. First, the study will use data-driven methods to identify risk and protective factors that contribute to overall student absence. Second, the study will utilize a focus-group interview of licensed school nurses to validate the results of data-driven methods. Third, the study will implement a mixed-method approach utilizing causal inference method, incorporating themes saturated from the focus-group interview, with factors identified from data-driven methods to examine how different types of school absences are distinguished.

Chapter II: Review of the Literature

There are a variety of reasons why students miss school originating not only from students themselves but their surrounding environments including, a) absence due to an illness or medical appointment, b) trouble with school assignments or with certain subjects, c) family attitude toward school, d) transportation, e) school disciplinary or suspension, and so on (Gottfried, 2017; Henderson et al., 2014; Melvin et al., 2019; Teasley, 2004; Woodman et al., 2015). Teasley (2004) states that those risk and protective factors can be described in six groups which are a) school factors, b) personal factors, c) developmental factors, d) family and parental factors, e) neighborhood and community factors, and f) ethnic minority status. Kearney (2016) also mentions school absenteeism is covered by a wide spectrum of factors which are individual, parental, familial, and environmental. The objective of this dissertation study is to incorporate those risk and protective factors affecting school absenteeism and conduct an analysis to identify the interconnectivity of school absences and compare how factors relate to each other. It requires a theoretical foundation as a guidance to lead the study mainly for a) selecting and preparing the dataset, b) organizing the results of the analysis, and c) interpreting the results and implications.

This dissertation research, a secondary data analysis utilizing the Minnesota Student Survey (MSS) conducted in 2016 and 2019, examined the relationships between school absences with risk and protective factors among secondary school students. This chapter first describes the theoretical framework which is the foundation of this study. The chapter then highlights empirical evidence supporting each component mentioned in the theoretical framework with its implication in addressing gaps in knowledge.

Theoretical Foundation

The Bioecological Model of Human Development

Bronfenbrenner (1977) states that an understanding of human development is more than observation of a person in a setting, but rather is a multisystemic approach which requires consideration of the components of the environment and the interactions within it. The bioecological model of human development explains the interactions between the individual and the components of the surrounding environment as a proximal process which is essential to human development (Bronfenbrenner & Morris, 2007). For school students, proximal processes are interactions with their friends, family, teachers, and with activities they engage in including extracurriculars (Melvin et al., 2019). The impact of proximal processes varies greatly depending on the person's characteristics, the context where the person is situated, and time affecting the response to such interactions (Bronfenbrenner & Morris, 2007; Melvin et al., 2019).

Bronfenbrenner (1977) also introduced a system of hierarchy when describing the subject and their interactions with the surrounding environment. The hierarchical structure is organized as microsystem (i.e. the environment immediately surrounding the child where the interactions occur such as family and school), mesosystem (i.e. interactions between a child's microsystems such as family's attitude toward school), exosystem (i.e. settings that affect micro and mesosystem but not experienced directly by the child), and macrosystem (i.e. surroundings that affect a child's environment in macro-level including socioeconomic resources, cultural values, government policy). There is also a chronosystem where time is introduced as a component that may influence the response of a child to proximal processes (e.g. school year, cohort effects) as well as macrosystem factors including socioeconomic resources, cultural values, and economic stability (Bronfenbrenner & Morris, 2007).

As described above, the current knowledge shows a variety of factors associated with school absenteeism not only from an individual aspect but also expanding to a number of environmental factors (Kearney, 2016; Melvin et al., 2019; Teasley, 2004). Some researchers studying the field of school absenteeism, being aware of such complexity, have argued for applying a multisystemic approach utilizing the bioecological model (Doren et al., 2014; Gottfried & Gee, 2017; Guralnick, 2017). For example, Doren et al. (2014) utilized a comprehensive set of predictors which are theoretically derived from the bioecological model and associated with school dropout among students with learning disabilities. While the study of Doren shows how the bioecological model can be associated with school absences, applying the theory to school absenteeism leaves a room for an improvement as 1) only a few studies on school absenteeism have implemented such an approach which leaves the interpretation of the theory to authors, and 2) lack of guideline or framework of the bioecological model that focuses itself to school absenteeism. The need for identifying a framework that is tailored to school absenteeism is apparent in order to proceed the study then also to interpret the results in this case.

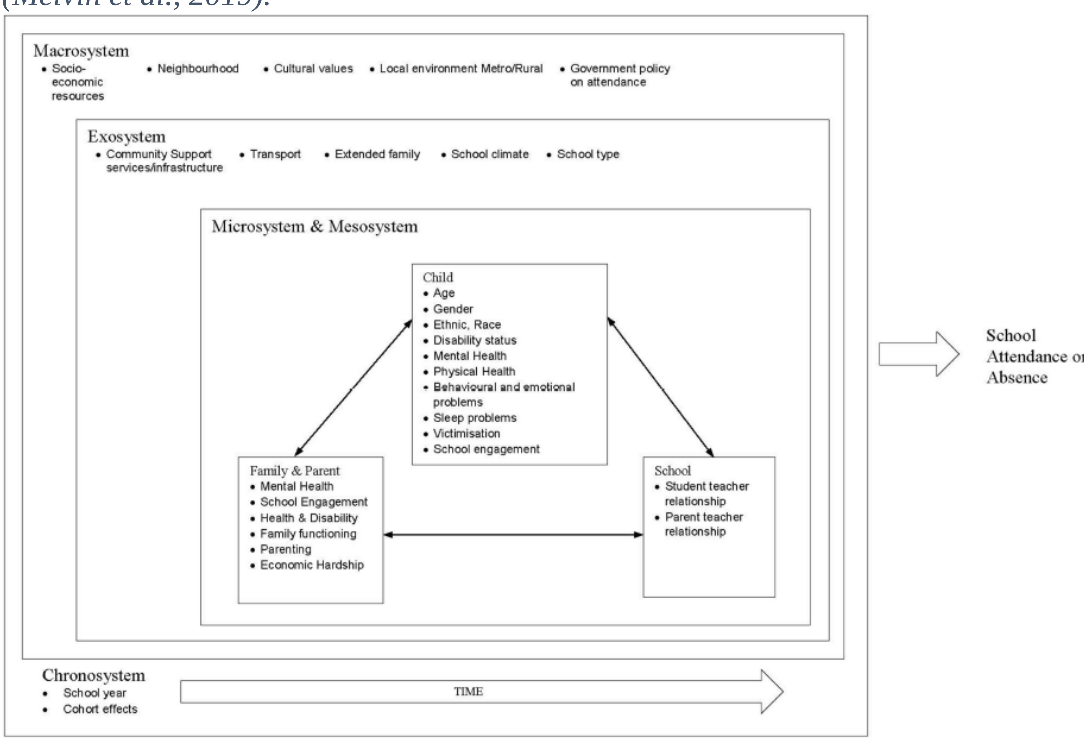
The Kids and Teens at School (KiTeS) Framework

The KiTeS framework is a comprehensive bioecological system approach specifically tailored to school absenteeism. Melvin et al. (2019) acknowledged the components of ‘proximal processes’ and hierarchical systems explained in Bronfenbrenner’s bioecological models, and utilized them while developing a framework tailored to school absenteeism and problems associated with school attendance (Bronfenbrenner & Morris, 2007; Melvin et al, 2019).

As indicated in Figure 2.1, the KiTeS framework centers the characteristics of youth laid out at the micro- and meso-system levels due to the inherent impact to the individual and their surrounding environment. It also locates factors of family, parent, and school at the micro- and meso- levels due to their relevance with youth characteristics in the context of school absenteeism. For the exo- and macro- levels, predictors of absenteeism supported by empirical evidences are grouped by system level, congruent with Bronfenbrenner's bioecological system. The chronosystem shows the influence of time depicted by school year and cohort effects as evidenced by the pattern of absenteeism which changes over time (Melvin et al., 2019).

Figure 2.1.

The KiTeS Bioecological Systems Framework for School Attendance and Absences (Melvin et al., 2019).



Application of KiTeS to this Study

The current dissertation study seeks to identify and compare the inter-connectivity of various risk and protective factors that are associated with school absences. As such, the study inherently requires theoretical guidance in order to organize the dataset and interpret the results of the analysis. Without such a framework, the results of the study could only be confirming a simple “A with B” associations among a number of risk and protective factors. With the purpose of examining interconnectivity among those factors, it is a necessity to proceed with adequate theoretical grounding.

The KiTeS bioecological systems framework provides a theoretical foundation originating from Bronfenbrenner’s bioecological model and supports the study including, a) decision making when choosing variables for the analysis, b) arrangement of interconnectedness using different levels of systems from the framework, and c) interpretation of the results identified from the analysis.

It is important to note that the framework is tailored to ‘all’ student populations which transcends across both disadvantaged and non-disadvantaged groups and this may provide important comprehensiveness when analyzing a variety of school absences related data. Melvin et al. (2019) stated the “KiTeS Framework is applicable to all student populations and fosters examination of the interacting factors that may underlie increased risk for different groups of students… offers a comprehensive context for exploring risk and protective factors to help explain absenteeism” (p. 6). The statement implies the framework’s suitability for the study as it examines the interconnectivity between risk and protective factors of absenteeism utilizing a large sample of data.

Empirical Evidence for Risk and Protective Factors Affecting School Absences

Teasley (2004) posits that risk and protective factors related to absenteeism are categorized into six groups including, a) school factors, b) personal factors, c) developmental factors, d) family and parental factors, e) neighborhood and community factors, and f) ethnic minority status. The KiTeS framework - the theoretical framework applied in this study - originates from the bioecological systems theory focused on proximal interactions among various environmental factors rather than a singular predictor (Bronfenbrenner, 1977; Melvin et al., 2019). Thus, the hierarchical levels described in the KiTeS framework inherently covers all the groups of risk and protective factors mentioned. Therefore, this section aims to provide the current empirical evidence related to absenteeism-related risk and protective factors using four levels of hierarchical systems (i.e. micro & meso-, exo-, macro-, chronosystem) based on the KiTeS framework. The section also provides evidence for the current state of school absenteeism studies based on characteristics including, a) sample size, b) level of data, and c) scope of the study.

Micro and Mesosystems

Child and School Absences. There are numerous child-centered factors that affect school attendance or absence including age, gender, ethnicity, and behavior (Melvin et al., 2019). Regarding age, the U.S Department of Education (2016) conducted a study utilizing nationwide chronic absenteeism data from the 2015-16 Civil Rights Data Collection (CRDC) and discovered that 21.1% (n = 2,982,704) of students in high school are chronically absent, which is the highest rate compared to middle (14.1%, n = 1,333,376) or elementary school students (13.6%, n = 3,115,540). A study utilizing classification and regression tree (CART) analysis for students from the Clark County School District of Nevada also identified a consistent impact of age, grade level, ethnicity, and female gender as a risk factor for problematic school absenteeism (Skedgell & Kearney, 2018). The U.S. Department of Education (2016) posited that in high school, female gender (girls) is slightly more chronically absent than male gender (boys). The results also showed a clear disparity of absenteeism ratio by ethnicity. Specific rates were White 15.7%, Black 21.0%, Hispanic 18.4%, Asian 10.4%, American Indian 24.2%, Pacific Islander 19.6%, and two or more races 18.6% (U.S Department of Education, 2016).

A child's physical and mental health status affects school absenteeism as well. A variety of common health issues including asthma (Basch, 2011; Meng et al., 2012), influenza (Graitcer et al., 2012; King et al., 2012), diabetes (Parent et al., 2009), obesity (Li et al., 2012; Rappaport et al., 2011), dental health (Thikkurissy et al., 2012), and seizure disorders (Aguiar et al., 2007) result in students not being able to attend school consistently. Mental health issues such as depression (Gase et al., 2014) and anxiety (Egger et al., 2003) have been identified as potential leverage points for school absenteeism as well. It has been shown that four DSM-IV related disorders (panic/somatic, depression, conduct problems, hyperactivity) are significantly associated with students' behavior of school absences (Ingul et al., 2012). While the student populations with intellectual/developmental disabilities (IDD) showed a similar association to school absenteeism as well (Black & Zablotsky, 2018; Hancock et al., 2013), other issues including sleep and substance use were also identified as risk factors (Gakh et al., 2019; Gase et al., 2014; Hysing et al., 2015).

Family, Parents, and School Absences. Familial functioning and parent factors (e.g. physical and mental health of parents, parenting style, stress from the parents) are known to be risk and protective factors for school absenteeism (Bahali et al., 2011; Carless et al., 2015; Woodman et al., 2015). For example, Carless et al. (2015) examined the relations between parenting self-efficacy and school-refusal (i.e. unwilling to attend school due to anxiety or emotional distress despite familial efforts) for adolescents aged 12-17 years. The study compared school-refusing students with school-attending students and the results indicated school-refusing students showed a lower level of parental efficacy which was inversely associated with family dysfunction. Bahali et al. (2011) conducted a study assessing psychological symptoms and familial risk factors for parents of students ($n=55$ pairs) who exhibited school refusal compared with a control group. A series of risk factors affecting school refusal initiating from family members were identified as significant including, a) punishment by the parents ($p < .0001$), b) disease history of the parents ($p < .01$), and c) history of mental disorder in the parents or other relatives ($p < 0.03$). In addition, Stempel (2017) examined the relations between CA and the role of adverse childhood experiences (ACE) (i.e. traumatic events in childhood related to abuse, neglect, and family dysfunction) and identified having one or more ACE was significantly associated with CA (i.e. missing more than 15 days per year). Empirical evidence presented here shows the significant impact the family members bring (especially parents) to child's behavior of missing school.

The School Context and School Absences. Factors such as student-teacher relationship, parent-teacher relationship, attitude toward school, and parents' involvement in a child's education were identified to be school absenteeism-related risk and protective factors (Balkis et al., 2016; Doren et al., 2012; Gottfried, 2017; Green et al., 2012). For example, Balkis et al. (2016) conducted the study examining relations between students' personal, family, and academic achievement factors with school absenteeism for high school students in two public schools ($n = 423$). Results of a structural equation model (SEM) suggested that personal factors including attitudes towards teachers and school were negatively associated with school absenteeism ($p < .001$). In a second study, Green et al. (2012) examined three theoretically-driven longitudinal models of academic processes leading to academic performance and the model with the superior heuristic value showed students' attitude toward school negatively affected school absenteeism. In a third study, Cook et al., (2017) conducted a project called the Early Truancy Prevention Project (ETPP) designed to improve attendance of primary school students by facilitating the communication between parents and teachers (e.g. home visits, texts). The results indicated that treatments showed a significant impact on absences among students with 4+ and 6+ absences ($p < .05$). The results here function as empirical evidences by showing how school absenteeism is affected by school-related factors such as teachers and attitudes toward school.

Exosystem

An exosystem includes the settings that affect components mentioned in the micro and mesosystem but not directly linked to the child (Bronfenbrenner, 1977; Melvin et al., 2019). A number of settings were identified to be risk and protective factors of school absences including school climate, school start times, school type, and classroom setting (Bowers & Moyer, 2017; Gottfried, 2017; Gottfried et al. 2019; Hendron & Kearney, 2016; Lenhoff & Pogodzinski, 2018; Van Eck et al., 2017). Also, the infrastructure of transportation for students was pointed out to be a protective factor as not having an adequate level of infrastructure in term of transportation would risk the students to be present at school (Gottfried, 2017). It is important to note that these factors are different with school factors shown in the micro- and mesosystem as these are components of settings or infrastructure that are not experienced directly by child (Melvin et al., 2019).

The U.S Department of Education (as cited in Van Eck et al., 2017) conceptualizes the school climate as “domains of safety, engagement, and environment, which encompass constructs such as perception of safety, incidents of delinquent or aggressive behavior, school connectedness, relationships with teachers, parental involvement, school resources, and perceptions of the physical and learning environment” (p. 91). Hendron and Kearney (2016) examined the relationship between absenteeism severity and school climate variables (i.e. 42-item scale from the School Climate Survey) among youth with problematic attendance (n = 398). Further, the study examined relations between school climate factors and school absenteeism using SEM for youths with problematic attendance (n=398). The authors used a SEM model to demonstrate that the school climate was inversely related to the severity of absenteeism.

Bowers and Moyer (2017) conducted a meta-analysis study examining relations between school start times and students’ sleep duration and attendance. The results indicated later school start times improved students’ sleep duration and students’ attendance. Lenhoff and Pogodzinski (2018) implemented an exploratory study that examined the relations between school organizational effectiveness and CA. The findings indicated that school organizational effectiveness focusing on five key “essential” areas including, a) effective leadership, b) collaborative teachers, c) ambitious instruction, d) supportive environment, and e) involved families affected school absenteeism for public schools ($n = 90$) in Detroit excluding charter schools ($n = 75$). Regarding transportation, Gottfried (2017) found that students who took the bus to school experiences fewer days absent (a decrease of 0.39 days, $p <.01$), with corresponding 3% decrease in CA ($p<0.001$). The studies presented in this section acts as empirical evidence that show how the factors within exosystem (e.g. transport, school climate, community support and infrastructure) are associated with school absences.

Macrosystem

School absence risk and protective factors in a macrosystem include the surroundings that affect a child's environment including socioeconomic resources, cultural values, and government policy (Bronfenbrenner, 1977; Melvin et al., 2019). Macrolevel factors including education policy, socio-economic status of the family such as housing instability, employment, education level, as well as neighborhood characteristics are identified to be risk and protective factors for school absenteeism (Balkis et al., 2016; Childs & Lofton, 2021; Deck, 2016; Gottfried, 2014a; Ingul et al., 2012).

Childs and Lofton (2021) points out how education policies have been established without a consideration of school absences and thus, current policies are not acting as a solution but more of a distraction for solving the issue. Deck (2016) conducted a quasi-experimental study comparing school outcomes (school mobility, school attendance, academic achievements) among three groups of children who are situated within different degrees of homelessness (sheltered, doubled-up, poor but housed). Sheltered students showed significant high levels of school mobility ($p < .05$) and low levels of school attendance ($p < .05$) compared with other two groups. Balkis et al. (2012) utilized SEM to examine the relations among factors including absenteeism, personal and family factors, and academic achievement. The SEM model showed family factors affecting absenteeism, specifically higher education levels for both parents and higher incomes of the family (i.e. socio-economic status) showed negative association with absenteeism. Another study also identified significant differences between three groups of youth (no absence, normal absence, high absence) in terms of both parents education level with mother's employment status ($p < .01$) (Ingul et al., 2012). The results from these studies show how the factors within macrosystem (e.g. socio-economic status, home mobility, education policies) are associated with school absences.

Chronosystem

In chronosystem, 'age' is used in the course of the biological development that represents the continuity of time (Simpkins et al., 2012). Prevalence of absenteeism differs by student age (Skedgell & Kearney, 2018; U.S Department of Education, 2016). For instance, transportation to school changes as students take the bus, are picked up by family members, or start driving themselves to school. Relatedly, Last and Strauss's (as cited in Melvin et al. 2019) study on school refusal in anxiety-disordered students also suggests the impact of separation anxiety from parents on missing school may be more influential for the younger students.

Current State of School Absenteeism Studies

This section, as mentioned, provides evidence for the current state of school absenteeism studies based on characteristics including, a) sample size, b) level of data, and c) scope of the study. Acquiring a large number of samples (i.e. Big data) is inherently beneficial in data-driven research. Also, focusing on sample size with the level of data is mainly to assess past studies' data characteristics based on the perspective of machine learning approach. Assessing sample size and the dimensions of data used in the past studies provides the contrast between the past approaches from big data analytic approach. It is not to discuss the whether the characteristics of sample was adequate for studies reviewed but rather to assume the potential impact of data-driven approach in the field, which is the approach of this dissertation.

For the studies that examined intervention programs for students dealing with school absences, Eklund et al. (2020) conducted a meta-analysis of evidence-based interventions for addressing CA. For eight between-group randomized controlled trial design studies, sample size ranged from 27 to 500 students except for one. Bottini (2017) conducted a study examining the influence of a classroom-wide Student Success Skills program for a total of 2,175 students. For nine between-group quasi-experimental design studies, the study sample ranged from 66 to 1,278 except for one study which examined the physical security measure in school settings including metal detectors and security cameras on school attendance to a total of 38,707 students (Tanner-Smith & Fisher, 2016).

Studies utilizing secondary data inherently showed wider variability in terms of sample size, level, and scope of the study. For instance, Doren et al. (2014) conducted a study identifying salient predictors of school dropout for 11,000 high-school students receiving special education. The study included demographics, individual risk, and family and school factors examining the relevance with school absence. Freeman et al.

(2015) utilized publicly available state-wide school-level data from 600-800 high schools across seven years (2005-2011) from the Positive Behavior Interventions and Supports (PBIS) datacenter. Implementation of PBIS with fidelity was used as an independent factor to measure student attendance. Gase et al. (2014) utilized 915 samples of 90-minute face-to-face interviews. The study included environment, social and individual influences to measure relevance with truancy.

Gottfried et al. (2019) compared chronic absenteeism between students with and without disabilities utilizing a sample that included 654,736 students across 37,867 classrooms and 1,148 public elementary schools. The study incorporated a number of measures including race/ethnicity, gender, age, English proficiency, and free/reduced lunch. For students with disabilities, indicators from the Individual with Disabilities Educations Act (IDEA), 13 disability classifications and special education settings were included in the study. Skedgell and Kearney (2018) conducted a study examining multiple-levels of predictors associated with school absenteeism for 316,004 youth from the Clark County School District (CCSD) in Nevada during the 2015-2016 school year. The study identified multiple risk factors associated with school absenteeism among a series of academic (grade level, GPA, individualized education plan eligibility, participation in school sports) and demographic variables (youth age, gender, ethnicity).

Summary and Conclusions

The KiTeS framework based on the theory of Bronfenbrenner's bioecological systems provides an excellent foundation for the current dissertation study (Bronfenbrenner, 1977; Melvin et al. 2019). In particular, this research benefits from the concept of 'proximal processes' or the interaction across hierarchical systems described in the framework. This theoretical foundation contributes to the study specifically related to, a) decision making when choosing variables for the analysis, b) arrangement of interconnectedness using different level of systems from the framework, and c) interpretation of the results identified from the analysis.

Empirical literature reveals diverse risk and protective factors associated with school absences in multiple hierarchical levels with significance (i.e. microsystem, mesosystem, exosystem, macrosystem, and chronosystem). The review provides strong evidence to support the claim that the current school absenteeism risk and protective factors cannot be examined by a singular determinant and thus, a comprehensive multi-aspect approach is needed.

Addressing Gaps in Knowledge

A review on the current studies of school absences showed a wide spectrum of scholarship with regards to design, sample size, and variables used for the analyses. Further, it showed that there has been a variety of approaches applied to examine school absenteeism from intervention studies (RCT, quasi-) to secondary data analyses. Sample size varied from less than 100 to 600,000 students. Some studies used a broader scope to identify predictors of school absenteeism originating not only from the child or parent but also from the environments (Gottfried, 2014a, 2017, 2019; Skedgell & Kearney, 2018,). However, the review suggests that no studies as of yet utilized a full-comprehensive scope of incorporating the multiple layers explained in the KiTeS framework to a large dataset at the student-level. Conducting a study that incorporates such scope is integral to the area of school absenteeism as the phenomena cannot be explained nor understood by a few determinants. Also, a study that sought to incorporate multiple aspects of the phenomena would benefit the field. In data-driven research, a dataset that helps to fulfill a certain level of generalizability (size of data) and capacity to pinpoint the problem of interest (individualized data) is essential in order to present the results with a certain degree of competency. To summarize, studies utilizing a large dataset or incorporating a comprehensive scope were identified, but not a study with both aspects.

These gaps were addressed by current dissertation study. First, the study utilized a dataset that fulfills both large size and detailed level of information (i.e. student level) that pertains to the interest of the study. The individualized level of information (i.e. student-level) in the dataset helped to proceed the study and interpret the results with the capacity of capturing dynamics of school absences with their associated factors surrounding 'students'.

Second, this dissertation research utilized a theoretical foundation to incorporate various risk and protective factors associated with school absenteeism. In particular, the study utilized the KiTeS framework, a theoretical framework based on bioecological systems model, was utilized because it specifically focuses on school absences (Bronfenbrenner, 1977; Melvin et al. 2019). The dataset showed a capacity of including a wide range of student characteristics which helped to capture the factors described in the framework and the review of empirical literature. By addressing these two gaps specifically, this data-driven research study examined a comprehensive list of factors that are associated with school absences.

Chapter III: Methods

This study, a secondary analysis of data from the Minnesota Student Surveys from 2016 and 2019, examined the interconnectivity between risk and protective factors for school absences among secondary school students. The chapter also includes details of the method to address each aim (e.g. study design, variables used in the study, and analytic plan).

Overview of the Minnesota Student Survey

The Minnesota Student Survey (MSS) is a statewide surveillance system that invites all types of schools to participate, including regular public schools, charter schools, tribal schools, nonpublic schools, alternative learning centers, and juvenile correctional facilities. It has been conducted every three years since 1989 with students across Minnesota. Students in grades six, nine, and twelve had been participating until 2010. After 2010, the participating grades were changed to grades five, eight, nine, and eleven. The survey asks a series of questions including components of school climate, bullying, out-of-school activities, health and nutrition, emotional and mental health, relationships, substance use, school absences, and more. All responses from students are de-identified (Minnesota Department of Education, n.d.).

The MSS 2016 was administered in the first half of 2016 to students in grades five, eight, nine, and eleven, statewide. Of the 330 public operating districts, 282 agreed to participate (85% of public operating school districts). Out of 300 variables available from the survey on 168,733 de-identified student-level data, the data specific to eighth, ninth, and eleventh-grade students were used for Aim 1 of this study. Data of fifth-grade students was excluded as there are significant differences between primary and secondary school students which would impact the interpretation of the results if the study includes a wide range of ages from both primary and secondary schools.

The MSS 2019 was conducted in the spring of 2019 and released in October of 2019. A total of 81% school districts in Minnesota participated, with most districts having all designated grades participate (Minnesota Department of Health, 2019). Out of 383 variables from the survey to 170,128 de-identified student-level data, the data specific to eighth, ninth, and eleventh-grade students were used for Aim 2 of this study. Fifth-grade students were excluded for an identical reason of Aim 1.

Study Design Overview

This was retrospective observational study using two separate secondary data sets (MSS 2016 and MSS 2019) examined the interconnectivity between risk and protective factors associated with school absences for secondary school students. Aim 1 focused on identifying risk and protective factors contributing to the last month's school absences using data from MSS 2016. The Aim 2 centered on a mixed method study of examining the interconnectivity between school absences (i.e. partial-day and full-day absences) with risk and protective factors. The study utilized the design of sequential explanatory mixed methods to integrate both quantitative (i.e. causal inference analysis using data from the MSS 2019) and qualitative components (i.e. focus group interviews with school nurses) of the study (Wilkins & Woodgate, 2008).

Methodology and Analysis for Aim 1: Identify Secondary Student-reported Factors that are Associated with School Absence in 2016 using Data-driven Approaches

Aim 1 Design

This was a secondary data analysis utilizing the MSS 2016 to identify the risk and protective factors that are associated with school absence. Additionally, the student-PI investigated the combination of risk and protective factors that best predict school absence using three different lists of risk and protective factors identified in the study.

Aim 1 Sample

From the MSS 2016 (n = 168,733), the study sample was limited to students who were attending secondary school at the time of data collection. This is because there are significant differences between primary and secondary school students which would impact the interpretation of the results if the study included a wide range of ages from both primary and secondary schools.

Of the MSS 2016 sample, 126,868 participants were included, excluding the 5th graders (n= 41,865). Of these participants, 4,667 were missing data regarding race and ethnicity and thus were eliminated. After removing the race and ethnicity missing data, 268 participants were missing the “biological sex” data and were removed as well. The intention of omitting those participants with missing data was to correctly analyze the sample’s descriptive characteristics before the analysis. Finally, 928 samples with missing values in the data of either full-day or partial-day absence were removed. With the eliminations described above, this analysis utilized student-level data with a sample size of 121,005 participants.

Variables and Measurement for Aim 1

Outcome variable – Unexcused School Absence. The survey used the questions ‘During the last 30 days, how many times have you skipped school or cut classes but NOT a full day of school, without being excused?’ and ‘During the last 30 days, how many times have you skipped school or cut a FULL day of school, without being excused?’ to measure school absences. For the purpose of capturing all the factors associated with school absence in Aim 1, the student PI combined responses from two questions asked from the 2016 MSS. Both school absences survey questions were coded into binary variables with the value of 0 and 1 then summed together to have value of 0, 1, and 2. Then those values were dichotomized by converting the value 0 into “Never”, and the value of 1 and 2 into “Any.”

Independent variables. From the MSS 2016, eligible variables were selected using the KiTeS framework (Melvin et al. 2019). As noted in Chapter 2, a total of four hierarchical systems mentioned in the KiTeS framework were used including micro-, meso-, exo-, and macrosystems. Chronosystem was not included in this Aim 1 analysis as the data used in the study are cross-sectional and thus do not include the component of ‘time’ (e.g. having data from the same student but at a different point in time).

Microsystem. According to the KiTeS framework, the environments where the proximal processes occur are those immediately surrounding a child, or the microsystem (Melvin et al. 2019). In order to differentiate the factors categorized as either micro- or mesosystem clearly, this dissertation study regarded the risk and protective factors categorized as ‘factors within the microsystem’ that pertain to the child, and the child only. Such factors include, for example, a child’s age, biological sex, race/ethnicity, substance use, and so on. The detailed list of factors are described in Table A1 at Appendix A.

Mesosystem. The system surrounding the microsystems of the child is defined as mesosystem. It is comprised of interactions that converse between the microsystems surrounding the child (Melvin et al. 2019). This study categorized the risk and protective factors as 'factors within the mesosystem' as those that describe the interaction between the child and the immediate surrounding systems such as family and school. Such factors used in this study include a child's perception of parent caring, adversarial childhood experiences, a child being sent out of class due to a disciplinary issue, and so on. More details are described in Table A2 at Appendix A.

Exosystem. The system surrounding that which encompasses both micro- and mesosystems, but not directly influencing the child, is defined as the exosystem (Melvin et al. 2019). Examples of the exosystem include transportation, classroom setting, school type, and school organizational factors. More details are described in Table A3 at Appendix A.

Macrosystem. According to Bronfenbrenner (1977), a macrosystem is defined as "a system that differs in a fundamental way from the preceding forms in that it refers not to the specific contexts affecting the life of a particular person but to general prototypes, existing in the culture or subculture" (p. 515). Therefore, the macrosystem in this study covers broader cultural and institutional norms with socio-economic resources that affect the systems covered by macrosystem (e.g. cultural values, government policy, neighborhood). The detailed list of risk and protective factors within macrosystem are described in Table A4 at Appendix A.

Data Preprocessing

Analytic Plan. The analysis plan for the Aim 1 is described below. Data preparation including data preprocessing and transformation was performed using R version 3.5.0 – an open source language for statistical computing and graphics. For implementing data analysis including feature selection, the Waikato Environment for Knowledge Analysis (Weka) developed by a machine learning group at the University of Waikato in New Zealand was used.

Data Imputation. The Classification and Regression Trees (CART) imputation method was used for imputing the missing data. CART originates from the method of Multiple Imputation through Chained Equations (MICE). According to Van Buuren (2018), CART models seek predictors and cut points in the predictors that are used to split the sample. The cut points divide the sample into more homogeneous subsamples. The splitting process is repeated on both subsamples, so that a series of splits defines a binary tree. The target variable can be discrete or continuous (p. 82). This method imputes the data by using the homogenous subsamples information divided by cut points in the predictors and replace the missing data to a value that are most likely to be explained by regression trees originated from the original data before the imputation.

As CART methods are robust against outliers, can deal with multicollinearity, and flexible enough to fit nonlinear interactions, it shows a number of properties which makes the method ideal for imputation (Burgette and Reiter, 2010). CART does not require any assumptions for underlying variable distribution which is also ideal for imputing the raw data. In addition, CART imputation is ideal for this study specifically as it excels in dealing with large datasets. It is known to perform well when dealing with ordinal data compared with other multiple imputation methods (Wongkamthong & Akande, 2020). This Aim 1 analysis utilized CART with the parameter of five imputation. Also, due to the matrix size of the study, with the hardware capacity to handle the multicollinearity imputation method, five times of iteration was selected. An imputation number of five times is regarded as normal when using CART to impute data. Burgette and Reiter (2010) discovered that an iteration of five times was acceptable compared with 10 iterations or more. The package 'mice 3.13.0' in R 3.5.0 was used for the CART imputation process.

Created Scales using the MSS 2016 variables. The data were prepared with the mixture of both ordinal and continuous data. As mentioned in Carifio and Perla (2007), Likert scales are not only ordinal but also can be treated as interval scales. The survey questions similar to Likert scale were treated as either ordinal or interval in the analysis. This allowed for the summary statistics including frequency distribution; mode, median, and mean; standard deviation, and variance of a dataset (Hillier, 2021). A number of scales were made in order to incorporate the risk and protective factors based on the KiTeS framework. For example, a total of five questions asking about students' school engagement from Appleton et al. (2006) (e.g. How often do you care about doing well in school?, How often do you pay attention in class?) answered in four items ranging from 'None of the time - 1' to 'All of the time - 4' were summed up, followed by a mean calculation which resulted in a composite scale that ranged from 1 to 4. The Cronbach's alpha values were calculated to ensure the consistency and reliability of created scales.

Also, answers that would not necessarily provide additional information were dichotomized. For example, the question 'When was the last time you saw a doctor or nurse for a check-up or physical exam when you were not sick or injured?' was answered by 4 items (1 – During the last year, 2 – Between 1 and 2 years ago, 3 – More than 2 years ago, 4 – Never). As the difference between students visiting a doctor or nurse from the last year to more than two years ago would not make a significant difference for this study, the study dichotomized the question from 'Never – 0' to 'Any -1'. The details of how data were coded with additional descriptions are in the Tables A1, A2, A3, and A4.

Data Standardization and Class Imbalance. Standardization is seldomly done in a pre-processing stage to equal the weight distributed to all variables. This is especially true when the data being used has different range of measures such as the MSS data. For example, when the variable A has a range of 1 to 5 and B with 1 to 100, the weight of the variable B will be immensely higher than that of the variable A due to its numeric value regardless of how those variables are measured. In these cases, standardizing data improves machine learning's performance by reducing overall variance throughout the variables used for the analysis (Dy & Brodley, 2004). As this study is using the MSS data that comes with diverse range of measures which could be interpreted as various weights, all the variables were standardized before the implementation of the analysis. The z-score standardization was used which is obtained by subtracting the value by the mean of the variable and dividing the result by the variable's standard deviation.

When the difference between majority and minority classes are significant (i.e. those who are absent vs. those who aren't), implementing a machine learning analysis with feature selection could lead to acquiring questionable results mainly due to a bias that comes from imbalanced data. For example, if you have a data with majority and minority ratio of 95 : 5, having a 95% accuracy prediction model wouldn't necessarily mean you have a good performing model as you will simply achieve 95% accuracy by just classifying all data as majority. This would mean that using imbalanced data to train the prediction model would make the measure accuracy meaningless and potentially acquire an oversimplified model that classifies most cases as majority (Fernández et al., 2018). In other words, when training the prediction model, a skewed outcome variable could train the model to result in false positives (i.e. high accuracy but not befitted to the real-world data).

Due to the potential caveats mentioned, the analysis could benefit from using resampling techniques to complement the class imbalance in the outcome variable in the case of evaluation/training of prediction model. And thus, the student-PI prepared two datasets which are 1. Original dataset without any resampling techniques, 2. Imbalance-countered dataset using synthetic minority oversampling technique (SMOTE) and compared their performances when evaluating the performance of prediction models (See Table 3.1 that illustrates the difference of with and without SMOTE). SMOTE merged datapoints from the existing pool of the 'school absence' class and added them to the dataset which created new 'minority' instances. In this case, 'minority' is equal to students who had high-tendency of missing school for a partial or full-day in the MSS 2016. The method ensures the minimum amount of data leakage when training the model by additionally creating new datapoints (Fernández, 2018). The SMOTE enables both 'up-sampling' and 'down-sampling' for the 'minority' instances. And both methods were used to up-sample the minority instances (i.e. high tendency of being absent from the school for either partial or a full-day) while down-sampling the majority instances (i.e. low tendency of being absent from school for either partial or a full-day). The outcome instances ratio for SMOTE sample was 1.7:1 (n = 91,560 : 53,410) compared to the original data ratio of 15:1 (n = 113375 : 7630).

Table 3.1

Dataset Preprocessed with and without SMOTE

Original Data	Resampled data using SMOTE
Standardized data without any resampling techniques with an outcome variable ratio of 15:1	Standardized data with a resampling technique (SMOTE) resulting in an outcome variable ratio of 1.7:1

Feature Selection

This study utilizes the process of 'feature selection,' which is widely used in the field of machine learning analytics. In the past few decades, the dimensionality of data has increased exponentially which results in 'overfitting' the model. When the model is overfitted, it means that the model is too tailored to dimensions shown in training data which results in subpar performance with real-world data. Therefore, the field of data-driven research had to pre-process the data by eliminating of the dimensions (i.e. variables) that were least likely to affect the outcome variable in order to train the model to perform well with the real-world data (Tang et al., 2014).

Feature selection methods are generally divided into two categories, supervised and unsupervised. Supervised feature selection methods fall into two broad categories, which are the filter model and the wrapper model. The filter model relies on the general characteristics of the training data to select some features without involving any learning algorithm (Yu & Liu, 2003). The wrapper model requires one predetermined learning algorithm in feature selection and uses its performance to evaluate and determine which features are selected. As for each new subset of features, the wrapper model needs to learn a hypothesis. It tends to find features better suited to the predetermined learning algorithm resulting in superior learning performance, but it also tends to be more computationally expensive than the filter model (Langley, 1994). Unsupervised machine learning techniques generally focus on jobs such as clustering, representation learning, and density estimation that does not require explicit labels provided from user. For the exploratory purpose of this study with labelling the input and output before the analysis, supervised feature selection method was used.

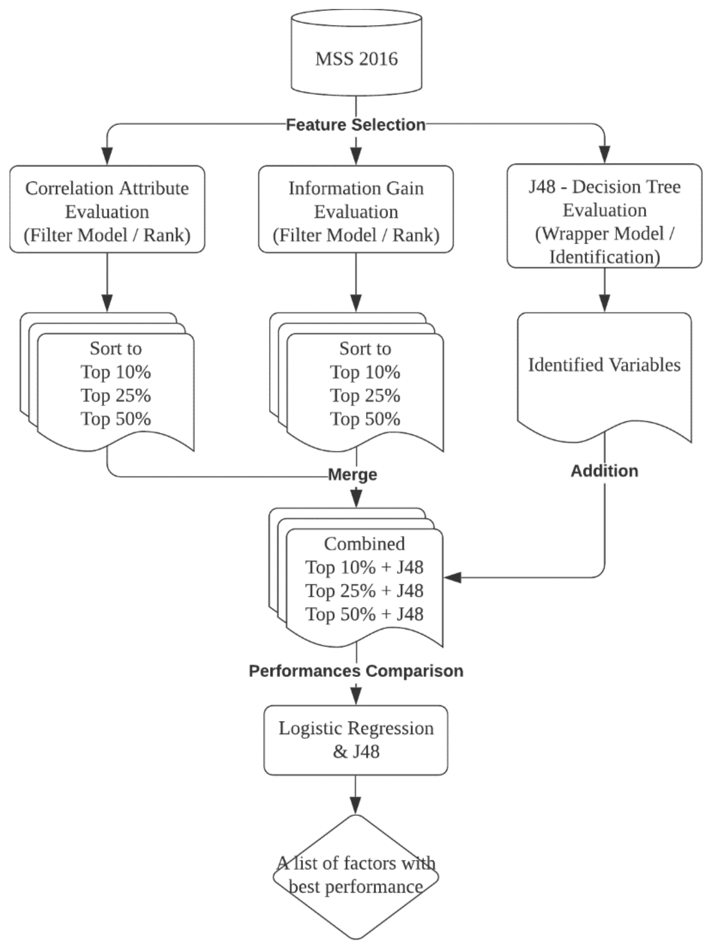
To identify factors that are associated with school absence in Aim 1, this dissertation used both filter models (correlation attribute evaluation, information gain attribute evaluation) and a wrapper model (J48 – decision tree method). The filter methods used are univariate statistical measures that calculate one input variable at a time with the outcome variable. This means that any interaction between input variables are not considered in the filtering process. This feature is complemented by the wrapper model. The wrapper model utilizes predetermined algorithms to reflect the general aspects of variables combined, which are labelled as non-linear. Wrapper method used the 'best first' method as a search method with the search direction of 'forward' terminating after five node expansions. By applying two steps of feature selection, the study was able to identify a complete list of school-absences-associated factors. All feature selection methods used incorporated a 10-fold cross-validation to ensure the quality of the results. 10-fold cross validation is regarded to work ideally to the similar size of samples used in this analysis regardless of the computational capacity (Kohavi, 1995).

Evaluation of Identified Factors

The Figure 3.1 below explains the flow of how the results from the feature selection methods are separated then used for the evaluation.

Figure 3.1

Evaluation Flow of the Aim 1.



The study was intended to acquire information regarding the factors that are most likely to be associated with students missing school, then relay that knowledge to the Aim 2. It is important to note that the filter models used for the study rank 'all' the independent variables as a result opposed to identifying a certain number of variables that associate with an outcome variable. A list of factors (i.e. X number of factors from 113 variables used in the study) that are most likely to be associated with school absences is required for Aim 2. And therefore, the results of feature selection methods (i.e. ranked lists including all 113 variables) need to be transformed into a validated list of factors that best represent the association with an outcome. As such, the student-PI prepared subsets of data from the results of the three feature selection methods used, then compared the performances of prediction models while feeding those subsets as data. A comparison of such performances helped to evaluate which list of factors best represents the outcome of interest (i.e. school absences).

Subsets from the Filter Models. The filter models used for the study rank all the independent variables used in the feature selection. The results let us know which variables are more likely to be associated with school absences. However, the result does not show by what portion (e.g. top 10%, top 50%) of those important factors would best represent the outcome of interest. Therefore, the study uses three cutoff values (top 10%, 25%, 50%) to compare three separate lists to measure which list best predict the outcome and thus, be most informative. As this study aims to gain the knowledge from both filter models, the three cutoff values were used to each of the ranked list. Then those lists from two different filter models (top 10%, 25%, 50% from correlation attribute evaluation and top 10%, 25%, and 50% from information gain evaluation) were grouped and merged by the same cutoff value. For example, the top 10% list from correlation attribute evaluation and the top 10% list from information gain evaluation were merged to acquire the list that represents the top 10% list of both filter models. With such a process, the study acquired three different lists of factors that represent the top 10%, 25%, and 50% ranked lists of both filter models.

A Subset from the Wrapper Model. As opposed to the filter models, the J48 used in this study returns the identified attributes among the variables and yields the most informative gains when used to explain (classify) the outcome variable. In addition, it also returns how many times the attributes repeatedly appeared during the 10-cross fold validation process. For example, 'out-of-school suspension' variables are identified as very informative when classifying the variable of 'school absence.' This is because the J48 model identified the 'out-of-school suspension' attribute to be informative when classifying the outcome variable in eight out of 10 times (80%) during the 10-fold cross validation process. Variables repeatedly appeared more than once in the J48 model, were all selected, and separated into four different subsets (all variables repeated for more than twice; three times; four times; and five times). The four subsets were then evaluated using logistic regression and J48 decision tree prediction modelling to see which performed the best. For J48, the confidence threshold of 0.25 was set for pruning with the minimum of two instances per leaf. The parameter of Accuracy, f-score (weighted average) and Area under a Curve (AUC) were used to assess the performance of prediction rates of each model (Hossin and Sulaiman, 2015).

Combining the Subsets. A subset with the best performance from the wrapper model was added to all three lists of factors acquired from two filter models (i.e. top 10%, 25%, and 50% of ranked list from both filter models), as the wrapper method was conducted in the study to complement the linearity characteristics the filter models. When there was a variable that repeatedly appeared in any of the lists from three models, it was counted in the resulting list as one variable (i.e. no duplication involved). This is because the aim of combining the subsets here was to prepare the list of variables for the performance comparison which does not involve the element of duplication.

Performance Comparison. To identify the list that best predicts the outcome variable (i.e. school absence) among three lists of factors, logistic regression and J48 decision tree prediction model methods were used. For J48, the confidence threshold of 0.25 was set for pruning with the minimum of two instances per leaf. The parameter of Accuracy, f-score (weighted average) and Area under a Curve (AUC) were used to assess the performance of prediction rates of each model.

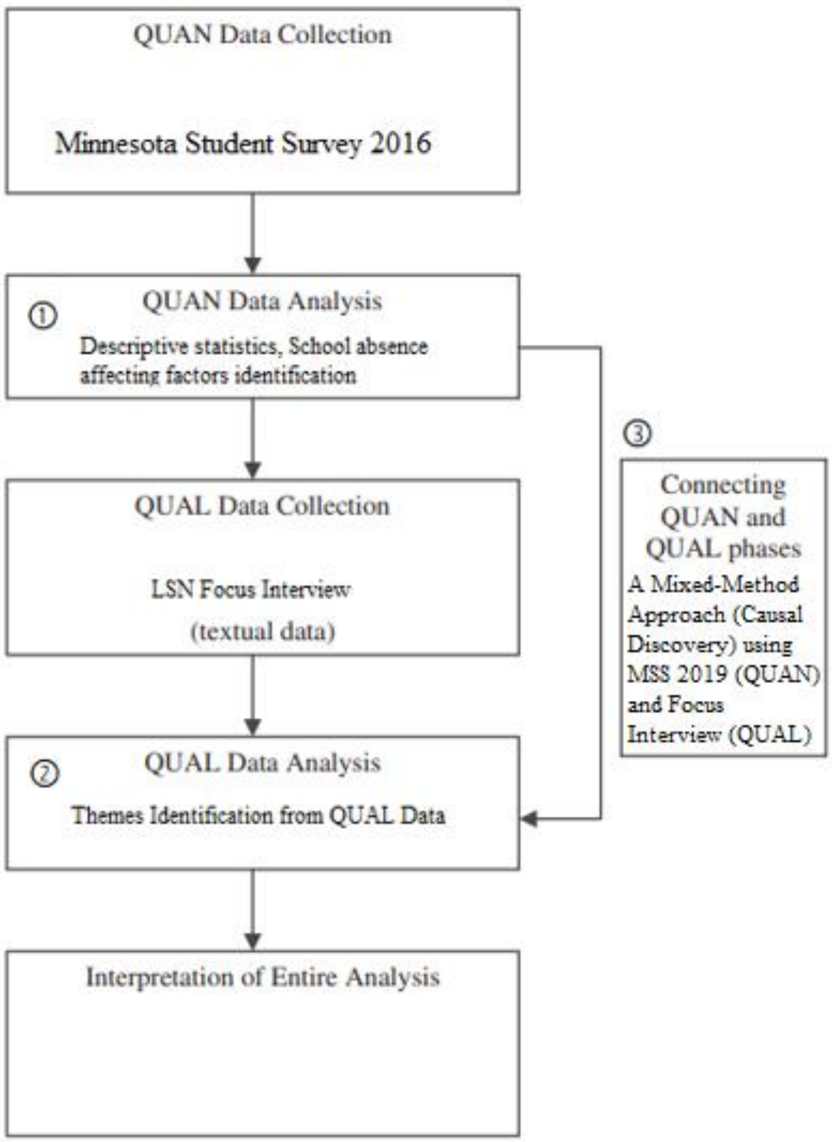
Human Subject Protection, Data Management and Security

This secondary data analysis utilizing the MSS 2016 was approved by the Institutional Review Board (IRB) at the University of Minnesota (See Appendix C1). The data of the MSS 2016 was uploaded and operated in the secured platform ‘Box’ provided by the University of Minnesota. The principal investigator for this study did not have access to any MSS 2016 participant identifiers.

Methodology and Analyses for Aim 2: Examine the Relations of School Absences and Associated Risk and Protective Factors by Identifying Factors that Distinguish Students' Reports of Past Month Partial-day Absence and Full-day Absence in 2019, using a Mixed-method Approach

Figure 3.2.

Visual Model for Sequential Explanatory Mixed Methods Design including Aim 1 and 2 of this Study



Aim 2 Design

The Aim 2 portion of this dissertation study is comprised of two phases, 1) A focus group interview with licensed school nurses (qualitative design), and 2) Three-steps of causal discovery analysis. This study incorporated the concept and flow of Sequential Explanatory Mixed Methods Design (Wilkins & Woodgate, 2008). Figure 3.2 describes the flow of this mixed-method design including the Aims 1 and 2. The figure starts with depicting the Aim 1 of the study which is noted as ①, followed by *Phase 1* (i.e. ②) and *Phase 2* (i.e. ③) of the Aim 2. *Phase 1* explains the qualitative data collection with analysis and *Phase 2* continues the study utilizing the causal discovery method. Utilizing causal discovery method enables the study to find the causal relations using the identified factors from Aim 1 in addition to qualitative approach (focus group interview) conducted in the Aim 2.

Phase 1: Focus Group Interviews with License School Nurses

The qualitative phase focused on capturing the voices of professionals with expertise in school absenteeism. Licensed School Nurses (LSN) are an integral part of the school. LSNs work closely with students with school absence behaviors (Jacobsen et al., 2016; Weismuller et al., 2007). The data collected from LSNs were analyzed by the primary investigator of the study and Dr. Camille Brown who is a nurse and a qualitative methodologist. The focus group interviews were conducted as a part of a larger project, the Minnesota Youth Trading Sex (MYST) project. The focus groups substudy lasted from June through November of 2020 and the primary investigator of this study participated in the MYST project as a study team member. The implementation of focus interview was reviewed and approved by the University of Minnesota IRB (see Appendix C3).

Recruitment and Sample. In order to recruit LSNs, we initiated the recruiting process by partnering with the professional organization ‘School Nurses of Minnesota’ (SNOM). A series of three emails each containing a link to the eligibility screener were sent to potential participants using the listserv from SNOM. Required criteria for the participant eligibility included, 1) participants must have practiced as an LSN in the state of Minnesota in the past two years, and 2) participants must have worked with either middle or high school aged students in this role. After the recruitment, participating LSNs were divided into a total of six groups (three to five participants each) by region and school type. For example, we conducted a focus group with LSN participants who were from Greater Minnesota and another group of participants from Alternative Learning Centers. As a result, a total of 21 LSNs participated.

Two participants (9.5%) from 21 candidates did not return the questionnaire used for initial demographics survey. This left 19 participants. Out of 19, only one participant identified as a ‘male’. 18 out of 19 participants checked ‘White or Caucasian’ when asked to identify their ‘Ethnicity and Race.’ One participant declined to answer this question. Given the high percentage of Caucasians in the state Minnesota (83.8%, according to the United States Census Bureau, 2019), a high percentage on Caucasians would be considered a norm in this case as well. Regarding education, 12 participants had a bachelor’s degree (63%) compared to seven participants with a master’s degree (37%). The distribution of school location among the participants was more even across the participants, ‘urban’ (n = 6; 32%), ‘suburban’ (n = 6; 32%), and ‘rural areas’ (n = 6;

32%). Finally, age range of students LSNs worked with were elementary school (n =7; 37%), middle school (n =4; 21%), and high school (n =7; 37%). One participant had mentioned she worked with adult populations. The LSNs experience with students' age aligns with the aim of this study (absenteeism for secondary school students) as more than half of the participants (58%) had experience with students from secondary schools.

Data Collection.

Meeting Preparation. As the focus groups were conducted during the COVID-19 global pandemic, we utilized the software Zoom – a telecommunications platform – to conduct six focus group interviews with LSNs over the course of a month. The focus groups were scheduled with at least two to three days between each one and no more than two focus groups were scheduled in a week. When a date for a focus group was confirmed, we sent a confirmation email with the meeting time/date, a copy of the consent form for review, a link to a demographics form, and instructions on how to use Zoom. A reminder email was sent to participants a day before the focus group. Fifteen minutes prior to the meeting, an email with directions on how to connect to the focus group was sent to participants.

Interview Process. Two study team members facilitated the focus group interviews. Team Member One (Dr. Camille Brown) facilitated the focus group interviews leading the participants with prepared questions and probes, while Team Member Two (Emily Singerhouse, Coordinator of the project MYST) was the contact point for technical aid outside of the space (for phone calls and emails). All of the focus group interviews were recorded. Team Member One reviewed consent forms with the group as all participants verbally consented prior to data collection. Verbal consents were documented and recorded. The focus groups began with general introductions of each participant to build trust and comfort, then questions regarding CA, partial-day and full-day absences and youth trading sex were asked in presented order (See Appendix D). The interviews ranged from 50 to 80 minutes. After the interviews were complete, the participants were informed of next steps, and reminded to fill out the demographics form if they had not already done so. The team members had a short debrief meeting to compare notes including observations, unusual events, and to plan any necessary follow-up with the participants. LSNs who participated were provided with a $50 gift card as compensation. The funding was supported from 'Sophia Grant Award – School of Nursing, University of Minnesota' that was awarded to the student-PI of this dissertation.

Data Analysis. All of the focus group interviews were transcribed by the student PI utilizing the Zoom closed caption program. Then transcriptions were reviewed and discussed with the second study member to ensure accuracy and correct any errors.

After reviewing the transcription, the student-PI prepared the structural codes for the content analysis; which are structurally organized labels that describe condensed meaning of interview contents (Erlingsson and Brysiewicz, 2017). The structural codes were then reviewed and discussed with the nursing Ph.D. (Dr. Camille Brown – Team Member One) to ensure that the codes were well organized and captured the topic of interest from the transcripts. As mentioned above, the Aim 2 analysis sought to understand how LSNs perceives chronic absenteeism, and how they perceive the different types of school absence including full and partial-day absence. In order to answer those questions, structural codes were categorized into three broad groups which were, 1) chronic absenteeism, 2) partial-day absence, and 3) full-day absence. In each group, we prepared the codes to capture three main areas of interest which were, a) factors influencing the ‘specific type of absence,’ b) LSN role in supporting children with the ‘specific type of absence,’ and c) barriers and facilitators to support the children with ‘the specific type of absence.’ For example, in the partial-day absence group, all codes were prepared under three main areas of interest (factors influencing partial-day absences, LSN role in supporting partial-day absent children, and barriers and facilitators to supporting partial-day absent children).

To capture the factors which transcends the level of student or the family, the study utilized the KiTeS framework which describes factors that affect children from multi-level hierarchical systems (Melvin et al., 2019). For example, the structural codes of the ‘factors’ not only included the level of student and the family which are within micro- and meso-systems (e.g. Student’s family intra-familial relationships, student mental health antecedent factors), but also social determinant factors known to be within macro- and exo-systems (e.g. special education involvement, housing stability, race or ethnicity) (Bronfenbrenner, 1977; Melvin et al., 2019). This bioecological concept was also utilized when preparing the structural codes within the group ‘barriers and facilitators supporting the specific type of school absence.’ Barriers and facilitators to providing care were divided into four levels which were, 1) individual LSN, 2) student or family, 3) school level, and 4) systemic level. The intention here was to capture the barriers and facilitators which are dispersed throughout the environments surrounding the student while focusing on an individual LSN as well. The detailed list of structural codes with description are in Appendix E.

Using NVivo 1.0 (March, 2020), transcripts were uploaded then two study members (student-PI, Dr. Brown) inductively coded the transcripts based on the structural codes prepared. Thematic inductive analysis was conducted based on methodology for thematic analysis from Braun and Clarke (2006). The student-PI met with the nursing Ph.D. (Dr. Camille Brown) to discuss any discrepancies found during the analysis and reached consensus if there were any disagreements in the process of inductive coding. After the discussions, emerging themes with exemplar quotes were identified which the two researchers agreed upon.

Reflexivity. In qualitative analysis, researcher reflexivity is useful because it can "not only increase the creditability of the findings but also [deepen] our understanding of the work" (Dodgson, 2019, p. 220).

The student-PI identifies as a foreign student who had spent most of his life in the Republic of Korea. He realizes that his capacity for understanding cultural and systemic aspects in the U.S. context could be limited. For example, a question such as "How are LSNs operating within schools and how are those functions different when comparing urban and rural areas?" is something the student-PI only can learn from reading or by having conversations with someone who has been through the education systems in the U.S. Dr. Brown's qualitative expertise and experience as a nurse in U.S. schools helps to round-out the research team's qualifications to conduct this qualitative portion of the dissertation.

The student-PI's cultural and racial background helps him to better understand the meaning of being a 'minority', and also the gap which exists between being 'minority' and 'majority.' Even though the student-PI (identifies as East-Asian and male) categorizes himself as a minority in the state of Minnesota and U.S, it is essential to note that he's been living most of his life within the category of being in the 'majority' for over 30 years in the Republic of Korea. Republic of Korea is demographically homogenous with only 5% of population being foreign born ("Demographics of South Korea," 2021). The interviews with LSNs regarding school absence had components that relate to cultural differences among 'minority' students who reside in suburban and rural regions of the state Minnesota. Therefore, the perspectives the student-PI has are beneficial to the interpretation of finding from the focus groups.

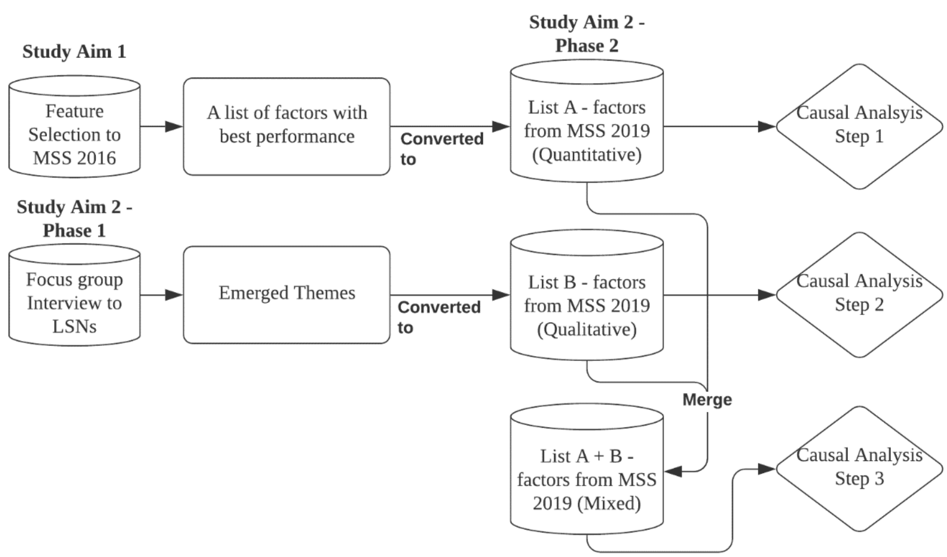
Phase 2: Causal Discovery

Study Design. Phase 2 of the study connected both quantitative and qualitative components of Aim 2. This phase was the last part of the Sequential Explanatory Mixed Methods Design right before the interpretation of the entire analysis (see the Figure 3.2). Phase 2 consisted of three steps while utilizing the causal discovery methods to conduct the analyses. Step 1 only used the list 'A' - factors resulted from Aim 1 (quantitative analysis), whereas Step 2 utilized the list 'B' - factors identified from emerged themes of the LSN focus-group interviews (qualitative analysis). Step 3 used a combined list of 'A' and 'B' then the causal discovery analysis was conducted. Having three separate steps of analyses helped to 1) acquire additional analytic results in order to compare and validate gained knowledge, and 2) benefit from having a diverse of scope (e.g. enables to recoup what could have been overlooked by utilizing only quantitative or qualitative approaches). The study converted results from both the Aim 1 (quantitative) and the focus group analysis (qualitative) into the list of factors in the MSS 2019 to acquire the list 'A' and 'B.' This is because the *Phase 2* utilized a different dataset - MSS 2019 -

compared to that of the Aim 1 or the *Phase 1* of the Aim 2. Utilizing different datasets was essential as the knowledge resulting from them validated the robustness of the knowledge gained from data-driven analytics. Figure 3.3 below depicts the detailed flow of *Phase 2* – three-step analysis.

Figure 3.3

Detailed Flowchart of the Three Step Analysis



Causal Discovery. While utilizing ‘big data’ for data-driven research, we used causal inference (in particular, causal discovery) method to infer how factors were interconnected with each other in the study. Causal inference is defined as inferring the causal relations from data based on the supporting assumptions. Based on the unclearness of causal relations, using a directed graph to represent the causal structure is deemed as effective. Eberhardt (2017) describes the causal discovery method as models showing

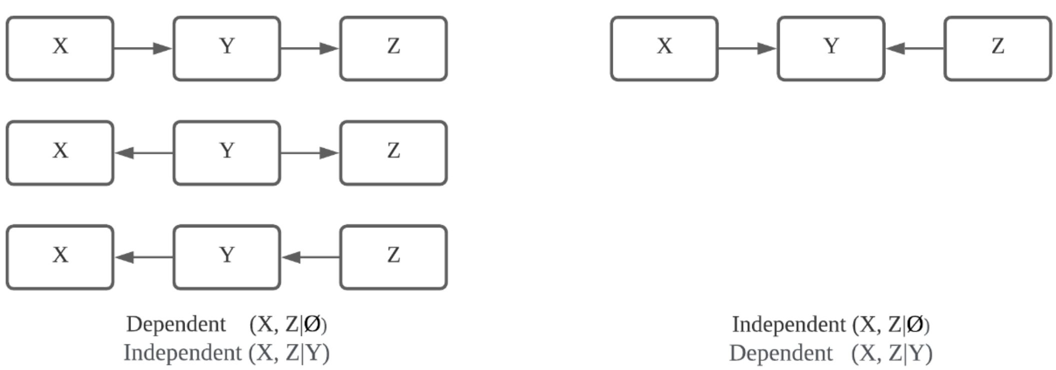
> the relationships among a number of variables: For a given set of variables V {X1, . . . Xn}, a causal graph G = {V, E} represents the causal relations over the

set of variables V, in the sense that for any directed edge e = Xi → Xj in E, Xi is a direct cause of Xj relative to variables in V. (p. 82)

There are assumptions in causal discovery that need to be met before implementing the methodology. The assumption of Causal Markov (every vertex X in the graph G is probabilistically independent of its non-descendants given its parents) and Causal faithfulness (if a variable X is independent of Y given a conditioning set C in the probability distribution P(V), then X is d-separated from Y given C in the graph G) together establish that d-separation correlates to probabilistic independence (Eberhardt, 2017). With these assumptions, the causal discovery method enables inferring causal relations which help to better understand how factors of interest are interconnected with surrounding factors.

Figure 3.4

Conditional Independence Test Graph

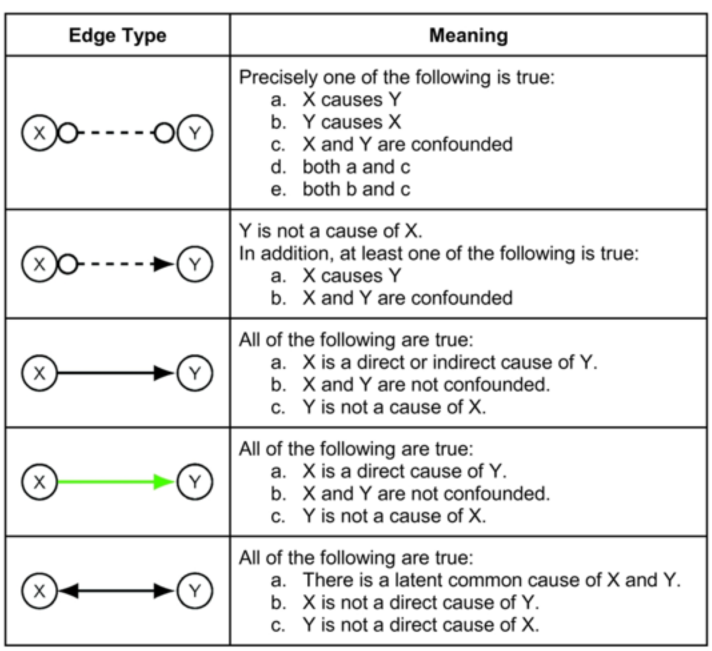


Among the methods of causal discovery, this study utilized the Greedy Fast Causal Inference (GFCI) method in using the platform TETRAD 6.7.0 for all three steps of Phase 2. GFCI proceeds the analysis with a two-step process (first, a preliminary assessment of causal relations among factors via space search of penalized likelihood score; and second, refinement of the preliminary search utilizing a series of conditional indepedence tests), which enables the results to be presented in a partial ancestral graph (Anker et al., 2018). Figure 3.4 explains how conditional independence tests are conducted. By conditioning on the node 'Y' in the middle and assessing whether X and Z are independent, one can differentiate the collider graph (X and Z directing to Y) from the other three graphs. See Anker et al. (2018) for more detailed information on GFCI.

Partial Ancestral Graph (PAG) is the representation of data flow using edges with arrowheads pointed from the node A to B mainly used in the field of Markov modelling. The type of vertex with arrowheads varies from directed, bi-directed, and undirected edge. Detailed definitions of edges are described in Figure 3.5 below (Kummerfeld et al., 2019).

Figure 3.5

Edge Types in a PAG (Kummerfeld et al., 2019)



Edge Type	Meaning
X o- - - -o Y	Precisely one of the following is true: a. X causes Y b. Y causes X c. X and Y are confounded d. both a and c e. both b and c
X o- - - -> Y	Y is not a cause of X. In addition, at least one of the following is true: a. X causes Y b. X and Y are confounded
X ——> Y	All of the following are true: a. X is a direct or indirect cause of Y. b. X and Y are not confounded. c. Y is not a cause of X.
X ——> Y	All of the following are true: a. X is a direct cause of Y. b. X and Y are not confounded. c. Y is not a cause of X.
X <——> Y	All of the following are true: a. There is a latent common cause of X and Y. b. X is not a direct cause of Y. c. Y is not a direct cause of X.

In order to verify the model stability, bootstrapping (1,000 repetitions) was used. Only the edges identified in a majority of the bootstrapping process were included in the results. As the sample size over 100,000 inferred causal relations in between almost every node used in the analysis, the value of effect sizes was used as a cut-off value to filter only the edges with a significant effect. Using the cut-off value was inevitable as without it, analyzing all the edges connected with each other is not viable. Penalty discount of 5 was used after assessing the number of edges between each node testing a range of penalty discount values. A default value of penalty discount used in GFCI is 4 ("Greedy Fast Causal Inference Algorithm for Continuous Variables", 2016). Usage of higher penalty discount '5' when scaling up the analysis due to over 100,000 samples is supported by another study as it used penalty discount of 8 for 1,000,000 samples and 4 for other samples between the size of 10,000 to 500,000 (Ramsey, 2015). Default values were used for all the other parameters while using GFCI. Effect sizes were found by fitting a structural equation model based on the causal relations inferred from GFCI. The package 'lavaan 0.6-8' in R 3.5.0 was used for the effect size calculations.

Step 1

Sample. The MSS 2019 was used for Phase 2 of the analysis. Out of 383 variables from the survey to 170,128 de-identified student-level data, the data specific to 8th, 9th, and 11th grade students were used for Aim 2 of the study. A total of 125,375 participants were included after excluding the 5th graders (n= 44,753). There were questions revised or removed compared to the MSS 2016 (e.g. a series of questions asking about the specific reasons for students missing school was added, questions regarding school nurse office / medical doctor visits were removed) as it is the updated version, however the similarity of the formatting and the majority of the questions between the two makes the comparison feasible.

PDA – First Outcome Variable. Unlike the MSS 2016, the MSS 2019 used the question 'During the last 30 days, how many times did you miss part of a day of school such as coming late, leaving early or missing class time during the day? (Do not include school-sponsored activities like field trips, sports, academic or music events)' for partial-day absence. The question operationalizes partial-day absences (i.e. multiple forms of PDA which correctly represents inconsistent usage of PDA measurement) with direction for readers to exclude 'excused absences'. It was answered using 5-options which were 'None', 'Once or twice,' '3 to 5 times,' '6 to 9 times,' or '10 or more times.' The responses 'None' and 'Once or twice' were dichotomized as 'not-PDA = 0' and the rest as 'PDA = 1'.

FDA – Second Outcome Variable. Similar to the partial-day absence question mentioned above, the MSS 2019 used the question 'During the last 30 days, how many times did you miss a full day of school? (Do not include school-sponsored activities like field trips, sports, academic or music events)' for full-day absence. The question correctly addresses what should be regarded as full-day absence (i.e. missing a full day of school excluding excused absences). It was answered using 5-options which included 'None,' 'Once or twice,' '3 to 5 times,' '6 to 9 times,' or '10 or more times'. 'None' and 'Once or twice' responses were dichotomized as 'not-FDA = 0' and the rest as 'FDA = 1'.

Independent Variables. As mentioned above, this study utilized a quantitative analysis (Aim 1) identifying a list of factors that are associated with school absences, in order to utilize the MSS 2019 based on the knowledge gained from the Aim 1 (i.e. feature selection / prediction modelling analysis using the MSS 2016). A list of 18 factors that best predict the outcome of school absence in Aim 1 was identified. Those selected variables in the list were then converted to corresponding variables existing in the MSS 2019. For example, Aim 1 identified the ‘school engagement scale’ to be one of the factors that best represents the outcome of school absences. The school engagement scale in the MSS 2016 were based on six questions and/or statements (“How often do you care about doing well in school?”, “How often do you pay attention in class?”, “How often do you go to class unprepared?”, “If something interests me, I try to learn more about it.”, “I think things I learn in school are useful.”, “Being a student is one of the most important parts of who I am.”). The study then identified the exact six corresponding items for questions in the MSS 2019 and used them to acquire the ‘school engagement (SE) scale.’ Details of how the variables were prepared is in Table A5 at Appendix A.

Data Preprocessing. Similar to Aim 1, CART was used to handle the missing data in this first step of Aim 2 quantitative analyses. Due to the size of the matrix and the hardware capacity to handle the multicollinearity imputation method, 5 times of iteration was selected. The package ‘mice 3.13.0’ in R 3.5.0 was used for the CART imputation process. Data were then standardized (z-score standardization) using the platform TETRAD 6.7.0 for the analysis.

Causal Discovery Analysis. As mentioned above, GFCI was used for the analysis. Pre-existing knowledge (i.e. race and ethnicity cannot be caused by other factors, PDA and FDA cannot cause reasons of why students missed school) was set before the analysis to avoid possible confusion in the interpretation process. In order to verify the model stability, bootstrapping (1,000 repetitions) was used (See Table A6 in Appendix A for frequency of each edge appearing in the bootstraps). Only the edges identified in a majority of a bootstrapping process were included in the results. An effect size value of 0.1 was initially used as a cut-off value to filter only the edges with significant effect. Then the effect size value of 0.05 was used as a cut-off value to capture detailed differences of causal relations between PDA and FDA. Effect sizes were found by fitting a structural equation model based on the causal relations inferred from GFCI. The package 'lavaan 0.6-8' in R 3.5.0 was used for the effect size calculations. The hierarchical multisystem framework based on the KiTeS study was used to categorize factors used in the analysis in order to understand the dynamics of school absence associated with a bioecological approach (Bronfenbrenner, 1977; Melvin et al., 2019).

Additionally, the results identified from the MSS 2019 were compared with the causal discovery analysis done for the MSS 2016 for the purposes of 'findings validation.' In other words, the identical causal discovery implementation used in this study was done for the MSS 2016 as well (identical parameters), then it was compared with the results from the MSS 2019 (i.e. Step 1 of Phase 2) for validation purposes. Two datasets were not identical in terms of questionnaire but similar enough for this validation process to be feasible. This extra step of validation is recommended in the field of data-driven research including causal discovery analysis to measure the consistency of results that lead to validity of the study itself.

Step 2

Study Sample. The MSS 2019 was used for Step 2 - Phase 2 of the analysis which is identical to Step 1. Details of the study sample are described at Step 1.

PDA – First Outcome Variable. Unlike the MSS 2016, the MSS 2019 used the question 'During the last 30 days, how many times did you miss part of a day of school such as coming late, leaving early or missing class time during the day? (Do not include school-sponsored activities like field trips, sports, academic or music events)' for partial-day absence. The question conveys what the survey considered as partial-day absence (i.e. multiple forms of PDA which correctly represents inconsistent usage of PDA measurement) with directions for readers to exclude 'excused absences'. It was answered with five options including 'None.' 'Once or twice,' '3 to 5 times,' '6 to 9 times,' and '10 or more times.' The responses 'None' and 'Once or twice' were dichotomized as 'not-PDA = 0' and the rest as 'PDA = 1.'

FDA – Second Outcome Variable. Similar with the partial-day absence question mentioned above, the MSS 2019 used the question 'During the last 30 days, how many times did you miss a full day of school? (Do not include school-sponsored activities like field trips, sports, academic or music events)' for full-day absence. The question correctly addresses what should be regarded as full-day absence (i.e. missing a full day of school excluding excused absences). It could be answered with five options including 'None,' 'Once or twice,' '3 to 5 times,' '6 to 9 times,' or '10 or more times'. 'None' and 'Once or twice' were dichotomized as 'not-FDA = 0' and the rest as 'FDA = 1.'

Independent Variables. As mentioned previously, this study utilized the knowledge gained from the *Phase 2* of Aim 2 – a qualitative analysis of LSN focus group interviews about chronic absenteeism and the differences between FDA and PDA. In order to utilize the MSS 2019 based on the knowledge gained from the focus group interviews, four emerging themes were identified and utilized, 1) Messages about mentally-ill children and unsupportive families, 2) Impact of parents' attitude and behavior toward school attendance, 3) Role of school nurse in students' school absence, and 4) Role of family and social system when caring school absences. Specified factors that associate with school absences from these themes (i.e. free-reduced lunch, mental issues, sleep issues, transportation, having a job, taking care of family, housing stability, behind in school work, perception of caring from parents) were then converted to corresponding variables existing in the MSS 2019. For example, the focus-group interviews identified 'messages about mentally-ill children and unsupportive families' to be one of the themes that best represents the current status of school absences according to the LSNs' perspectives. In order to capture the issue of mental illness of school children, the study identified questionnaire "Have you ever been treated for a mental health, emotional or behavioral problem?" with the responses options of "No," "Yes, during the last year," and "Yes, more than a year ago." from the MSS 2019. A scale was created and used that assessed whether the student had the experience of suffering from mental illness from never to any. In addition, the questionnaire asked "What are the reasons you missed a full or part of a day of school in the last 30 days?" The response options were "Felt very sad," " hopeless," "anxious," "stressed," or "angry." The question pinpoints the crossroads of school absences with mental illness. Details of how the variables were prepared are in Table A7 at Appendix A.

Data Preprocessing. Similar to Aim 1, CART was used to handle the missing data. Due to the matrix size of the study with the hardware capacity to handle the multicollinearity imputation method, 5 times of iteration was selected. The package 'mice

3.13.0’ in R 3.5.0 was used for the CART imputation process. Data were then standardized (z-score standardization) using the platform TETRAD 6.7.0.

Causal Discovery Analysis. As mentioned above, GFCI was used for the analysis. Pre-existing knowledge (i.e. race and ethnicity cannot be caused by other factors, PDA and FDA cannot cause reasons of why students missed school) was set before the analysis to avoid possible confusion in the interpretation process. Bootstrapping (1,000 repetitions) was used to assess and maintain the graph stability (See Table A8 in the Appendix A for frequency of each edge appearing in the bootstraps). Only the edges identified in a majority of the bootstrapping processes were included in the results. An effect size value of 0.1 was initially used as a cut-off to filter only the edges with significant effects. Then the effect size value of 0.05 was used as a cut-off value to capture detailed differences of causal relations between PDA and FDA. The package ‘lavaan 0.6-8’ in R 3.5.0 was used for the effect size calculations. The hierarchical multisystem framework based on the KiTeS study were used to categorize factors to better understand the dynamics of school absence associated with the bioecological approach (Bronfenbrenner, 1977; Melvin et al., 2019).

Additionally, a ‘validation’ process using the MSS 2016 was conducted. In other words, the causal discovery implementation with the same algorithm used in this step was done for the MSS 2016 as well, then compared with the results from the MSS 2019 (i.e. Step 2 of Phase 2) for the validation purposes. This extra step of validation is recommended to measure the consistency of the results which lead to validity of the study itself.

Step 3

Sample. Identical to Step 1 and 2, the MSS 2019 was used for Phase 2 of the analysis.

Outcome Variables. Identical to Step 1 and 2, a dichotomized version of PDA and FDA measurement was used as outcome variables.

Independent Variables. As mentioned in the 'study design' section of the Phase 2, Step 3 included a causal discovery analysis using data gained from the both quantitative (i.e. Aim 1 – feature selection analysis) and qualitative (i.e. focus group interview) components. Therefore, this study established a list of factors merging variables used in Step 1 with Step 2. For example, the quantitative analysis revealed students' use of tobacco products associated with school absences, whereas the qualitative analysis identified that students' mental illness was greatly associated with school absences. For the purpose of the causal discovery analysis or Step 3, both mental-illness and tobacco products use are included in the list of independent variables. Detailed descriptions of independent variables are in Tables A5 and A7 at Appendix A.

Data Pre-processing. The imputation method of CART was used to handle the missing data. Due to the matrix size and the hardware capacity to handle the multicollinearity imputation method, 5 times of iteration was selected. The package 'mice 3.13.0' in R 3.5.0 was used for the CART imputation process. Data were then standardized (z-score standardization) using the platform TETRAD 6.7.0.

Causal Discovery Analysis. As mentioned previously, GFCI was used for the analysis. For the graph stability, bootstrapping (1,000 repetitions) was used. Pre-existing knowledge (i.e. race and ethnicity cannot be caused by other factors, PDA and FDA cannot cause reasons for why students missed school) was set before the analysis to avoid possible confusion in the interpretation process. Only the edges identified in a majority of a bootstrapping processes were included in the results (See Table A9 in Appendix A for frequency of each edge appearing in the bootstraps). An effect size value of 0.1 was initially used as a cut-off to filter only the edges with significant effects. Then the effect size value of 0.07 was used as a cut-off value to capture detailed differences of causal relations between PDA and FDA. The package ‘lavaan 0.6-8’ in R 3.5.0 was used for the effect size calculations. The hierarchical multisystem framework based on the KiTeS study were used to categorize factors to better understand the dynamics of school absence associated with the bioecological approach (Bronfenbrenner, 1977; Melvin et al., 2019).

Aligning with the approaches of Step 1 and Step 2, the results identified from the MSS 2019 were compared with the causal discovery analysis done for the MSS 2016 for the purposes of validating the findings. This extra step of validation is recommended in the field of data-driven research including causal discovery analysis to measure the consistency of the results that lead to validity of the study itself.

Human Subject Protection, Data Management and Security

Both the focus groups interviews with LSNs and the secondary data analysis utilizing the MSS 2016 and MSS 2019 were approved by the Institutional Review Board (IRB) at the University of Minnesota (See Appendix C). The data from the focus group interviews, MSS 2016, and MSS 2019 were uploaded and operated in the secured platform ‘Box’ provided by the University of Minnesota. The principal investigator for this study did not have access to any participant identification.

Chapter IV: Results

This study, a secondary data analysis based on the MSS 2016 and 2019, examines the interconnectivity between school absences risk and protective factors among secondary school students. The purpose of this dissertation study was to examine the interconnectivity between risk and protective factors that are linked to school absences among secondary school students.

Aim 1: Identify secondary student-reported factors that are associated with school absences in 2016 utilizing data-driven approaches.

Sample Characteristics

Participants The study included 121,005 participants for the analysis; students in Grades 8, 9, and 11 participated in the MSS 2016. The percentage of eighth (*n* = 42,791; 35.4%), ninth (*n* = 43,246; 35.7%), and 11th graders (*n* = 34,968; 28.9%) were roughly similar across all three grades. The sample was evenly divided between students who identified as female and male (M = 50.4%, F = 49.6%). Participants were primarily White followed by multiple race groups, Black, Hispanic, Asian, and American Indian or Alaskan Native. The participants who were Native Hawaiian or Pacific Islanders were compared with other races even though they comprised less than 1% (*n* = 555; 0.5%), as these students are known to miss school more than students of any other race (U.S. Department of Education, 2016b). Additional demographic characteristics of the study (race or ethnicity, free/reduced-cost lunch, and school region) with outcomes and primary independent variables (i.e. variables identified to be highly associated with an outcome variable from the results of Aim 1) are summarized in Table B1 at Appendix B.

Feature Selection

Correlation and information gain attribute evaluation were used for the filter method and J48 (i.e. decision tree) attribute evaluation was used for the wrapper method. All three methods were applied to the identical list of independent variables from the 2016 MSS in order to acquire the results that represents both linearity and non-linearity of the feature selection approach.

Correlation Attribute Evaluation The feature selection method "correlation attribute evaluation" assesses the worth of an attribute by measuring the correlation (Pearson's) between independent factors and the dependent factor (i.e. unexcused school absences). This attribute evaluator evaluated each attribute's correlation (i.e. merit) independently of the others for each fold, sums these values, then calculate the average merit by dividing the summation by 10 (i.e. number of the folds). The method ordered the 113 independent variables utilized for the analysis by average merit and average rank. The merit and rank in the results are described as averages, as the study utilized the 10-fold cross-validation process. The ranked results indicated that the majority of the top factors were comprised of attributes either from the microsystem or the mesosystem, based on the KiTeS framework (Melvin et al., 2019). When evaluating the top 10% of the factors (n = 12), 58% (n = 7) were from the microsystem, followed by 42% (n = 6) from the mesosystem. Factors regarding substance usage (tobacco product usage, alcohol consumption, marijuana usage, and other substance usage) and the social competency scale (SCS) were found to be the most relevant to students missing school among factors within the microsystem. For factors within the mesosystem, school disciplinary issues (e.g. in- and out-of-school suspension or office visits due to disciplinary issues), the school engagement scale, and friends' approval of substance use were the most relevant to school absences. A factor from the exosystem (out-of-school activity: sports) first appeared in the list at the 58th position and then the rest were distributed behind the 58th position. The majority of the factors from the macrosystem were found to rank higher than factors within the exosystem. The factor "free and reduced cost lunch" placed 28th (the highest rank among the factors in the macrosystem), followed by "neighborhood safety" (41st), "skipping meals due to financial issues" (50th), and "perception of caring from community adults" (56th). A list of top 10 variables is described in Table 4.1. A detailed list of all factors with average ranking and average merit is described in Table B2 at Appendix B.

Table 4.1

Correlation Attribute Evaluation (Top 10 Variables)

Attribute	Average merit	Average rank
Tobacco Product Use (TBP)	0.24 +- 0.001	1 +- 0
Sent to office for disciplinary issue	0.235 +- 0.003	2 +- 0
Marijuana use past year	0.213 +- 0.002	3 +- 0
Substances Use (Methamphetamine, Cocaine, etc.)	0.204 +- 0.001	4.2 +- 0.4
Marijuana use frequency	0.202 +- 0.001	5.3 +- 1
Prescription drug usage to get high (Vicodin, Valium, etc.)	0.2 +- 0.003	6.2 +- 0.6
In-school suspension	0.2 +- 0.002	6.3 +- 0.78
Social competency Scale (SCS)	0.193 +- 0.001	8.6 +- 0.8
Binge drinking – 2 (5 or more drinks in a row)	0.193 +- 0.001	9 +- 0.89
School engagement (SE)	0.192 +- 0.001	9.4 +- 0.49

Information Gain Attribute Evaluation This method evaluates the worth of an attribute by measuring the information gained with respect to the class (i.e. class of student with high and low tendency of missing school). This attribute evaluator evaluated each attribute's information gain (i.e. merit) independently of the others for each fold, sums these values, then calculate the average merit by dividing the summation by 10 (i.e. number of the folds). The method ordered the 113 independent variables utilized for the analysis by average merit and average rank. Similar to the correlation attribute evaluation, the ranked results indicated that the majority of the top factors were comprised of attributes either from the microsystem or the mesosystem. When evaluating the top 10% of the factors (n = 12), 58% (n = 7) of the factors were from the microsystem, followed by 42% (n = 5) from the mesosystem. Factors regarding substance use (tobacco product use, marijuana use, and other substance use), the SCS, and staying home due to sickness were found to be most relevant among factors within the microsystem to students missing school. For factors within the mesosystem, being sent to the school office due to disciplinary issues, the school engagement scale, friends' approval of substance use, teacher-student relationships (TSR), and ACEs were found to be the most relevant to school absences. In the top 10% list, 75% (n = 9) of factors from this method also appeared in the top 10% list of the correlation attribute evaluation but in a different order as presented utilizing information gain evaluation. A factor from the exosystem (out-of-school activity: sports) first appeared in the list in the 46th position and then the rest of the factors from the exosystem were distributed behind the factor 'out-of-school activity: sports'. Three factors from the macrosystem were found to be ranked higher than factors from the exosystem. The factor perception of caring from community adults placed 25th (the highest rank among the factors in the macrosystem), followed by free and reduced lunch (32nd), and neighborhood safety (45th). A list of top 10 variables is described in Table 4.2. A detailed list of all factors with average ranking and average merit is described in Table B3 at Appendix B.

Table 4.2

Information Gain Attribute Evaluation (Top 10 Variables)

Attribute	Average merit	Average rank
Social Competency Scale (SCS)	0.031 +- 0	1 +- 0
Tobacco Product Use (TBP)	0.03 +- 0	2 +- 0
School Engagement (SE)	0.027 +- 0	3 +- 0
Friends approval of substance use (FAS)	0.025 +- 0	4 +- 0
Sent to office for disciplinary issue	0.024 +- 0	5 +- 0
Non-medical marijuana use frequency	0.023 +- 0	6 +- 0
Marijuana use frequency	0.021 +- 0	7.2 +- 0.6
Staying home due to sickness	0.021 +- 0	8.3 +- 0.78
Teacher Student Relationship (TSR)	0.02 +- 0	8.6 +- 0.49
Substance use – 1	0.02 +- 0	9.9 +- 0.3

J48 Decision Tree Wrapper Evaluation The wrapper method utilized the J48 decision tree algorithm to identify recurring factors that yield optimal information when dividing samples with an outcome factor. Among the 113 variables (n = 51), 45% appeared in the process of 10-fold cross validation of J48 at least once. The factor out-of-school suspension appeared in the model eight of 10 times in the 10-fold cross validation (80%), which was the most repetition among all the variables. This means that when the J48 algorithm was attempting to find the factor that yields the most information for 10 times, the factor out-of-school suspension was found to be one of those informative factors eight times. The factor Non-Hispanic American Indian was repeated six times

(60%), followed by Native Hawaiian or Pacific Islanders (five times, 50%). Both in-school suspension and visiting doctor or nurse for physical checkup appeared four times, followed by being bullied: LGB (three times, 30%) and treated for alcohol or drug problem (three times, 30%). A total of 10 variables – age, feeling safe at home, perception of peer caring, non-suicidal self-injury, binge drinking, perception of family caring, suicidal attempt, perpetrator, measure of homelessness, and other substance use (including LSD, glue, spray can, or coke) appeared in the J48 results two times (20%). A list of top 10 variables is described in Table 4.3. A detailed, ranked list of J48 results including all 113 variables with the number of folds (%) are described in Table B4 at Appendix B.

Table 4.3

J48 Results during the Wrapper Method Feature Selection (Top 10 Variables)

Attribute	Total Instance (%)
Out of school suspension	8(80 %)
Race & Ethnicity: American Indian Non-Hispanic	6(60 %)
Race: Native Hawaiian or Pacific Islander only	5(50 %)
In school suspension	4(40 %)
Physical checkup	4(40 %)
ACE	4(40 %)
Harassed by peers: LGB	3(30 %)
Substance use treatment history	3(30 %)
Grade	2(20 %)
Home safety	2(20 %)

Prediction Models Comparison

J48 Subset Evaluation The goal of this evaluation was to acquire the subset of essential factors that best represent the outcome of interest (i.e. unexcused school absences) that also has the minimal number of factors that present the best results (accuracy, AUC, and f-score). A total of four subsets (all variables repeated more than twice, three times, four times, or five times) were prepared for the performance comparison. For readability in this section, the study utilizes the names Subset A (all variables repeated more than twice), B (more than three times), C (more than four times), and D (more than five times). Table 4.3 includes all the variables used for each subset. The study utilized accuracy, the f-score (average), and the AUC from two prediction modeling methods (logistic regression and J48) to evaluate the prediction performance.

Subset C presented the best prediction accuracy (93.78%) compared to either Subset B (93.77%) or D (93.73%); the f-score did not make a significant difference (B – 0.913; C – 0.912; D – 0.909) in the logistic regression results. Subset B performed better on the AUC (0.705) than Subset C (0.695) and D (0.553), but the differences were not significant compared to Subset C. J48 demonstrated similar results to Subset C with the best accuracy (93.79%) over Subset B (93.77%) and D (93.76%) but did not present significant differences in f-score or AUC (B – 0.911, 0.547; C – 0.911, 0.548; D – 0.910, 0.541). Comparing Subset A to Subset C, the accuracy of Subset A was identical in logistic regression (93.78%) and lower in J48 (93.69%). Subset A was higher for the AUC in both methods (0.776 in logistic regression and 0.622 in J48) and slightly higher for f-score (0.916 in logistic regression and 0.914 in J48) compared to Subset C, which was not significant. Evaluating the results, the study included variables utilized in Subset

C to be used for the next process (i.e. performance comparison), as it best aligned with the goal of acquiring the minimal number of factors that best represent the outcome variable, as the number of variables utilized in Subset C (*n* = 6) was significantly lower than in Subset A (*n* = 18). The details of the measures (i.e. accuracy, f-score and AUC) that utilize both methods for the four subsets are described in Table 4.4.

Table 4.4

J48 Subsets Evaluation

	Subset A*	Subset B	Subset C	Subset D
		Accuracy / f-score (average) / AUC		
Logistic Regression	93.78 / 0.916 / 0.776	93.77 / 0.913 / 0.705	93.78 / 0.912 / 0.695	93.73 / 0.909 / 0.553
J48 – Decision Tree	93.69 / 0.914 / 0.622	93.77 / 0.911 / 0.547	93.79 / 0.911 / 0.548	93.76 / 0.910 / 0.541

*Subset A – All the variables repeated in the J48 wrapper method for more than a time (Subset B – more than twice; C – three times; D – four times)

Performance Comparison As mentioned in the methods section, the study prepared three subsets (Top 10% + J48; Top 25% + J48; Top 50% + J48) for the performance comparison. Similar to the section above, subsets include Subset E (Top 10% + J48, *n* = 18), F (Top 25% + J48, *n* = 37), and G (Top 50% + J48, *n* = 64). Table 4.6 describes all 18 variables included in Subset E (See Table B9, B10 at the Appendix B for Subset F and G). The study utilized accuracy, the f-score (average), and the AUC from two prediction modeling methods (logistic regression and J48) to evaluate the prediction performance. In logistic regression, Subset G portrayed the best results in all measures (accuracy, f-score, and AUC – 93.85, 0.921, and 0.860); however, the differences were not significant (Subset E – 93.84, 0.919, 0.839; Subset F – 93.82, 0.920, 0.853), with the exception of the AUCs (Subset E – 0.839; Subset F – 0.853; Subset G –

0.860). In J48, Subset E outperformed the other two in accuracy and the AUC (Subset E – 93.61, 0.917, 0.687; Subset F – 93.06, 0.916, 0.678; Subset G – 92.83, 0.917, 0.676); the f-score was identical to Subset G (0.917).

As mentioned in the methods section, the study also implemented the same methods (logistic regression and J48) with another version of the dataset that countered the imbalance of an outcome variable (i.e. unexcused school absences) by utilizing SMOTE for the reference, whereby the outcome instances ratio was changed to 1.7:1 (*n* = 61,560:53,410) compared to the original data ratio of 15:1 (*n* = 113,375:7,630). The logistic regression results revealed that Subset G outperformed the other two (Subset E – 76.21, 0.760, 0.848; Subset F – 77.49, 0.774, 0.858; Subset G – 78.36, 0.783, 0.865). The accuracy level dropped significantly (original logistic regression – 92-93%; SMOTE sample logistic regression – 76-78%) when compared with the results from the original samples. In J48, Subset G was slightly better in accuracy (92.86%) compared with the other two (Subset E – 92.80%; F – 92.60%), which was not significant. Subset E outperformed in the AUC (Subset E – 0.957; Subset F – 0.941; Subset G – 0.865) while having an identical f-score with Subset G (0.928), both of which outperformed the f-score of Subset F (0.926). The details of the measures that utilized both methods with the four subsets are described in Table 4.5.

Table 4.5

Subsets E, F and G from Feature Selection Methods Performance Comparison

	Subset E - Top 10% + J48 (18/113)	Subset F - Top 25% + J48 (37/113)	Subset G - Top 50% + J48 (64/113)
	Accuracy / f-score (avg) / AUC		

Logistic Regression	93.84 / 0.919 / 0.839	93.82 / 0.920 / 0.853	93.85 / 0.921 / 0.860
J48	93.61 / 0.917 / 0.687	93.06 / 0.916 / 0.678	92.83 / 0.917 / 0.676
Logistic Regression (SMOTE)	76.21 / 0.760 / 0.848	77.49 / 0.774 / 0.858	78.36 / 0.783 / 0.865
J48 (SMOTE)	92.80 / 0.928 / 0.957	92.60 / 0.926 / 0.941	92.86 / 0.928 / 0.936

The results indicated that Subset E (Top 10% + J48, *n* = 18/113) best represented the outcome of interest (i.e. unexcused school absence) with the minimum number of factors. Utilizing the original data, Subset E outperformed the other two subsets in its J48 performance while utilizing the smallest number of variables. The differences in the logistic regression model were insignificant, while the number of variables utilized in Subsets F and G were significantly higher than in Subset E. Utilizing the SMOTE data, Subset E outperformed Subset F and was slightly better or worse than Subset G in J48 (i.e. Subset F better than G in AUC bur worse in accuracy). Subset E was deemed the most appropriate in this case, as the number of variables utilized in Subset E (*n* = 18) for the prediction modeling was significantly lower than in Subset G (*n* = 64). Performances portrayed in the logistic regression modeling of the SMOTE data indicated that Subset G yielded the highest performance. The results of utilizing logistic regression and J48 prediction modeling compared to both original and SMOTE data indicated that two yielded the best performances utilizing Subset E (J48 – original sample, J48 – SMOTE sample), while one performed the best by utilizing Subset G and another achieved inconclusive results. Therefore, the results demonstrated that Subset E best represented the outcome of unexcused school absences with the minimum number of factors.

Table 4.6

A list of Attributes from Subset E - Top 10% + J48 (n = 18/113)

Attribute	System in the KiTeS Framework
Social competency Scale (SCS)	Microsystem
Tobacco Product Use (TBP)	Microsystem
School engagement (SE)	Mesosystem
Friends approval of substance use (FAS)	Mesosystem
Sent to office for disciplinary issue	Mesosystem
Marijuana use past year	Microsystem
Marijuana use frequency	Microsystem
Staying home due to sickness	Microsystem
Teacher-student relationship (TSR)	Mesosystem
Substance use – 1	Microsystem
Substance use – 2	Microsystem
ACEs	Mesosystem
In-school suspension	Mesosystem
Binge drinking – 2 (5 or more drinks in a row)	Microsystem
Out-of-school suspension	Mesosystem
Race & Ethnicity: American Indian Non-Hispanic	Microsystem
Race: Native Hawaiian or Pacific Islander only	Microsystem
Physical Checkup	Microsystem

Aim 2 – Phase 1: LSN Focus Group Interview

The focus group interview study from the six LSN groups resulted in identifying four emerging, saturated themes about factors in school absenteeism and provided detailed dynamics of how those factors are functioning within the school setting. Four themes correspond to families and health, families and school, families and systems, and the role of school nursing. Structural codes not saturated (only appeared once or twice) were not included as emerging themes.

Absenteeism at the Intersection of Family and Health

Physical health, such as chronic physical illness (e.g. asthma and diabetes), was mentioned briefly by participants as a factor that influenced CA in their schools. For instance, one participant noted, *"A lot of my chronic absences have been supposedly medically related. So you hear a lot of effort, some just wild diagnosis, and they've went around and found a doctor that will document something and pretty much has excused them from life. And so therefore, the parents see no reason to send them to school because they're sick."* In cases where physical health was impacting CA, participants tended to feel that students were adequately supported by families, schools, and their medical teams.

The majority of the discussion about chronically absent youths' health focused on mental health, particularly on how untreated or uncontrolled mental health conditions drove much of the CA that participants were aware of. Participants described how specific mental health conditions (e.g. depression), psychosomatic symptoms (e.g. stomachaches) from unspecified mental health issues, and anxiety related to incidences of personal or family crises led to regularly missed instructional time. Participant A from Group 6 posited, *"I would say the majority [of CA], 90%, it's due to mental health. School refusal [related to a student's] known depression, known anxiety."*

While participants tended to feel students were usually well supported in their physical health, they extensively described that students' mental health was not being adequately addressed. Often this was discussed as related to misalignments between parental understanding of how best to manage youths' mental health and the nurses' attempted guidance toward evidence-based mental healthcare. This misalignment was described as difficult and often distressing to participants. Participant A from Group Suburban became emotional when she relayed, *"the hard part is the student is screaming, 'I want help, I want help' and the family's like 'nope, nope, nope.'"* Participant perceptions of parental attitudes of avoidance when confronting their child's mental healthcare could complicate certain situations in which participants felt that mental health treatment would readily address issues. Participant D from Group Greater Minnesota described, *"A lot of times, parents do not want to go down that road; they would much rather hear that it is an abdominal migraine than to hear that there could be an anxiety piece to this."* This quote specifically reflects on how some families sought diagnosis of a physical ailment rather than accept that their child may be experiencing psychosomatic symptoms related to mental health issues. Echoing similar sentiments, the participant reflected on how a parent's own mental health may play a role in avoiding care:

> *One of my really chronic absentee kids last year, I would say he's probably got some undiagnosed mental health [condition], but so do the parents, and they don't want to go down that road with their son; they don't want to deal with that themselves.*

Additional factors influencing family decisions to forgo mental health treatment, which in turn leads to students missing more school time, were noted by participants to include: specific cultural beliefs, miseducation about mental health treatments, or limited access to mental healthcare. For instance, Participant C from Group Suburban noted,

> *By far and large, I think it's the mental health component [that drives CA], and I really feel like there are a dearth of resources available families and kids....It's huge, and there's a huge gap in meeting the needs of these kids and the family members as well.*

Families' abilities to afford the high cost of mental healthcare was particularly concerning to some participants. This was highlighted by Participant A in Group Suburban, who recounted parent reactions to phone calls she made after a student engaged in self-harm:

> *I am calling the crisis line for your student because they harmed themselves this morning. They are still feeling this way; they need to be brought in immediately. And the parents get mad at me; they've yelled at me, saying, "Who's going to pay for this?*

Absenteeism at the Intersection of Family and School

Aside from the family's lack of support in students experiencing mental health issues, family attitudes and behavior toward school attendance were stressed repeatedly by a number of participants as affecting students to engage in CA. The attitudes of families who proactively do not want their kids to go to school can been seen in the first (from Participant C from Group ALC) and second quotes (from Participant C from Group 6):

> *[The] parents really didn't want them to go to school, so they would develop reasons why the kids couldn't go. They had, you know, one of them said that they had heart issues and different things that you know through the years. I saw the kids go all the way through to high school, [and it] turns out mom really just wanted the kids at home.*

> *In rare cases, we've had a couple of [times] where the parents, I think, really just want them home and kind of have a codependency situation, and their anxiety really rises when their children aren't there with them.*

Additionally, parents' perceptions that school is not the priority in the kids' lives can be found in the first (Participant A from Group Greater Minnesota) and second quotes (Participant E from Group Greater Minnesota):

> *Parents, I think, are doing the best they can just to keep the roof on the house and food on the table or in the fridge, and so they just don't have a lot of, doesn't seem like school is the priority for their parents. So if their child says, yeah, I don't feel good, I don't want to go, they don't question; they just let them stay.*

> *I think a lot of the kids that we have that are chronically absent are kids whose parents, you know, don't see the importance of making them come to school regularly. One of the things that we see is that we'll have a parent arrive at school to pick up their kid when they haven't been out of the health office; they haven't communicated with anybody. They just texted their parents and said, I want to go home come, and pick me up. And we'll talk to them and will say, you know, the health office really needs to be involved in the decision for your child to go home. They're not supposed to contact you; we are supposed to contact you, but it's frustrating when it just bypass[es] the whole system, and they go home.*

Such attitudes and behaviors shown by family can also act as a barrier when attempting to address absenteeism. That is, the parents' low prioritization of school attendance leads to attitudes or behaviors that lead to students missing school, which becomes a family barrier, as exhibited in the quote Participant B from Group Suburban:

> *I'll have students miss school multiple times in a month, even for various funerals and that sort of thing, or the family may not tell us why they're absent; just a family emergency, and that's excused.*

To encourage students to avoid being chronically absent from school, aligning with the results from above, family impact was mentioned to be important and would facilitate students attending school, which is indicated in this quote: "*The fact that mom is still involved, you know, keeps the student from being dropped.*"

The results in this section emphasize the crucial role of family and their attitudes and behaviors that impact students' absenteeism. Family approval or failure to address the absence might be perceived by the child as the belief that it is okay to not go to school, which eventually transpires into CA.

Absenteeism at the Intersection of Family and Systems

The interviews also revealed the role of family and social systems when caring for the students who miss school, which aligns with the environmental interaction that exists within the exosystem and macrosystem seen in the KiTeS bioecological framework (Melvin et al., 2019). CA-causing factors mentioned by multiple participants were transportation and housing instability. When the school bus is their only transportation to school (the family does not have a car), it was mentioned as, *"Our three- and four-year-olds that often is absent because they miss their only transportation to school" (Participant B from Group ALC)*. The lack of transportation when the family moves to another place also becomes a problem, as seen in the following quote:

> *There can be big lags in transport to [school]; it can take a week to get the transportation once they move, which can be really problematic (Participant C from Group 5).*

Additionally, housing stability as an obstacle for the students to be present at school arose repeatedly. Students who are homeless and highly mobile have challenges coming to school, as seen in these quotes: "*We also have a lot of the homeless. We have had a couple of them that lived in hotels, and for some reason, they had trouble getting to school" (Participant C from Group ALC).* And *"I would say of the students that I have become aware, a lot of them might be also special ed students or homeless and highly mobile students" (Participant C from Group 5).*

Furthermore, participants could identify the problem of inadequate sleep affecting students' functioning even when they are present in school, as how Participant A from Group ALC described "*I think a lot of them that we see are the homeless or highly* *mobile; their challenges can be many. You know, [if] they haven't slept the night before, they can't get there".*

Partial day absences also had factors categorized in the exosystem and the macrosystem. Participants mentioned that these are caused by having a limitation in access to resources such as food. Students can be encouraged to come to school by being offered free lunch. However, this promotion makes students arrive only right before the lunch, which would make their appearance be counted as partially present which is shown by Participant C from Group Urban as *"We've had groups of kids that come in at lunch time...they go and hang out, and then when they get hungry, they come to school, around lunch. Pretty typical pattern for some kids".*

Additionally, students themselves having a job to support their own family leads them to be partially absent from school.

> *And the other reason for coming late is some of them work the night shift, so [they] sleep or worked late or maybe did other things too, but some of them do have a job. And so they end up being late."*
> *(Participant A from Group ALC)*

The participant also mentioned that being mandated to come for legal reasons also promotes students being present at school, as the participant described "*And I would say some that also come late are mandated to come, so if they don't show up, they don't get whatever".*

Family working environment was mentioned, as when the parents are working before the students go to school. It is problematic, as mentioned in the following quote:

> *With my middle school age, none of them can drive, so if maybe they're not feeling well in the morning and they might be feeling fine an hour later, but the parent let them stay home... a kid oversleeps, and again, if the parents, already in work by the time their kid is going off to school,*

then the kid gets kind of a free pass and has to stay home all day." (Participant B from Group Suburban)

These results demonstrate the pattern that factors such as access to resources and low socioeconomic status cause student absences for multiple reasons (e.g. students working to support family or family having a job that cannot care for the child adequately). The lack of childcare from such factors can be overcome through a variety of social infrastructure and systems (transportation, free lunch, and community support), and the interviewees revealed the ability of environmental factors from the exosystem and macrosystem to complement the family's lack of caring when students are absent.

Similar results were identified in the barrier that prevents students from coming to school. The interviewees revealed a barrier that leads students to miss school for a full day: the family's socioeconomic status, which includes parents who are excessively busy working, which can be seen in the quote below:

> *So I would like to add that my students, a lot of their moms and dads go to work before they have to go to school. And so when our secretary will pull up, you know, they'll do a list and say, hey, you know, you want to check in with this family; this kid's been out for X number of days, and the parents are like, What do you mean, you know, and they're completely clueless about that." (Participant C from Group Suburban)*

School Nurse Roles in Supporting Chronically Absent Students

Students who miss school for a partial day demonstrated a tendency to intentionally visit the school nurse's office to avoid certain classes they do not prefer, which is revealed in the quote from Participant B from Group Greater Minnesota:

> *I kind of have a couple high school kids that really don't like the sciences or the math or whatever subjects, so, um, it's very easy to pick*

up on their patterns with our electronic charting and coming into the health office.

Mental health-related issues were described by participants as well. Psychosomatic symptoms that students describe in the school nurse's office provide LSNs no choice but to allow them to be absent from the school, as there is a limit of capacity when assessing psychosomatic symptoms. This finding is illustrated in the quote by Participant D from Group Urban:

> *So I'll have kids that will, like, cycle at certain times of the day, and like, they are suddenly, like, ill, but they're, you know that they're not actually physically ill, but you can't really pinpoint it, so sometimes those kids are going home. And then I have kids if, for instance, that will work themselves up into, like, respiratory symptoms that really doesn't seem like asthma, and they're not wheezing, but I really can't keep a kid there with respiratory distress in my office.*

Furthermore, students who cannot have quality sleep visit the nurse's office to sleep, which leads to missing classes partially and is included in the quote by Participant A from Group Urban:

> *Every year there seems to be two or three students, who, they're not sleeping at night because they're coming into the health office, and they're wanting to lay down, and they are sound asleep like, like, they didn't sleep.*

A variety of PDA-related factors are detected by LSNs, as those students who suffer from the issues mentioned (e.g. class preference, psychosomatic symptoms, and lack of sleep) come to the school nurse's office to describe what is wrong. This pattern emphasizes the role of LSNs in school absences, as school nurses are able to detect students' problems that lead to students being partially absent.

Multiple participants mentioned that LSNs attempt to intervene in students being absent from school by collaborating with others (i.e. local law enforcement, Student assistant program team, guidance counselors, administration, and social workers), which is indicated in the quotes below:

> *Our local law enforcement is great to try to do to help us and tracking some of those kids down, and saying, okay, you kids, you need to get to school and that type of thing (Participant B from Group ALC).*

> *I had the guidance counselor involved. At least, I bounced the idea off with the guidance counselor. Well, we also have a drug treatment program…that was very close to the high school. And I suggested that she utilize them for a resource. (Participant D from Group Greater Minnesota)*

While mentioning the collaborations, the participants also mentioned the limitations regarding collaboration.

> *We do have social workers in our building, and we have these floor offices, we call them, and they run attendance reports every day, and the health offices are not in our, in our buildings, not responsible for attendance, so we do run reports every day. And the social workers actually manage that a lot closer than I do, but I do get involved when there is a medical situation, reaching out to the doctor to ask, you know, to a doctor, "Is this medically necessary for this student to be absent," or something like that. I have done that in the past. (Participant A from Group 5)*

> *I would try to get more school staff involved; the social worker, we have one in the building, but you know, so much of that is, I need [the student's] consent to do that. And that makes it real difficult, um. So unless they are truly harming themselves, it's kinda stuck. (Participant D from Group Greater Minnesota)*

Furthermore, the limitation of applying any type of intervention when the student disappears from school was identified, as presented in the following quotes:

> *[The] social worker does a nice job of keeping up with them in making sure they're not in crisis, or if they are, how we can help them. But some of them just go up off the grid, and you don't see him till you see him again. (Participant A from Group ALC)*

> *Last year, kids were often gone. When they did come to school, they would sleep a lot, and they were just closed; [I]couldn't really figure*

> *out how to help them. And then they just up and left one day, so that's always sad and frustrating. (Participant C from Group 5)*

The results signify the current below-satisfactory level of attention to students who are experiencing school absenteeism. Comments from participants demonstrated evident limitations that LSNs are encountering with other collaborators as well. To summarize, four saturated themes emerged from the LSN focus interviews including lack of familial, systemic support and the role of LSN taking initiative to support the students who are chronically absent.

Aim 2 – Phase 2: Causal Discovery Analysis

The phase 2 of aim 2 aims to answer three research questions:

1. Based on knowledge gained from Aim 1 (quantitative analysis), how are the risk and protective factors associated with partial and full-day absences?
2. Informed by perceptions of licensed school nurses (qualitative analysis), how are the risk and protective factors associated with partial-day absence, different or similar to, full-day absences?
3. How are the risk and protective factors associated with partial and full-day absence based on knowledge gained from Aim 1 (quantitative analysis) with perceptions of licensed school nurses (qualitative analysis) combined and how are they distinguished compared to the Aim 2 - research questions 2 and 3?

Sample Characteristics

Participants The study utilized 125,375 participants from the MSS 2019 (*n* = 170,128), excluding students in Grade 5 (*n* = 44,753), as the dynamics of students missing school are different between primary and secondary schools. The distribution of all three grades were similar as shown here: eighth (*n* = 44,919; 36%), ninth (*n* = 45,232; 36%), and 11th graders (*n* = 34,968; 28%). The portion of biological sex was also evenly distributed (M = 62,375: 50%; F = 62,709: 50%). Participants were primarily White followed by multiple race groups, Black, Asian, Hispanic, and American Indian or Alaskan Native. The participants who were Native Hawaiian or Pacific Islanders were also compared with other races, although the portion of those were less than 1% (*n* = 271; 0.002%), as these students are known to miss school more than any other race (U.S. Department of Education, 2016b). All missing data that pertains to demographic characteristics were below 10% which is acceptable (Walczak and Massart, 2001).

In addition, as mentioned in the methods section, all missing data including demographics variables were imputed using CART method before the analysis. Additional demographic characteristics of the study (race or ethnicity, free or reduced-cost lunch, and school region) with outcome and primary independent variables (i.e. variables identified to be highly associated with school absences from the results of Aim 1) are summarized in Table B5 at Appendix B.

Causal Analysis – Step 1

Figure 4.1 is a graph with directed edges with the effect size (ES) higher than 0.1; it utilizes factors identified in Aim 1 – quantitative approach (See Table B6 at Appendix B for a detailed information of ES for each node). The SEM fit measurement for the model utilized in the analysis is described in Table 4.7.

Figure 4.1

Causal Discovery Graph for the MSS 2019 (ES > |.1|) - Causal Analysis Step 1

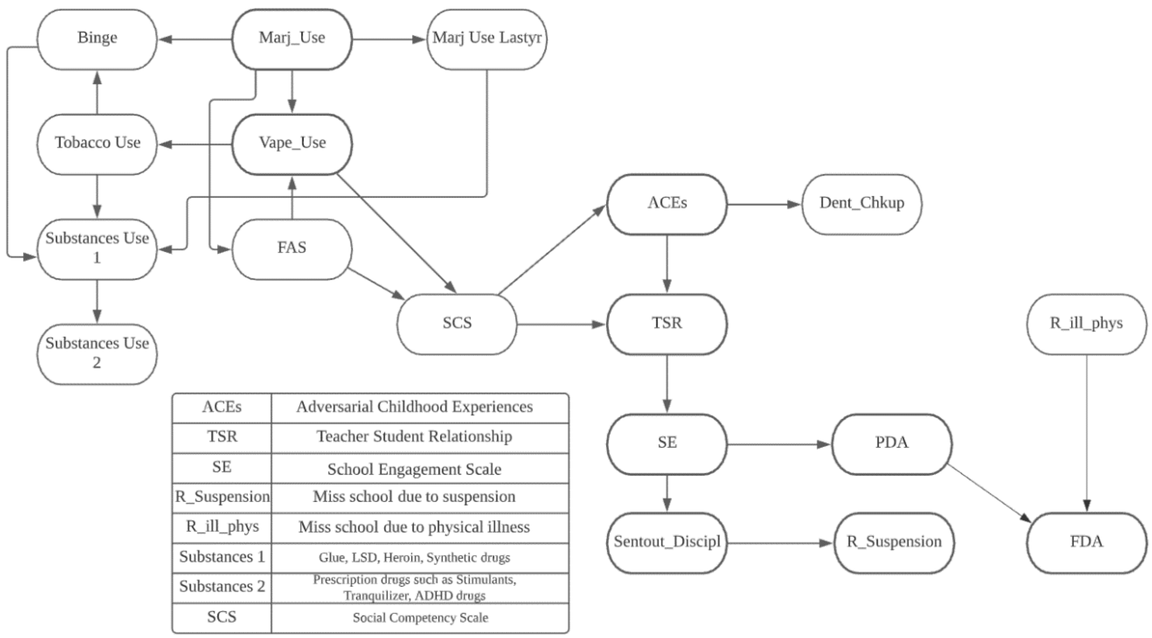


Table 4.7

SEM Fit Measurement for the MSS 2019 - Causal Analysis Step 1

SEM Fit Measurement	Value	Acceptable*
Model Chi-square	1779 (p<.01), df = 84	
Comparative Fit Index (CFI)	0.98	CFI >= 0.90
Root Mean Square Error of Approximation (RMSEA)	0.01	RMSEA < 0.08
Standardized Root Mean Square Residual (SRMR)	0.01	SRMR < 0.08

*Hooper et al., 2008.

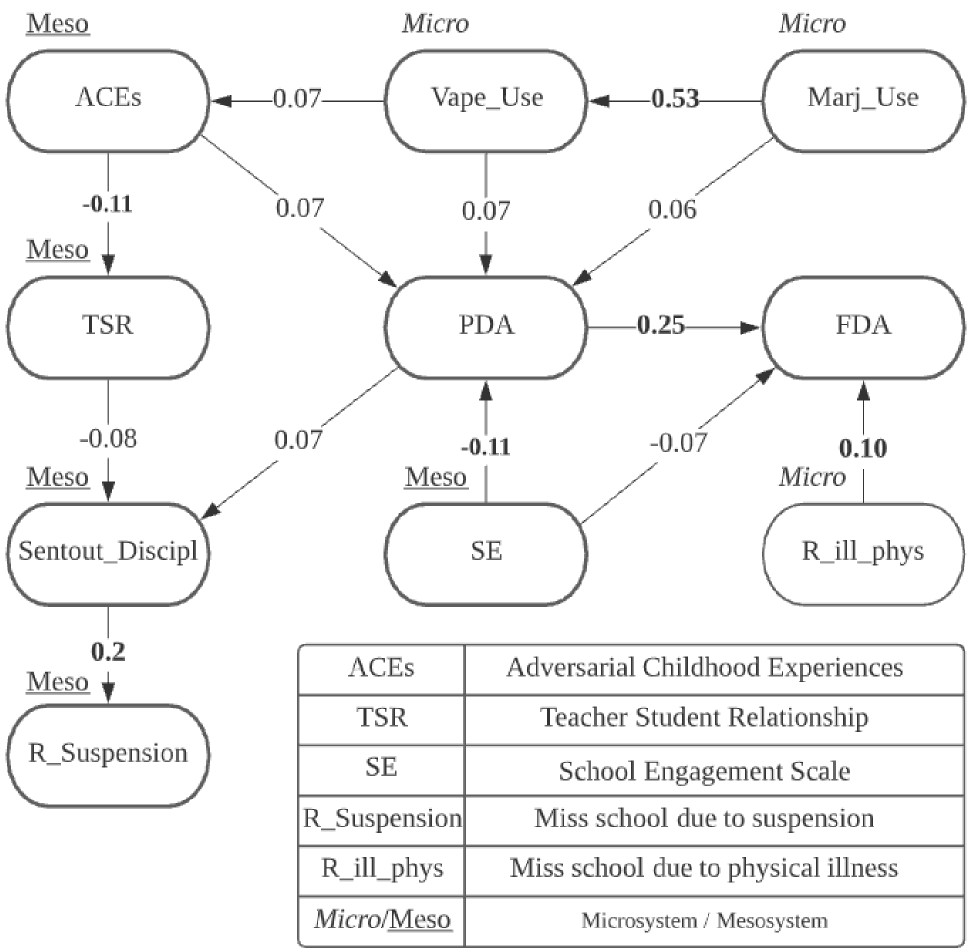
The graph portrays substance-related factors on the upper-left (binge drinking, marijuana frequency, marijuana use frequency last year, tobacco products use, vape use, other substance use, and friends' approval of substance use) forming a web of interconnection to each other, then reaching out to other school-related factors (lower-right) through the SCS. Focusing on school absences, the graph features PDAs, which are causally related to the school engagement (SE) but directly affect FDAs along with students missing school due to physical illness. The race and ethnicity of Native Hawaiian or Pacific Islanders and American Indian or American Natives did not present any causal relations with other factors that exceeded the effect size of 0.1.

To focus on the interconnectivity between PDAs and FDAs, the cutoff effect size value of 0.05 was utilized to populate more edges. Figure 4.2 is a graph from the same SEM model that features a lower cutoff ES value (> |.05|). Only the edges that 1. Directly affect PDA or FDA, 2. Children nodes of PDA and FDA are described in the graph to focus on factors directly related to school absences that either cause or are caused by PDA or FDA. This graph (ES < |.05|) reaffirms the role of PDA acting as a precursor to FDA. Additionally, the causal discovery inferred that students' experiences of ACEs, SE, and marijuana or vape use directly affects PDA, which then directly develop to FDA and students being sent out of the classroom due to a disciplinary issue. Students' levels of SE cause both PDA and FDA but the effect sizes are higher for PDA compared with FDA which means SE has a stronger relationship with PDA compared to FDA. Consequently, FDA presents itself as an outcome rather than a predictor.

In terms of the hierarchical multisystemic approach from the KiTeS framework (Melvin et al., 2019), both PDA and FDA were directly influenced by factors within the microsystem and the mesosystem. SE (mesosystem) and students missing school due to physical illness (microsystem) directly affected FDA while PDA was directly affected by ACEs (mesosystem), vape and marijuana usage (microsystem).

Figure 4.2

Causal Discovery Graph (ES > |.05|) - Causal Analysis Step 1



The validation process that utilized data (ES > |.05|) from the MSS 2016 corroborated the causal relations identified in the Causal Analysis – Step 1 from the MSS 2019. Specifically, the results from the MSS 2016 were identical to MSS 2019 in students' SE directly affecting PDA, substance and alcohol usage (tobacco products and binge drinking) directly affecting PDA, and PDA directly affecting FDA. Consistent with the MSS 2019, both PDA and FDA were directly influenced by factors within the microsystem and the mesosystem. Figure 4.3 describes the direct causal relations of both

PDA and FDA from the MSS 2016. The SEM fit measurement for the model utilized in the analysis is described in Table 4.8.

Figure 4.3

Causal Discovery Graph (ES > |.05|) from the MSS 2016 - Causal Analysis Step 1

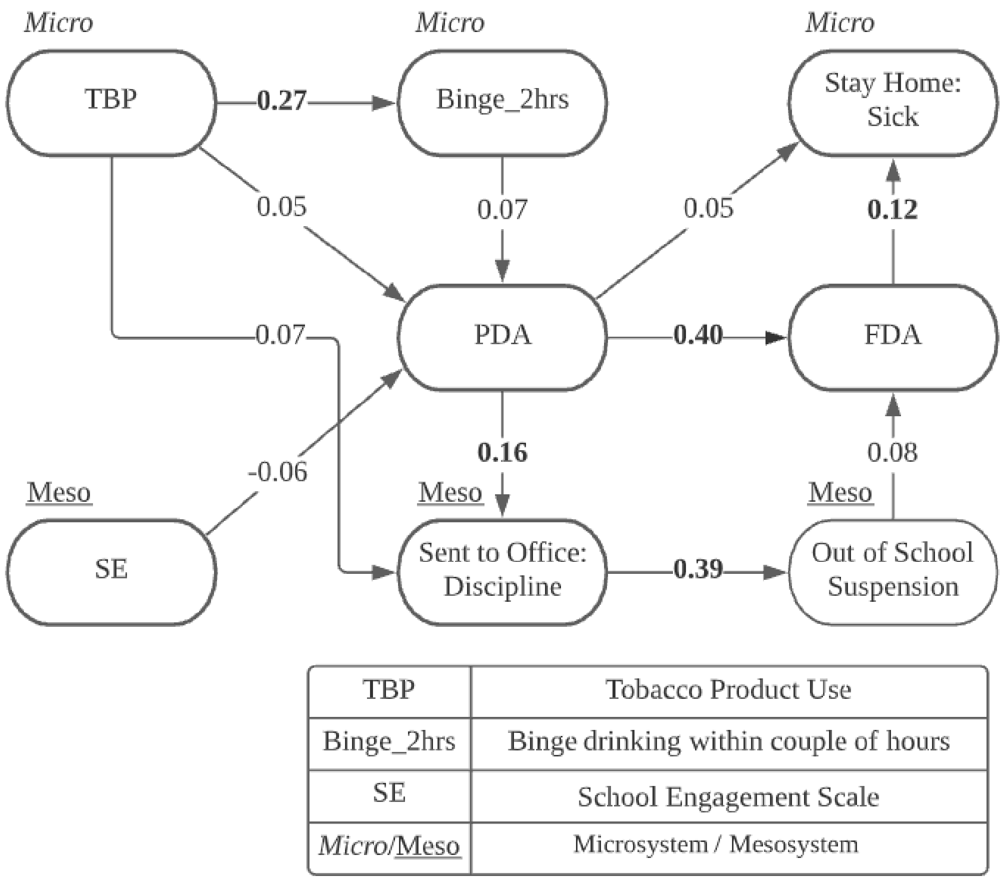


Table 4.8

SEM Fit Measurement for the MSS 2016 - Causal Analysis Step 1

SEM Fit Measurement	Value	Acceptable*
Model Chi-square	1780 (p<.01), df = 109	
Comparative Fit Index (CFI)	0.99	CFI >= 0.90
Root Mean Square Error of Approximation (RMSEA)	0.01	RMSEA < 0.08
Standardized Root Mean Square Residual (SRMR)	0.01	SRMR < 0.08

*Hooper et al., 2008.

Causal Analysis – Step 2

Figure 4.4 is a graph with directed edges (ES > |.1|); it utilizes factors identified in Phase 1 of Aim 2 (qualitative approach). The SEM fit measurement for the model utilized in the analysis is described in Table 4.9.

Figure 4.4

Causal Discovery Graph (ES > |.1|) - Causal Analysis Step 2

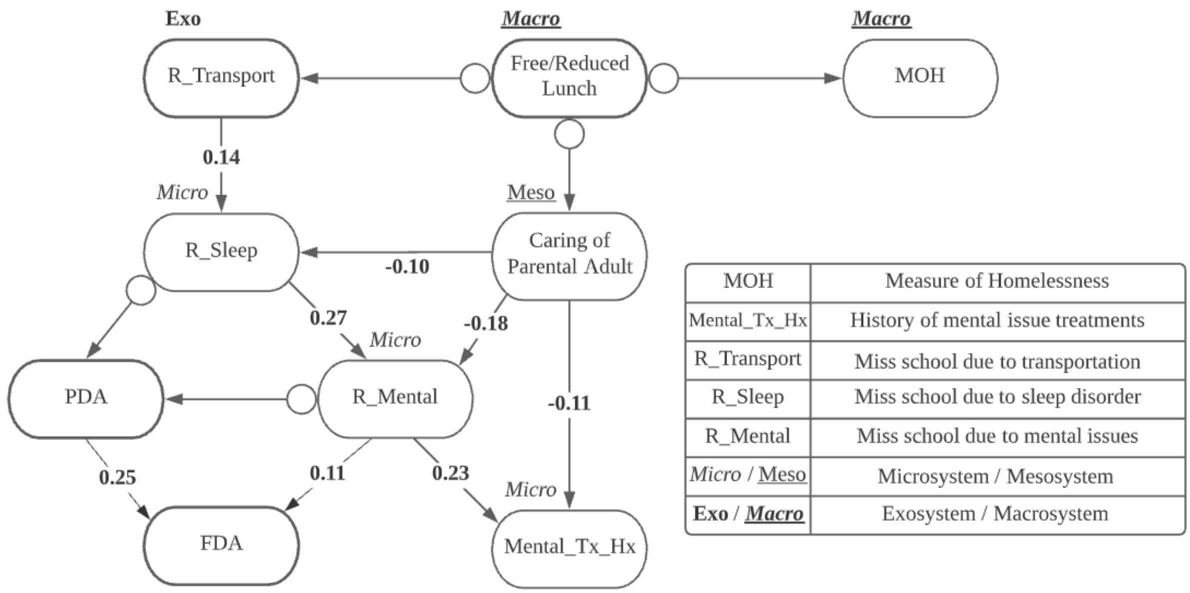


Table 4.9

SEM Fit Measurement for the MSS 2019 - Causal Analysis Step 2

SEM Fit Measurement	Value	Acceptable*
Model Chi-square	896 (p<.05), df = 25	
Comparative Fit Index (CFI)	0.99	CFI >= 0.90
Root Mean Square Error of Approximation (RMSEA)	0.02	RMSEA < 0.08
Standardized Root Mean Square Residual (SRMR)	0.01	SRMR < 0.08

*Hooper et al., 2008.

The graph presents a) the reasons for students missing school leading to FDA; b) students' perceptions of caring from parental adults that affect mental health; and c) PDA directly causing FDA.

Causal discovery inferred that the close association between lack of transportation to school, sleep disorders, mental issues (e.g. feeling excessively sad, hopeless, anxious, stressed, or angry) with the implication of the factors leading to increasing FDA. Notably, causal discovery linked three reasons for students missing school (transportation, sleeping, and mental issues) with FDA either directly or indirectly. However, in terms of PDA, it only managed to present that PDA is not the cause of students missing school due to sleep disorders or mental issues. The edge with an arrow and a circle on the other end indicates that both sleeping disorders and mental issues could imply the causation of PDA or that there is a confounding variable between the factors (between sleeping disorders and PDA or between mental issues and PDA). Therefore, while the reasons for students missing school revealed high associations to FDA, that was not the case for PDA.

Notably, the graph also exemplifies the role of caring from parental adults, as it directly affected mental health treatment history, students missing school due to sleep problems, and students missing school due to mental issues. Free or reduced-cost lunch provisions exhibited relations to the measure of homelessness, students missing school due to transportation, and caring of parental adults. However, similar to PDA, causal relations could not be inferred, as the causal discovery could not exclude the potential existence of the confounding variable between the provision of free or reduced-cost lunch with all three factors. Finally, Causal Discovery – Step 2 portrayed a direct causal relation from PDAs that points to FDA aligning with the results of Step 1. The graph that utilized a lower cutoff value (ES > |.05|) was not utilized in this step, as the original cutoff value (ES > |.1|) managed to reveal causal relations among the factors utilized in the analysis (See Table B7 in Appendix B for an information of ES for each node).

In terms of the hierarchical multisystemic approach from the KiTeS framework (Melvin et al., 2019), FDA was directly influenced by a factor within the microsystem (i.e. students missing school due to mental health issues). There were no factors directly causing PDA, as the existence of a confounding variable could not be excluded from the analysis. A factor from the exosystem (i.e. students missing school due to transportation) indirectly affected FDA through sleep problems and mental issues, but no direct pathways with ESs of .1 or above were identified.

The validation process (i.e. using a data that resembles an original data; MSS 2016 while not identical, is similar with the MSS 2019 for the validation to be implemented) that utilized data from the MSS 2016 was conducted to compare the results of the MSS 2019 with 2016. The results from the MSS 2016 were identical to the MSS 2019 in PDA directly affecting FDA. The MSS 2016 also revealed that sleep disorders directly cause PDA which implies that the result of the MSS 2019 (sleep disorders directly causing PDA with the potential existence of a confounding variable) could be true. All the relations with other variables were shown to be bidirected edges, which implies the existence of a latent variable between those according to how bidirected edges occur. Figure 4.5 describes the direct causal relations of both PDA and FDA from the MSS 2016. The SEM fit measurement for the model utilized in the analysis is described in Table 4.10.

Figure 4.5

Causal Discovery Graph (ES > |.1|) from the MSS 2016 - Causal Analysis Step 2

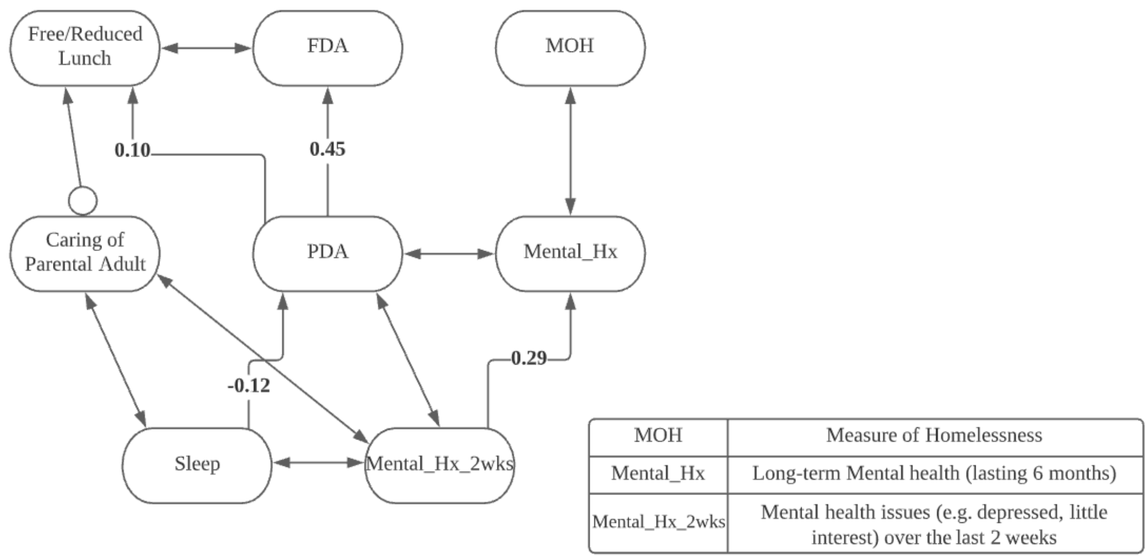


Table 4.10

SEM Fit Measurement for the MSS 2016 - Causal Analysis Step 2

SEM Fit Measurement	Value	Acceptable*
Model Chi-square	150 (p<.05), df = 8	
Comparative Fit Index (CFI)	0.99	CFI >= 0.90
Root Mean Square Error of Approximation (RMSEA)	0.01	RMSEA < 0.08
Standardized Root Mean Square Residual (SRMR)	0.01	SRMR < 0.08

*Hooper et al., 2008.

Causal Analysis – Step 3

Figure 4.6 is a graph with directed edges (ES > |.1|); it utilizes factors identified in Phase 2 of Aim 2, a mixed-methods approach (See Table B8 in Appendix B for a detailed information of ES for each node). The SEM fit measurement for the model utilized in the analysis is described in Table 4.11.

Figure 4.6

Causal Discovery Graph (ES > |.1|) - Causal Analysis Step 3

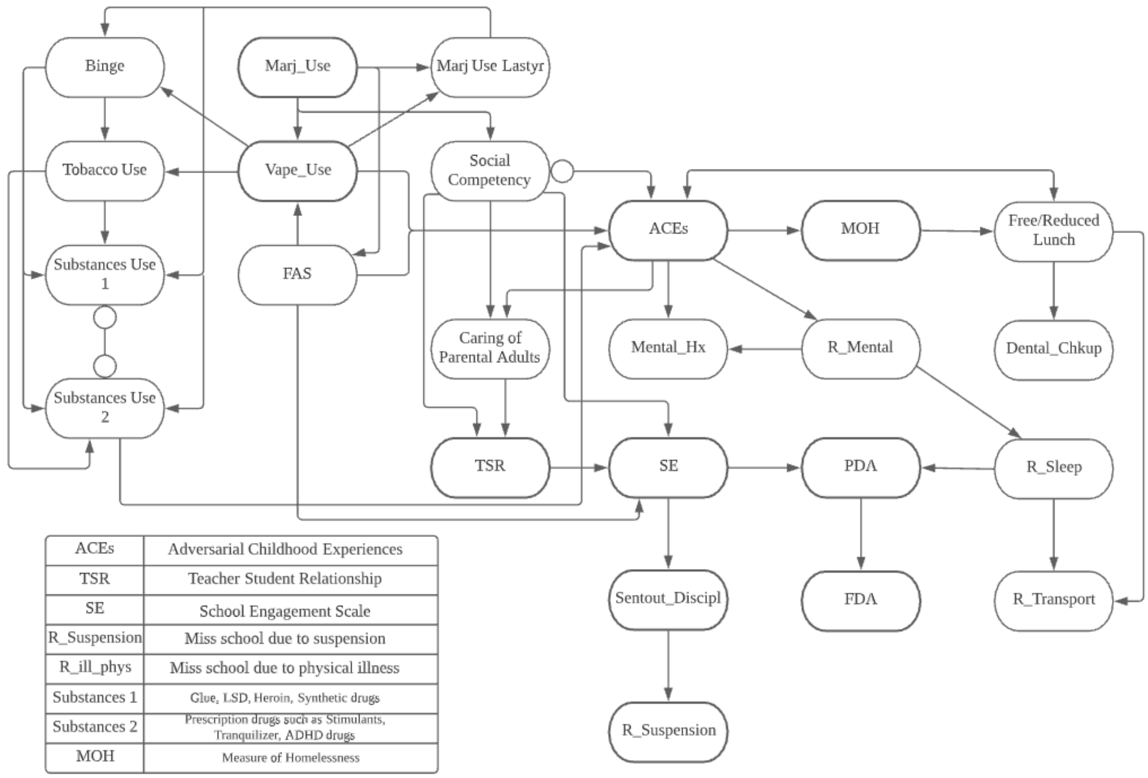


Table 4.11

SEM Fit Measurement for the MSS 2019 - Causal Analysis Step 3

SEM Fit Measurement	Value	Acceptable*
Model Chi-square	6927 (p<.01)	
Comparative Fit Index (CFI)	0.99	CFI >= 0.90
Root Mean Square Error of Approximation (RMSEA)	0.01	RMSEA < 0.08
Standardized Root Mean Square Residual (SRMR)	0.01	SRMR < 0.08

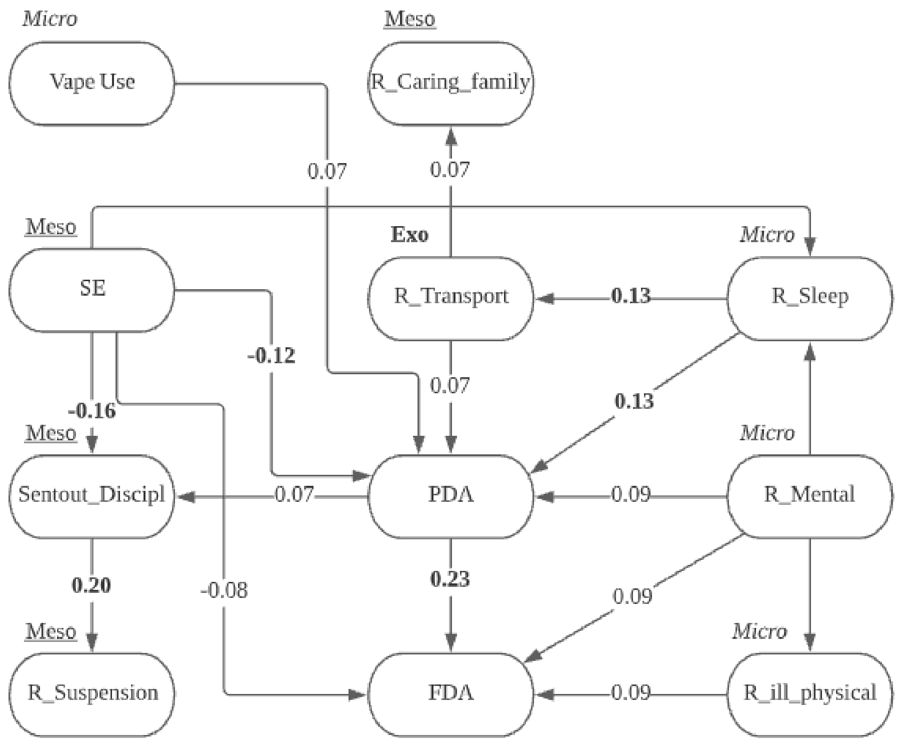
*Hooper et al., 2008.

The graph presents a) nodes of substance usage (e.g. marijuana, prescription drug usage without a prescription, tobacco usage, or alcohol consumption) linked with other school-setting factors utilized in the analysis via the SCS and ACEs; b) mental health issues that transpired from ACEs and led to students' lack of sleep, which directly affects students' behavior of missing school for a partial day (i.e. ACEs → mental health issues → sleep disorder → PDAs); c) both the SCS and TSR leading to students' levels of SE that affect PDAs; d) caring from parental adults (affected directly by both ACEs and the SCS) leading to TSRs that follow the route of SE and ultimately low tendency of PDAs (i.e. ACEs and the SCS → caring of parental adult → TSR → SE → low PDAs); and e) PDA directly causing FDA.

To focus on the interconnectivity between PDA and FDA, the cutoff effect size value of 0.05 was utilized to populate more edges. Figure 4.7 presents a graph from the same SEM model but with a lower cutoff ES value (> |.05|). Only the nodes and edges that a) directly affect PDA or FDA, and b) are children nodes of PDA and FDA are described in the graph to focus on factors directly related to school absences that either cause or are caused by PDA or FDA. This graph (ES < |.05|) reaffirms the role of PDA as a precursor to FDA. Additionally, the causal discovery managed to infer that students' levels of SE, usage of vape, transportation, sleep disorders, and mental issues directly affect PDA. FDA was directly caused by SE, mental health, PDA, and physical illness. Furthermore, the result revealed that students' mental health issues burst into a multitude of factors (PDA, FDA, and students missing school due to sleep, physical illness, or to care for family members), which emphasizes the impact of mental health management in school absenteeism.

Figure 4.7

Causal Discovery Graph (ES > |.05|) - Causal Analysis Step 3

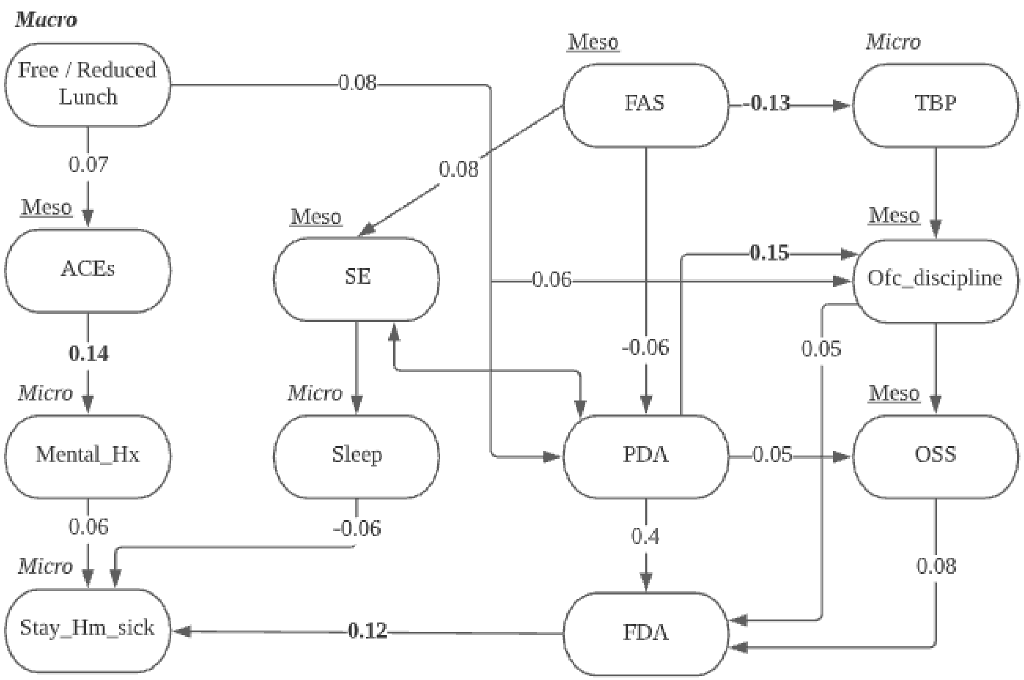


R_Sleep	Miss school due to sleep disorder
R_Mental	Miss school due to mental issues
R_ill_physical	Miss school due to physically ill
R_caring_family	Miss school due to taking care of family
R_ill_physical	Miss school due to physical illness
R_Suspension	Miss school due to suspension
R_Transport	Miss school due to transportation
SE	School Engagement
Micro / Meso	Microsystem / Mesosystem
Exo / ***Macro***	Exosystem / Macrosystem

In terms of the hierarchical multisystemic approach from the KiTeS framework (Melvin et al., 2019), the Step 3 analysis revealed that PDA was directly influenced by factors within the exosystem (transportation), mesosystem (SE), and microsystem (vape usage, sleeping disorders, and mental health issues). Similar to previous steps, FDA was directly influenced by factors within the microsystem (mental and physical health) and mesosystem (SE).

Figure 4.8

Causal Discovery Graph (ES > |.05|) from the MSS 2016 - Causal Analysis Step 3



ACEs	Miss school due to sleep disorder
Stay_Hm_sick	Miss school due to mental issues
Sleep	Sleep hours during a typical school night
FAS	Friends approval of substances use
TBP	Tobacco Products Use
OSS	Out-of-school suspension
Mental_Hx	Long-term Mental health (lasting 6 months)
SE	School Engagement
Micro / Meso	Microsystem / Mesosystem
Exo / ***Macro***	Exosystem / Macrosystem

The validation process, utilizing data (ES > |.05|) from the MSS 2016, corroborated the causal relations identified in the Causal Analysis – Step 3 from the MSS 2019. Specifically, the results from the MSS 2016 revealed that PDA directly affect FDA.

The results that were inconsistent with the MSS 2019 were as follows: PDA was affected by factors within the exosystem (free or reduced-cost lunch), mesosystem, and microsystem, whereas FDA was affected by factors within the mesosystem; factors of mental health and sleep disorders were not related to PDA but led to students staying home due to sickness, which was directly related to FDA. Figure 4.8 describes the direct causal relations of both PDAs and FDAs. The SEM fit measurement for the model utilized in the analysis is described in Table 4.12.

Table 4.12

SEM Fit Measurement for the MSS 2016 - Causal Analysis Step 3

SEM Fit Measurement	Value	Acceptable*
Model Chi-square	3619 (p<.01)	
Comparative Fit Index (CFI)	0.99	CFI >= 0.90
Root Mean Square Error of Approximation (RMSEA)	0.01	RMSEA < 0.08
Standardized Root Mean Square Residual (SRMR)	0.01	SRMR < 0.08

*Hooper et al., 2008.

Chapter V: Discussion

The purpose of this dissertation study, a secondary data analysis of the Minnesota Student Surveys administered in 2016 and 2019, was to examine the interconnectivity between risk and protective factors that are linked to school absenteeism among secondary school students. The study addressed two specific aims and their related research questions, and results are discussed for each in the following sections.

This chapter discusses the findings of the study according to the aims and research question, followed by the limitations, future directions for research, and implications for the nursing discipline.

Aim 1: Research Question – Which risk and protective factors are associated with school absences among secondary school students in 2016?

The focus of Aim 1 was to identify risk and protective factors associated with school absences. It is important to note that the outcome variable measured in the Aim 1 was 'unexcused' school absence, due to the limitation of the 2016 MSS. Feature selection methods followed by a comparison of prediction models were used to conduct analysis to fulfill Aim 1. In the process of the prediction models comparison between the original and the SMOTE data, there was a significant decrease in accuracy for the 'logistic regression prediction model' with the SMOTE data compared with the original, which implied that the accuracy shown with original data could have been as a result of 'overfitting'. In other words, skewed data distribution in the original data could have yielded high accuracy in prediction since the testing data were skewed from the outset. To complement such a caveat, an f-score was used as it represents weighted precision and recall. As a result, 18 out of 113 risk and protective predictors were identified; these are the predictors most associated with an outcome variable of ‘unexcused school absences’. All 18 variables were within either the micro- or the mesosystems (10 variables in the microsystem, 8 variables in the mesosystem).

Key Microsystem Predictors

The microsystem in the KiTeS framework used in this study focused on the components related to children, including age, gender, race, mental and physical health, sleeping problems, and behavioral and emotional problems (Melvin et al., 2019). Ten variables in the microsystem that were identified to be most associated with school absenteeism were 1) social competency, 2) staying home due to sickness, 3) physical health checkup, 4) race-ethnicity (American-Indian Non-Hispanic, Native Hawaiian or Pacific Islander), 5) marijuana use during past year, 6) marijuana use frequency, 7) tobacco product use, 8) substance use (methamphetamine, cocaine, etc.), 9) prescription drugs use without prescription (Benzedrine, Ritalin, Oxycontin, Valium, Xanax, etc.), and 10) binge drinking (i.e. five or more drinks consecutively).

The results revealed that over half of these microsystem variables could operationalize behavioral problems reported by students, especially substance use. These findings of substance use being associated with school absenteeism aligns with an integrative literature review, confirming the connection between substance use and school absenteeism (Gakh et al., 2019). Other studies also mention that a long-term effect of school absenteeism may be substance abuse in adulthood, which implies an association of substance abuse with school absenteeism (Gottfried, 2014; Henry et al. 2012). Two race- and ethnicity-related variables (i.e. American-Indian Non-Hispanic, Native

Hawaiian or Pacific Islanders) aligned with a national study conducted by the U.S Department of Education (2016a). This previous study revealed that students who identified as American Indians and Pacific Islanders were the most chronically absent from school (26% American Indian, 22.6% Pacific Islander). However, this national study did not specify whether the chronic absenteeism was comprised of only full-day absences or both FDAs and PDAs. Therefore, results from this Aim 1 analysis supplement current knowledge by showing that the disparities in absenteeism shown in both races are not limited to only full-day but also to partial-day absences since the outcome variable used incorporated data for both full and partial-day absences, labelled as 'school absences'.

Key Mesosystem Predictors

The mesosystem in the KiTeS framework is comprised of interactions children have with two proximal surrounding environments namely 'family & parent' and 'school' (Melvin et al., 2019). Seven factors were identified as associated with unexcused school absences from the Aim 1 analysis. These were 1) school engagement, 2) teacher-student relationship, 3) friends' approval of substance use, 4) adverse childhood experiences (ACEs), 5) being sent to school office for disciplinary issues, 6) in-school suspension, and 7) out-of-school suspension. Among these seven, only one factor (i.e. ACEs) was from the environment of 'family & parent,' while the other six were interactions with the 'school' environment. The association of the results between ACEs and school absences aligns with previous knowledge that having had one or more ACEs is significantly associated with CA (Stempel, 2017). However, Stempel (2017) did not specify whether their data were comprised of only either FDAs or PDAs, or both.

Therefore, this result of Aim 1 also adds to previous knowledge by showing that this association (i.e. chronic absenteeism and ACEs) applies to PDAs as well as FDAs, as the outcome variable incorporated both FDAs and PDAs. Hendron and Kearney (2016) found school scales (School Climate Survey Revised Edition) including 'order and discipline' were inversely associated with absenteeism which aligns with the association of school disciplinary issues and school absences found in this study. Student-teacher relationships and school engagement were negatively associated with absenteeism; this finding confirms results from a previous study that found that personal factors, including attitudes towards teachers and negativity towards school, were associated with absenteeism (Balkis et al., 2016). Finally, the finding on friends' approval of substance use being associated with absenteeism was not revealed in any previous studies. And additional studies are required to corroborate this result and to identify 'how' friends' approval lead to a student missing school.

In summary, the Aim 1 analysis identified risk and protective factors associated with school absenteeism using a data-driven approach. The student PI compared the performances of prediction models to compile a list of factors that were most associated with school absences from 113 variables. This approach to identify not only the order of importance but also a list of factors that are associated with an outcome variable could be an important step that complements what we know from empirical evidence. Having an additional step that used a data-driven result tailored into a specific dataset can include/exclude factors that could have been overlooked. With this approach, the findings from Aim 1 reaffirmed the risk and protective factors associated with school absenteeism identified from a number of school absenteeism-related studies (Gottfried, 2014; Henry et al. 2012; Stempel, 2017; US Department of Education, 2016a). In addition, the results acknowledge the proximity between school absenteeism with risk and protective factors within the micro- and mesosystems compared to factors in the exo- and macrosystems.

Aim 2: Research Question 1 – What are licensed school nurses' perceptions of chronic absenteeism and the differences between partial- and full-day absences?

Qualitative analyses revealed that a major portion of the participants' comments related to CA were closely associated with 'family,' which emphasizes how important the role of family is when dealing with school absenteeism. A total of three categories that involve family and chronic absenteeism were found, namely 1) family and health, 2) family and school, and 3) family and system.

In the theme of 'family and health,' LSNs focused on 'mental' health issues as being a main causal factor of chronic absenteeism. A negative impact of mental health issues (e.g. depression, anxiety) on chronic absenteeism has been reported in several studies (Henderson et al., 2018, Wood et al., 2012), and the focus group participants reaffirmed the existence of such association among students in their school setting. In addition, LSNs mentioned the disparities between care for mental health issues compared to care for physical health issues. Based on the participants' comments, the study implies that when students are physically ill or challenged, adequate alternative ways to participate in the education system are forthcoming, but students who suffer from mental issues are neglected by family, the school, and the system. In addition, the participants noted possible discouragement from family members when students have symptoms of mental health issues, in the form of 'neglect' or a 'preference' of physical over mental illness. Family members often provide critical support when living with person who is dealing with a mental illness (Pernice-Duca, 2010). In addition, the closest interactions the student will have is with family members, which precedes school or community systems. Therefore, results from this analysis imply that, if family members are receptive, responsive, and supportive of the mental health issues, chronic absenteeism could be alleviated.

Several studies have shown that the attitude and behavior of parents towards school is known to affect absenteeism (Sexson and Madan-Swain, 1995, Ross et al., 2019). Sexson and Madan-Swain (1995) posited that 'overprotective' parents and parents who do not recognize the importance of absence from school leads to their children missing school, while Ross et al. (2019) showed that parent attitude is positively associated with 'active transportation to school,' which is one of the reasons why students miss school. The focus group participants corroborated these findings and mentioned that the attitude of parents, either not wanting their children to go to school or having the perception that school is not a priority in their children's lives. The results of this section emphasize the crucial role of family and their attitudes and behaviors that impact the behavior of absenteeism. The family's approval of or failure to address absenteeism might be perceived by the child as 'it is okay not to go to school,' which eventually leads to 'chronic' absenteeism.

Under 'family and systems,' factors that cause CA, according to multiple LSN focus group participants, were 'transportation' and 'housing instability.' When the school bus is the only available transportation or students are homeless, attending school is a challenge. Not having transportation, while also not staying in one place, results in a negative synergic effect of missing school, as school bus allocation to students' locations takes at least a week to be implemented. This means that students will be absent for that period if they have moved. In addition, the participants mentioned that the students who are homeless or highly mobile have challenges even when they are at school, as they often so not have good quality sleep during the night.

Features of partial-day absences were shown in factors categorized in the exo- and macrosystems from the KiTeS framework (Melvin et al., 2019). Having limited access to resources lead to students being absent for a part of a day. For example, students come to school for a free lunch only, meaning that they are only there just before lunchtime. In addition, students who have jobs might miss school partially due to working a night shift or working late. Being legally mandated to attend school, although encouraging school attendance, is, however, not ideal as students might attend without being motivated to want to be in the school system. Low socio-economic status of the family also affect full-day absences, as parents leave for work early and there are no other options for transport if the children miss the school bus. These results show that limited access to resources and low socio-economic status can cause both partial and full-day absences. Given that such risk factors could be mitigated by a number of social services and systems (e.g. transportation, free lunch, community support), the interviews revealed the ability of environmental factors from the exo- and macrosystem to complement or exacerbate family's lack of caring that cause students to miss school.

Aim 2: Research Question 2 - Based on knowledge gained from Aim 1 (quantitative analysis), how are risk and protective factors associated with partial- and full-day absences?

The Aim 1 (quantitative analysis – feature selection) analysis resulted in the identification of 18 variables most associated with unexcused school absences from the 2016 MSS. For Aim 2, those variables were converted into 18 corresponding variables from the 2019 MSS, and then a causal discovery analysis was conducted. The analysis revealed that 1) school engagement affects PDAs directly, 2) PDAs affects FDAs directly, and 3) missing school due to physical illness affects FDAs directly.

The level of the students' engagement with school was measured on a scale that assesses both cognitive and psychological engagement of students based on the six-factors model (Appleton et al., 2006). Better school engagement reduces PDAs and FDAs, which indicates that the level of engagement in school activities affects the time that students are willing to spend in school. For example, a student with a low level of engagement with school will be more likely to leave school early (or arrive late) by actively looking for a reason to miss school than a student who is highly engaged. Furthermore, students who are less engaged and missing school for a part of a day are more likely to be absent for full days, as the results show that PDAs affect FDAs directly. This flow of causal relationships exemplifies the role of PDAs as a 'pre-cursor' of FDAs, as PDAs are a phenomenon that occurs before students start missing full days.

Physically ill students tend to miss more school days than their healthy peers (Weitzman, 1986), which emphasizes the role of schools and school nurses reaching out to those students who are at risk of being chronically absent as they fail to attend the school for an entire day. However, for this analysis, only two out of 15 factors (i.e. school missed due to suspension and physical illness) why students missed school during the last 30 days were used, as the other reasons were not identified in the 18 corresponding variables. Therefore, there is a possibility of other reasons affecting a causal relationship between physically ill students and school absences.

Utilizing the KiTeS framework, the study sought to categorize the factors based on multisystemic hierarchical environments (i.e. micro-, meso-, exo-, macrosystem). The study identified that a factor that directly causes PDAs (i.e. school engagement) is in the mesosystem, while a factor causing FDAs (i.e. missing school due to physical illness) is in the microsystem. An additional step of identifying causal relationships of effect size above 0.05 was conducted to focus and augment the factors interacting between PDAs and FDAs, as the cutoff effect size of 0.1 only revealed one factor that causes each type of absence directly (school engagement causing PDAs, physical illness causing FDAs). The results showed additional causal relationship leading directly to PDAs including 1) ACEs, 2) vape use, and 3) marijuana use along with school engagement, which affects FDA directly. Due to the low ES (< 0.1) of these analyzed causal discovery edges, it could only be assumed that both FDAs and PDAs might be caused directly by factors in the micro- and mesosystems. However, none of the factors from either the exo- or the macrosystems were identified to be directly associated with either type of absence.

Aim 2, results gained from the 2019 MSS were validated with the corresponding 18 variables from the 2016 MSS. The results showed that PDAs directly affect FDAs, with an ES of 0.4, which aligns with what was shown in the 2019 MSS results. The 2016 MSS also showed that FDAs causing ‘students staying home due to sickness,’ which is a reversal of what was shown in the 2019 MSS (i.e. students missing school due to physical illness causing FDA). The cause of this reversal could be as a result of the tailoring of the questions. In the 2016 MSS, the wording of the question (i.e. during the last 30 days, how many times did you stay at home because you were sick?) focused on students missing school for a full day. In contrast, in the 2019 MSS, the question specifically asks that both partial- and full-day absences be included in the consideration (i.e. what are the reasons why you missed a full or part of a day of school in the last 30 days? Illness [feeling physically sick] includes problems with breathing or your teeth).

Aim 2: Research Question 3 - Informed by the perceptions of licensed school nurses, how are the risk and protective factors associated with partial-day absence different or similar to those related to full-day absences?

Phase 1 of Aim 2, which consisted of a qualitative analysis – LSN focus group interview, resulted in four emerging themes related to CA and differences between PDAs and FDAs. These four themes were converted to 10 corresponding variables in the 2019 MSS and used in the causal discovery analysis.

In the result graph (ES > 0.1), a route that starts from 'students missing school due to transportation' to 'FDAs' explains one of the routes of how FDAs are initialized (i.e. transportation problem → sleep disorder → mental issues → FDA). For example, a student might need to rise early because of a lack of transportation to school, which would lead to the student suffering from sleep deprivation during school time and even afterwards. A continued pattern of sleep deprivation could cause students to be unmotivated and indifferent to school activities, which then could eventually lead to the student missing school for a full day. In addition, the graph shows that the 'caring of parental adults' is a direct protective factor for both 'sleep disorder' and 'mental issues.' Notably, among the factors that lead to FDAs, 'caring of parental adults' shows a direct causal relationship to 'sleep disorder' and 'mental issues' but not to 'transportation.' A more detailed explanation of how the caring of parental adults alleviate the students' sleep and mental health issues would need further research, but this result emphasizes the role of family in terms of school absenteeism. Importantly, it also shows that 'parental support' could be an intervention for students missing school for full and possibly partial days (as the results show that PDAs could be caused by both sleep and mental health issues), which would require less local or city expenditure, such as infrastructure rebuilding for problems such as 'transportation.'

In the result graph, a number of directed edges were identified, including 'free and reduced lunch' dispersing to 1) housing instability, 2) students missing school due to transportation, and 3) students' perception of caring of parental adults. In addition, students missing school due to sleep deprivation and mental issues showed an inconclusive causal relation directed to PDAs (See Figure 3.5 and 4.4 for detailed explanation). Therefore, it was not possible to identify definitive factors that directly affect PDAs or factors that are directly affected by 'free and reduced lunch'.

Compared with PDAs, which had no direct causal associated factors, 'students missing school due to mental issues' causally affected FDAs. However, even though the analysis revealed that PDAs are not caused by students missing school for either sleep or mental issues, there is still a possibility that either of these could be a factor causing PDAs given the inconclusive directed edges shown in the Figure 4.4. Conducting an analysis with more variables that related to PDAs and the reasons why students miss school could help to uncover the direct causal relationship between the factors. The PDAs and FDAs showed direct relationships with factors in the microsystem only (i.e. mental issues causing FDA, sleep disorder and mental issues potentially causing PDA), excluding any direct influences of factors in other systems.

Aim 2: Research Question 4 - How are the risk and protective factors associated with partial- and full-day absences based on knowledge gained from Aim 1 (quantitative analysis) combined with the perceptions of licensed school nurses (qualitative analysis), and how are they distinguished compared to research questions 2 and 3 of Aim 2?

The last step of the causal discovery analysis for Aim 2 was to implement an analysis of a combined list of variables used in both the step 1 (using factors from quantitative analysis), and step 2 analyses (using factors from qualitative analysis) from the three-steps causal analysis. This mixed-method approach complemented the study by adding possibly overlooked factors when utilizing data gathered from only qualitative or quantitative studies. Therefore, the approach enabled the study to 1) validate the results of causal discovery analyses, using a different version of the dataset from both the quantitative and qualitative studies, 2) discover any implicit relationship between nodes that were not revealed when using the datasets separately, and 3) better understand and explain connections discovered in the previous steps regarding newly discovered relationships. From the graph portraying the causal relationship of the factors used (ES > 0.1), three factors (i.e. ACEs, school engagement, social competency) were identified to link factors related to substance use (e.g. binge drinking, tobacco product use, marijuana use, vape use, friends' approval of substance use) with all the other school absence-related risk and protective factors used in the analysis. This result contradicts what was revealed in the step 1 analysis of the causal discovery, as step 1 revealed only 'social competency' as the sole link between substance use and other school-related factors. However, to assume that the factor of 'social competency' is the sole link between students' substance use and the school-related factors could be seen as an over-simplification, given the complexity of school dynamics with the diverse bioecological factors known to exist in school settings (Bronfenbrenner & Morris, 2007; Melvin et al., 2019). Therefore, the result seen here illustrates how the mixed method approach complements such potential caveats as opposed to conducting an analysis using only quantitatively or qualitatively derived knowledge.

Focusing on school absences, the graph (ES > 0.1) showed that sleep disorder and school engagement directly affect PDAs, which then escalates to FDAs. Similar with what was mentioned above, a mixed-method approach in this instance complements the results of an approach utilizing only quantitative-derived factors (i.e. step 1 of the three-steps causal analysis), as the step 1 analysis identified only school engagement to directly affect PDAs. The step 2 analysis failed to connect any factors that directly affect PDAs but found that mental illness affects only FDAs directly. In addition to factors directly affecting PDAs, the analysis revealed routes of how those factors affect PDAs.

The factor of 'school engagement' was directly affected by teacher-student relationship, social competency, and friends' approval of substance use. While the associations between substance use and peer influences are well-known topic (Simons-Morton and Chen, 2009), a friends' approval that is specific to 'substance use' directly affecting school engagement is not a causal relationship that is commonly known. Likewise, while the impact of peer approval of the use of substance such as alcohol and marijuana has been assessed (Merianos et al. 2017), the collateral impact of such approval has not been thoroughly studied. And thus, it might be a causal relationship worth investigating. A route of sleep disorder leading to PDAs is also worth noting. An analysis identified that mental issues (caused by ACEs) directly affect students' sleep patterns, which then leads to PDAs. A child who experienced ACEs is highly likely to suffer from mental illness such as depression (Singer et al., 1995; Warne et at., 2017). Studies also suggest that there are strong bidirectional correlations between sleep and mental health issues, such as negative moods or anxiety disorders (Dahl & Harvey, 2007; Short et al., 2019). This study took a further step by providing data-driven evidence of such factors leading to PDAs. Of the many reasons why students miss school found in the 2019 MSS, only the reason of 'sleep issue' directly affects PDAs, with an ES > 0.1, which implies that there is a potential correlation between sleep disorders and PDAs that needs investigating.

Last, all three steps used in the causal discovery analysis consistently and unanimously confirmed the causal relationship of PDAs directly affecting FDAs. This causal relationship was also supported by three separate validation processes (i.e. a process utilizing the 2016 MSS to evaluate the validity of the results from the 2019 MSS) applied to all three steps of the causal discovery analysis, which means that the study strongly implies that students' behavior of missing school for part of a day is likely to lead to them missing school for a full day. This recognition of the role of PDAs in the issue of school absenteeism in relation to the perception of FDAs has been limited in the research to date. To the best of the author's knowledge, the acknowledgement of PDAs as a causing factor of FDAs among all the other school-related factors has not been presented anywhere else. In addition, the results from steps 1 and 3 firmly establish PDAs as the ‘receptor’ for FDAs. That is, PDAs were shown to be the sole link between FDAs and all other school-absences-related risk and protective factors used in the analysis. Since the PDAs affect the FDAs, the results show PDAs act as a ‘receptor’ that advances to FDA, and its impact on the student’s behavior results in FDAs.

Utilizing the KiTeS framework, an additional step of identifying causal relationships of effect sizes greater than 0.05 was conducted to focus and augment the factors that interact between PDAs and FDAs, as the cutoff effect size of 0.1 only revealed two factors that directly cause PDAs and none for FDAs, other than PDAs. The results showed additional causal relationships directly leading to PDA from 1) transportation, 2) mental health issues, and 3) vape use. Three factors, 1) school engagement, 2) mental health issues, and 3) physical illness, directly affected FDAs. Due to the low ES (< 0.1) for these analyzed directed edges, the implication seems to be that both FDAs and PDAs might be mostly directly caused by factors in the micro- and mesosystems. However, it is worth mentioning that a factor within the exosystem (i.e. transportation) directly influenced PDAs (ES = 0.07), whereas none of the factors from either the exo- or macrosystems were directly associated with the FDA. While the extent of the interpretation is limited due to the low ES, the results imply that PDAs cover a wider spectrum of school-absence-related factors past micro- and mesosystems (i.e. factors from exo- or macrosystem such as transportation) then relays that information to FDA (i.e. school-absence-related factors → PDA → FDA).

Limitations

Despite all of its strengths, this mixed method dissertation study has several limitations. It is important to note that, while the study presents causal relationships of factors relating to school absences and other surrounding factors, the methods this study utilized (i.e. causal discovery) ‘infers’ the causal relationship between factors ‘A’ and ‘B,’ given the characteristics and features of the data used. In other words, the causal discovery method calculates and provides an output that ‘A’ shows a high probability of causing ‘B,’ given that the data represents the real world correctly which would lead to an inherent problem when using cross-sectional data. Due to such caveat, a step of validation using other data (i.e. using the 2016 MSS to validate the causal analysis results of the 2019 MSS) is necessary to confirm the relations found. However, an inherent limitation of the data still exists. For example, the structure of questions implemented in the MSS could have affected a pattern of response. Both MSS 2016 and 2019 asks a question regarding school absences but MSS 2019 specifically points out which occasional absences shouldn’t be counted whereas for MSS 2016 it’s comparatively vague. These subtle differences could lead to student reporting the data slightly different based on how they’ve understood the questions.

In addition, the potential risk of misleading results (e.g. overfitting the problem while training machine learning predictive models) from imperfect data, such as class imbalance and missing data, were considered in advance, followed by pre-emptive steps (SMOTE resampling method for class imbalance, CART imputation for missing data). However, there is an inherent limitation in these steps, as imputed and resampled data cannot completely substitute the quality of the real-world data. Using resampled data only in the prediction model performance comparison could also be seen as an additional limitation. Original data without the resampling method was used for the feature selection method, followed by the prediction model performances comparison with both the original and resampled data. While this was because a skewed ratio of an outcome variable would mostly impact the performance of the prediction model, conducting an additional analysis in the feature selection process utilizing the resampled dataset along with the original dataset might yield additional knowledge.

It is important to acknowledge that of the components used in this dissertation study, unexcused absence was used as an outcome in Aim 1 whereas both unexcused and excused absences combined was used (i.e. CA) in Aim 2. This was because of how the data (i.e. MSS 2016, 2019) were gathered. Both MSS 2016 and 2019 asked question to students using questionnaire specifically excluding excused absences which means using those data would pertain to study about unexcused absences. Using term 'truancy – unexcused absences' also was not ideal as focus interview and the mixed-method approach were conducted with the usage of the term 'CA'. Only the Aim 1 of the study and causal discovery analysis step 1 used the outcome variable of unexcused absences while the rest utilized the definition of CA and knowledge gained from such (i.e. focus interview) in the causal discovery analyses. In addition, PDA and FDA are not related to any specific definitions used in the field of school absences (i.e. unexcused and excused absences) which lead the study to generally describe students missing school as 'school absences' or 'school absenteeism' except for the Aim 1 and the Step 1 of causal discovery analyses.

Differences in content between the two MSS datasets also affected the results, especially when examining the impact of 'school nurse' in the school setting. Even though the factors identification process in Aim 1 and the focus group interview in Aim 2 both included the component of 'school nurse,' this question was removed in the 2019

MSS. Therefore, the nursing implication from the causal discovery part (i.e. Phase 2 of Aim 2) were contextual, supported by results of Aim 1 (factors identification using the MSS 2016) and phase 1 of Aim 2 (LSN focus group interview).

In the causal discovery analysis ES, the study used a cutoff value 0.1 of Pearson's *r*. The value of 0.1 represents a small correlational effect size between A and B (Cohen, 1992). In other words, a directional edge between A and B with an ES higher than 0.1 would mean that there is at least a 'small' effect size of correlation between two factors. It is important to note that with the size of data used for this dissertation study, it still only revealed small effect sizes which implies the limitation of the dataset used.

As this study has the limitations mentioned above, such as the data characteristics and analytic caveats, future studies focusing on PDAs with data from elsewhere will help to validate and expand the boundary of knowledge that pertains to PDAs. For example, data with different timelines to corroborate causal relationships identified in this study is needed to validate this claim. In addition, locating data with a specific variable of interest (e.g. school-nurse-related variable) would help to analyze PDAs and their relationship to surrounding factors, with the variable depending on the focus of the study.

Implications

Recommendations for Future Research

This study has identified a number of factors affecting school absences with a number of interconnections among risk and protective factors including PDAs directly driving FDAs. The recognition of PDAs in the field of school absenteeism is difficult to detect; therefore, most of the CA studies were conducted without distinguishing PDAs from FDAs, treating both as school absences in general (i.e. no further definition than student missing school for X days). This vague conceptualization of school absences in previous research creates a potential risk for the future studies to 1) misinterpret the result of school absences related exploratory or intervention efficacy studies, 2) miscommunicate to readers and other researchers conducting school-absences-related studies, and 3) ineffectively apply the knowledge identified in the real world.

Furthermore, additional studies focusing on PDAs are needed to fully understand the phenomena of PDAs compared to FDAs. One study brought attention to PDAs by conducting a descriptive analysis with demographic variables (age, minority status) to both PDAs and FDAs, which emphasized how the effects of PDAs and FDAs are noticeably different in terms of age, minority status, and whether the student missing school has been excused or not (Whitney and Liu, 2017). However, to the author's best knowledge, the PDAs' causal relationship to FDAs, in addition to a different pattern of interconnection with other school-related factors, has not been acknowledged before.

Finally, utilizing a data that adhere to a definition of CA (unexcused and excused absences combined) will help to better understand the interconnections of risk and protective factor and CA but not limiting itself to unexcused absences. This study complemented such limitation by conducting a focus-interview with the definition of CA and using those knowledges in the course of causal analyses. However, focusing solely on CA by having such a data will help to better corroborate what's identified in this study.

Recommendations for School Nurses

The focus-group interview findings indicated current limitations of LSN's recognition in school. Specifically, LSNs described their role mostly dealings with physically ill or injured students. When applied to school absenteeism, results of the LSN focus-group interview pointed out that LSNs deal with chronically ill students who miss school but not the students who miss school chronically in general. Studies on school nurses and students chronically absent due to chronic 'health' issues (e.g. diabetes, asthma) are widely available (Allen et al., 2018; Kearney & Bensaheb, 2006; Rodriguez et al., 2013). But such availability is not the case for studies on correlations between school nurse and school absences-related 'risk and protective' factors (e.g. substance use, suspension, socio-economic status) which validates a current recognition on school nurse in terms of school absenteeism. In addition, the role of the school nurse has not been recognized as being 'essential' for at-risk students who are dealing with school absenteeism, which limits their involvement to students with chronic illnesses only (Jacobsen et al., 2016). Data characteristics used for the study reflects such limitation as well (i.e. LSN-related questionnaire removed in the MSS 2019).

However, data from the focus group interviews with LSNs shows the impact of school nurses transcending their current focus of supporting only 'physically ill' children. In the interviews, LSN participants mentioned that they were actively assessing at-risk children during their work day, using their offices as a 'safe zone' for students who do not want to be in the classroom, or if they need food because of their limited resources. Furthermore, some studies show that the impact of school nurses in terms of absenteeism is critical (Allen et al., 2018; Jacobsen et al., 2016). In addition, a prior study implemented by the student PI, a causal discovery analysis to the MSS 2016, discovered the school nurse office visit to be directly caused by suspension, PDA, substance usage, and mental health issue then transpiring to a student staying home due to sickness (Lee et al., 2021). These findings demonstrate the linkage between LSN and risk and protective factors related to school absences from a data-driven approach standpoint. They also emphasize the crucial role of LSN in terms of school absenteeism.

There is a shortage of school nurses (i.e. nurse to student ratio) while the demand for advanced nursing care in school setting is increasing for the students who attend school with chronic health conditions (Dolatowski et al., 2015). Such trends limit LSNs current role to mainly supporting physically ill students; however, it might be worth doing an initial assessment of how a full-time school nurse with an adequate student to nurse ratio could function as a point person in terms of school absenteeism. Also, the results of this dissertation study suggest that future researcher focus on locating or acquiring school absences data such as the MSS that has a sufficient amount of LSN-related questions or measures to implement data-driven research that captures the linkage between LSN and school absences with its surrounding risk and protective factors.

Conclusions

This dissertation study is unique and innovative because it utilized quantitative (i.e. feature selection, prediction models comparison, causal discovery) and qualitative (i.e. focus-group interview) methods to conduct a mixed-method study, which enabled the identification of factors and interconnections regarding school absenteeism. Quantitative methods identified factors closely associated with school absences and causal interconnections between those identified factors. The LSNs focus group interview also revealed factors associated with CA and the difference between PDAs and FDAs. In addition, the interviews revealed the important role of the school nurse in terms of alleviating school absenteeism. In the mixed-method approach, causal discovery analysis was used to connect both quantitative and qualitative approaches, and, as a result, this dissertation research was able to identify unique connections between factors and school absences.

To date, the studies in the field of CA have failed to acknowledge PDAs as a key component. The results from this dissertation highlights the potential risk of misinterpreting the CA-related results when PDAs are not incorporated and thus exemplifies the need for greater appreciation on PDAs. In addition, the KiTeS hierarchical multisystem approach helped the study to confirm in which part of social determinants of health factors are related with school absences. This study revealed the proximity between school absenteeism with risk and protective factors within the micro- and mesosystems, compared to factors in the exo- and macrosystems which calls the need for attention on factors within micro- and mesosystem to alleviate students missing school.

Results of the focus group interview and previous study done by student-PI showed the importance of LSNs which transcends their current focus of supporting only ‘physically ill’ children. Such findings call the attention for a research focusing on potential impact of LSNs in schools to school absenteeism, especially when current obstacles LSNs facing (i.e. workload, nurse to student ratio, inadequate level of attention to the role of LSNs) are alleviated.

Finally, the study acknowledged the crucial role of the causal discovery method. When analyzing a list of 113 ranked factors associated with school absenteeism, a causal discovery method returned a purely data-driven output that helped to locate the pivotal chain-links where stakeholders can intervene. As the quality and characteristics of data is critical in any data-driven study, utilizing a causal discovery method with different data with the component of PDAs will help to identify, validate, and expand knowledge gleaned from this study.

References

Aguiar, B. V. K., Guerreiro, M. M., McBrian, D., & Montenegro, M. A. (2007). Seizure impact on the school attendance in children with epilepsy. *Seizure*, *16*(8), 698-702.

Allen, C. W., Diamond-Myrsten, S., & Rollins, L. K. (2018). School absenteeism in children and adolescents. *American family physician*, *98*(12), 738-744.

Anker, J. J., Kummerfeld, E., Rix, A., Burwell, S. J., & Kushner, M. G. (2019). Causal network modeling of the determinants of drinking behavior in comorbid alcohol use and anxiety disorder. *Alcoholism: clinical and experimental research*, *43*(1), 91-97.

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of school psychology*, *44*(5), 427-445.

Archambault, I., Janosz, M., Goulet, M., Dupéré, V., & Gilbert-Blanchard, O. (2019). Promoting student engagement from childhood to adolescence as a way to improve positive youth development and school completion. In *Handbook of Student Engagement Interventions* (pp. 13-29). Academic Press.

Attendance Works and Everyone Graduates Center. (2018). *Strategies for school sites*. http://www.attendanceworks.org/chronic-absence/addressing-chronic-absence/strategies-for-school-sites/

Bahali, K., Tahiroglu, A. Y., Avci, A., & Seydaoglu, G. (2011). Parental psychological symptoms and familial risk factors of children and adolescents who exhibit school refusal. *East Asian Archives of Psychiatry*, *21*(4), 164.

Balkis, M., Arslan, G., & Duru, E. (2016). The School Absenteeism among High School Students: Contributing Factors. *Educational Sciences: Theory and Practice*, *16*(6), 1819-1831.

Basch, C. E. (2011). Asthma and the achievement gap among urban minority youth. *Journal of School Health*, *81*(10), 606-613.

Black, L. I., & Zablotsky, B. (2018). Chronic School Absenteeism among Children with Selected Developmental Disabilities: National Health Interview Survey, 2014-2016. National Health Statistics Reports. Number 118. *National Center for Health Statistics*.

Bottini, C. L. (2017). *The effects of the student success skills classroom program on self-regulation, school attendance, and test anxiety on Hispanic fifth-grade students*. [Doctoral dissertation, Florida Atlantic University]. Proquest Dissertations and Theses Global.

Bowers, J. M., & Moyer, A. (2017). Effects of school start time on students' sleep duration, daytime sleepiness, and attendance: a meta-analysis. *Sleep health*, *3*(6), 423-431.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101.

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). Classification and Regression Trees, new edition [cart93]. *Monterey, CA: Wadsworth and Brooks*.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American psychologist*, *32*(7), 513.

Bronfenbrenner, U., & Morris, P. A. (2007). The bioecological model of human development. *Handbook of child psychology*, *1*.

Burgette, L. F., & Reiter, J. P. (2010). Multiple imputation for missing data via sequential regression trees. *American journal of epidemiology*, *172*(9), 1070-1076.

Carifio, J., & Perla, R. J. (2007). Ten common misunderstandings, misconceptions, persistent myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of social sciences*, *3*(3), 106-116.

Carless, B., Melvin, G. A., Tonge, B. J., & Newman, L. K. (2015). The role of parental self-efficacy in adolescent school-refusal. *Journal of Family Psychology*, *29*(2), 162.

Center for Causal Discovery. (2016). Greedy Causal Inference (GFCI) Algorithm for Continuous Variables. Retrieved from URL https://www.ccd.pitt.edu/wiki/index.php?title=Greedy_Fast_Causal_Inference_(GFCI)_Algorithm_for_Continuous_Variables&oldid=665

Childs, J., & Lofton, R. (2021). Masking attendance: How education policy distracts from the wicked problem (s) of chronic absenteeism. *Educational Policy*, *35*(2), 213-234.

Cicutto, L., To, T., & Murphy, S. (2013). A randomized controlled trial of a public health nurse-delivered asthma program to elementary schools. *Journal of School Health, 83*(12), 876- 884.

Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155.

Cook, P. J., Dodge, K. A., Gifford, E. J., & Schulting, A. B. (2017). A new program to prevent primary school absenteeism: Results of a pilot study in five schools. *Children and Youth Services Review*, *82*, 262-270.

Dahl, R. E., & Harvey, A. G. (2007). Sleep in children and adolescents with behavioral and emotional disorders. *Sleep medicine clinics*, *2*(3), 501-511.

Dahlgren, G., & Whitehead, M. (1993, September). Tackling inequalities in health: what can we learn from what has been tried. Working paper prepared for the king's fund international seminar on tackling inequalities in health.

Deck, S. M. (2017). School outcomes for homeless children: Differences among sheltered, doubled-up, and poor, housed children. *Journal of Children and Poverty*, *23*(1), 57-77.

Demographics of South Korea. (2021, May 16). In *Wikipedia*. Retrieved from URL https://en.wikipedia.org/w/index.php?title=Demographics_of_South_Korea&oldid=1023431833

Dodgson, J. E. (2019). Reflexivity in qualitative research. *Journal of Human Lactation*, *35*(2), 220-222.

Dolatowski, R., Endsley, P., Hiltz, C., Johansen, A., Maughan, E., Minchella, L., & Trefry, S. (2015). School Nurse Workload: Staffing for Safe Care. Position Statement. *National Association of School Nurses (NASN)*.

Doren, B., Murray, C., & Gau, J. M. (2014). Salient predictors of school dropout among secondary students with learning disabilities. *Learning Disabilities Research & Practice*, *29*(4), 150-159.

Dy, J. G., & Brodley, C. E. (2004). Feature selection for unsupervised learning. *Journal of machine learning research*, *5*(Aug), 845-889.

Eberhardt, F. (2017). Introduction to the foundations of causal discovery. *International Journal of Data Science and Analytics*, *3*(2), 81-91.

Egger, H. L., Costello, J. E., & Angold, A. (2003). School refusal and psychiatric disorders: A community study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*(7), 797-807.

Eklund, K., Burns, M. K., Oyen, K., DeMarchena, S., & McCollom, E. M. (2020). Addressing chronic absenteeism in schools: A meta-analysis of evidence-based interventions. *School Psychology Review*, 1-17.

Erlingsson, C., & Brysiewicz, P. (2017). A hands-on guide to doing content analysis. *African Journal of Emergency Medicine*, *7*(3), 93-99.

Fernández, A., Garcia, S., Herrera, F., & Chawla, N. V. (2018). SMOTE for learning from imbalanced data: progress and challenges, marking the 15-year anniversary. *Journal of artificial intelligence research*, *61*, 863-905.

Fornander, M. J., & Kearney, C. (2019). Family environment variables as predictors of school absenteeism severity at multiple levels: ensemble and classification and regression tree analysis. *Frontiers in Psychology*, *10*, 2381.

Freeman, J., Simonsen, B., McCoach, D. B., Sugai, G., Lombardi, A., & Horner, R. (2015). An analysis of the relationship between implementation of school-wide positive behavior interventions and supports and high school dropout rates. *The High School Journal*, 290-315.

Gakh, M., Coughenour, C., Assoumou, B. O., & Vanderstelt, M. (2020). The relationship between school absenteeism and substance use: An integrative literature review. *Substance use & misuse*, *55*(3), 491-502.

Gase, L. N., Kuo, T., Coller, K., Guerrero, L. R., & Wong, M. D. (2014). Assessing the connection between health and education: identifying potential leverage points for public health to improve school attendance. *American journal of public health*, *104*(9), e47-e54.

Gottfried, M. A. (2014a). Can neighbor attributes predict school absences?. *Urban Education*, *49*(2), 216-250.

Gottfried, M. A. (2014b). Chronic absenteeism and its effects on students' academic and socioemotional outcomes. *Journal of Education for Students Placed at Risk (JESPAR)*, *19*(2), 53-75.

Gottfried, M. A. (2017). Linking getting to school with going to school. *Educational Evaluation and Policy Analysis*, *39*(4), 571-592.

Gottfried, M. A., & Gee, K. A. (2017). Identifying the determinants of chronic absenteeism: A bioecological systems approach. *Teachers College Record*, *119*(7), 1-34.

Gottfried, M. A. (2019). Chronic absenteeism in the classroom context: Effects on achievement. *Urban Education*, *54*(1), 3-34.

Gottfried, M., Stiefel, L., Schwartz, A., & Hopkins, B. (2019). Showing up: Disparities in chronic absenteeism between students with and without disabilities in traditional public schools. *Teachers College Record*, *121*(8), 1-34.

Graitcer, S. B., Dube, N. L., Basurto-Davila, R., Smith, P. F., Ferdinands, J., Thompson, M., ... & Shay, D. K. (2012). Effects of immunizing school children with 2009 influenza A (H1N1) monovalent vaccine on absenteeism among students and teachers in Maine. *Vaccine*, *30*(32), 4835-4841.

Green, J., Liem, G. A. D., Martin, A. J., Colmar, S., Marsh, H. W., & McInerney, D. (2012). Academic motivation, self-concept, engagement, and performance in high school: Key processes from a longitudinal perspective. *Journal of adolescence*, *35*(5), 1111-1122.

Guralnick, M. J. (2017). Early intervention for children with intellectual disabilities: An update. *Journal of Applied Research in Intellectual Disabilities*, *30*(2), 211-229.

Hancock, K. J., Shepherd, C. C., Lawrence, D., & Zubrick, S. R. (2013). Student attendance and educational outcomes: Every day counts. *Canberra: Department of Education, Employment and Workplace Relations*.

Henderson, T., Hill, C., & Norton, K. (2014). The connection between missing school and health: A review of chronic absenteeism. *Portland, OR: Upstream Public Health*.

Hendron, M., & Kearney, C. A. (2016). School climate and student absenteeism and internalizing and externalizing behavioral problems. *Children & Schools*, *38*(2), 109-116.

Henry, K. L., Knight, K. E., & Thornberry, T. P. (2012). School disengagement as a predictor of dropout, delinquency, and problem substance use during adolescence and early adulthood. *Journal of youth and adolescence*, *41*(2), 156-166.

Hillier, W. (2021, March 23). What Is Interval Data?. Careerfoundry. Retrieved from URL https://careerfoundry.com/en/blog/data-analytics/what-is-interval-data/

Hooper, D., Coughlan, J., & Mullen, M. (2008, June). Evaluating model fit: a synthesis of the structural equation modelling literature. In *7th European Conference on research methodology for business and management studies* (pp. 195-200).

Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, *5*(2), 1.

Hysing, M., Haugland, S., Stormark, K. M., Bøe, T., & Sivertsen, B. (2015). Sleep and school attendance in adolescence: results from a large population-based study. *Scandinavian journal of public health*, *43*(1), 2-9.

Ingul, J. M., Klöckner, C. A., Silverman, W. K., & Nordahl, H. M. (2012). Adolescent school absenteeism: modelling social and individual risk factors. *Child and Adolescent Mental Health*, *17*(2), 93-100.

Iorfino, F., Ho, N., Carpenter, J. S., Cross, S. P., Davenport, T. A., Hermens, D. F., ... & Hickie, I. B. (2020). Predicting self-harm within six months after initial presentation to youth mental health services: A machine learning study. *PLoS one*, *15*(12), e0243467.

Jacobsen, K., Meeder, L., & Voskuil, V. R. (2016). Chronic student absenteeism: The critical role of school nurses. *NASN School Nurse*, *31*(3), 178-185.

Kearney, C. A. (2016). *Managing school absenteeism at multiple tiers: An evidence-based and practical guide for professionals*. Oxford University Press.

Kearney, C. A., & Bensaheb, A. (2006). School absenteeism and school refusal behavior: A review and suggestions for school-based health professionals. *Journal of school health*, *76*(1), 3-7.

King Jr, J. C., Beckett, D., Snyder, J., Cummings, G. E., King, B. S., & Magder, L. S. (2012). Direct and indirect impact of influenza vaccination of young children on school absenteeism. *Vaccine*, *30*(2), 289-293.

Kohavi, R. (1995, August). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai* (Vol. 14, No. 2, pp. 1137-1145).

Krawczyk, B. (2016). Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence*, *5*(4), 221-232.

Kummerfeld, E., Rix, A., Anker, J. J., & Kushner, M. G. (2019). Assessing the collective utility of multiple analyses on clinical alcohol use disorder data. *Journal of the American Medical Informatics Association*, *26*(10), 1046-1055.

Langley, P. (1994, November). Selection of relevant features in machine learning. In *Proceedings of the AAAI Fall symposium on relevance* (Vol. 184, pp. 245-271).

Lenhoff, S. W., & Pogodzinski, B. (2018). School organizational effectiveness and chronic absenteeism: Implications for accountability. *Journal of Education for Students Placed at Risk (JESPAR)*, *23*(1-2), 153-169.

Lee, K., Kummerfeld, E., McMorris, B. J., Delaney, C. W. Comparison of Interconnections between Full-Day and Partial-Day Absences in School Children: A Causal Discovery Analysis. [Poster Presentation]. Midwest Nursing Research Society 2021 Annual Conference, Virtual due to COVID-19. 2021.

Li, Y., Raychowdhury, S., Tedders, S. H., Lyn, R., Lòpez-De Fede, A., & Zhang, J. (2012). Association between increased BMI and severe school absenteeism among US children and adolescents: findings from a national survey, 2005–2008. *International Journal of Obesity*, *36*(4), 517-523.

Maughan, E. (2003). The impact of school nursing on school performance: A research synthesis. *The Journal of School Nursing*, *19*(3), 163-171.

Melvin, G. A., Heyne, D., Gray, K. M., Hastings, R. P., Totsika, V., Tonge, B. J., & Freeman, M. M. (2019, June). The Kids and Teens at School (KiTeS) framework: An inclusive bioecological systems approach to understanding school absenteeism and school attendance problems. In *Frontiers in Education* (Vol. 4, p. 61). Frontiers.

Merianos, A. L., Rosen, B. L., Montgomery, L., Barry, A. E., & Smith, M. L. (2017). Impact of perceived risk and friend influence on alcohol and marijuana use among students. *The Journal of School Nursing*, *33*(6), 446-455.

Meng, Y. Y., Babey, S. H., & Wolstein, J. (2012). Asthma-related school absenteeism and school concentration of low-income students in California. *Preventing chronic disease*, *9*.

Milton, C. L. (2017). The ethics of big data and nursing science. *Nursing science quarterly*, *30*(4), 300-302.

Minnesota Department of Education. (n.d.). *Minnesota Student Survey.* Retrieved from URL https://education.mn.gov/MDE/dse/health/mss/

Minnesota Department of Health. (October 2019*). 2019 Minnesota Student Survey Results Released.* Retrieved from URL https://www.health.state.mn.us/news/pressrel/2019/studentsurvey101719.html

Mittelstadt, B., & Floridi, L. (2016). The ethics of big data: Current and foreseeable issues in biomedical contexts. *Science of Engineering Ethics, 22,* 303-341.

Parent, K. B., Wodrich, D. L., & Hasan, K. S. (2009). Type 1 diabetes mellitus and school: a comparison of patients and healthy siblings. *Pediatric diabetes*, *10*(8), 554-562.

Pernice-Duca, F. (2010). Family network support and mental health recovery. *Journal of Marital and Family Therapy*, *36*(1), 13-27.

Rafa, A. (2017). Chronic Absenteeism: A Key Indicator of Student Success. Policy Analysis. *Education Commission of the States*.

Ramsey, J. D. (2015). Scaling up greedy causal search for continuous variables. *arXiv preprint arXiv:1507.07749*.

Rappaport, E. B., Daskalakis, C., & Andrel, J. (2011). Obesity and other predictors of absenteeism in Philadelphia school children. *Journal of School Health*, *81*(6), 341-344.

Richardson, K. L., Weiss, N. S., & Halbach, S. (2018). Chronic school absenteeism of children with chronic kidney disease. *The Journal of pediatrics*, *199*, 267-271.

Robert Wood Johnson Foundation. (*2016). The relationship between school attendance and health.* Retrieved from URL https://www.rwjf.org/content/dam/farm /reports/issue_briefs/2016/rwjf431726

Rodriguez, E., Rivera, D. A., Perlroth, D., Becker, E., Wang, N. E., & Landau, M. (2013). School nurses' role in asthma management, school absenteeism, and cost savings: A demonstration project. *Journal of School Health*, *83*(12), 842-850.

Rogers, T., & Feller, A. (2018). Reducing student absences at scale by targeting parents' misbeliefs. *Nature Human Behaviour*, *2*(5), 335-342.

Ross, A., Kwon, J. Y., Kulinna, P. H., & Searle, M. (2019). Active transportation: the role of parent attitude, the physical environment, and social capital. *Journal of Physical Activity and Health*, *16*(1), 60-67

Sexson, S., & Madan-Swain, A. (1995). The chronically ill child in the school. *School Psychology Quarterly*, *10*(4), 359.

Short, M. A., Bartel, K., & Carskadon, M. A. (2019). Sleep and mental health in children and adolescents. In *Sleep and health* (pp. 435-445). Academic Press.

Simons-Morton, B., & Chen, R. (2009). Peer and parent influences on school engagement among early adolescents. *Youth & society*, *41*(1), 3-25.

Simpkins, S. D., Vest, A. E., Delgado, M. Y., & Price, C. D. (2012). Do school friends participate in similar extracurricular activities?: Examining the moderating role of race/ethnicity and age. *Journal of Leisure Research*, *44*(3), 332-352.

Singer MI, Anglin TM, Song Ly, Lunghofer L: Adolescents' exposure to violence and associated symptoms of psychological trauma. JAMA. 1995, 273: 477-482. 10.1001/jama.273.6.477.

Skedgell, K., & Kearney, C. A. (2018). Predictors of school absenteeism severity at multiple levels: A classification and regression tree analysis. *Children and Youth Services Review*, *86*, 236-245.

Spirito, A., Hernandez, L., Cancilliere, M. K., Graves, H. R., Rodriguez, A. M., Operario, D., ... & Barnett, N. P. (2018). Parent and adolescent motivational enhancement intervention for substance-using, truant adolescents: a pilot randomized trial. *Journal of Clinical Child & Adolescent Psychology*, *47*(sup1), S467-S479

South Carolina Department of Education. (n.d.). *Chronic Absenteeism*. Retrieved from: https://ed.sc.gov/districts-schools/student-intervention-services/chronic-absenteeism/

State of New Jersey Department of Education. (May 2020). *Guidance for Reporting Student Absences and Calculating Chronic Absenteeism.* Retrieved from: https://www.state.nj.us/education/students/safety/behavior/attendance/

State of Pennsylvania Department of Education. (n.d.). *Guidelines for Reporting Regular Attendance and Chronic Absenteeism.* Retrieved from: https://www.education.pa.gov/K-12/ESSA/FutureReady/Pages/Attendance.aspx

Stempel, H., Cox-Martin, M., Bronsert, M., Dickinson, L. M., & Allison, M. A. (2017). Chronic school absenteeism and the role of adverse childhood experiences. *Academic pediatrics*, *17*(8), 837-843.

Tang, J., Alelyani, S., & Liu, H. (2014). Feature selection for classification: A review. *Data classification: Algorithms and applications*, 37.

Tanner-Smith, E. E., & Fisher, B. W. (2016). Visible school security measures and student academic performance, attendance, and postsecondary aspirations. *Journal of youth and adolescence*, *45*(1), 195-210.

Teasley, M. L. (2004). Absenteeism and truancy: Risk, protection, and best practice implications for school social workers. *Children & Schools*, *26*(2), 117-128.

Thikkurissy, S., Glazer, K., Amini, H., Casamassimo, P. S., & Rashid, R. (2012). The comparative morbidities of acute dental pain and acute asthma on quality of life in children. *Pediatric dentistry*, *34*(4), 77E-80E.

United States Census Bureau. (2019). *U.S. Census Bureau Quickfacts: MN.* Retrieved from https://www.census.gov/quickfacts/MN

U.S Department of Education. (2016a). *Chronic Absenteeism in The Nation's Schools.* https://www2.ed.gov/datastory/chronicabsenteeism.html#intro

U.S Department of Education. (2016b). *New Data Show Chronic Absenteeism is Widespread and Prevalent Among All Student Groups.* https://www.ed.gov/news/press-releases/new-data-show-chronic-absenteeism-widespread-and-prevalent-among-all-student-groups

Walczak, B., & Massart, D. L. (2001). Dealing with missing data: Part I. *Chemometrics and intelligent laboratory systems*, *58*(1), 15-27.

Warne, D., Dulacki, K., Spurlock, M., Meath, T., Davis, M. M., Wright, B., & McConnell, K. J. (2017). Adverse childhood experiences (ACE) among American Indians in South Dakota and associations with mental health conditions, alcohol use, and smoking. *Journal of health care for the poor and underserved*, *28*(4), 1559-1577.

Weismuller, P. C., Grasska, M. A., Alexander, M., White, C. G., & Kramer, P. (2007). Elementary school nurse interventions: Attendance and health outcomes. *The Journal of School Nursing*, *23*(2), 111-118.

Weitzman, M. (1986). School absence rates as outcome measures in studies of children with chronic illness. *Journal of chronic diseases*, *39*(10), 799-808.

Wilkins, K., & Woodgate, R. (2008). Designing a mixed methods study in pediatric oncology nursing research. *Journal of Pediatric Oncology Nursing*, *25*(1), 24-33.

Whitney, C. R., & Liu, J. (2017). What we're missing: A descriptive analysis of part-day absenteeism in secondary school. *AERA Open*, *3*(2), 2332858417703660.

Wongkamthong, C., & Akande, O. (2020). A Comparative Study of Imputation Methods for Multivariate Ordinal Data. *arXiv preprint arXiv:2010.10471*.

Wood, J. J., Lynne-Landsman, S. D., Langer, D. A., Wood, P. A., Clark, S. L., Mark Eddy, J., & Ialongo, N. (2012). School attendance problems and youth psychopathology: Structural cross-lagged regression models in three longitudinal data sets. *Child development*, *83*(1), 351-366.

Woodman, A. C., Mawdsley, H. P., & Hauser-Cram, P. (2015). Parenting stress and child behavior problems within families of children with developmental disabilities: Transactional relations across 15 years. *Research in developmental disabilities*, *36*, 264-276.

Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PloS one*, *14*(11), e0224365.

Van Buuren, S. (2018). *Flexible imputation of missing data*. CRC press.

Van Eck, K., Johnson, S. R., Bettencourt, A., & Johnson, S. L. (2017). How school climate relates to chronic absence: A multi–level latent profile analysis. *Journal of School Psychology*, *61*, 89-102.

Yu, L., & Liu, H. (2003). Feature selection for high-dimensional data: A fast correlation-based filter solution. In *Proceedings of the 20th international conference on machine learning (ICML-03)* (pp. 856-863).

Appendix A

Tables in the Chapter III

Table A1

Measures of 2016 Minnesota Student Survey Factors within the Microsystem

Variable	Item(s)	Response options	Description (Cronbach's α)
Grade	One item "What is your grade in school right now?"	8 – grade 8 9 – grade 9 11 – grade 11	
Biological sex	One item "What is your biological sex?"	1 – Male 2 – Female	Converted into two dummy variables with 0 and 1 answer
Special education	One item "Do you have an IEP or special education services?"	1 – Yes 0 – No	
Staying home due to sickness	One item "During the last 30 days, how many times have you stayed home because you were sick?"	1 – None 2 – Once or twice 3 – 3 to 5 times 4 – 6 to 9 times 5 – 10 or more times	
General health	One item "How would you describe your health in general?"	1 – Excellent 2 – Very good 3 – Good 4 – Fair 5 – Poor	
Medical checkup	One item "When was the last time you saw a doctor or nurse for a check-up or physical exam when you were not sick or injured?"	1 – During the last year 2 – 1~2 years ago 3 – More than 2 years ago 4 – Never	Dichotomized 1 – Any 0 – None
Dental checkup	One item "When was the last time you saw a dentist or dental hygienist for a regular check-up, exam or teeth cleaning or other dental work?"	1 – During the last year 2 – 1~2 years ago 3 – More than 2 years ago 4 – Never	Dichotomized 1 – Any 0 – None
Physical disabilities	One item "Do you have any physical disabilities, or long-term health problems (such as asthma, cancer, diabetes, epilepsy or something else)? Long-term means lasting 6 months or more."	1 – Yes 0 – No	

Variable	Item(s)	Response options	Description (Cronbach's α)
Long-term mental health history	One item "Do you have any long-term mental health, behavioral or emotional problems? Long-term means lasting 6 months or more"	1 – Yes 0 – No	
Mental health treatment history	Three items "Have you ever been treated for a mental health, emotional or behavioral problem: No?"	1 – Checked, 0 – Not checked	Dichotomized 1 – Any 0 – None
	"Have you ever been treated for a mental health, emotional or behavioral problem: Yes, during the last year?"	1 – Checked, 0 – Not checked	
	"Have you ever been treated for a mental health, emotional or behavioral problem: Yes, more than a year ago?"	1 – Checked, 0 – Not checked	
Substance use treatment history	Three items "Have you ever been treated for an alcohol or drug problem: No?"	1 – Checked, 0 – Not checked	Dichotomized 1 – Any 0 – None
	"Have you ever been treated for an alcohol or drug problem: Yes, during the last year?"	1 – Checked, 0 – Not checked	
	"Have you ever been treated for an alcohol or drug problem: Yes, more than a year ago?"	1 – Checked, 0 – Not checked	
Physically Active	One item "During the last 7 days, on how many days were you physically active for a total of AT LEAST 60 MINUTES PER DAY?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 days 5 – 4 days 6 – 5 days 7 – 6 days 8 – 7 days	
History of asthma	One item "Has a doctor or nurse ever told you that you have asthma?"	1 – Yes 0 – No	
History of diabetes	One item "Has a doctor or nurse ever told you that you have an allergy that requires you to carry an epi-pen?"	1 – Yes 0 – No	
Sleep during school day	One item "During a typical school night, how many hours of sleep do you get?"	1 – 4 hours or less 2 – 5 hours 3 – 6 hours 4 – 7 hours 5 – 8 hours 6 – 9 hours	

Variable	Item(s)	Response options	Description (Cronbach's α)
		7 – 10 or more hours	
Positive Identity Scale	Six items "I feel in control of my life and future." "I feel good about myself." "I feel good about my future." "I deal with disappointment without getting too upset." "I find good ways to deal with things that are hard in my life." "I am thinking about what my purpose is in life."	1 – Not at all or rarely 2 – Somewhat or sometimes 3 – Very or often 4 – Extremely or almost always	Converted to mean value of all items ranging from 1 to 4 (0.84)
Social competency scale (SCS)	Eight items "I say no to things that are dangerous or unhealthy." "I build friendships with other people." "I express my feelings in proper ways." "I plan ahead and make good choices." "I stay away from bad influences." "I resolve conflicts without anyone getting hurt." "I accept people who are different from me." "I am sensitive to the needs and feelings of others."	1 – Not at all or rarely 2 – Somewhat or sometimes 3 – Very or often 4 – Extremely or almost always	Converted to mean value of all items ranging from 1 to 4 (0.84)
Empowerment	Three items "I feel valued and appreciated by others." "I am included in family tasks and decisions. " "I am given useful roles and responsibilities."	1 – Not at all or rarely 2 – Somewhat or sometimes 3 – Very or often 4 – Extremely or almost always	Converted to mean value of all items ranging from 1 to 4 (0.81)
Risky behavior while driving	Three items "When driving a car, how often do you wear a seat belt?" "When driving a car, how often do you send or read text messages or emails?" "When driving a car, how often do you make or answer a phone call?"	1 – I don't drive a car 2 – I never do this 3 – Sometimes 4 – Often 5 - Always	Converted to mean value of all items ranging from 1 to 5 (0.83)
Patient health questionnaire-2 (PHQ2)	Two items "Over the last 2 weeks, how often have you been bothered by little interest or pleasure in doing things?"	1 – Not at all 2 – Several days 3 – More than half the days	Dichotomized 0 – Never to several days 1 – More than half the days to

Variable	Item(s)	Response options	Description (Cronbach's α)
	"Over the last 2 weeks, how often have you been bothered by feeling down, depressed or hopeless?"	4 – Nearly every day	nearly every day (0.71)
Global appraisal of individual needs (GAIN)	Five items "During the last 12 months, did you do any of the following TWO OR MORE TIMES: lie or con to get things you wanted or to avoid having to do something?" "During the last 12 months, did you do any of the following TWO OR MORE TIMES: have a hard time paying attention at school, work or home?" "During the last 12 months, did you do any of the following TWO OR MORE TIMES: have a hard time listening to instructions at school, work or home?" "During the last 12 months, did you do any of the following TWO OR MORE TIMES: be a bully or threaten other people?" "During the last 12 months, did you do any of the following TWO OR MORE TIMES: start fights with other people?"	1 – Yes 0 – No	Counted Five items ranging from 0 to 6 (0.5)
Global appraisal of individual needs-1 (GAIN-1)	Four items "During the last 12 months, did you do any of the following TWO OR MORE TIMES: lie or con to get things you wanted or to avoid having to do something?" "During the last 12 months, did you do any of the following TWO OR MORE TIMES: have a hard time paying attention at school, work or home?" "During the last 12 months, did you do any of the following TWO OR MORE TIMES: be a bully or threaten other people?" "During the last 12 months, did you do any of the following TWO OR MORE TIMES: start fights with other people?"	1 – Yes 0 – No	Dichotomized 1 – Any 0 – None (0.5)
Non-suicidal self-injury	One item "During the last 12 months, how many times did you do something to	1 – 0 times 2 – 1 or 2 times 3 – 3 to 5 times	Dichotomized 1 – Any 0 – None

Variable	Item(s)	Response options	Description (Cronbach's α)
	purposely hurt or injure yourself without wanting to die, such as cutting, burning, or bruising yourself on purpose?"	4 – 6 to 9 times 5 – 10 to 19 times 6 – 20 or more times	
Suicidal ideation	Three items "Have you ever seriously considered attempting suicide: No?" "Have you ever seriously considered attempting suicide: Yes, during the last year?" "Have you ever seriously considered attempting suicide: Yes, more than a year ago?"	1 – Checked, 0 – Not checked 1 – Checked, 0 – Not checked 1 – Checked, 0 – Not checked	Dichotomized 1 – During the last year 0 – None or a year ago
Suicidal attempt	Three items "Have you ever actually attempted suicide: No?" "Have you ever actually attempted suicide: Yes, during the last year?" "Have you ever actually attempted suicide: Yes, more than a year ago?"	1 – Checked, 0 – Not checked 1 – Checked, 0 – Not checked 1 – Checked, 0 – Not checked	Dichotomized 1 – During the last year 0 – None or a year ago
Perpetrator	Three items "Have YOU ever done any of the following to a boyfriend or girlfriend in a dating or serious relationship: called him/her names or put him/her down verbally?" "Have YOU ever done any of the following to a boyfriend or girlfriend in a dating or serious relationship: hit, slapped or physically hurt him/her on purpose?" "Have YOU ever done any of the following to a boyfriend or girlfriend in a dating or serious relationship: pressured him/her into having sex when he/she did not want to?"	1 – Checked, 0 – Not checked 1 – Checked, 0 – Not checked 1 – Checked, 0 – Not checked	Dichotomized 1 – Any 0 – None
Gambling	Four items "During the last 12 months, how often have you done the following gambling/betting activities: Played cards, bet on sports teams or games of personal skill like video gaming, pool, golf or bowling?" "During the last 12 months, how often have you done the following	1 – Not at all 2 – Less than once a month 3 – About once a month 4 – About once a week 5 – 2 to 6 times a week 6 – Daily	Dichotomized 1 – Any 0 – None (0.6)

Variable	Item(s)	Response options	Description (Cronbach's α)
	gambling/betting activities: Bought lottery tickets or scratch offs?" "During the last 12 months, how often have you done the following gambling/betting activities: Gambled in a casino?" "During the last 12 months, how often have you done the following gambling/betting activities: Gambled for money online?"		
Brief Adolescent Gambling Screen	Three items "During the last 12 months, how often have you hidden your gambling/betting from your parents, other family members or teachers?" "During the last 12 months, how often have you felt that you might have a problem with gambling/betting?" "During the last 12 months, how often have you skipped hanging out with friends who do not gamble/bet to hang out with friends who do gamble/bet?"	1 – Never 2 – Sometimes 3 – Many times 4 – All of the time	Dichotomized 0 – Low 1 - High (0.74)
Crime / violence subscription	Three items "During the last 12 months, how often have you damaged or destroyed property?" "During the last 12 months, how often have you hit or beat up another person?" "During the last 12 months, how often have you taken something from a store without paying for it?"	1 – Never 2 – Once or twice 3 – 3 to 5 times 4 – 6 to 9 times 5 – 10 or more times	Dichotomized 0 – None 1 – Any (0.63)
Tobacco product usage	Five items "During the last 30 days, on how many days did you smoke a cigarette?" "During the last 30 days, on how many days did you smoke cigars, cigarillos or little cigars?" "During the last 30 days, on how many days did you use chewing tobacco, snuff or dip?" "During the last 30 days, on how many days did you use an electronic cigarette (e-cigarette, e-hookah, vaping pen)?"	1 – 0 days 2 – 1 to 2 days 3 – 3 to 9 days 4 – 10 to 19 days 5 – 20 to 29 days 6 – All 30 days	Dichotomized 0 – None 1 – Any (0.72)

Variable	Item(s)	Response options	Description (Cronbach's α)
	"During the last 30 days, on how many days did you use a hookah or a waterpipe to smoke tobacco?"		
Alcohol consumption frequency	One item "During the last 12 months, on how many occasions (if any) have you had alcoholic beverages to drink?"	1 – 0 2 – 1~2 3 – 3~5 4 – 6~9 5 – 10~19 6 – 20~39 7 – 40+	Dichotomized 1 – Any 0 – None
Binge drinking-1	One item "If you drink beer/wine/wine coolers/liquor, generally, how much (if any) do you drink at one time?"	1 – I don't drink 2 – 1 glass/can/ drink 3 – 2 glasses/cans/ drinks 4 – 3 glasses/cans/ drinks 5 – 4 glasses/cans/ drinks 6 – 5 or more glasses/cans/drinks	Dichotomize 1 – Any 0 – None
Binge drinking-2	One item "During the past 30 days, on how many days did you have 5 or more drinks in a row, that is, within a couple of hours?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 5 days 5 – 6 to 9 days 6 – 10 to 19 days 7 – 20 or more days	Dichotomized 1 – Any 0 – None
Non-medical marijuana use frequency	One item "During the last 12 months, on how many occasions (if any) have you used marijuana or hashish? (Do NOT count medical marijuana prescribed for you by a doctor.)"	1 – 0 2 – 1~2 3 – 3~5 4 – 6~9 5 – 10~19 6 – 20~39 7 – 40+	Dichotomized 1 – Any 0 – None
Marijuana use frequency	One item "How often do you use each of the following: Marijuana (pot, hash, hash oil)?"	1 – Never 2 – Once or twice 3 – Once or twice a year 4 – Once a month 5 – Twice a month 6 – Once a week 7 – Daily	Dichotomized 1 – Any 0 – None
Substance use – 1	Eight items "During the last 12 months, on how many occasions (if any) have you sniffed glue or huffed or inhaled the	1 – 0 2 – 1~2 3 – 3~5 4 – 6~9 5 – 10~19	Dichotomized each question into 1 – Any, 0 – None, then counted 8 items

Variable	Item(s)	Response options	Description (Cronbach's α)
	contents of aerosol spray cans or other gases to get high?" "During the last 12 months, on how many occasions (if any) have you used LSD (acid), PCP (wet sticks or dipped joints), or other psychedelics (mushrooms, angel dust)?" "During the last 12 months, on how many occasions (if any) have you used MDMA (E, X, ecstasy), GHB (G, Liquid E, Liquid X, roofies) or Ketamine (Special K)?" "During the last 12 months, on how many occasions (if any) have you used crack, coke or cocaine in any other form?" "During the last 12 months, on how many occasions (if any) have you used heroin?" "During the last 12 months, on how many occasions (if any) have you used methamphetamine (meth, glass, crank, crystal meth, ice)?" "During the last 12 months, on how many occasions (if any) have you used over-the-counter drugs such as cough syrup, cold medicine or diet pills that you took only to get high?" "During the last 12 months, on how many occasions (if any) have you used synthetic drugs such as bath salts (Ivory Wave, White Lightning) or synthetic marijuana (K2, Gold) that you took only to get high?"	6 – 20~39 7 – 40+	ranging from 0 to 8 (0.83)
Substance use - 2 (prescription substances usage not prescribed to user)	Four items "During the last 12 months, on how many occasions have you used any of the following prescription drugs that were NOT prescribed for you or that you took ONLY to get high: Stimulants such as Benzedrine (bennies, speed, uppers) or diet pills?" "During the last 12 months, on how many occasions have you used any of the following prescription drugs that were NOT prescribed for you or that you took ONLY to get high:	1 – 0 2 – 1~2 3 – 3~5 4 – 6~9 5 – 10~19 6 – 20+	Dichotomized each question into 1 – Any, 0 – None, then counted 4 items ranging from 0 to 4 (0.77)

Variable	Item(s)	Response options	Description (Cronbach's α)
	ADHD or ADD drugs like Ritalin (hyper pills)?" "During the last 12 months, on how many occasions have you used any of the following prescription drugs that were NOT prescribed for you or that you took ONLY to get high: Pain relievers such as OxyContin, Percocet, Vicodin or others?" "During the last 12 months, on how many occasions have you used any of the following prescription drugs that were NOT prescribed for you or that you took ONLY to get high: Tranquilizers such as Valium, Xanax or sedatives or barbiturates?"		
Substance use frequency	One item "During the last 30 days, on how many days did you use prescription drugs not prescribed for you?"	1 – 0 days 2 – 1 to 2 days 3 – 3 to 5 days 4 – 6 to 9 days 5 – 10 to 19 days 6 – 20 to 29 days 7 – All 30 days	Dichotomize 1 – Any 0 – None
Perceptions of substance use risk	Four items "How much do you think people risk harming themselves physically or in other ways if they smoke one or more packs of cigarettes per day?" "How much do you think people risk harming themselves physically or in other ways if they have five or more drinks of an alcoholic beverage once or twice per week?" "How much do you think people risk harming themselves physically or in other ways if they smoke marijuana once or twice per week?" "How much do you think people risk harming themselves physically or in other ways if they use prescription drugs not prescribed for them?"	1 – No risk 2 – Slight risk 3 – Moderate risk 4 – Great risk	Converted to mean value of all items ranging from 1 to 4 (0.9)
Tobacco use frequency	One item "How often do you use each of the following: Tobacco (cigarettes, chew)?"	1 – Never 2 – Once or twice 3 – Once or twice a year 4 – Once a month 5 – Twice a month 6 – Once a week	Dichotomized 1 – Any 0 – None

Variable	Item(s)	Response options	Description (Cronbach's α)
		7 – Daily	
Alcohol consumption frequency	One item "How often do you use each of the following: Alcohol (beer, wine, liquor)?"	1 – Never 2 – Once or twice 3 – Once or twice a year 4 – Once a month 5 – Twice a month 6 – Once a week 7 – Daily	Dichotomized 1 – Any 0 – None
Overweight	Two items "How tall are you? (Write in whole numbers; no fractions or decimals) "How much do you weigh? (Write in whole numbers; no fractions or decimals)"	0 – Normal or underweight 1 – Overweight 2 – Obese	
Race	One item "What is your race? (mark all that apply)	1 - American Indian only 2 - Asian only 3 - Black, African or African American only 4 - Native Hawaiian or Pacific Islander only 5 - White only 6 - Multiple Races (checked more than one)	Converted into six dummy variables with 0 and 1 answer
Ethnicity	Are you Hispanic or Latino/a?	1 – Yes 0 – No	Seven groups* specified from the question "Race" and "Ethnicity", then converted into seven dummy variables with 0 and 1 answer

*American Indian Non-Hispanic, Asian Non-Hispanic, Black Non-Hispanic, Pacific Islander Non-Hispanic, White Non-Hispanic, Multiple Races Non-Hispanic

Table A2

Measures of Factors of 2016 Minnesota Student Survey within the Mesosystem

Variable	Item(s)	Response options	Description (Cronbach's α)
Family structure (two-parent household versus else)	Four items "Which adults do you live with: Biological mother (the woman who gave birth to me)?" "Which adults do you live with: Biological father?" "Which adults do you live with: Adoptive mother?" "Which adults do you live with: Adoptive father?"	1 – Checked 0 – Not checked	Either both biological parents OR adoptive parents exist – 1 else – 0
Relationship with father	One item "Can you talk to your father about problems you are having?"	1 – My father is not around 2 – No, not at all 3 – No, not very often 4 – Yes, some of the time 5 – Yes, most of the time	
Relationship with mother	One item "Can you talk to your mother about problems you are having?"	1 – My mother is not around 2 – No, not at all 3 – No, not very often 4 – Yes, some of the time 5 – Yes, most of the time	
Transient student	One item "Since the beginning of this school year, how many times have you changed schools?"	1 – 0 times 2 – 1 time 3 – 2 times 4 – 3 or more times	
School nurse office visit	One item "During the last 30 days, how many times have you gone to the nurses office?"	1 – None 2 – Once or twice 3 – 3 to 5 times 4 – 6 to 9 times 5 – 10 or more times	
Sent to office for discipline	One item "During the last 30 days, how many times have you been sent to the office for discipline?"	1 – None 2 – Once or twice 3 – 3 to 5 times 4 – 6 to 9 times 5 – 10 or more times	

Variable	Item(s)	Response options	Description (Cronbach's α)
In school suspension	One item "During the last 30 days, how many times have you had in-school suspension (ISS)?"	1 – None 2 – Once or twice 3 – 3 to 5 times 4 – 6 to 9 times 5 – 10 or more times	
Out of school suspension	One item "During the last 30 days, how many times have you been suspended from school (out-of-school suspension-OSS)?"	1 – None 2 – Once or twice 3 – 3 to 5 times 4 – 6 to 9 times 5 – 10 or more times	
School Engagement Scale (SE)	Six items "How often do you care about doing well in school?" "How often do you pay attention in class?" "How often do you go to class unprepared?" "If something interests me, I try to learn more about it." "I think things I learn in school are useful." "Being a student is one of the most important parts of who I am."	1 – Strongly disagree 2 – Disagree 3 – Agree 4 – Strongly agree	Converted to mean value of all items ranging from 1 to 4 (0.68)
Teacher Student Relationship (TSR)	Five items "Overall, adults at my school treat students fairly." "Adults at my school listen to the students." "The school rules are fair." "At my school, teachers care about students." "Most teachers at my school are interested in me as a person."	1 – Strongly disagree 2 – Disagree 3 – Agree 4 – Strongly agree	Converted to mean value of all items ranging from 1 to 4 (0.85)
Home safety	One item "I feel safe at home."	1 – Strongly disagree 2 – Disagree 3 – Agree 4 – Strongly agree	
Harassed by peers: race, ethnicity or national origin	One item "During the last 30 days, how often have other students harassed or bullied you for any of the following reasons: Your race, ethnicity or national origin?"	1 – Never 2 – Once or twice 3 – About once a week 4 – Several times a week 5 – Every day	
Harassed by peers: religion	One item	1 – Never 2 – Once or twice	

Variable	Item(s)	Response options	Description (Cronbach's α)
	"During the last 30 days, how often have other students harassed or bullied you for any of the following reasons: Your religion?"	3 – About once a week 4 – Several times a week 5 – Every day	
Harassed by peers: gender	One item "During the last 30 days, how often have other students harassed or bullied you for any of the following reasons: Your gender (being male, female, transgender, etc.)?"	1 – Never 2 – Once or twice 3 – About once a week 4 – Several times a week 5 – Every day	
Harassed by peers: LGB	One item "During the last 30 days, how often have other students harassed or bullied you for any of the following reasons: Because you are gay, lesbian, or bisexual or because someone thought you were?"	1 – Never 2 – Once or twice 3 – About once a week 4 – Several times a week 5 – Every day	
Harassed by peers: disability	One item "During the last 30 days, how often have other students harassed or bullied you for any of the following reasons: A physical or mental disability?"	1 – Never 2 – Once or twice 3 – About once a week 4 – Several times a week 5 – Every day	
Harassed by peers: size or weight	One item "During the last 30 days, how often have other students harassed or bullied you for any of the following reasons: Your size or weight?"	1 – Never 2 – Once or twice 3 – About once a week 4 – Several times a week 5 – Every day	
Harassed by peers: physical appearance	One item "During the last 30 days, how often have other students harassed or bullied you for any of the following reasons: Your physical appearance?"	1 – Never 2 – Once or twice 3 – About once a week 4 – Several times a week 5 – Every day	
Online bullying	One item "During the last 30 days, how often have you been bullied through e-mail, chat rooms, instant messaging, websites or texting?"	1 – Never 2 – Once or twice 3 – About once a week 4 – Several times a week 5 – Every day	

Variable	Item(s)	Response options	Description (Cronbach's α)
Hostile school climate by peers	Five items "During the last 30 days, how often have other students at school pushed, shoved, slapped, hit or kicked you when they weren't kidding around?" "During the last 30 days, how often have other students at school threatened to beat you up?" "During the last 30 days, how often have other students at school spread mean rumors or lies about you?" "During the last 30 days, how often have other students at school made sexual jokes, comments or gestures towards you?" "During the last 30 days, how often have other students at school excluded you from friends, other students or activities?"	1 – Never 2 – Once or twice 3 – About once a week 4 – Several times a week 5 – Every day	Converted to mean value of all items ranging from 1 to 5 (0.76)
Hostile school climate by respondent	Five items "During the last 30 days, how many times at school have YOU pushed, shoved, slapped, hit or kicked someone when you weren't kidding around?" "During the last 30 days, how many times at school have YOU threatened to beat someone up?" "During the last 30 days, how many times at school have YOU spread mean rumors or lies about someone else?" "During the last 30 days, how many times at school have YOU made sexual jokes, comments or gestures towards someone else?" "During the last 30 days, how many times at school have YOU excluded someone from friends, other students or activities?"	1 – Never 2 – Once or twice 3 – About once a week 4 – Several times a week 5 – Every day	Converted to mean value of all items ranging from 1 to 5 (0.74)
Physical education frequency	One item "During a typical school week, on how many days do you go to physical education (PE or gym) classes?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 days 5 – 4 days 6 – 5 days	
Perception of family caring	Two items "How much do you feel your parents care about you?"	1 – Not at all 2 – A little 3 – Some	Converted to mean value of all items

Variable	Item(s)	Response options	Description (Cronbach's α)
	"How much do you feel other adult relatives care about you?"	4 – Quite a bit 5 – Very much	ranging from 1 to 5 (0.79)
Perception of peer caring	One item "How much do you feel friends care about you?"	1 – Not at all 2 – A little 3 – Some 4 – Quite a bit 5 – Very much	
Victim of Intimate Partner Violence (IPV)	Three items "Have you ever had a boyfriend or girlfriend in a dating or serious relationship who called you names or put you down verbally?" "Have you ever had a boyfriend or girlfriend in a dating or serious relationship who hit, slapped or physically hurt you on purpose?" "Have you ever had a boyfriend or girlfriend in a dating or serious relationship who pressured you into having sex when you did not want to?	1 – Yes 0 – No	Three items sum calculated ranging from 0 to 3, then dichotomized 1 – Any 0 – None (0.68)
Incarcerated parents	Three items "Have any of your parents or guardians ever been in jail or prison?: None of my parents or guardians has ever been in jail or prison" "Have any of your parents or guardians ever been in jail or prison?: Yes, I have a parent or guardian in jail or prison right now?" "Have any of your parents or guardians ever been in jail or prison?: Yes, I have had a parent or guardian in jail or prison in the past"	1 – Yes 0 – No	Dichotomized 1 – Any 0 – None
Adverse Childhood Experiences (ACEs)	Seven items "Do you live with anyone who drinks too much alcohol?" "Do you live with anyone who uses illegal drugs or abuses prescription drugs?" "Does a parent or other adult in your home regularly swear at you, insult you or put you down?" "Has a parent or other adult in your household ever hit, beat, kicked or physically hurt you in any way?"	1 – Yes 0 – No	Sum of all items used as a measure of count - Range from 0 to 7

Variable	Item(s)	Response options	Description (Cronbach's α)
	"Have your parents or other adults in your home ever slapped, hit, kicked, punched or beat each other up?" "Has any adult or other person outside of the family ever touched you sexually against your wishes or forced you to touch them sexually? "Has any older or stronger member of your family ever touched you or had you touch them sexually?		
Runaway	One item "During the last 12 months, how often have you run away from home?"	1 – Never 2 – Once or twice 3 – About once a week 4 – Several times a week 5 – Every day	Dichotomized 1 – Any 0 – None
Parents' approval of substance use	Four items "How wrong do your parents feel it would be for you to smoke cigarettes?" "How wrong do your parents feel it would be for you to have one or more drinks of alcoholic beverage nearly every day?" "How wrong do your parents feel it would be for you to smoke marijuana?" "How wrong do your parents feel it would be for you to use prescription drugs not prescribed for you?"	1 – Not at all wrong 2 – A little big wrong 3 – Wrong 4 – Very wrong	Average of four items ranging from 1 to 4 used as a scale (0.92)
Friends' approval of substance use	Four items "How wrong do your friends feel it would be for you to smoke cigarettes?" "How wrong do your friends feel it would be for you to have one or more drinks of alcoholic beverage nearly every day?" "How wrong do your friends feel it would be for you to smoke marijuana?" "How wrong do your friends feel it would be for you to use prescription drugs not prescribed for you?"	1 – Not at all wrong 2 – A little big wrong 3 – Wrong 4 – Very wrong	Average of four items ranging from 1 to 4 used as a scale (0.93)
Attitudes toward drinking	Two items "How do you feel about each of the following statements: Parents and other adults should clearly	1 – Strongly disagree 2 – Disagree	Average of two items ranging from 1 to 4

Variable	Item(s)	Response options	Description (Cronbach's α)
	communicate with their children about the importance of not using alcohol?" "How do you feel about each of the following statements: Drinking alcohol is never a good thing for anyone my age to do?"	3 – Neither agree nor disagree 4 – Agree 5 – Strongly agree	used as a scale (0.86)
Peers' Attitudes toward drinking	Two items "How do you think MOST STUDENTS in your school feel about each of the following statements: Parents and other adults should clearly communicate with their children about the importance of not using alcohol?" "How do you think MOST STUDENTS in your school feel about each of the following statements: Drinking alcohol is never a good thing for anyone my age to do?"	1 – Strongly disagree 2 – Disagree 3 – Neither agree nor disagree 4 – Agree 5 – Strongly agree	Average of two items ranging from 1 to 4 used as a scale (0.88)
Tobacco use frequency of peers	"How often do you think MOST STUDENTS in your school use each of the following: Tobacco (cigarettes, chew)?"	1 – Never 2 – Tried once or twice 3 – Once or twice a year 4 – Once a month 5 – Twice a month 6 – Once a week 7 – Daily	Dichotomized 1 – Any 0 – None
Alcohol consumption frequency	"How often do you think MOST STUDENTS in your school use each of the following: Alcohol (beer, wine, liquor)?"	1 – Never 2 – Tried once or twice 3 – Once or twice a year 4 – Once a month 5 – Twice a month 6 – Once a week 7 – Daily	Dichotomized 1 – Any 0 – None
Marijuana use frequency of peers	How often do you think MOST STUDENTS in your school use each of the following: Marijuana (pot, hash, hash oil)?	1 – Never 2 – Tried once or twice 3 – Once or twice a year 4 – Once a month 5 – Twice a month 6 – Once a week 7 – Daily	Dichotomized 1 – Any 0 – None

Table A3

Measures of Factors of 2016 Minnesota Student Survey within the Exosystem

Variable	Item(s)	Response options	Description (Cronbach's α)
Perception of safety while commuting	One item "I feel safe going to and from school."	1 – Strongly disagree 2 – Disagree 3 – Agree 4 – Strongly agree	
Perception of school safety	One item "I feel safe at school."	1 – Strongly disagree 2 – Disagree 3 – Agree 4 – Strongly agree	
Afterschool activity: In-school	One item "During a typical week, how often do you go to the following places after school: I stay at my school or go to another school?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 days	
Afterschool activity: In-home	One item "During a typical week, how often do you go to the following places after school: Your home or another home such as a friend's, relative's or neighbor's?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 days	
Afterschool activity: Youth center	One item "During a typical week, how often do you go to the following places after school: A rec, community or other youth center?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 days	
Afterschool activity: Outdoor	One item "During a typical week, how often do you go to the following places after school: A park or other outdoor space?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 days	
Afterschool activity: Library	One item "During a typical week, how often do you go to the following places after school: A library?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 days	
Afterschool activity: Religion	One item "During a typical week, how often do you go to the following places after school: A church, synagogue, mosque, or other spiritual/religious place?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 days	
Afterschool activity: Job	One item	1 – 0 days 2 – 1 day	

Variable	Item(s)	Response options	Description (Cronbach's α)
	"During a typical week, how often do you go to the following places after school: A job?"	3 – 2 days 4 – 3 to 4 days 5 – 5 days	
Out-of-school activity: Sports	One item "During a typical week, how often do you participate in each of the following activities outside of the regular school day: Sports teams, such as park and rec teams, school teams, in-house teams or traveling teams?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 or more days	
Out-of-school activity: School sponsored activities (not sports)	One item "During a typical week, how often do you participate in each of the following activities outside of the regular school day: School sponsored activities or clubs that are not sports, such as drama, music, chess or science club?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 or more days	
Out-of-school activity: Academic program including tutoring	One item "During a typical week, how often do you participate in each of the following activities outside of the regular school day: Tutoring, homework help or academic programs?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 or more days	
Out-of-school activity: Leadership activities	One item "During a typical week, how often do you participate in each of the following activities outside of the regular school day: Leadership activities such as student government, youth councils or committees?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 or more days	
Out-of-school activity: Artistic lessons	One item "During a typical week, how often do you participate in each of the following activities outside of the regular school day: Artistic lessons, such as music or dance?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 or more days	
Out-of-school activity: Physical activity lessons	One item "During a typical week, how often do you participate in each of the following activities outside of the regular school day: Physical activity lessons, such as tennis or karate?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 or more days	
Out-of-school activity: Other	One item "During a typical week, how often do you participate in each of the	1 – 0 days 2 – 1 day 3 – 2 days	

Variable	Item(s)	Response options	Description (Cronbach's α)
community programs	following activities outside of the regular school day: Other community clubs and programs such as 4-H, Scouts, Y-clubs or Community Ed?"	4 – 3 to 4 days 5 – 5 or more days	
Out-of-school activity: Religious activities	One item "During a typical week, how often do you participate in each of the following activities outside of the regular school day: Religious activities such as religious services, education or youth groups?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 4 days 5 – 5 or more days	
Positive Youth Development Scale	Seven items "When you spend time doing activities outside of the regular school day, how often do you feel safe?" "When you spend time doing activities outside of the regular school day, how often do you learn skills like teamwork or leadership?" "When you spend time doing activities outside of the regular school day, how often do you develop trusting relationships with peers your age?" "When you spend time doing activities outside of the regular school day, how often do you develop trusting relationships with adults?" "When you spend time doing activities outside of the regular school day, how often do you help make decisions?" "When you spend time doing activities outside of the regular school day, how often do you do something that gives you joy and energy?" "When you spend time doing activities outside of the regular school day, how often do you learn skills that you can use in a future job?"	1 – Rarely or never 2 – Sometimes 3 – Often 4 – Very often	Average of seven items ranging from 1 to 4 used as a scale (0.85)

Table A4

Measures of Factors of 2016 Minnesota Student Survey within the Macrosystem

Variable	Item(s)	Response options	Description (Cronbach's α)
Free or reduced-price lunch at school	One item "Do you currently get free or reduced-price lunch at school?"	1 – Yes 0 – No	
Neighborhood safety	One item "I feel safe in my neighborhood."	1 – Strongly disagree 2 – Disagree 3 – Agree 4 – Strongly agree	
Skipping meal due to financial issues	One item "During the last 30 days, have you had to skip meals because your family did not have enough money to buy food?"	1 – Yes 0 – No	
Perceptions of caring from adults in the community	One item "How much do you feel adults in your community care about you?"	1 – Not at all 2 – A little 3 – Some 4 – Quite a bit 5 – Very much	
Region	Five items of responses about combined location and district size	1 – Twin Cities Metro 2 – Greater MN-district of 5,000 or more 3 – Greater MN-district of 2,000 - 4,999 4 – Greater MN-district of 1,000 – 1,999 5 – Greater MN-district of 999 or less	
Homelessness	Three items "During the past 12 months, have you stayed in a shelter, somewhere not intended as a place to live, or someone else's home because you had no other place to stay: No?" "During the past 12 months, have you stayed in a shelter, somewhere not intended as a place to live, or someone else's home because you had no other place to stay: Yes, with	1 – Yes 0 – No	Dichotomized 1 – Any 0 – None

Variable	Item(s)	Response options	Description (Cronbach's α)
	my parents or an adult family member?" "During the past 12 months, have you stayed in a shelter, somewhere not intended as a place to live, or someone else's home because you had no other place to stay: Yes, on my own without any adult family members?"		

Table A5

Measures of 2019 Minnesota Student Survey Factors – Causal Discovery Analysis Step 1

Variable	Item(s)	Response options	Description (Cronbach's α)
Race: American Indian or Alaskan Native	One item "How do you describe yourself? American Indian or Alaskan Native"	1 – Yes 0 – No	
Race: Native Hawaiian or Other Pacific Islander	One item "How do you describe yourself? Native Hawaiian or Other Pacific Islander"	1 – Yes 0 – No	
Reason of school absence: Illness	One item "What are the reasons you missed a full or part of a day of school in the last 30 days? Illness (feeling physically sick), including problems with breathing or your teeth"	1 – Yes 0 – No	
Reason of school absence: Suspension	One item "What are the reasons you missed a full or part of a day of school in the last 30 days? Suspended from school"	1 – Yes 0 – No	
Suspension: sent out of the classroom	One item "During the last 30 days, how many times did you get sent out of the classroom for discipline?"	1 – None 2 – Once or twice 3 – 3 to 5 times 4 – 6 to 9 times 5 – 10 or more times	
School Engagement Scale	Six items "How often do you care about doing well in school?" "How often do you pay attention in class?"	1 – None of the time 2 – Some of the time 3 – Most of the time	(0.7)

Variable	Item(s)	Response options	Description (Cronbach's α)
	"How often do you go to class unprepared?" "If something interests me, I try to learn more about it." "I think things I learn at school are useful." "Being a student is one of the most important parts of who I am."	4 – All of the time Response order for the third question opposite as it's the only negative questions.	
Teacher Student Relationship	Five items "Overall, adults at my school treat students fairly." "Adults at my school listen to the students." "The school rules are fair." "At my school, teachers care about students." "Most teachers at my school are interested in me as a person."	1 – None of the time 2 – Some of the time 3 – Most of the time 4 – All of the time	(0.85)
Dental Checkup	One item "When was the last time you saw a dentist for a check-up, exam or teeth cleaning or other dental work?"	1 – During the last year 2 – Between 1 and 2 years ago 3 – More than 2 years ago 4 - Never	
Social Competency Scale	Eight items "I say no to things that are dangerous or unhealthy." "I build friendships with other people." "I express my feelings in proper ways." "I plan ahead and make good choices." "I stay away from bad influences." "I resolve conflicts without anyone getting hurt." "I accept people who are different from me." "I am sensitive to the needs and feelings of others."	1 – Not at all or rarely 2 – Somewhat or sometimes 3 – Very or often 4 – Extremely or almost always	(0.84)
Adversarial Childhood Experiences	Eight items "Parent or Guardian has ever been in prison or jail?" "Experienced sexual abuse from person within or outside family?" "Do you live with anyone who drinks too much alcohol?"	1 – Yes 0 – No	Sum of all items used as a measure of count - Range from 0 to 4 0 – None 1 – One 2 – Two

Variable	Item(s)	Response options	Description (Cronbach's α)
	"Do you live with anyone who uses illegal drugs or abuses prescription drugs?" "Do you live with anyone who is depressed or has any other mental health issues?" "Does a parent or other adult in your home regularly swear at you, insult you or put you down?" "Has a parent or other adult in your home ever hit, beat, kicked or physically hurt you in any way?" "Have your parents or other adults in your home ever slapped, hit, kicked, punched or beat each other up?"		3 – Three 4 – Four or more
Tobacco product use	Four items "During the last 30 days, on how many days did you smoke a cigarette?" "During the last 30 days, on how many days did you smoke cigars, cigarillos or little cigars?" "During the last 30 days, on how many days did you use chewing tobacco, snuff or dip?" "During the last 30 days, on how many days did you use a hookah or a waterpipe to smoke tobacco?"	1 – 0 days 2 – 1 to 2 days 3 – 3 to 9 days 4 – 10 to 19 days 5 – 20 to 29 days 6 – All 30 days	All questions dichotomized 1 – Any 0 – None Then sum calculated ranging from 0 to 4 (0.74)
Vape use	One item "During the last 30 days, on how many days did you vape or use an e-cigarette like JUUL, suorin, blu, VUSE, or logic?"	1 – 0 days 2 – 1 to 2 days 3 – 3 to 9 days 4 – 10 to 19 days 5 – 20 to 29 days 6 – All 30 days	
Binge Drinking	Two items "(Female) During the past 30 days, on how many days did you have 4 or more drinks of alcohol in a row, that is, within a couple of hours?" "(Male) During the past 30 days, on how many days did you have 5 or more drinks of alcohol in a row, that is, within a couple of hours?"	1 – 0 days 2 – 1 day 3 – 2 days 4 – 3 to 5 days 5 – 6 to 9 days 6 – 10 to 19 days 7 – 20 or more days	Binge drinking (4 or more drinks in a row (females) or 5 or more drinks in a row (males) within a couple of hours
Non-medical marijuana use	One item "During the last 12 months, on how many occasions (if any) have you used marijuana or hashish? (Do NOT count medical marijuana prescribed for you by a doctor)"	1 – 0 2 – 1 to 2 3 – 3 to 5 4 – 6 to 9 5 – 10 to 19 6 – 20 to 39	

Variable	Item(s)	Response options	Description (Cronbach's α)
		7 – 40 or more	
Substance use – 1	Nine items "During the last 12 months, on how many occasions (if any) have you sniffed glue or huffed or inhaled the contents of aerosol spray cans or other gases to get high?" "During the last 12 months, on how many occasions (if any) have you used LSD (acid), PCP (wet sticks or dipped joints), or other psychedelics (mushrooms, angel dust)?" "During the last 12 months, on how many occasions (if any) have you used MDMA (E, X, ecstasy), GHB (G, Liquid E, Liquid X, roofies) or Ketamine (Special K)?" "During the last 12 months, on how many occasions (if any) have you used crack, coke or cocaine in any other form?" "During the last 12 months, on how many occasions (if any) have you used heroin(smack, junk, China White)?" "During the last 12 months, on how many occasions (if any) have you used methamphetamine (meth, glass, crank, crystal meth, ice)?" "During the last 12 months, on how many occasions (if any) have you used over-the-counter drugs such as cough syrup, cold medicine or diet pills that you took only to get high?" "During the last 12 months, on how many occasions (if any) have you used synthetic marijuana (K2, Gold) that you took only to get high? "During the last 12 months, on how many occasions (if any) have you used synthetic drugs such as bath salts (Ivory Wave, White Lightning) or synthetic marijuana (K2, Gold) that you took only to get high?"	1 – 0 2 – 1~2 3 – 3~5 4 – 6~9 5 – 10~19 6 – 20~39 7 – 40+	Dichotomized each question into 1 – Any, 0 – None, then counted 9 items ranging 0 to 9 (0.85)
Prescription Substances Use (not prescribed to user)	Four items "During the last 12 months, on how many occasions (if any) have you used the following prescription drugs without a doctor's prescription or	1 – 0 2 – 1 to 2 3 – 3 to 5 4 – 6 to 9 5 – 10 to 19	Dichotomized each question into 1 – Any, 0 – None, then

Variable	Item(s)	Response options	Description (Cronbach's α)
	differently than how a doctor told you to use it? Stimulants such as Amphetamines or diet pills" "During the last 12 months, on how many occasions (if any) have you used the following prescription drugs without a doctor's prescription or differently than how a doctor told you to use it? ADHD or ADD drugs (Ritalin, Adderall, hyper pills) " "During the last 12 months, on how many occasions (if any) have you used the following prescription drugs without a doctor's prescription or differently than how a doctor told you to use it? Pain relievers such as OxyContin, Percocet, Vicodin or others" "During the last 12 months, on how many occasions (if any) have you used the following prescription drugs without a doctor's prescription or differently than how a doctor told you to use it? Tranquilizers such as Valium, Xanax, Klonopin or others"	6 – 20+	counted 4 items ranging 0 to 4 (0.7)
Friends' approval of substance use	Five items "How wrong do your friends feel it would be for you to smoke cigarettes?" "How wrong do your friends feel it would be for you to have one or more drinks of alcoholic beverage nearly every day?" "How wrong do your friends feel it would be for you to smoke marijuana?" "How wrong do your friends feel it would be for you to use prescription drugs not prescribed for you?" "How wrong do your friends feel it would be for you to	1 – Not at all wrong 2 – A little big wrong 3 – Wrong 4 – Very wrong	Average of five items ranging from 1 to 4 used as a scale (0.92)
Marijuana use frequency	One item "How often do you use the following? Marijuana (pot, hash, hash oil)"	1 – Never 2 – Once or twice 3 – Once or twice a year 4 – Once a month 5 – Twice a month 6 – Once a week 7 – Daily	

Table A6

The Step 1 Causal Discovery Analysis Bootstrap Re-sampling Analysis (the MSS 2019)

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	Reason of school absence: Suspension	Suspension: sent out of the classroom	0	1	0	0	0	0	0
←	FDA	School Engagement	0.003	0.997	0	0	0	0	0
←	Suspension: sent out of the classroom	Tobacco Product Use	0.005	0.989	0	0.001	0	0.005	0
←	School Engagement	Social Competency Scale	0.001	0.996	0	0	0	0.003	0
←	Suspension: sent out of the classroom	Teacher Student Relationship	0.002	0.998	0	0	0	0	0
←	School Engagement	Teacher Student Relationship	0.033	0.957	0	0	0.008	0.002	0
←	School Engagement	Friends' Approval of Substance Use	0.004	0.9391	0.002	0.017	0.008	0.03	0
←	Dental Checkup	Social Competency Scale	0	1	0	0	0	0	0
←	Teacher Student Relationship	Friends' Approval of Substance Use	0	0.9321	0	0	0.01	0.0559	0.002
←	Suspension: sent out of the classroom	School Engagement	0.001	0.998	0	0	0	0.001	0
←	FDA	PDA	0.005	0.994	0	0	0	0.001	0
←	PDA	School Engagement	0	0.997	0	0	0.001	0.002	0
←	Suspension: sent out of the classroom	Substance use – 1	0.014	0.9231	0	0.02	0	0.042	0.001
←	PDA	Vape Use	0.006	0.7882	0	0.019	0	0.1868	0
←	PDA	Marijuana use frequency	0.007	0.7802	0	0	0	0	0.2128
←	Binge	Vape Use	0.001	0.9191	0.001	0.0679	0.004	0.007	0
←	Non-medical marijuana use	Marijuana use frequency	0.014	0.984	0.001			0.001	0

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	Tobacco Product Use	Marijuana use frequency	0.014	0.7423	0.014	0.002	0.003	0.002	0.2228
←	Teacher Student Relationship	Social Competency Scale	0	0.999	0	0	0	0.001	0
←	FDA	Reason of school absence: Illness	0	0.6803	0	0.3197	0	0	0
←	FDA	Reason of school absence: Suspension	0	0.6613	0	0	0	0	0.3387
←	Reason of school absence: Illness	Teacher Student Relationship	0	0.6613	0	0	0	0	0.3387
←	Reason of school absence: Suspension	Non-medical marijuana use	0	0.6314	0	0	0	0	0.3686
←	Vape Use	Marijuana use frequency	0.1049	0.7982	0.008	0.004	0.007	0.0779	0
←	Social Competency Scale	Prescription Substances Use (not prescribed to user)	0.0629	0.6464	0	0	0	0.005	0.2857
←	School Engagement	Marijuana use frequency	0	0.5145	0	0	0.001	0	0.4845
←	ACEs	Prescription Substances Use (not prescribed to user)	0.018	0.981	0	0.001	0	0	0
←	Social Competency Scale	Marijuana use frequency	0.014	0.5155	0.004	0.1309	0.0819	0.005	0.2488
←	Binge	Tobacco Product Use	0.4086	0.5774	0	0.006	0	0.008	0
←	Reason of school absence: Suspension	Social Competency Scale	0	0.4785	0	0	0	0.001	0.5205
←	Social Competency Scale	Friends' Approval of Substance Use	0.1409	0.7172	0	0.0669	0.004	0.0709	0
←	Friends' Approval of Substance Use	Marijuana use frequency	0.03	0.6213	0.006	0.005	0.3267	0.011	0
←	Tobacco Product Use	Vape Use	0.1229	0.4635	0.006	0.043	0.3327	0.032	0
←	ACEs	Vape Use	0.017	0.3776	0	0.3586	0	0.2398	0.007
←	Substance use – 1	Marijuana use frequency	0.1179	0.3836	0.002	0	0	0.0529	0.4436

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	Social Competency Scale	Vape Use	0.1029	0.3686	0	0.3457	0.004	0.1788	0
←	Reason of school absence: Suspension	Marijuana use frequency	0	0.3586	0	0	0	0	0.6414
←	Tobacco Product Use	Friends' Approval of Substance Use	0.4406	0.4016	0.007	0.016	0.022	0.005	0.1079
←	ACEs	Non-medical marijuana use	0.014	0.3537	0	0.001	0	0.005	0.6264
←	Vape Use	Friends' Approval of Substance Use	0.042	0.3707	0.1608	0.019	0.0799	0.3277	0
←	School Engagement	Vape Use	0.003	0.3127	0	0.003	0.007	0.027	0.6474
←	ACEs	Social Competency Scale	0.3177	0.3567	0	0.2977	0.004	0.024	0
←	FDA	Suspension: sent out of the classroom	0.005	0.3496	0	0	0	0	0.6454
←	Prescription Substances Use (not prescribed to user)	Marijuana use frequency	0.003	0.2977	0	0	0	0	0.6993
←	Binge	Marijuana use frequency	0.043	0.2318	0.001	0	0	0.008	0.7163
←	Reason of school absence: Suspension	Tobacco Product Use	0	0.2198	0	0	0	0.004	0.7762
←	PDA	Non-medical marijuana use	0	0.1938	0	0	0	0.019	0.7872
←	Teacher Student Relationship	Vape Use	0	0.1508	0	0	0	0	0.8492
←	Reason of school absence: Suspension	Substance use – 1	0	0.0829	0	0.001	0	0.005	0.9111
←	Social Competency Scale	Substance use – 1	0.024	0.1378	0	0	0	0.001	0.8372
←	Binge	Social Competency Scale	0	0.0589	0	0	0	0	0.9411
←	Dental Checkup	Substance use – 1	0	0.053	0	0	0	0	0.9471
←	Reason of school absence: Suspension	Prescription Substances Use (not prescribed to user)	0	0.042	0	0	0	0	0.958
←	School Engagement	Non-medical marijuana use	0	0.032	0	0	0	0.001	0.967

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	School Engagement	Tobacco Product Use	0	0.022	0	0	0	0	0.978
←	Dental Checkup	Tobacco Product Use	0.001	0.023	0.001	0	0	0	0.975
←	Reason of school absence: Illness	Vape Use	0	0.006	0	0	0	0	0.994
←	Binge	Friends' Approval of Substance Use	0	0.006	0	0	0	0.002	0.992
←	Reason of school absence: Suspension	Vape Use	0	0.004	0	0	0	0	0.996
←	Suspension: sent out of the classroom	Vape Use	0	0.002	0	0	0	0	0.998
←	Binge	PDA	0	0.001	0	0	0	0	0.999
←	PDA	Teacher Student Relationship	0.001	0.001	0	0	0	0	0.998
←	Tobacco Product Use	Non-medical marijuana use	0	0.001	0	0	0	0	0.999
←	Prescription Substances Use (not prescribed to user)	Friends' Approval of Substance Use	0	0.001	0	0	0	0	0.999
←	Social Competency Scale	Non-medical marijuana use	0	0.002	0	0	0	0	0.998
←	Suspension: sent out of the classroom	Prescription Substances Use (not prescribed to user)	0	0.001	0	0	0	0	0.999
←0	ACEs	American Indian or Alaskan Native	0	0	0	1	0	0	0
←0	ACEs	Marijuana use frequency	0.017	0.1759	0.001	0.4396	0.002	0.005	0.3596
←0	ACEs	Friends' Approval of Substance Use	0.035	0.4845	0	0.4006	0	0.0739	0.006
←0	ACEs	Native Hawaiian or Other Pacific Islander	0	0	0	0.011	0	0	0.989
←0	FDA	American Indian or Alaskan Native	0	0	0	0.006	0	0	0.994
←0	Suspension: sent out of the classroom	American Indian or Alaskan Native	0	0	0	0.001	0	0	0.999

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
→	ACEs	PDA	0.99	0.007	0	0	0	0	0.003
→	ACEs	FDA	0.9471	0	0	0	0	0	0.0529
→	PDA	Suspension: sent out of the classroom	0.8691	0.1279	0.001	0	0	0	0.002
→	ACEs	Dental Checkup	1	0	0	0	0	0	0
→	Binge	Prescription Substances Use (not prescribed to user)	0.8482	0	0.002	0	0.002	0	0.1479
→	Tobacco Product Use	Prescription Substances Use (not prescribed to user)	0.997	0	0.002	0	0	0.001	0
→	ACEs	Teacher Student Relationship	0.9791	0.021	0	0	0	0	0
→	FDA	Dental Checkup	0.5555	0.001	0	0	0	0	0.4436
→	ACEs	Reason of school absence: Illness	0.8892	0	0	0	0	0	0.1109
→	Vape Use	Non-medical marijuana use	0.8542	0.014	0.006	0.039	0.0719	0.015	0
→	Tobacco Product Use	Substance use – 1	0.8412	0.1249	0.006	0.001	0.002	0.025	0
→	Binge	Non-medical marijuana use	0.5235	0.4566	0	0.001	0	0.019	0
→	Binge	Substance use – 1	0.4576	0.1209	0.001	0.1349	0.002	0.2837	0
→	Non-medical marijuana use	Prescription Substances Use (not prescribed to user)	0.7422	0	0.001	0	0	0.006	0.2507
→	Non-medical marijuana use	Substance use – 1	0.5414	0.4166	0.008	0	0.001	0.033	0
→	Substance use – 1	Prescription Substances Use (not prescribed to user)	0.6074	0.0899	0.024	0	0.1139	0.1648	0
→	Suspension: sent out of the classroom	Dental Checkup	0.2108	0	0	0	0	0	0.7892
→	Social Competency Scale	Tobacco Product Use	0.0879	0.032	0.001	0	0.001	0	0.8781
→	Vape Use	Prescription Substances Use (not prescribed to user)	0.025	0.001	0	0	0.001	0.006	0.967

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
→	ACEs	School Engagement	0.02	0	0	0	0	0	0.98
→	ACEs	Tobacco Product Use	0.016	0.001	0	0.001	0	0	0.982
→	ACEs	Substance use – 1	0.007	0	0	0	0	0.001	0.992
→	PDA	Friends' Approval of Substance Use	0.005	0.003	0	0	0	0	0.992
→	ACEs	Suspension: sent out of the classroom	0.002	0	0	0	0	0	0.998
→	FDA	Social Competency Scale	0.001	0	0	0	0	0.001	0.998
→	Suspension: sent out of the classroom	Friends' Approval of Substance Use	0.001	0	0	0	0	0	0.999
→	Non-medical marijuana use	Friends' Approval of Substance Use	0.001	0	0	0	0	0.001	0.998
→	American Indian or Alaskan Native	Marijuana use frequency	0.004	0	0.9421	0	0	0.003	0.0509
0→	American Indian or Alaskan Native	Friends' Approval of Substance Use	0	0	0.3497	0	0	0	0.6503
0→	American Indian or Alaskan Native	Social Competency Scale	0.009	0	0.2697	0	0	0.003	0.7183
0→	American Indian or Alaskan Native	Tobacco Product Use	0	0	0.1089	0	0	0	0.8911
0→	American Indian or Alaskan Native	Non-medical marijuana use	0	0	0.011	0	0	0	0.989
0→	Native Hawaiian or Other Pacific Islander	Substance use – 1	0	0	0.003	0	0	0	0.997
0→	American Indian or Alaskan Native	Substance use – 1	0	0	0.001	0	0	0	0.999

Table A7

Crosswalk between 2019 MSS Measures and Themes Noted in the LSN Focus Group

Variable	Item(s)	Response options	Description (Cronbach's α)
Free/reduced cost lunch	One item "Do you currently get free or reduced-price lunch at school?"	1 – Yes 2 – No 3 – Not sure	The response 'not sure' treated as 'NA' then imputed resulting in dichotomized variable of '1 – Yes' and '0 – No'.
Missed school: mental issues	One item "What are the reasons you missed a full or part of a day of school in the last 30 days? Felt very sad, hopeless, anxious, stressed or angry"	1 – Yes 0 – No	
Missed school: sleep issue	One item "What are the reasons you missed a full or part of a day of school in the last 30 days? Didn't get enough sleep"	1 – Yes 0 – No	
Missed school: transportation	One item "What are the reasons you missed a full or part of a day of school in the last 30 days? Missed your ride or didn't have a way to get to school"	1 – Yes 0 – No	
Missed school: had to work	One item "What are the reasons you missed a full or part of a day of school in the last 30 days? Had to work"	1 – Yes 0 – No	
Missed school: taking care of family or friend	One item "What are the reasons you missed a full or part of a day of school in the last 30 days? Had to take care of or help a family member or friend"	1 – Yes 0 – No	
Missed school: Housing instability	One item "What are the reasons you missed a full or part of a day of school in the last 30 days? Had no place to shower or wash clothes"	1 – Yes 0 – No	
Parent support	Five items "How much do you feel your parents care about you?"	1 – Not at all 2 – A little 3 – Some 4 – Quite a bit 5 – Very Much 1 – Yes	First item dichotomized to "low caring – 0" "high caring – 1" Rest three items dichotomized into "1 – replied yes to parent or

Variable	Item(s)	Response options	Description (Cronbach's α)
	"Which of these adults can you talk to about problems you are having? Parent or guardian" "Which of these adults can you talk to about problems you are having? Adult at school" "Which of these adults can you talk to about problems you are having? Some other adult"	0 – No	guardian", "0 – replied else" Then two items summed with the value of "0,1,2". Higher means better support from the parents (0.6)
Housing stability	Three items "During the past 12 months, have you stayed in a shelter, somewhere not intended as a place to live, or someone else's home because you had no other place to stay? No" "During the past 12 months, have you stayed in a shelter, somewhere not intended as a place to live, or someone else's home because you had no other place to stay? Yes, I was with my parents or adult family member" "During the past 12 months, have you stayed in a shelter, somewhere not intended as a place to live, or someone else's home because you had no other place to stay? Yes, I was on my own without any adult family members"	1 – Yes 0 – No	Dichotomized 1 – Any 0 – None
Mental health treatment history	Three items "Have you ever been treated for a mental health, emotional or behavioral problem? No" "Have you ever been treated for a mental health, emotional or behavioral problem? Yes, during the last year" "Have you ever been treated for a mental health, emotional or behavioral problem? Yes, more than a year ago"	1 – Yes 0 – No	Dichotomized 1 – Any 0 – None

Table A8

The Step 2 Causal Discovery Analysis Bootstrap Re-sampling Analysis (the MSS 2019)

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	FDA	PDA	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	FDA	Missed school: mental issues	0.000	0.983	0.000	0.017	0.000	0.000	0.000
←	FDA	Missed school: taking care of family or friend	0.000	0.990	0.000	0.005	0.000	0.000	0.005
←	Missed school: mental issues	Missed school: sleep issue	0.028	0.951	0.020	0.000	0.001	0.000	0.000
←	Mental health treatment history	Parent support	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	Missed school: mental issues	Parent support	0.111	0.854	0.030	0.000	0.000	0.005	0.000
←	Missed school: Housing instability	Missed school: transportation	0.000	0.813	0.000	0.000	0.000	0.000	0.187
←	Missed school: sleep issue	Missed school: transportation	0.039	0.784	0.011	0.000	0.166	0.000	0.000
←	Missed school: sleep issue	Parent support	0.010	0.769	0.017	0.006	0.079	0.000	0.119
←	Missed school: Housing instability	Missed school: mental issues	0.000	0.762	0.000	0.010	0.000	0.063	0.165
←	PDA	Parent support	0.097	0.666	0.000	0.000	0.000	0.000	0.237
←	Missed school: Housing instability	Housing stability	0.232	0.743	0.000	0.000	0.000	0.000	0.025
←	Missed school: had to work	Missed school: taking care of family or friend	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	Missed school: Housing instability	Missed school: had to work	0.488	0.509	0.000	0.003	0.000	0.000	0.001
←	Missed school: Housing instability	Missed school: taking care of family or friend	0.028	0.677	0.000	0.000	0.000	0.000	0.295
←	Missed school: Housing instability	Missed school: sleep issue	0.000	0.026	0.000	0.000	0.001	0.001	0.972

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	Missed school: mental issues	Missed school: transportation	0.000	0.001	0.000	0.000	0.000	0.000	0.999
←0	PDA	Mental health treatment history	0.097	0.245	0.000	0.657	0.000	0.001	0.000
→	Missed school: sleep issue	Missed school: had to work	0.872	0.000	0.000	0.000	0.000	0.000	0.128
→	Missed school: mental issues	Missed school: taking care of family or friend	0.885	0.005	0.048	0.000	0.000	0.062	0.000
→	Missed school: mental issues	Mental health treatment history	0.991	0.000	0.000	0.000	0.004	0.005	0.000
→	Missed school: transportation	PDA	0.828	0.000	0.172	0.000	0.000	0.000	0.000
→	Mental health treatment history	Housing stability	0.568	0.012	0.003	0.000	0.000	0.000	0.417
→	Missed school: sleep issue	Missed school: taking care of family or friend	0.347	0.005	0.097	0.000	0.000	0.012	0.540
→	FDA	Housing stability	0.054	0.000	0.000	0.000	0.000	0.000	0.946
→	Missed school: mental issues	Housing stability	0.021	0.008	0.001	0.000	0.000	0.003	0.967
→	Missed school: taking care of family or friend	PDA	0.004	0.000	0.000	0.000	0.000	0.000	0.996
0→	Free or reduced-price lunch at school	FDA	0.000	0.000	1.000	0.000	0.000	0.000	0.000
0→	Free or reduced-price lunch at school	Missed school: transportation	0.000	0.000	1.000	0.000	0.000	0.000	0.000
0→	Free or reduced-price lunch at school	Missed school: taking care of family or friend	0.000	0.000	1.000	0.000	0.000	0.000	0.000
0→	Free or reduced-price lunch at school	Parent support	0.000	0.000	1.000	0.000	0.000	0.000	0.000
0→	Free or reduced-price lunch at school	Housing stability	0.008	0.000	0.992	0.000	0.000	0.000	0.000
0→	Missed school: transportation	Missed school: taking care of family or friend	0.038	0.000	0.950	0.000	0.005	0.007	0.000

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
0→	Parent support	Housing stability	0.164	0.000	0.717	0.000	0.119	0.000	0.000
0→	Missed school: mental issues	PDA	0.297	0.000	0.703	0.000	0.000	0.000	0.000
0→	Missed school: sleep issue	PDA	0.354	0.000	0.646	0.000	0.000	0.000	0.000
0→	Free or reduced-price lunch at school	Missed school: sleep issue	0.000	0.000	0.216	0.000	0.000	0.000	0.784
0→	Free or reduced-price lunch at school	PDA	0.000	0.000	0.175	0.000	0.000	0.000	0.825
0→	Free or reduced-price lunch at school	Missed school: mental issues	0.000	0.000	0.001	0.000	0.000	0.000	0.999
0—0	Missed school: transportation	Parent support	0.013	0.007	0.035	0.000	0.758	0.000	0.187
0—0	Missed school: transportation	Housing stability	0.000	0.007	0.002	0.005	0.009	0.000	0.977

Table A9

The Step 3 Causal Discovery Analysis Bootstrap Re-sampling Analysis (the MSS 2019)

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	FDA	Reason of school absence: Illness	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	Reason of school absence: Illness	Missed school: sleep issue	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	FDA	Missed school: mental issues	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	Reason of school absence: Illness	Missed school: mental issues	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	FDA	School Engagement	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	PDA	School Engagement	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	Suspension: sent out of the classroom	Teacher Student Relationship	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	School Engagement	Social Competency Scale	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	Teacher Student Relationship	Social Competency Scale	0.000	0.999	0.000	0.001	0.000	0.000	0.000
←	Missed school: had to work	Tobacco Product Use	0.000	0.992	0.000	0.000	0.000	0.007	0.001
←	Suspension: sent out of the classroom	School Engagement	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	Suspension: sent out of the classroom	Tobacco Product Use	0.003	0.959	0.000	0.000	0.000	0.004	0.034
←	School Engagement	Friends’ Approval of Substance Use	0.040	0.954	0.000	0.000	0.000	0.006	0.000
←	Mental health treatment history	Prescription Substances Use (not prescribed to user)	0.000	0.996	0.000	0.000	0.000	0.001	0.003
←	Teacher Student Relationship	Friends’ Approval of Substance Use	0.001	0.957	0.036	0.000	0.000	0.002	0.004
←	Dental Checkup	Social Competency Scale	0.003	0.996	0.000	0.000	0.000	0.001	0.000

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	School Engagement	Parent support	0.002	0.974	0.000	0.000	0.000	0.000	0.024
←	Missed school: sleep issue	School Engagement	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	School Engagement	Teacher Student Relationship	0.002	0.998	0.000	0.000	0.000	0.000	0.000
←	Housing stability	Substance use – 1	0.003	0.988	0.000	0.003	0.000	0.006	0.000
←	Suspension: sent out of the classroom	Substance use – 1	0.005	0.961	0.000	0.003	0.000	0.008	0.023
←	Reason of school absence: Suspension	Marijuana use frequency	0.000	0.875	0.000	0.000	0.000	0.002	0.123
←	Missed school: mental issues	Parent support	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	Parent support	Social Competency Scale	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	Missed school: had to work	Missed school: taking care of family or friend	0.025	0.975	0.000	0.000	0.000	0.000	0.000
←	Missed school: Housing instability	Missed school: had to work	0.159	0.839	0.000	0.000	0.000	0.000	0.002
←	Missed school: mental issues	Teacher Student Relationship	0.075	0.875	0.000	0.000	0.000	0.000	0.050
←	Reason of school absence: Suspension	Suspension: sent out of the classroom	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	Tobacco Product Use	Marijuana use frequency	0.002	0.792	0.001	0.012	0.002	0.010	0.181
←	ACEs	Prescription Substances Use (not prescribed to user)	0.046	0.953	0.000	0.000	0.000	0.001	0.000
←	Social Competency Scale	Marijuana use frequency	0.070	0.751	0.000	0.025	0.000	0.005	0.149
←	Binge	Vape Use	0.006	0.934	0.000	0.037	0.007	0.016	0.000
←	Non-medical marijuana use	Marijuana use frequency	0.030	0.965	0.001	0.000	0.001	0.003	0.000
←	Social Competency Scale	Prescription Substances Use (not prescribed to user)	0.025	0.796	0.000	0.000	0.000	0.001	0.178

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	Free or reduced-price lunch at school	FDA	0.286	0.714	0.000	0.000	0.000	0.000	0.000
←	FDA	PDA	0.000	1.000	0.000	0.000	0.000	0.000	0.000
←	PDA	Vape Use	0.000	0.689	0.000	0.140	0.000	0.097	0.074
←	PDA	Marijuana use frequency	0.001	0.633	0.000	0.001	0.000	0.005	0.360
←	Free or reduced-price lunch at school	Housing stability	0.365	0.633	0.000	0.000	0.000	0.002	0.000
←	Missed school: mental issues	Marijuana use frequency	0.000	0.660	0.000	0.000	0.000	0.004	0.336
←	Vape Use	Marijuana use frequency	0.097	0.779	0.007	0.014	0.003	0.100	0.000
←	Dental Checkup	Housing stability	0.006	0.881	0.000	0.000	0.000	0.000	0.113
←	Tobacco Product Use	Vape Use	0.057	0.619	0.000	0.051	0.238	0.035	0.000
←	Friends' Approval of Substance Use	Marijuana use frequency	0.068	0.625	0.000	0.004	0.296	0.007	0.000
←	Teacher Student Relationship	Parent support	0.053	0.525	0.000	0.385	0.038	0.000	0.000
←	Missed school: Housing instability	Missed school: transportation	0.000	0.491	0.000	0.000	0.000	0.000	0.510
←	Missed school: Housing instability	Missed school: mental issues	0.000	0.470	0.000	0.000	0.000	0.000	0.531
←	ACEs	Friends' Approval of Substance Use	0.048	0.613	0.000	0.332	0.000	0.007	0.000
←	Parent support	Prescription Substances Use (not prescribed to user)	0.001	0.455	0.000	0.000	0.000	0.000	0.545
←	Social Competency Scale	Friends' Approval of Substance Use	0.103	0.873	0.000	0.012	0.000	0.012	0.000
←	Missed school: had to work	Suspension: sent out of the classroom	0.000	0.573	0.000	0.000	0.000	0.000	0.428
←	Missed school: Housing instability	Missed school: taking care of family or friend	0.005	0.548	0.000	0.000	0.000	0.000	0.448

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	ACEs	Vape Use	0.003	0.413	0.000	0.342	0.000	0.191	0.052
←	Missed school: sleep issue	Vape Use	0.000	0.397	0.000	0.000	0.000	0.123	0.481
←	Missed school: Housing instability	Housing stability	0.025	0.570	0.000	0.000	0.000	0.000	0.405
←	Housing stability	Tobacco Product Use	0.004	0.386	0.000	0.001	0.000	0.000	0.609
←	Missed school: Housing instability	Suspension: sent out of the classroom	0.000	0.548	0.000	0.000	0.000	0.000	0.452
←	ACEs	Non-medical marijuana use	0.030	0.348	0.000	0.002	0.000	0.029	0.591
←	Binge	Non-medical marijuana use	0.306	0.627	0.000	0.003	0.001	0.063	0.000
←	Missed school: sleep issue	Non-medical marijuana use	0.000	0.343	0.000	0.000	0.000	0.034	0.623
←	Vape Use	Friends' Approval of Substance Use	0.095	0.325	0.281	0.008	0.013	0.279	0.000
←	Missed school: Housing instability	Substance use – 1	0.000	0.300	0.000	0.000	0.000	0.003	0.697
←	PDA	Non-medical marijuana use	0.000	0.286	0.000	0.001	0.000	0.008	0.705
←	Prescription Substances Use (not prescribed to user)	Marijuana use frequency	0.002	0.280	0.000	0.000	0.000	0.003	0.715
←	Missed school: mental issues	Vape Use	0.000	0.278	0.000	0.000	0.000	0.061	0.661
←	Reason of school absence: Suspension	Tobacco Product Use	0.000	0.238	0.000	0.000	0.000	0.004	0.758
←	Housing stability	Prescription Substances Use (not prescribed to user)	0.000	0.203	0.000	0.000	0.000	0.001	0.796
←	Binge	Marijuana use frequency	0.032	0.193	0.000	0.000	0.000	0.002	0.773
←	FDA	Reason of school absence: Suspension	0.000	0.179	0.000	0.000	0.000	0.000	0.821
←	FDA	Suspension: sent out of the classroom	0.000	0.161	0.000	0.000	0.000	0.000	0.839

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	Substance use – 1	Marijuana use frequency	0.048	0.150	0.002	0.000	0.000	0.065	0.735
←	Reason of school absence: Suspension	Non-medical marijuana use	0.000	0.120	0.000	0.001	0.000	0.000	0.879
←	Reason of school absence: Suspension	Substance use – 1	0.000	0.115	0.000	0.000	0.000	0.001	0.884
←	Missed school: sleep issue	Marijuana use frequency	0.000	0.097	0.000	0.000	0.000	0.000	0.903
←	School Engagement	Marijuana use frequency	0.002	0.095	0.000	0.002	0.000	0.000	0.901
←	Reason of school absence: Suspension	Prescription Substances Use (not prescribed to user)	0.000	0.092	0.000	0.000	0.000	0.000	0.908
←	Reason of school absence: Suspension	Missed school: transportation	0.002	0.097	0.000	0.000	0.000	0.000	0.901
←	Teacher Student Relationship	Vape Use	0.000	0.058	0.000	0.000	0.000	0.000	0.942
←	Missed school: transportation	Dental Checkup	0.005	0.053	0.000	0.000	0.000	0.000	0.942
←	Reason of school absence: Suspension	Social Competency Scale	0.000	0.042	0.000	0.000	0.000	0.000	0.958
←	Social Competency Scale	Substance use – 1	0.002	0.070	0.000	0.000	0.000	0.000	0.928
←	Binge	Social Competency Scale	0.000	0.043	0.000	0.000	0.000	0.000	0.957
←	Missed school: mental issues	School Engagement	0.000	0.036	0.000	0.000	0.000	0.000	0.964
←	Parent support	Substance use – 1	0.000	0.024	0.000	0.001	0.000	0.000	0.975
←	FDA	Missed school: taking care of family or friend	0.000	0.012	0.000	0.000	0.000	0.000	0.988
←	Suspension: sent out of the classroom	Prescription Substances Use (not prescribed to user)	0.000	0.012	0.000	0.000	0.000	0.000	0.988
←	Parent support	Friends' Approval of Substance Use	0.000	0.009	0.000	0.000	0.000	0.000	0.991
←	Missed school: Housing instability	Missed school: sleep issue	0.000	0.006	0.000	0.000	0.000	0.000	0.994

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←	Missed school: had to work	School Engagement	0.000	0.005	0.000	0.000	0.000	0.000	0.995
←	Binge	Friends' Approval of Substance Use	0.000	0.008	0.000	0.001	0.000	0.000	0.991
←	Missed school: transportation	Social Competency Scale	0.000	0.003	0.000	0.000	0.000	0.000	0.997
←	Missed school: Housing instability	Prescription Substances Use (not prescribed to user)	0.000	0.004	0.000	0.000	0.000	0.000	0.996
←	Tobacco Product Use	Non-medical marijuana use	0.001	0.002	0.000	0.000	0.000	0.000	0.997
←	Mental health treatment history	Non-medical marijuana use	0.000	0.002	0.000	0.000	0.000	0.000	0.998
←	Dental Checkup	Substance use – 1	0.000	0.002	0.000	0.000	0.000	0.000	0.998
←	Missed school: transportation	Suspension: sent out of the classroom	0.000	0.001	0.000	0.000	0.000	0.000	0.999
←	Binge	PDA	0.000	0.001	0.000	0.000	0.000	0.000	0.999
←	Reason of school absence: Suspension	Teacher Student Relationship	0.000	0.001	0.000	0.000	0.000	0.000	0.999
←	Missed school: Housing instability	Tobacco Product Use	0.000	0.001	0.000	0.000	0.000	0.000	0.999
←	Dental Checkup	Tobacco Product Use	0.001	0.001	0.000	0.000	0.000	0.000	0.998
←	Missed school: mental issues	Non-medical marijuana use	0.000	0.001	0.000	0.000	0.000	0.000	0.999
←	Social Competency Scale	Non-medical marijuana use	0.000	0.002	0.000	0.000	0.000	0.000	0.998
←→	Free or reduced-price lunch at school	Social Competency Scale	0.058	0.125	0.000	0.002	0.000	0.814	0.001
←→	ACEs	Free or reduced-price lunch at school	0.225	0.000	0.000	0.009	0.000	0.766	0.000
←0	ACEs	American Indian or Alaskan Native	0.000	0.005	0.000	0.995	0.000	0.000	0.000
←0	Free or reduced-price lunch at school	American Indian or Alaskan Native	0.000	0.224	0.000	0.776	0.000	0.000	0.000

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
←0	Social Competency Scale	Vape Use	0.010	0.243	0.000	0.503	0.000	0.245	0.000
←0	ACEs	Social Competency Scale	0.173	0.384	0.000	0.411	0.018	0.015	0.000
←0	ACEs	Marijuana use frequency	0.045	0.199	0.000	0.373	0.000	0.006	0.378
←0	ACEs	Native Hawaiian or Other Pacific Islander	0.000	0.000	0.000	0.010	0.000	0.000	0.990
→	ACEs	Missed school: mental issues	1.000	0.000	0.000	0.000	0.000	0.000	0.000
→	ACEs	Mental health treatment history	1.000	0.000	0.000	0.000	0.000	0.000	0.000
→	Missed school: mental issues	Missed school: taking care of family or friend	1.000	0.000	0.000	0.000	0.000	0.000	0.000
→	Missed school: mental issues	PDA	1.000	0.000	0.000	0.000	0.000	0.000	0.000
→	Free or reduced-price lunch at school	Missed school: taking care of family or friend	1.000	0.000	0.000	0.000	0.000	0.000	0.000
→	ACEs	Missed school: transportation	0.996	0.000	0.000	0.000	0.000	0.000	0.004
→	ACEs	Dental Checkup	1.000	0.000	0.000	0.000	0.000	0.000	0.000
→	Missed school: transportation	Missed school: taking care of family or friend	0.999	0.001	0.000	0.000	0.000	0.000	0.000
→	ACEs	Missed school: sleep issue	0.965	0.000	0.000	0.000	0.000	0.000	0.035
→	ACEs	Teacher Student Relationship	0.974	0.000	0.000	0.000	0.000	0.000	0.026
→	Missed school: mental issues	Missed school: sleep issue	1.000	0.000	0.000	0.000	0.000	0.000	0.000
→	Missed school: mental issues	Mental health treatment history	1.000	0.000	0.000	0.000	0.000	0.000	0.000
→	Binge	Prescription Substances Use (not prescribed to user)	0.907	0.000	0.001	0.000	0.002	0.000	0.090
→	PDA	Mental health treatment history	0.875	0.047	0.000	0.000	0.000	0.000	0.078

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
→	Free or reduced-price lunch at school	Missed school: transportation	1.000	0.000	0.000	0.000	0.000	0.000	0.000
→	Tobacco Product Use	Prescription Substances Use (not prescribed to user)	0.995	0.004	0.000	0.000	0.000	0.001	0.000
→	ACEs	Parent support	1.000	0.000	0.000	0.000	0.000	0.000	0.000
→	Binge	Substance use – 1	0.797	0.033	0.000	0.026	0.002	0.142	0.000
→	Missed school: sleep issue	Missed school: transportation	1.000	0.000	0.000	0.000	0.000	0.000	0.000
→	Missed school: sleep issue	PDA	0.982	0.000	0.018	0.000	0.000	0.000	0.000
→	Non-medical marijuana use	Substance use – 1	0.769	0.119	0.000	0.000	0.000	0.112	0.000
→	Free or reduced-price lunch at school	Dental Checkup	0.993	0.003	0.000	0.001	0.000	0.003	0.000
→	Non-medical marijuana use	Prescription Substances Use (not prescribed to user)	0.736	0.001	0.000	0.000	0.000	0.005	0.258
→	ACEs	Housing stability	0.986	0.008	0.002	0.001	0.000	0.003	0.000
→	Tobacco Product Use	Substance use – 1	0.884	0.106	0.002	0.000	0.000	0.008	0.000
→	Binge	Tobacco Product Use	0.569	0.416	0.000	0.003	0.000	0.013	0.000
→	Tobacco Product Use	Friends' Approval of Substance Use	0.576	0.335	0.001	0.006	0.015	0.003	0.064
→	Free or reduced-price lunch at school	Suspension: sent out of the classroom	0.783	0.004	0.000	0.000	0.000	0.000	0.213
→	PDA	Suspension: sent out of the classroom	0.955	0.045	0.000	0.000	0.000	0.000	0.000
→	Vape Use	Non-medical marijuana use	0.619	0.042	0.029	0.037	0.220	0.053	0.000
→	Missed school: transportation	PDA	0.286	0.000	0.000	0.000	0.000	0.000	0.714
→	ACEs	Missed school: taking care of family or friend	0.282	0.000	0.000	0.000	0.000	0.000	0.718

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
→	Missed school: Housing instability	Reason of school absence: Suspension	0.238	0.175	0.000	0.000	0.000	0.000	0.587
→	Missed school: sleep issue	Missed school: taking care of family or friend	0.195	0.000	0.000	0.000	0.000	0.000	0.805
→	Missed school: sleep issue	Missed school: had to work	0.130	0.000	0.000	0.000	0.000	0.000	0.870
→	Suspension: sent out of the classroom	Housing stability	0.113	0.001	0.000	0.000	0.000	0.000	0.886
→	Social Competency Scale	Tobacco Product Use	0.049	0.014	0.000	0.000	0.000	0.001	0.936
→	Binge	Missed school: Housing instability	0.035	0.000	0.000	0.000	0.000	0.000	0.965
→	School Engagement	Vape Use	0.035	0.028	0.000	0.000	0.000	0.000	0.937
→	Dental Checkup	Parent support	0.031	0.002	0.000	0.000	0.000	0.000	0.967
→	Reason of school absence: Suspension	Missed school: had to work	0.032	0.005	0.000	0.000	0.000	0.000	0.963
→	ACEs	Tobacco Product Use	0.021	0.000	0.000	0.000	0.000	0.001	0.978
→	Binge	Missed school: had to work	0.012	0.000	0.000	0.000	0.000	0.000	0.988
→	Vape Use	Prescription Substances Use (not prescribed to user)	0.008	0.000	0.001	0.000	0.000	0.002	0.989
→	Free or reduced-price lunch at school	Reason of school absence: Suspension	0.009	0.000	0.000	0.000	0.000	0.000	0.991
→	ACEs	Substance use – 1	0.004	0.001	0.000	0.000	0.000	0.000	0.995
→	Non-medical marijuana use	Friends' Approval of Substance Use	0.001	0.000	0.000	0.000	0.000	0.000	0.999
0→	American Indian or Alaskan Native	Marijuana use frequency	0.000	0.000	0.889	0.000	0.000	0.001	0.110
0→	American Indian or Alaskan Native	Friends' Approval of Substance Use	0.000	0.000	0.320	0.000	0.000	0.000	0.680
0→	American Indian or Alaskan Native	Housing stability	0.004	0.000	0.083	0.000	0.000	0.000	0.913

Interaction	Nodes		Proportion of 1,000 bootstrap resamples						
	Node 1	Node 2	→	←	0→	←0	0—0	←→	None
0→	American Indian or Alaskan Native	Social Competency Scale	0.002	0.000	0.074	0.000	0.000	0.001	0.923
0→	American Indian or Alaskan Native	Tobacco Product Use	0.001	0.000	0.045	0.000	0.000	0.000	0.954
0→	American Indian or Alaskan Native	Non-medical marijuana use	0.000	0.000	0.024	0.000	0.000	0.000	0.976
0—0	American Indian or Alaskan Native	Native Hawaiian or Other Pacific Islander	0.001	0.000	0.000	0.000	0.997	0.000	0.002
0—0	Substance use – 1	Prescription Substances Use (not prescribed to user)	0.455	0.068	0.008	0.001	0.368	0.101	0.000

Appendix B

Tables For Chapter IV

Table B1

Sample Characteristics and Distribution of Study Variables from the MSS 2016

	n	%
Sex		
Female	60,074	50
Male	60,931	50
Grade		
8th grade	42,791	35
9th grade	43,246	36
11th grade	34,968	29
Race and ethnicity		
American Indian or Alaskan Native only	2,190	2
Asian only	7,627	6
Black, African, or African American only	8,317	7
Native Hawaiian, or Pacific Islander only	555	1
White only	91,940	76
Multiple racial/ethnic groups	10,376	9
Ethnicity: Hispanic or Latino	8,327	7
Free/Reduced Cost Lunch		
Yes	32254	27
No	88751	73
School Region		
7-County Twin cities Metro Area	64,165	53
Greater Minnesota	56,840	47
Outcome Variable		
School absences – High	7,630	6
School absences – Low	113,375	94
Independent Variables		
Last time seeing a doctor or nurse		
Any	115,019	95
None	3,190	3
NA	2,796	2

	N	*M*	*SD*	Range
Stayed home due to sickness	120,555	1.55	0.75	1-5
Sent to office for discipline	121,330	1.11	0.43	1-5
In-school suspension	120,508	1.04	0.26	1-5
Out-of-school suspension	120,434	1.02	0.19	1-5
School Engagement Scale (SE)	119,439	3.15	0.47	1-4
Teacher-student Relationship Scale (TSR)	118,191	2.04	0.59	1-4
Social Competency Scale (SCS)	113,254	3.07	0.60	1-4
Adversarial Childhood Experiences (ACEs)	112,690	0.51	1.01	0-7
Tobacco product use (TBP)	112,999	0.23	0.72	0-5

Table B2

Correlation Attribute Evaluation

Attribute	Average merit	Average rank
Tobacco Product Use	0.24 +- 0.001	1 +- 0
Sent to office for disciplinary issue	0.235 +- 0.003	2 +- 0
Marijuana use past year	0.213 +- 0.002	3 +- 0
Substances Use (Methamphetamine, Cocaine, etc.)	0.204 +- 0.001	4.2 +- 0.4
Marijuana use frequency	0.202 +- 0.001	5.3 +- 1
Prescription drug usage to get high (Vicodin, Valium, etc.)	0.2 +- 0.003	6.2 +- 0.6
In-school suspension	0.2 +- 0.002	6.3 +- 0.78
Social competency Scale (SCS)	0.193 +- 0.001	8.6 +- 0.8
Binge drinking – 2 (5 or more drinks in a row)	0.193 +- 0.001	9 +- 0.89
School engagement (SE)	0.192 +- 0.001	9.4 +- 0.49
Friends approval of substance use (FAS)	0.171 +- 0.001	11.2 +- 0.4
Out-of-school suspension	0.169 +- 0.002	12.3 +- 1.19
Staying home due to sickness	0.167 +- 0.001	13.1 +- 0.7
Teacher-student relationship (TSR)	0.167 +- 0.001	13.6 +- 0.49
Crime/violence subscription	0.164 +- 0.001	14.9 +- 0.7
ACEs	0.162 +- 0.001	16.1 +- 0.54
Binge drinking – 1 (how much do you drink at one time)	0.16 +- 0.001	17.2 +- 0.87
Tobacco use frequency	0.159 +- 0.001	17.8 +- 0.6
Substances use frequency	0.156 +- 0.002	19.1 +- 0.7
Parents approval of drugs	0.155 +- 0.001	19.7 +- 0.46

Attribute	Average merit	Average rank
Perception of family caring	0.15 +- 0.001	21.4 +- 0.66
Alcohol consumption frequency	0.149 +- 0.001	22 +- 0.63
Runaway	0.147 +- 0.002	22.6 +- 0.66
Global Appraisal of Individual Needs	0.142 +- 0.001	24 +- 0
Perpetrator	0.137 +- 0.001	25 +- 0
Positive Identity Scale	0.129 +- 0.001	26.6 +- 0.66
Substances use treatment history	0.13 +- 0.002	26.6 +- 0.66
Free/reduced lunch	0.128 +- 0.001	28 +- 0.63
Empowerment	0.127 +- 0.001	29.4 +- 0.66
Alcohol consumption frequency	0.125 +- 0.001	30.4 +- 1.02
Positive youth development scale	0.125 +- 0.001	30.5 +- 1.12
Incarcerated parents	0.124 +- 0.002	31.5 +- 0.92
Relationship with mother	0.12 +- 0.001	33.7 +- 1
School nurse office visit	0.118 +- 0.001	35.5 +- 2.11
Perception of safety while a commuting	0.118 +- 0.001	35.5 +- 1.43
Hostile school climate by peers	0.118 +- 0.001	35.9 +- 1.92
Perception of school safety	0.118 +- 0.001	36.3 +- 1.55
General health	0.117 +- 0.001	37.9 +- 1.58
Perceptions of substance use risk	0.116 +- 0.001	38.4 +- 1.96
Race & Ethnicity: White Non-Hispanic	0.116 +- 0.001	39.1 +- 1.3
Neighborhood safety	0.115 +- 0.001	40.9 +- 0.83
Attitudes toward drinking	0.113 +- 0.001	41.9 +- 0.54
Suicidal attempt	0.111 +- 0.001	43 +- 0.45
Perpetrator	0.11 +- 0.001	44.8 +- 1.17
Sleep during school day	0.11 +- 0.001	45.5 +- 1.12
Intimate partner violence	0.109 +- 0.002	46.4 +- 2.06
Home safety	0.109 +- 0.001	46.5 +- 1.75
Race: White only	0.107 +- 0.001	48.9 +- 1.22
Online bullying	0.107 +- 0.002	49 +- 1.67
Skipping meal due to financial issues	0.106 +- 0.002	49 +- 2
Relationship with father	0.105 +- 0.001	51 +- 1.18
Perceptions of caring from adults in the community	0.104 +- 0.001	51.1 +- 1.14
None-suicidal self-injury	0.103 +- 0.001	52.8 +- 0.6
Transient student	0.101 +- 0.002	53.9 +- 0.3
Suicidal Ideation	0.097 +- 0.001	55 +- 0
Long-term mental health history	0.094 +- 0.001	56.1 +- 0.3

Attribute	Average merit	Average rank
Global Appraisal of Individual Needs-1	0.093 +- 0	57.1 +- 0.5
Out-of-school activity: Sports	0.091 +- 0.002	58.5 +- 1.02
Harassed by peers: disability	0.09 +- 0.002	58.7 +- 0.78
Patient health questionnaire-2 (PHQ2)	0.088 +- 0.001	60.4 +- 0.92
Harassed by peers: race, ethnicity or national origin	0.088 +- 0.002	60.4 +- 0.92
Harassed by peers: LGB	0.086 +- 0.001	62 +- 0.63
Mental health treatment history	0.084 +- 0.001	62.9 +- 0.54
Harassed by peers: physical appearance	0.081 +- 0.001	64 +- 0.45
Perception of peer caring	0.078 +- 0.001	65.4 +- 0.66
Harassed by peers: size or weight	0.078 +- 0.002	66.1 +- 1.04
Race: Black, African or African American only	0.075 +- 0.001	67.5 +- 0.92
Harassed by peers: gender	0.074 +- 0.001	68.9 +- 1.04
Special education	0.072 +- 0.002	70.1 +- 0.83
Grade	0.071 +- 0.001	70.8 +- 0.6
Race & Ethnicity: Black Non-Hispanic	0.068 +- 0.001	72.3 +- 0.64
Race & Ethnicity: Hispanic	0.067 +- 0.002	72.7 +- 0.9
Family structure (two-parent household VS else)	0.064 +- 0.001	74.1 +- 0.83
Harassed by peers: religion	0.062 +- 0.001	75.5 +- 1.12
Marijuana use frequency	0.063 +- 0.001	75.7 +- 0.64
Afterschool activity: Youth center	0.062 +- 0.001	76.5 +- 0.67
Race: America Indian only	0.057 +- 0.001	78.3 +- 0.46
Brief adolescent gambling screen	0.056 +- 0.002	78.9 +- 0.83
Out-of-school activity: Academic program including tutoring	0.052 +- 0.001	80.5 +- 0.81
Peers' Attitudes toward drinking	0.051 +- 0.001	80.7 +- 0.9
Race & Ethnicity: American Indian Non-Hispanic	0.05 +- 0.001	81.6 +- 0.49
Risky behavior while driving	0.048 +- 0.001	83.8 +- 0.98
Physically Active	0.047 +- 0.001	83.9 +- 0.7
Multiple Races (checked more than one)	0.046 +- 0.002	85 +- 1.1
Dental checkup	0.045 +- 0.001	85.5 +- 1.02
Afterschool activity: Outdoor	0.044 +- 0.001	86.8 +- 0.4
Tobacco use frequency	0.04 +- 0.001	88 +- 0
Gambling experience	0.037 +- 0.001	89.9 +- 0.94
Afterschool activity: Library	0.036 +- 0.001	90.4 +- 1.2
Measure of Homelessness	0.036 +- 0.001	91.1 +- 1.97

Attribute	Average merit	Average rank
Out-of-school activity: School sponsored activities (not sports)	0.035 +- 0.001	92.3 +- 0.9
Race & Ethnicity: Multiple Races Non-Hispanic	0.035 +- 0.002	93 +- 2.49
Out-of-school activity: Physical activity lessons	0.034 +- 0.001	94.2 +- 1.94
Alcohol consumption frequency	0.033 +- 0.001	94.6 +- 1.28
Afterschool activity: In-home	0.033 +- 0.001	95.5 +- 1.2
Medical checkup	0.032 +- 0.002	97.3 +- 1.95
Race: Native Hawaiian or Pacific Islander only	0.031 +- 0.001	98.1 +- 1.45
Region	0.031 +- 0.001	98.6 +- 0.92
Overweight	0.03 +- 0.001	99.2 +- 1.08
Physical education frequency	0.029 +- 0.001	100.8 +- 0.4
Afterschool activity: Youth center	0.023 +- 0.001	102.2 +- 0.4
Physical disabilities	0.022 +- 0.001	102.9 +- 0.54
Afterschool activity: In-school	0.02 +- 0.001	104.3 +- 0.64
History of asthma	0.018 +- 0.001	105.4 +- 0.8
Race & Ethnicity: Pacific Islander Non-Hispanic	0.019 +- 0.002	105.6 +- 1.28
History of diabetes	0.017 +- 0.001	106.9 +- 0.83
Race: Asian only	0.015 +- 0.001	107.8 +- 0.4
Race & Ethnicity: Asian Non-Hispanic	0.014 +- 0.001	109 +- 0.45
Community programs	0.012 +- 0.001	110.5 +- 0.81
Out-of-school activity: Artistic lessons	0.012 +- 0.001	110.7 +- 0.78
Out-of-school activity: Leadership activities	0.011 +- 0.001	111.7 +- 0.46
Female	0.004 +- 0.001	113 +- 0
Male	0.004 +- 0.001	114 +- 0

Table B3

Information Gain Attribute Evaluation

Attribute	Average merit	Average rank
Social Competency Scale (SCS)	0.031 +- 0	1 +- 0
Tobacco Product Use	0.03 +- 0	2 +- 0
School Engagement	0.027 +- 0	3 +- 0
Friends approval of substance use (FAS)	0.025 +- 0	4 +- 0
Sent to office for disciplinary issue	0.024 +- 0	5 +- 0
Non-medical marijuana use frequency	0.023 +- 0	6 +- 0
Marijuana use frequency	0.021 +- 0	7.2 +- 0.6

Attribute	Average merit	Average rank
Staying home due to sickness	0.021 +- 0	8.3 +- 0.78
Teacher Student Relationship (TSR)	0.02 +- 0	8.6 +- 0.49
Substance use – 1	0.02 +- 0	9.9 +- 0.3
ACEs	0.018 +- 0	11.6 +- 0.49
Substance use – 2	0.018 +- 0	11.6 +- 0.66
Parents approval of drugs	0.018 +- 0	12.9 +- 0.7
Global Appraisal of Individual Needs	0.018 +- 0	13.9 +- 0.3
Empowerment	0.017 +- 0	15.8 +- 0.87
Perceptions of substance use risk	0.017 +- 0	16.1 +- 0.83
Binge drinking – 2 (5 or more drinks in a row)	0.017 +- 0	16.4 +- 1.02
Positive Identity Scale	0.016 +- 0	17.9 +- 0.83
Perception of family caring	0.016 +- 0	19.1 +- 0.7
Crime / Violence Subscription	0.016 +- 0	19.7 +- 0.46
Attitudes toward drinking	0.015 +- 0	21.7 +- 0.9
In-school suspension	0.015 +- 0	22 +- 1.18
Binge drinking – 1 (how much do you drink at one time)	0.015 +- 0	23.1 +- 1.22
Sleep during school day	0.015 +- 0	23.7 +- 0.78
Perceptions of caring from adults in the community	0.014 +- 0	25 +- 0.89
Positive Youth Development Scale	0.014 +- 0	25.5 +- 0.67
Alcohol consumption frequency	0.014 +- 0	27 +- 0
Tobacco use frequency	0.013 +- 0	28 +- 0
Perpetrator	0.012 +- 0	29 +- 0
General health	0.011 +- 0	30.2 +- 0.4
Substance use frequency	0.011 +- 0	30.8 +- 0.4
Free or reduced-price lunch at school	0.011 +- 0	32.4 +- 0.66
Out of school suspension	0.01 +- 0	33.5 +- 0.92
Runaway	0.01 +- 0	34.2 +- 1.33
School nurse office visit	0.01 +- 0	34.8 +- 1.17
Alcohol consumption frequency	0.01 +- 0	35.5 +- 1.43
Perception of peer caring	0.01 +- 0	36.7 +- 0.46
Hostile school climate by peers	0.009 +- 0	38.2 +- 0.75
Relationship with mother	0.009 +- 0	39.9 +- 1.22
Perception of school safety	0.009 +- 0	40.1 +- 0.94
Incarcerated parents	0.009 +- 0	41.1 +- 1.92
Perception of safety while commuting	0.009 +- 0	41.7 +- 1

Attribute	Average merit	Average rank
Race & Ethnicity: White Non-Hispanic	0.009 +- 0	42.1 +- 1.22
Relationship with father	0.008 +- 0	44 +- 0.63
Neighborhood safety	0.008 +- 0	44.8 +- 0.4
Out-of-school activity: Sports	0.008 +- 0	46.3 +- 0.64
Global Appraisal of Individual Needs-1	0.008 +- 0	46.9 +- 0.54
Race: White only	0.007 +- 0	48.7 +- 0.78
Online bullying	0.007 +- 0	49 +- 1.18
Home safety	0.007 +- 0	49.1 +- 0.7
Out-of-school activity: Religious activities	0.007 +- 0	51.6 +- 0.8
Substance use treatment history	0.007 +- 0	52.7 +- 1.42
Intimate partner violence	0.006 +- 0	54 +- 1.18
Suicidal attempt	0.006 +- 0	54.3 +- 1.1
Non-suicidal self-injury	0.006 +- 0	55.6 +- 0.66
Peers’ Attitudes toward drinking	0.006 +- 0	57.4 +- 0.49
Perpetrator	0.006 +- 0	58.1 +- 0.83
Skipping meal due to financial issues	0.006 +- 0	59 +- 1.34
Harassed by peers: physical appearance	0.006 +- 0	60.9 +- 1.22
Long-term mental health history	0.006 +- 0	61 +- 1.1
Suicidal Ideation	0.006 +- 0	61.2 +- 1.25
Physically Active	0.005 +- 0	62.9 +- 1.04
Transient student	0.005 +- 0	64.4 +- 1.62
Harassed by peers: race, ethnicity or national origin	0.005 +- 0	65.7 +- 1.19
Harassed by peers: size or weight	0.005 +- 0	66.4 +- 1.43
Patient health questionnaire-2 (PHQ2)	0.005 +- 0	66.8 +- 1.72
Risky behavior while driving	0.005 +- 0	66.9 +- 1.76
Harassed by peers: disability	0.005 +- 0	68.8 +- 0.98
Harassed by peers: LGB	0.005 +- 0	69.5 +- 0.67
Mental health treatment history	0.004 +- 0	71.2 +- 0.4
Grade	0.004 +- 0	72.2 +- 0.75
Harassed by peers: gender	0.004 +- 0	73.1 +- 0.83
Physical education frequency	0.004 +- 0	73.5 +- 0.67
Afterschool activity: In-home	0.003 +- 0	75.5 +- 0.67
Race: Black, African or African American only	0.003 +- 0	77.1 +- 1.45
Harassed by peers: religion	0.003 +- 0	77.7 +- 1.27
Out-of-school activity: Academic program including tutoring	0.003 +- 0	78.1 +- 1.7

Attribute	Average merit	Average rank
Marijuana use frequency	0.003 +- 0	78.4 +- 2.01
Special education	0.003 +- 0	78.7 +- 1.85
Out-of-school activity: School sponsored activities (not sports)	0.003 +- 0	81.6 +- 0.8
Afterschool activity: Outdoor	0.003 +- 0	82.3 +- 1.42
Out-of-school activity: Physical activity lessons	0.003 +- 0	82.9 +- 1.45
Family structure (two-parent household VS else)	0.003 +- 0	83.1 +- 1.3
Race & Ethnicity: Black Non-Hispanic	0.003 +- 0	85.5 +- 0.92
Race & Ethnicity: Hispanic	0.003 +- 0	85.6 +- 1.11
Out-of-school activity: Artistic lessons	0.003 +- 0	86.5 +- 0.67
Community programs	0.002 +- 0	88.6 +- 0.66
Afterschool activity: Library	0.002 +- 0	89.1 +- 0.83
Out-of-school activity: Leadership activities	0.002 +- 0	89.3 +- 0.78
Afterschool activity: In-school	0.002 +- 0	91.1 +- 0.3
Afterschool activity: Youth center	0.002 +- 0	92.1 +- 0.3
Race: American Indian only	0.002 +- 0	92.8 +- 0.6
Brief adolescent gambling screen	0.001 +- 0	95 +- 1.18
Race: Multiple Races	0.001 +- 0	95 +- 1.1
Race & Ethnicity: American Indian Non-Hispanic	0.001 +- 0	95.8 +- 0.87
Tobacco use frequency	0.001 +- 0	96.2 +- 0.75
Dental Checkup	0.001 +- 0	98.1 +- 0.3
Gambling experience	0.001 +- 0	99 +- 0.45
Alcohol consumption frequency	0.001 +- 0	100.2 +- 0.75
Region	0.001 +- 0	101 +- 0.63
Race & Ethnicity: Multiple Races Non-Hispanic	0.001 +- 0	101.8 +- 0.6
Medical checkup	0.001 +- 0	103.5 +- 0.67
Overweight	0.001 +- 0	103.6 +- 0.66
Measure of Homelessness	0.001 +- 0	105 +- 0.77
Race: Native Hawaiian or Pacific Islander only	0 +- 0	105.8 +- 0.4
Physical disabilities	0 +- 0	107 +- 0
History of asthma	0 +- 0	108.1 +- 0.3
History of diabetes	0 +- 0	109.6 +- 0.92
Race & Ethnicity: Pacific Islander Non-Hispanic	0 +- 0	110.3 +- 1.35
Race: Asian only	0 +- 0	110.4 +- 0.66
Race & Ethnicity: Asian Non-Hispanic	0 +- 0	111.6 +- 0.49
Female	0 +- 0	113 +- 0

Attribute	Average merit	Average rank
Male	0 +- 0	114 +- 0

Table B4

Ranked List of J48 Results during the Wrapper Method Feature Selection to MSS 2016

Attribute	Total Instance (%)
Out of school suspension	8(80 %)
Race & Ethnicity: American Indian Non-Hispanic	6(60 %)
Race: Native Hawaiian or Pacific Islander only	5(50 %)
In school suspension	4(40 %)
Physical checkup	4(40 %)
ACE	4(40 %)
Harassed by peers: LGB	3(30 %)
Substance use treatment history	3(30 %)
Grade	2(20 %)
Home safety	2(20 %)
Perception of peer caring	2(20 %)
Non-suicidal self-injury	2(20 %)
Binge drinking - 2	2(20 %)
Perception of family caring	2(20 %)
Suicidal attempt	2(20 %)
Perpetrator	2(20 %)
Homelessness	2(20 %)
Substance use - 1	2(20 %)
Relationship with father	1(10 %)
Staying home due to sickness	1(10 %)
Sent to office for discipline	1(10 %)
Perception of safety while commuting	1(10 %)
Perception of school safety	1(10 %)
Afterschool activity: In-home	1(10 %)
Afterschool activity: Youth center	1(10 %)
Out-of-school activity: Sports	1(10 %)
Out-of-school activity: Academic program including tutoring	1(10 %)
Out-of-school activity: Artistic lessons	1(10 %)
Out-of-school activity: Religious activities	1(10 %)
Physical disabilities	1(10 %)

Attribute	Total Instance (%)
Physical education frequency	1(10 %)
Skipping meal due to financial issues	1(10 %)
History of diabetes	1(10 %)
Sleep during school day	1(10 %)
Perceptions of caring from adults in the community	1(10 %)
Substance use frequency	1(10 %)
Alcohol consumption frequency	1(10 %)
Alcohol consumption frequency	1(10 %)
School Engagement	1(10 %)
Positive Youth Development Scale	1(10 %)
Risky behavior while driving	1(10 %)
Suicidal Ideation	1(10 %)
Runaway	1(10 %)
Substance use - 2	1(10 %)
Perceptions of substance use risk	1(10 %)
Parents' approval of substance use	1(10 %)
Family structure (two-parent household versus else)	1(10 %)
Race: American Indian only	1(10 %)
Race: Asian only	1(10 %)
Race & Ethnicity: Pacific Islander Non-Hispanic	1(10 %)
Race & Ethnicity: White Non-Hispanic	1(10 %)
Rest of the variables (*n* = 62)	0(0 %)

Table B5

Sample Characteristics and Distribution of Study Variables from the MSS 2019

	n	%
Sex		
Female	62,709	50.0
Male	62,375	49.8
NA	291	0.2
Grade		
8th grade	44,919	36
9th grade	45,232	36
11th grade	35,224	28
Race and ethnicity		

	n	%
American Indian or Alaskan Native only	1,519	1
Asian only	8,261	7
Black, African, or African American only	9,731	8
Hispanic or Latino only	7,650	6
Native Hawaiian, or Pacific Islander only	271	0.2
White only	86,077	69
Multiple racial/ethnic groups	10,865	9
NA	1001	0.8
Free/Reduced Cost Lunch		
Yes	29,007	23
No	79,638	64
Not sure	14,360	11
NA	2,375	2
School Region		
7-County Twin cities Metro Area	66,917	53.0
Greater Minnesota	58,458	47.0
Outcome Variable		
FDA High	19,899	16
Low	102,548	82
NA	2,928	2
PDA High	23,561	19
Low	99,114	79
NA	2,700	2
Independent Variables		
Last time seeing a doctor or nurse		
Any	117,795	94
None	1,660	1
NA	5,920	5
Missed school due to physical illness in the last 30 days.		
Yes	63,438	51
No	57,587	46
NA	4,350	3
Missed school due to suspension in the last 30 days.		

	n		%	
Yes	88,851		71	
No	1,412		1	
NA	35,112		28	
	N	*M*	*SD*	Range
Sent to office for discipline	122,229	1.13	0.47	1-5
School Engagement Scale (SE)	119,476	3.11	0.45	1-4
Teacher-student Relationship Scale (TSR)	120,619	2.88	0.58	1-4
Social Competency Scale (SCS)	111,106	3.00	0.60	1-4
Adversarial Childhood Experiences (ACEs)	108,507	0.94	1.24	0-4
Tobacco product use (TBP)	111,525	0.09	0.42	0-4

Table B6

The Step 1 Causal Discovery Analysis Effect Sizes (the MSS 2019)

Edges	Nodes		Standardized ES	Standard Error	95% CI		Z-score	p-value
	Node 1	Node 2			Lower	Upper		
←	Non-medical marijuana use	Marijuana use frequency	0.76	0.00	0.73	0.74	382.37	0.00
←	Vape Use	Marijuana use frequency	0.53	0.00	0.53	0.54	223.48	0.00
←	Prescription Substances Use (not prescribed to user)	Substance use – 1	0.52	0.00	0.40	0.41	151.79	0.00
←	Substance use – 1	Tobacco Product Use	0.38	0.00	0.39	0.40	150.90	0.00
←	Teacher Student Relationship	Social Competency Scale	0.34	0.00	0.38	0.39	142.94	0.00
←	School Engagement	Social Competency Scale	0.33	0.00	0.36	0.37	144.19	0.00
←	School Engagement	Teacher Student Relationship	0.33	0.00	0.33	0.34	122.39	0.00
←	Tobacco Product Use	Vape Use	0.31	0.00	0.33	0.34	128.14	0.00
←	Social Competency Scale	Friends' Approval of Substance Use	0.30	0.00	0.31	0.32	114.23	0.00
←	Binge	Tobacco Product Use	0.28	0.00	0.30	0.31	111.44	0.00
←	FDA	PDA	0.25	0.00	0.27	0.28	85.51	0.00
←	Binge	Vape Use	0.23	0.00	0.26	0.27	93.76	0.00
←	Reason of school absence: Suspension	Suspension: sent out of the classroom	0.20	0.00	0.20	0.22	49.54	0.00
←	Binge	Marijuana use frequency	0.15	0.00	0.19	0.21	68.75	0.00
←	Tobacco Product Use	Marijuana use frequency	0.14	0.01	0.18	0.19	67.42	0.00
←	Substance use – 1	Binge	0.13	0.00	0.16	0.18	54.75	0.00
←	Substance use – 1	Non-medical marijuana use	0.12	0.00	0.15	0.16	55.53	0.00
←	Non-medical marijuana use	Vape Use	0.10	0.00	0.15	0.16	53.30	0.00
←	Dental Checkup	ACEs	0.10	0.00	0.13	0.15	33.03	0.00

Edges	Nodes		Standardized	Standard	95% CI		Z-score	p-value
	Node 1	Node 2	ES	Error	Lower	Upper		
←	FDA	Reason of school absence: Illness	0.10	0.00	0.13	0.14	46.31	0.00
←	ACEs	Prescription Substances Use (not prescribed to user)	0.09	0.00	0.13	0.14	42.14	0.00
←	ACEs	Marijuana use frequency	0.08	0.01	0.12	0.13	66.66	0.00
←	Suspension: sent out of the classroom	PDA	0.07	0.00	0.11	0.13	45.14	0.00
←	ACEs	Vape Use	0.07	0.00	0.11	0.12	37.56	0.00
←	Teacher Student Relationship	Friends' Approval of Substance Use	0.07	0.00	0.11	0.12	38.10	0.00
←	Prescription Substances Use (not prescribed to user)	Tobacco Product Use	0.07	0.00	0.10	0.11	35.98	0.00
←	Suspension: sent out of the classroom	Tobacco Product Use	0.07	0.00	0.10	0.11	34.60	0.00
←	PDA	ACEs	0.07	0.00	0.09	0.10	27.45	0.00
←	PDA	Vape Use	0.07	0.00	0.08	0.09	31.49	0.00
←	Substance use – 1	Marijuana use frequency	0.07	0.00	0.07	0.08	27.90	0.00
←	Suspension: sent out of the classroom	Substance use – 1	0.06	0.00	0.06	0.07	20.91	0.00
←	Non-medical marijuana use	Binge	0.06	0.00	0.06	0.07	24.21	0.00
←	PDA	Marijuana use frequency	0.06	0.01	0.06	0.07	20.68	0.00
←	School Engagement	Friends' Approval of Substance Use	0.06	0.00	0.06	0.07	20.23	0.00
←	ACEs	Non-medical marijuana use	0.06	0.01	0.06	0.07	20.54	0.00
←	FDA	ACEs	0.05	0.00	0.05	0.07	17.31	0.00
←	FDA	Reason of school absence: Suspension	0.05	0.00	0.05	0.06	19.98	0.00
←	FDA	Suspension: sent out of the classroom	0.05	0.00	0.05	0.06	18.01	0.00

Edges	Nodes		Standardized ES	Standard Error	95% CI		Z-score	p-value
	Node 1	Node 2			Lower	Upper		
←	Prescription Substances Use (not prescribed to user)	Non-medical marijuana use	0.04	0.00	0.05	0.06	17.28	0.00
←	Prescription Substances Use (not prescribed to user)	Binge	0.04	0.00	0.05	0.06	15.80	0.00
←	Prescription Substances Use (not prescribed to user)	Marijuana use frequency	0.04	0.00	0.05	0.06	18.43	0.00
←	Dental Checkup	FDA	0.04	0.00	0.05	0.06	16.17	0.00
←	PDA	Non-medical marijuana use	0.04	0.01	0.05	0.06	18.26	0.00
←	Tobacco Product Use	Non-medical marijuana use	0.04	0.01	0.04	0.06	13.90	0.00
←	Suspension: sent out of the classroom	ACEs	0.03	0.00	0.04	0.05	18.34	0.00
←	Reason of school absence: Illness	ACEs	0.03	0.00	0.04	0.05	18.82	0.00
←	Prescription Substances Use (not prescribed to user)	Vape Use	0.03	0.00	0.04	0.05	15.18	0.00
←	Dental Checkup	Suspension: sent out of the classroom	0.03	0.00	0.03	0.05	8.54	0.00
←	Reason of school absence: Suspension	Non-medical marijuana use	0.03	0.01	0.04	0.05	13.17	0.00
←	Binge	PDA	0.03	0.00	0.03	0.05	12.88	0.00
←	Tobacco Product Use	ACEs	0.03	0.00	0.03	0.05	10.33	0.00
←	FDA	American Indian or Alaskan Native	0.03	0.00	0.03	0.04	13.96	0.00
←	Suspension: sent out of the classroom	Prescription Substances Use (not prescribed to user)	0.02	0.00	0.03	0.04	9.70	0.00
←	Reason of school absence: Suspension	Tobacco Product Use	0.02	0.00	0.02	0.04	6.92	0.00
←	Dental Checkup	Substance use – 1	0.02	0.00	0.03	0.04	10.97	0.00

Edges	Nodes		Standardized ES	Standard Error	95% CI		Z-score	p-value
	Node 1	Node 2			Lower	Upper		
←	Reason of school absence: Illness	Vape Use	0.02	0.00	0.03	0.04	13.39	0.00
←	Reason of school absence: Suspension	Marijuana use frequency	0.02	0.01	0.03	0.04	11.79	0.00
←	Dental Checkup	Tobacco Product Use	0.02	0.00	0.02	0.04	10.15	0.00
←	Reason of school absence: Suspension	Prescription Substances Use (not prescribed to user)	0.01	0.00	0.02	0.04	6.34	0.00
←	Reason of school absence: Suspension	Substance use – 1	0.01	0.00	0.02	0.04	5.71	0.00
←	Substance use – 1	ACEs	0.01	0.00	0.02	0.04	5.88	0.00
←	Reason of school absence: Suspension	Vape Use	0.01	0.00	0.02	0.03	9.60	0.00
←	Suspension: sent out of the classroom	Vape Use	0.01	0.00	0.02	0.03	8.06	0.00
←	School Engagement	Non-medical marijuana use	0.00	0.00	0.02	0.03	8.44	0.00
←	Binge	Social Competency Scale	-0.01	0.00	0.02	0.03	7.54	0.00
←	Binge	Friends' Approval of Substance Use	-0.01	0.00	0.02	0.03	8.32	0.00
←	Social Competency Scale	FDA	-0.01	0.00	0.02	0.03	8.76	0.00
←	Prescription Substances Use (not prescribed to user)	Friends' Approval of Substance Use	-0.02	0.00	0.02	0.03	8.03	0.00
←	School Engagement	Tobacco Product Use	-0.02	0.00	0.02	0.02	12.05	0.00
←	School Engagement	Vape Use	-0.02	0.00	0.01	0.02	7.38	0.00
←	Social Competency Scale	Non-medical marijuana use	-0.02	0.00	0.01	0.02	5.16	0.00
←	Social Competency Scale	Substance use – 1	-0.02	0.00	0.01	0.02	5.84	0.00
←	School Engagement	ACEs	-0.02	0.00	0.00	0.02	2.90	0.00
←	School Engagement	Marijuana use frequency	-0.02	0.00	0.01	0.02	3.96	0.00
←	Friends' Approval of Substance Use	Suspension: sent out of the classroom	-0.03	0.00	0.01	0.02	4.13	0.00
←	Tobacco Product Use	Social Competency Scale	-0.03	0.00	0.00	0.02	3.16	0.00

Edges	Nodes		Standardized	Standard	95% CI		Z-score	p-value
	Node 1	Node 2	ES	Error	Lower	Upper		
←	Friends' Approval of Substance Use	Non-medical marijuana use	-0.03	0.01	0.00	0.02	3.74	0.00
←	Reason of school absence: Illness	Teacher Student Relationship	-0.03	0.00	0.00	0.01	1.68	0.09
←	Reason of school absence: Suspension	Social Competency Scale	-0.03	0.00	-0.01	0.01	-0.29	0.77
←	Friends' Approval of Substance Use	PDA	-0.04	0.00	-0.01	0.00	-2.60	0.01
←	PDA	Teacher Student Relationship	-0.04	0.00	-0.01	0.00	-3.56	0.00
←	Teacher Student Relationship	Vape Use	-0.04	0.00	-0.02	-0.01	-4.44	0.00
←	Tobacco Product Use	Friends' Approval of Substance Use	-0.04	0.00	-0.02	-0.01	-4.11	0.00
←	Social Competency Scale	Prescription Substances Use (not prescribed to user)	-0.04	0.00	-0.02	-0.01	-4.01	0.00
←	Social Competency Scale	Marijuana use frequency	-0.06	0.00	-0.02	-0.01	-4.75	0.00
←	FDA	School Engagement	-0.07	0.00	-0.02	-0.01	-5.53	0.00
←	Suspension: sent out of the classroom	Teacher Student Relationship	-0.08	0.00	-0.02	-0.01	-5.27	0.00
←	Dental Checkup	Social Competency Scale	-0.08	0.00	-0.02	-0.01	-7.31	0.00
←	PDA	School Engagement	-0.11	0.00	-0.02	-0.01	-6.06	0.00
←	Teacher Student Relationship	ACEs	-0.11	0.00	-0.02	-0.02	-10.47	0.00
←	Social Competency Scale	Vape Use	-0.13	0.00	-0.03	-0.02	-7.52	0.00
←	Suspension: sent out of the classroom	School Engagement	-0.15	0.00	-0.03	-0.02	-7.16	0.00
←	Vape Use	Friends' Approval of Substance Use	-0.17	0.00	-0.03	-0.02	-13.39	0.00
←	ACEs	Social Competency Scale	-0.24	0.00	-0.03	-0.02	-10.22	0.00
←	Friends' Approval of Substance Use	Marijuana use frequency	-0.29	0.01	-0.03	-0.02	-11.50	0.00

Edges	Nodes: Node 1	Nodes: Node 2	Standardized ES	Standard Error	95% CI Lower	95% CI Upper	Z-score	p-value
←→	Marijuana use frequency	Marijuana use frequency	1.00	0.00	-0.04	-0.03	-11.15	0.00
←→	Native Hawaiian or Other Pacific Islander	Native Hawaiian or Other Pacific Islander	1.00	0.00	-0.04	-0.03	-11.21	0.00
←→	American Indian or Alaskan Native	American Indian or Alaskan Native	1.00	0.00	-0.04	-0.03	-9.82	0.00
←→	Reason of school absence: Illness	Reason of school absence: Illness	1.00	0.00	-0.04	-0.03	-14.88	0.00
←→	Dental Checkup	Dental Checkup	0.97	0.00	-0.05	-0.03	-11.85	0.00
←→	Reason of school absence: Suspension	Reason of school absence: Suspension	0.94	0.00	-0.05	-0.04	-13.61	0.00
←→	PDA	PDA	0.93	0.00	-0.07	-0.06	-18.72	0.00
←→	Suspension: sent out of the classroom	Suspension: sent out of the classroom	0.90	0.00	-0.07	-0.06	-24.81	0.00
←→	FDA	FDA	0.89	0.00	-0.07	-0.06	-21.25	0.00
←→	Friends' Approval of Substance Use	Friends' Approval of Substance Use	0.88	0.00	-0.08	-0.07	-23.72	0.00
←→	ACEs	ACEs	0.83	0.00	-0.08	-0.07	-25.46	0.00
←→	Social Competency Scale	Social Competency Scale	0.80	0.00	-0.10	-0.09	-35.06	0.00
←→	Teacher Student Relationship	Teacher Student Relationship	0.80	0.00	-0.11	-0.10	-39.74	0.00
←→	Tobacco Product Use	Tobacco Product Use	0.77	0.00	-0.13	-0.12	-47.91	0.00
←→	Binge	Binge	0.70	0.00	-0.19	-0.18	-64.86	0.00
←→	Substance use – 1	Substance use – 1	0.69	0.00	-0.25	-0.24	-83.99	0.00
←→	Vape Use	Vape Use	0.62	0.00	-0.26	-0.25	-94.27	0.00
←→	School Engagement	School Engagement	0.62	0.00	-0.36	-0.35	-141.32	0.00
←→	Prescription Substances Use (not prescribed to user)	Prescription Substances Use (not prescribed to user)	0.60	0.00	-0.37	-0.36	-124.22	0.00
←→	Non-medical marijuana use	Non-medical marijuana use	0.26	0.00	0.99	1.01	246.91	0.00
←→	ACEs	American Indian or Alaskan Native	0.08	0.00	0.99	1.01	246.91	0.00

Edges	Nodes		Standardized ES	Standard Error	95% CI		Z-score	p-value
	Node 1	Node 2			Lower	Upper		
←→	Marijuana use frequency	American Indian or Alaskan Native	0.07	0.00	0.99	1.01	246.91	0.00
←→	ACEs	Native Hawaiian or Other Pacific Islander	0.02	0.00	0.99	1.01	246.91	0.00
←→	Suspension: sent out of the classroom	American Indian or Alaskan Native	0.02	0.00	0.99	1.00	246.91	0.00
←→	Tobacco Product Use	American Indian or Alaskan Native	0.02	0.00	0.94	0.95	246.91	0.00
←→	Substance use – 1	Native Hawaiian or Other Pacific Islander	0.01	0.00	0.89	0.91	246.82	0.00
←→	Substance use – 1	American Indian or Alaskan Native	0.01	0.00	0.87	0.88	246.91	0.00
←→	Non-medical marijuana use	American Indian or Alaskan Native	0.00	0.00	0.82	0.84	246.60	0.00
←→	Friends' Approval of Substance Use	American Indian or Alaskan Native	-0.03	0.00	0.82	0.83	246.74	0.00
←→	Social Competency Scale	American Indian or Alaskan Native	-0.04	0.00	0.81	0.83	246.91	0.00
←→	ACEs	Friends' Approval of Substance Use	-0.06	0.00	0.78	0.79	246.48	0.00

Table B7

The Step 2 Causal Discovery Analysis Effect Sizes (the MSS 2019)

Edges	Nodes		Standardized	Standard	95% CI		Z-score	p-value
	Node 1	Node 2	ES	Error	Lower	Upper		
←	Missed school: mental issues	Missed school: sleep issue	0.27	0.00	0.27	0.28	101.03	0.00
←	FDA	PDA	0.25	0.00	0.25	0.26	92.28	0.00
←	Mental health treatment history	Missed school: mental issues	0.23	0.00	0.22	0.23	81.02	0.00
←	PDA	Missed school: sleep issue	0.14	0.00	0.14	0.15	50.15	0.00
←	Missed school: sleep issue	Missed school: transportation	0.14	0.00	0.14	0.15	50.52	0.00
←	Parent support	Free or reduced-price lunch at school	0.13	0.00	0.12	0.13	46.11	0.00
←	FDA	Missed school: mental issues	0.11	0.00	0.11	0.12	40.80	0.00
←	PDA	Missed school: mental issues	0.11	0.00	0.10	0.12	37.87	0.00
←	Missed school: taking care of family or friend	Missed school: mental issues	0.09	0.00	0.08	0.09	29.86	0.00
←	Missed school: taking care of family or friend	Missed school: transportation	0.08	0.00	0.07	0.08	27.13	0.00
←	Missed school: Housing instability	Missed school: had to work	0.08	0.00	0.07	0.08	27.05	0.00
←	PDA	Missed school: transportation	0.07	0.00	0.07	0.08	26.75	0.00
←	Missed school: had to work	Missed school: taking care of family or friend	0.07	0.00	0.06	0.07	23.95	0.00
←	Missed school: Housing instability	Housing stability	0.05	0.00	0.05	0.06	19.42	0.00
←	FDA	Missed school: taking care of family or friend	0.05	0.00	0.05	0.06	19.55	0.00
←	Missed school: taking care of family or friend	Missed school: sleep issue	0.05	0.00	0.04	0.05	16.01	0.00

Edges	Nodes		Standardized ES	Standard Error	95% CI		Z-score	p-value
	Node 1	Node 2			Lower	Upper		
←	Missed school: Housing instability	Missed school: taking care of family or friend	0.05	0.00	0.04	0.05	16.36	0.00
←	Missed school: had to work	Missed school: sleep issue	0.05	0.00	0.04	0.05	16.29	0.00
←	Missed school: Housing instability	Missed school: transportation	0.05	0.00	0.04	0.05	15.89	0.00
←	Missed school: Housing instability	Missed school: mental issues	0.04	0.00	0.04	0.05	13.97	0.00
←	Housing stability	FDA	0.04	0.00	0.03	0.04	13.49	0.00
←	PDA	Missed school: taking care of family or friend	0.03	0.00	0.03	0.04	12.44	0.00
←	Housing stability	Mental health treatment history	0.03	0.00	0.03	0.04	11.84	0.00
←	Housing stability	Missed school: transportation	0.03	0.00	0.03	0.04	11.98	0.00
←	Missed school: Housing instability	Missed school: sleep issue	0.03	0.00	0.02	0.03	8.64	0.00
←	Housing stability	Missed school: mental issues	0.02	0.00	0.02	0.03	7.90	0.00
←	Missed school: mental issues	Missed school: transportation	0.02	0.00	0.01	0.02	7.12	0.00
←	Missed school: sleep issue	Free or reduced-price lunch at school	-0.03	0.00	-0.04	-0.02	-10.57	0.00
←	PDA	Free or reduced-price lunch at school	-0.04	0.00	-0.04	-0.03	-14.13	0.00
←	PDA	Parent support	-0.06	0.00	-0.06	-0.05	-19.92	0.00
←	Missed school: taking care of family or friend	Free or reduced-price lunch at school	-0.09	0.00	-0.10	-0.09	-32.88	0.00
←	Missed school: sleep issue	Parent support	-0.10	0.00	-0.11	-0.10	-36.91	0.00
←	Mental health treatment history	Parent support	-0.11	0.00	-0.12	-0.11	-41.05	0.00
←	Missed school: transportation	Free or reduced-price lunch at school	-0.18	0.00	-0.18	-0.17	-63.72	0.00

Edges	Nodes		Standardized	Standard	95% CI		Z-score	p-value
	Node 1	Node 2	ES	Error	Lower	Upper		
←	Missed school: mental issues	Parent support	-0.18	0.00	-0.18	-0.17	-66.61	0.00
←→	Free or reduced-price lunch at school	Free or reduced-price lunch at school	1.00	0.00	0.99	1.01	250.38	0.00
←→	Missed school: had to work	Missed school: had to work	0.99	0.00	0.98	1.00	250.37	0.00
←→	Housing stability	Housing stability	0.99	0.00	0.98	1.00	250.12	0.00
←→	Parent support	Parent support	0.98	0.00	0.98	0.99	250.37	0.00
←→	Missed school: Housing instability	Missed school: Housing instability	0.98	0.00	0.97	0.99	250.37	0.00
←→	Missed school: transportation	Missed school: transportation	0.97	0.00	0.96	0.98	250.37	0.00
←→	Missed school: taking care of family or friend	Missed school: taking care of family or friend	0.97	0.00	0.96	0.97	250.37	0.00
←→	Missed school: sleep issue	Missed school: sleep issue	0.96	0.00	0.96	0.97	250.37	0.00
←→	PDA	PDA	0.93	0.00	0.92	0.94	250.37	0.00
←→	Mental health treatment history	Mental health treatment history	0.93	0.00	0.92	0.93	250.37	0.00
←→	FDA	FDA	0.90	0.00	0.90	0.91	250.34	0.00
←→	Missed school: mental issues	Missed school: mental issues	0.88	0.00	0.87	0.88	250.37	0.00
←→	Mental health treatment history	PDA	0.07	0.00	0.06	0.07	25.40	0.00
←→	Missed school: mental issues	Free or reduced-price lunch at school	-0.02	0.00	-0.03	-0.01	-7.30	0.00
←→	Missed school: transportation	Parent support	-0.06	0.00	-0.06	-0.05	-21.02	0.00
←→	Housing stability	Parent support	-0.07	0.00	-0.08	-0.07	-25.73	0.00
←→	FDA	Free or reduced-price lunch at school	-0.08	0.00	-0.08	-0.07	-28.66	0.00
←→	Housing stability	Free or reduced-price lunch at school	-0.12	0.00	-0.13	-0.12	-43.35	0.00

Table B8

The Step 3 Causal Discovery Analysis Effect Sizes (the MSS 2019)

Edges	Nodes		Standardized ES	Standard Error	95% CI		Z-score	p-value
	Node 1	Node 2			Lower	Upper		
←	Non-medical marijuana use	Marijuana use frequency	0.78	0.00	0.78	0.78	429.44	0.00
←	Vape Use	Marijuana use frequency	0.54	0.00	0.53	0.54	223.76	0.00
←	Substance use – 1	Tobacco Product Use	0.37	0.00	0.37	0.38	136.34	0.00
←	School Engagement	Social Competency Scale	0.33	0.00	0.32	0.33	121.04	0.00
←	School Engagement	Teacher Student Relationship	0.33	0.00	0.32	0.33	131.06	0.00
←	Social Competency Scale	Friends' Approval of Substance Use	0.32	0.00	0.31	0.32	117.26	0.00
←	Teacher Student Relationship	Social Competency Scale	0.31	0.00	0.30	0.31	103.19	0.00
←	Binge	Vape Use	0.29	0.00	0.28	0.30	90.73	0.00
←	Tobacco Product Use	Binge	0.28	0.00	0.27	0.28	100.76	0.00
←	Parent support	Social Competency Scale	0.27	0.00	0.26	0.27	97.35	0.00
←	Prescription Substances Use (not prescribed to user)	Tobacco Product Use	0.27	0.00	0.26	0.27	90.03	0.00
←	Missed school: sleep issue	Missed school: mental issues	0.26	0.00	0.25	0.27	93.13	0.00
←	FDA	PDA	0.23	0.00	0.23	0.24	84.71	0.00
←	Tobacco Product Use	Vape Use	0.22	0.00	0.21	0.23	68.58	0.00
←	Mental health treatment history	ACEs	0.21	0.00	0.20	0.22	73.50	0.00
←	Reason of school absence: Suspension	Suspension: sent out of the classroom	0.20	0.00	0.19	0.20	68.79	0.00
←	Binge	Non-medical marijuana use	0.19	0.00	0.18	0.20	40.05	0.00
←	Missed school: mental issues	ACEs	0.18	0.00	0.18	0.19	59.32	0.00

Edges	Nodes: Node 1	Node 2	Standardized ES	Standard Error	95% CI Lower	Upper	Z-score	p-value
←	Mental health treatment history	Missed school: mental issues	0.17	0.00	0.17	0.18	62.27	0.00
←	Substance use – 1	Non-medical marijuana use	0.14	0.00	0.13	0.14	29.87	0.00
←	Missed school: transportation	Missed school: sleep issue	0.13	0.00	0.13	0.14	47.42	0.00
←	Teacher Student Relationship	Parent support	0.13	0.00	0.12	0.13	43.60	0.00
←	PDA	Missed school: sleep issue	0.13	0.00	0.12	0.13	44.43	0.00
←	Substance use – 1	Binge	0.12	0.00	0.12	0.13	44.57	0.00
←	Non-medical marijuana use	Vape Use	0.12	0.00	0.12	0.13	66.69	0.00
←	Housing stability	ACEs	0.12	0.00	0.11	0.12	41.35	0.00
←	Prescription Substances Use (not prescribed to user)	Non-medical marijuana use	0.12	0.00	0.11	0.13	23.82	0.00
←	Prescription Substances Use (not prescribed to user)	Binge	0.11	0.00	0.11	0.12	37.36	0.00
←	ACEs	Prescription Substances Use (not prescribed to user)	0.11	0.00	0.10	0.11	37.92	0.00
←	ACEs	Vape Use	0.10	0.00	0.10	0.11	30.59	0.00
←	FDA	Missed school: mental issues	0.09	0.00	0.09	0.10	34.17	0.00
←	Reason of school absence: Illness	Missed school: mental issues	0.09	0.00	0.09	0.10	31.64	0.00
←	PDA	Missed school: mental issues	0.09	0.00	0.09	0.10	32.28	0.00
←	Tobacco Product Use	Marijuana use frequency	0.09	0.00	0.08	0.10	19.24	0.00
←	FDA	Reason of school absence: Illness	0.09	0.00	0.08	0.09	32.75	0.00
←	Housing stability	Substance use – 1	0.08	0.00	0.07	0.09	21.00	0.00
←	Missed school: taking care of family or friend	Missed school: mental issues	0.08	0.00	0.07	0.08	25.52	0.00

Edges	Nodes		Standardized ES	Standard Error	95% CI		Z-score	p-value
	Node 1	Node 2			Lower	Upper		
←	Missed school: taking care of family or friend	Missed school: transportation	0.07	0.00	0.07	0.08	26.26	0.00
←	Teacher Student Relationship	Friends' Approval of Substance Use	0.07	0.00	0.07	0.08	26.25	0.00
←	PDA	Missed school: transportation	0.07	0.00	0.07	0.08	27.19	0.00
←	Suspension: sent out of the classroom	Tobacco Product Use	0.07	0.00	0.07	0.08	23.40	0.00
←	Suspension: sent out of the classroom	PDA	0.07	0.00	0.07	0.08	26.59	0.00
←	PDA	Vape Use	0.07	0.00	0.06	0.08	20.87	0.00
←	Missed school: Housing instability	Missed school: had to work	0.07	0.00	0.06	0.08	24.88	0.00
←	School Engagement	Friends' Approval of Substance Use	0.07	0.00	0.06	0.07	27.93	0.00
←	Prescription Substances Use (not prescribed to user)	Marijuana use frequency	0.07	0.00	0.06	0.08	13.51	0.00
←	Dental Checkup	ACEs	0.06	0.00	0.06	0.07	21.60	0.00
←	Suspension: sent out of the classroom	Substance use – 1	0.06	0.00	0.06	0.07	17.50	0.00
←	ACEs	Marijuana use frequency	0.06	0.01	0.05	0.07	12.21	0.00
←	ACEs	Non-medical marijuana use	0.06	0.01	0.05	0.07	12.24	0.00
←	Missed school: had to work	Missed school: taking care of family or friend	0.06	0.00	0.06	0.07	22.10	0.00
←	Substance use – 1	Marijuana use frequency	0.06	0.00	0.05	0.07	12.89	0.00
←	Missed school: had to work	Tobacco Product Use	0.06	0.00	0.05	0.06	18.43	0.00
←	Mental health treatment history	Prescription Substances Use (not prescribed to user)	0.05	0.00	0.05	0.06	19.22	0.00
←	Binge	Marijuana use frequency	0.05	0.00	0.04	0.06	10.87	0.00
←	Missed school: sleep issue	ACEs	0.05	0.00	0.05	0.06	17.80	0.00

Edges	Nodes: Node 1	Nodes: Node 2	Standardized ES	Standard Error	95% CI Lower	95% CI Upper	Z-score	p-value
←	School Engagement	Parent support	0.05	0.00	0.04	0.05	20.14	0.00
←	FDA	Missed school: taking care of family or friend	0.05	0.00	0.04	0.05	18.45	0.00
←	Dental Checkup	Housing stability	0.05	0.00	0.04	0.05	17.05	0.00
←	Mental health treatment history	PDA	0.05	0.00	0.04	0.05	17.33	0.00
←	Missed school: Housing instability	Suspension: sent out of the classroom	0.05	0.00	0.04	0.05	16.42	0.00
←	Missed school: taking care of family or friend	ACEs	0.05	0.00	0.04	0.05	15.61	0.00
←	Missed school: transportation	ACEs	0.05	0.00	0.04	0.05	15.40	0.00
←	Missed school: mental issues	Marijuana use frequency	0.05	0.01	0.04	0.06	8.73	0.00
←	Housing stability	Prescription Substances Use (not prescribed to user)	0.05	0.00	0.04	0.05	12.88	0.00
←	FDA	Suspension: sent out of the classroom	0.05	0.00	0.04	0.05	16.03	0.00
←	Reason of school absence: Illness	Missed school: sleep issue	0.05	0.00	0.04	0.05	15.36	0.00
←	Missed school: Housing instability	Missed school: taking care of family or friend	0.05	0.00	0.04	0.05	15.91	0.00
←	Missed school: Housing instability	Housing stability	0.04	0.00	0.04	0.05	15.69	0.00
←	FDA	Reason of school absence: Suspension	0.04	0.00	0.04	0.05	16.23	0.00
←	Missed school: taking care of family or friend	Missed school: sleep issue	0.04	0.00	0.04	0.05	14.94	0.00
←	Missed school: mental issues	Vape Use	0.04	0.00	0.04	0.05	12.56	0.00
←	Housing stability	Tobacco Product Use	0.04	0.00	0.04	0.05	13.21	0.00

Edges	Nodes		Standardized ES	Standard Error	95% CI		Z-score	p-value
	Node 1	Node 2			Lower	Upper		
←	Missed school: Housing instability	Missed school: transportation	0.04	0.00	0.04	0.05	14.94	0.00
←	PDA	Marijuana use frequency	0.04	0.01	0.03	0.05	8.08	0.00
←	PDA	Non-medical marijuana use	0.04	0.01	0.03	0.05	7.82	0.00
←	Missed school: sleep issue	Vape Use	0.04	0.00	0.03	0.05	11.74	0.00
←	Reason of school absence: Suspension	Marijuana use frequency	0.04	0.01	0.03	0.05	7.57	0.00
←	Tobacco Product Use	Non-medical marijuana use	0.04	0.00	0.03	0.05	8.23	0.00
←	Missed school: transportation	Dental Checkup	0.04	0.00	0.03	0.04	13.34	0.00
←	Native Hawaiian or Other Pacific Islander	American Indian or Alaskan Native	0.04	0.00	0.03	0.04	13.02	0.00
←	Missed school: Housing instability	Missed school: mental issues	0.04	0.00	0.03	0.04	12.38	0.00
←	Housing stability	Suspension: sent out of the classroom	0.04	0.00	0.03	0.04	12.83	0.00
←	Reason of school absence: Suspension	Missed school: Housing instability	0.04	0.00	0.03	0.04	13.17	0.00
←	Missed school: had to work	Suspension: sent out of the classroom	0.03	0.00	0.03	0.04	11.56	0.00
←	Missed school: had to work	Missed school: sleep issue	0.03	0.00	0.03	0.04	12.04	0.00
←	Suspension: sent out of the classroom	Prescription Substances Use (not prescribed to user)	0.03	0.00	0.03	0.04	9.30	0.00
←	Mental health treatment history	Non-medical marijuana use	0.03	0.00	0.03	0.04	10.94	0.00
←	Missed school: sleep issue	Non-medical marijuana use	0.03	0.01	0.02	0.04	5.90	0.00
←	Prescription Substances Use (not prescribed to user)	Vape Use	0.03	0.00	0.02	0.04	10.28	0.00
←	Parent support	Friends' Approval of Substance Use	0.03	0.00	0.03	0.04	11.47	0.00

Edges	Nodes		Standardized	Standard	95% CI		Z-score	p-value
	Node 1	Node 2	ES	Error	Lower	Upper		
←	Missed school: Housing instability	Substance use – 1	0.03	0.00	0.02	0.04	7.54	0.00
←	Reason of school absence: Suspension	Missed school: transportation	0.03	0.00	0.02	0.03	10.28	0.00
←	Tobacco Product Use	ACEs	0.03	0.00	0.02	0.03	10.00	0.00
←	Missed school: had to work	Binge	0.03	0.00	0.02	0.03	8.45	0.00
←	Binge	PDA	0.03	0.00	0.02	0.03	10.27	0.00
←	Reason of school absence: Suspension	Tobacco Product Use	0.03	0.00	0.02	0.03	7.60	0.00
←	Missed school: had to work	Reason of school absence: Suspension	0.02	0.00	0.02	0.03	8.19	0.00
←	Dental Checkup	Tobacco Product Use	0.02	0.00	0.02	0.03	7.18	0.00
←	Missed school: transportation	Suspension: sent out of the classroom	0.02	0.00	0.02	0.03	7.99	0.00
←	Missed school: Housing instability	Missed school: sleep issue	0.02	0.00	0.02	0.03	7.35	0.00
←	Missed school: Housing instability	Binge	0.02	0.00	0.01	0.03	6.51	0.00
←	Dental Checkup	Substance use – 1	0.02	0.00	0.01	0.03	6.29	0.00
←	Reason of school absence: Suspension	Prescription Substances Use (not prescribed to user)	0.02	0.00	0.01	0.03	5.16	0.00
←	Missed school: sleep issue	Marijuana use frequency	0.02	0.01	0.01	0.03	3.37	0.00
←	Reason of school absence: Suspension	Non-medical marijuana use	0.02	0.01	0.01	0.03	3.23	0.00
←	Substance use – 1	ACEs	0.02	0.00	0.01	0.02	7.43	0.00
←	Reason of school absence: Suspension	Substance use – 1	0.01	0.00	0.01	0.02	3.67	0.00
←	Missed school: mental issues	Non-medical marijuana use	0.01	0.01	0.00	0.02	1.86	0.06
←	Missed school: Housing instability	Prescription Substances Use (not prescribed to user)	0.00	0.00	0.00	0.01	0.98	0.33

Edges	Nodes		Standardized ES	Standard Error	95% CI		Z-score	p-value
	Node 1	Node 2			Lower	Upper		
←	Missed school: Housing instability	Tobacco Product Use	0.00	0.00	-0.01	0.01	-0.33	0.74
←	Parent support	Substance use – 1	-0.01	0.00	-0.02	-0.01	-4.27	0.00
←	Social Competency Scale	Substance use – 1	-0.02	0.00	-0.02	-0.01	-4.53	0.00
←	Binge	Friends' Approval of Substance Use	-0.02	0.00	-0.02	-0.01	-5.38	0.00
←	Reason of school absence: Suspension	Teacher Student Relationship	-0.02	0.00	-0.02	-0.01	-5.43	0.00
←	Vape Use	School Engagement	-0.02	0.00	-0.03	-0.01	-7.28	0.00
←	Reason of school absence: Suspension	Social Competency Scale	-0.02	0.00	-0.03	-0.01	-6.56	0.00
←	Missed school: had to work	School Engagement	-0.02	0.00	-0.03	-0.01	-7.01	0.00
←	Reason of school absence: Suspension	Free or reduced-price lunch at school	-0.02	0.00	-0.03	-0.02	-7.38	0.00
←	Binge	Social Competency Scale	-0.02	0.00	-0.03	-0.02	-7.53	0.00
←	Tobacco Product Use	Social Competency Scale	-0.02	0.00	-0.03	-0.02	-8.26	0.00
←	Social Competency Scale	Non-medical marijuana use	-0.03	0.00	-0.03	-0.02	-5.05	0.00
←	Parent support	Prescription Substances Use (not prescribed to user)	-0.03	0.00	-0.03	-0.02	-8.20	0.00
←	Missed school: transportation	Social Competency Scale	-0.03	0.00	-0.03	-0.02	-9.66	0.00
←	Parent support	Dental Checkup	-0.03	0.00	-0.04	-0.03	-12.79	0.00
←	Missed school: mental issues	School Engagement	-0.03	0.00	-0.04	-0.03	-10.22	0.00
←	Friends' Approval of Substance Use	Non-medical marijuana use	-0.03	0.01	-0.04	-0.02	-6.24	0.00
←	Teacher Student Relationship	Vape Use	-0.04	0.00	-0.05	-0.03	-14.00	0.00
←	School Engagement	Marijuana use frequency	-0.04	0.00	-0.05	-0.04	-17.40	0.00
←	Missed school: mental issues	Teacher Student Relationship	-0.05	0.00	-0.05	-0.04	-14.33	0.00

Edges	Nodes		Standardized	Standard	95% CI		Z-score	p-value
	Node 1	Node 2	ES	Error	Lower	Upper		
←	Social Competency Scale	Prescription Substances Use (not prescribed to user)	-0.05	0.00	-0.06	-0.04	-15.37	0.00
←	Friends' Approval of Substance Use	Tobacco Product Use	-0.05	0.00	-0.06	-0.05	-16.21	0.00
←	Suspension: sent out of the classroom	Free or reduced-price lunch at school	-0.06	0.00	-0.06	-0.05	-21.42	0.00
←	Missed school: sleep issue	School Engagement	-0.06	0.00	-0.07	-0.06	-21.89	0.00
←	Teacher Student Relationship	ACEs	-0.07	0.00	-0.07	-0.06	-22.48	0.00
←	Free or reduced-price lunch at school	FDA	-0.07	0.00	-0.08	-0.07	-25.97	0.00
←	Dental Checkup	Social Competency Scale	-0.08	0.00	-0.08	-0.07	-25.41	0.00
←	FDA	School Engagement	-0.08	0.00	-0.09	-0.07	-28.46	0.00
←	Missed school: taking care of family or friend	Free or reduced-price lunch at school	-0.08	0.00	-0.09	-0.08	-28.20	0.00
←	Suspension: sent out of the classroom	Teacher Student Relationship	-0.08	0.00	-0.09	-0.08	-26.64	0.00
←	Missed school: mental issues	Parent support	-0.09	0.00	-0.10	-0.09	-29.81	0.00
←	Free or reduced-price lunch at school	Housing stability	-0.11	0.00	-0.11	-0.10	-38.88	0.00
←	PDA	School Engagement	-0.12	0.00	-0.12	-0.11	-42.12	0.00
←	Social Competency Scale	Marijuana use frequency	-0.12	0.00	-0.13	-0.11	-24.99	0.00
←	ACEs	Friends' Approval of Substance Use	-0.13	0.00	-0.14	-0.13	-47.71	0.00
←	Missed school: transportation	Free or reduced-price lunch at school	-0.14	0.00	-0.15	-0.14	-50.07	0.00
←	Suspension: sent out of the classroom	School Engagement	-0.16	0.00	-0.16	-0.15	-49.75	0.00
←	Vape Use	Friends' Approval of Substance Use	-0.16	0.00	-0.16	-0.15	-61.73	0.00

Edges	Nodes: Node 1	Nodes: Node 2	Standardized ES	Standard Error	95% CI Lower	95% CI Upper	Z-score	p-value
←	Dental Checkup	Free or reduced-price lunch at school	-0.17	0.00	-0.18	-0.17	-61.54	0.00
←	Friends' Approval of Substance Use	Marijuana use frequency	-0.28	0.01	-0.29	-0.27	-52.89	0.00
←	Parent support	ACEs	-0.33	0.00	-0.34	-0.33	-127.02	0.00
←→	Marijuana use frequency	Marijuana use frequency	1.00	0.00	0.99	1.01	250.37	0.00
←→	American Indian or Alaskan Native	American Indian or Alaskan Native	1.00	0.00	0.99	1.01	250.39	0.00
←→	Native Hawaiian or Other Pacific Islander	Native Hawaiian or Other Pacific Islander	1.00	0.00	0.99	1.01	250.37	0.00
←→	Reason of school absence: Illness	Reason of school absence: Illness	0.99	0.00	0.98	0.99	250.37	0.00
←→	Missed school: had to work	Missed school: had to work	0.98	0.00	0.97	0.99	250.37	0.00
←→	Missed school: Housing instability	Missed school: Housing instability	0.98	0.00	0.97	0.98	250.37	0.00
←→	Free or reduced-price lunch at school	Free or reduced-price lunch at school	0.97	0.00	0.96	0.98	250.09	0.00
←→	Missed school: taking care of family or friend	Missed school: taking care of family or friend	0.96	0.00	0.96	0.97	250.37	0.00
←→	Housing stability	Housing stability	0.95	0.00	0.95	0.96	250.36	0.00
←→	Missed school: transportation	Missed school: transportation	0.94	0.00	0.93	0.95	250.37	0.00
←→	Dental Checkup	Dental Checkup	0.94	0.00	0.93	0.94	250.37	0.00
←→	Reason of school absence: Suspension	Reason of school absence: Suspension	0.94	0.00	0.93	0.94	250.37	0.00
←→	Missed school: mental issues	Missed school: mental issues	0.90	0.00	0.90	0.91	250.37	0.00
←→	Suspension: sent out of the classroom	Suspension: sent out of the classroom	0.90	0.00	0.89	0.91	250.36	0.00
←→	PDA	PDA	0.90	0.00	0.89	0.91	250.37	0.00
←→	Missed school: sleep issue	Missed school: sleep issue	0.89	0.00	0.89	0.90	250.37	0.00

Edges	Nodes: Node 1	Nodes: Node 2	Standardized ES	Standard Error	95% CI Lower	95% CI Upper	Z-score	p-value
←→	FDA	FDA	0.88	0.00	0.88	0.89	250.36	0.00
←→	Friends' Approval of Substance Use	Friends' Approval of Substance Use	0.88	0.00	0.87	0.89	248.41	0.00
←→	Mental health treatment history	Mental health treatment history	0.88	0.00	0.87	0.89	250.37	0.00
←→	ACEs	ACEs	0.88	0.00	0.87	0.88	250.36	0.00
←→	Social Competency Scale	Social Competency Scale	0.82	0.00	0.81	0.82	249.92	0.00
←→	Teacher Student Relationship	Teacher Student Relationship	0.79	0.00	0.78	0.79	250.36	0.00
←→	Prescription Substances Use (not prescribed to user)	Prescription Substances Use (not prescribed to user)	0.79	0.00	0.78	0.79	250.37	0.00
←→	Binge	Binge	0.75	0.00	0.75	0.76	250.28	0.00
←→	Parent support	Parent support	0.73	0.00	0.72	0.73	250.37	0.00
←→	Tobacco Product Use	Tobacco Product Use	0.72	0.00	0.71	0.72	249.87	0.00
←→	Substance use – 1	Substance use – 1	0.69	0.00	0.68	0.70	249.91	0.00
←→	Vape Use	Vape Use	0.62	0.00	0.61	0.62	248.73	0.00
←→	School Engagement	School Engagement	0.62	0.00	0.61	0.62	250.37	0.00
←→	Prescription Substances Use (not prescribed to user)	Substance use – 1	0.36	0.00	0.35	0.36	153.67	0.00
←→	Non-medical marijuana use	Non-medical marijuana use	0.26	0.00	0.26	0.27	250.37	0.00
←→	Social Competency Scale	Free or reduced-price lunch at school	0.10	0.00	0.09	0.10	37.65	0.00
←→	ACEs	American Indian or Alaskan Native	0.08	0.00	0.08	0.09	31.23	0.00
←→	Marijuana use frequency	American Indian or Alaskan Native	0.07	0.00	0.06	0.07	24.15	0.00
←→	Housing stability	American Indian or Alaskan Native	0.04	0.00	0.03	0.04	13.45	0.00

Edges	Nodes		Standardized ES	Standard Error	95% CI		Z-score	p-value
	Node 1	Node 2			Lower	Upper		
←→	ACEs	Native Hawaiian or Other Pacific Islander	0.02	0.00	0.02	0.03	8.64	0.00
←→	Tobacco Product Use	American Indian or Alaskan Native	0.02	0.00	0.01	0.02	7.76	0.00
←→	Non-medical marijuana use	American Indian or Alaskan Native	0.01	0.00	0.00	0.01	4.08	0.00
←→	Mental health treatment history	Native Hawaiian or Other Pacific Islander	0.00	0.00	-0.01	0.00	-1.11	0.27
←→	Friends' Approval of Substance Use	American Indian or Alaskan Native	-0.03	0.00	-0.03	-0.02	-11.05	0.00
←→	Social Competency Scale	American Indian or Alaskan Native	-0.04	0.00	-0.05	-0.04	-16.37	0.00
←→	Social Competency Scale	Vape Use	-0.07	0.00	-0.08	-0.07	-32.25	0.00
←→	Free or reduced-price lunch at school	American Indian or Alaskan Native	-0.10	0.00	-0.10	-0.09	-34.93	0.00
←→	ACEs	Free or reduced-price lunch at school	-0.17	0.00	-0.17	-0.16	-62.33	0.00
←→	ACEs	Social Competency Scale	-0.19	0.00	-0.20	-0.19	-78.57	0.00

Table B9

Attributes from Subset F - Top 25% + J48 (n = 37/113)

Attribute
Free or reduced-price lunch at school
Staying home due to sickness
Sent to office for discipline
In school suspension
Out of school suspension
Medical checkup
Sleep during school day
Perceptions of caring from adults in the community
Alcohol consumption frequency
Binge drinking-1
Binge drinking-2
Non-medical marijuana use frequency
Substance use frequency
Tobacco use frequency
Marijuana use frequency
School Engagement (SE)
Teacher Student Relationship
Hostile school climate by respondent
Positive Youth Development Scale
Substance use treatment history
Perception of family caring
Positive Identity Scale
Social Competency Scale
Empowerment
Global appraisal of individual needs (GAIN)
ACEs
Runaway
Crime / violence subscription
Tobacco product usage
Substance use – 1
Substance use – 2
Perceptions of substance use risk
Parents' approval of substance use

Attribute
Friends' approval of substance use
Attitudes toward drinking
Race: Native Hawaiian or Pacific Islander only
Race & Ethnicity: American Indian Non-Hispanic

Table B10

Attributes from Subset G - Top 50% + J48 (n = 64/113)

Attribute
Relationship with father
Relationship with mother
Free or reduced-price lunch at school
Transient student
School nurse office visit
Staying home due to sickness
Sent to office for discipline
In school suspension
Out of school suspension
Perception of safety while commuting
Perception of school safety
Neighborhood safety
Home safety
Online bullying
Out-of-school activity: Sports
Out-of-school activity: Religious activities
General health
Medical checkup
Long-term mental health history
Skipping meal due to financial issues
Sleep during school day
Perception of peer caring
Perceptions of caring from adults in the community
Non-suicidal self-injury
Alcohol consumption frequency

Attribute
Binge drinking-1
Binge drinking-2
Non-medical marijuana use frequency
Substance use frequency
Tobacco use frequency
Alcohol consumption frequency
Marijuana use frequency
School Engagement (SE)
Teacher Student Relationship
Hostile school climate by peers
Hostile school climate by respondent
Positive Youth Development Scale
Substance use treatment history
Perception of family caring
Positive Identity Scale
Social Competency Scale
Empowerment
Suicidal Ideation
Suicidal Attempt
Global Appraisal of Individual Needs
Global Appraisal of Individual Needs_1
Intimate Partner Violence
Perpetrator
Incarcerated parents
ACEs
Runaway
Crime / violence subscription
Tobacco product usage
Substance use – 1
Substance use – 2
Perceptions of substance use risk
Parents' approval of substance use
Friends' approval of substance use
Attitudes toward drinking

Attribute
Attitudes toward drinking - 2
Race: Native Hawaiian or Pacific Islander only
Race: White only
Race & Ethnicity: American Indian Non-Hispanic
Race & Ethnicity: White Non-Hispanic

Appendix C

Institutional Review Board Supplements

Appendix C1

Institutional Review Board Exemption (Quantitative – MSS 2016)

NOT HUMAN RESEARCH

February 13, 2020

Barbara McMorris

Dear Barbara McMorris:

On 2/13/2020, the IRB reviewed the following submission:

Type of Review:	Initial Study
Title of Study:	A data-driven approach using the causal discovery method to MSS (Minnesota student survey) data.
Investigator:	Barbara McMorris
IRB ID:	STUDY00008941
Sponsored Funding:	None
Grant ID:	None
Internal UMN Funding:	None
Fund Management Outside University:	None
IND, IDE, or HDE:	None
Documents Reviewed with this Submission:	• IRB_Application_Data_Only_Protocol, Category: IRB Protocol;

The IRB determined that the proposed activity is not research involving human subjects as defined by DHHS and FDA regulations. To arrive at this determination, the IRB used "WORKSHEET: Human Research (HRP-310)." If you have any questions about this determination, please review that Worksheet in the HRPP Toolkit Library and contact the IRB office if needed.

Ongoing IRB review and approval for this activity is not required; however, this determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether IRB review is required, please submit a Modification to the IRB for a determination.

Sincerely,
Sara Dufour, MPH

IRB Analyst

Appendix C2

Institutional Review Board Exemption (Quantitative – MSS 2019)

NOT HUMAN RESEARCH

September 11, 2020

Barbara McMorris

952-649-8690
mcmo0023@umn.edu

Dear Barbara McMorris:

On 9/11/2020, the IRB reviewed the following submission:

Type of Review:	Initial Study
Title of Study:	A DATA-DRIVEN APPROACH USING THE CAUSAL DISCOVERY ANALYSIS METHOD TO MSS (Minnesota student survey) 2019 DATA.
Investigator:	Barbara McMorris
IRB ID:	STUDY00010810
Sponsored Funding:	None
Grant ID:	None
Internal UMN Funding:	None
Fund Management Outside University:	None
IND, IDE, or HDE:	None
Documents Reviewed with this Submission:	• HRP-595 - Data or Specimen Only Protocol_08312020.docx, Category: IRB Protocol;

IMPORTANT: All human research conducted at the University of Minnesota must adhere to the IRB guidance and requirements, Office of the Vice President for Research guidance, and the Medical School/Office of Academic Clinical Affairs Sunrise Implementation Plan in response to the COVID-19 pandemic. Non-medical school investigators should contact their Associate Dean for Research for information on the "sunrise" process.

Even with IRB approval, in-person research visits may not take place without documented approval by either the Medical School/OACA sunrise process or the Associate Dean for Research sunrise process. These reviews are intended to protect the health of all research participants and the broader University/Fairview communities

during the COVID-19 pandemic. Researchers must inform the IRB of their approved sunrise plans. The IRB will document the approval status on ETHOS via a comment in the study history section Please note that IRB approved COVID-19 related research is exempt from the sunrise requirements.
All researchers should review the guidance for the IRB, the medical school and their own departments as guidance is updated frequently.

The IRB determined that the proposed activity is not research involving human subjects as defined by DHHS and FDA regulations. To arrive at this determination, the IRB used "WORKSHEET: Human Research (HRP-310)." If you have any questions about this determination, please review that Worksheet in the HRPP Toolkit Library and contact the IRB office if needed.

Ongoing IRB review and approval for this activity is not required; however, this determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether IRB review is required, please submit a Modification to the IRB for a determination.

Sincerely,

Cynthia McGill CIP
IRB Analyst

We strive to provide clear, consistent and timely service to maintain a culture of respect, beneficence and justice in research. Complete a brief survey about your experience.

The University of Minnesota regularly updates COVID-19 Guidance for the Research Community, which includes the Latest IRB Guidance and FAQs. Please visit that website for up-to-date information.

Appendix C3

Institutional Review Board Approval (Qualitative – LSN Focus Group Interview)

NOT HUMAN RESEARCH

July 17, 2020

Lauren Martin
mart2114@umn.edu

Dear Lauren Martin:

On 7/17/2020, the IRB reviewed the following submission:

Type of Review:	Initial Study
Title of Study:	Minnesota Youth Sex Trading Study; Focus Groups with Stakeholders
Investigator:	Lauren Martin
IRB ID:	STUDY00010371
Sponsored Funding:	Sponsor Name: UNIVERSITY OF MINNESOTA FOUNDATION, Grant Title: Estimating the Prevalence of Sexual Exploitation Among A
Grant ID:	CON000000084067;
Internal UMN Funding:	None
Fund Management Outside University:	None
IND, IDE, or HDE:	None
Documents Reviewed with this Submission:	• HRP-503-Human-Research-Determination-Form - MYST_focus_grous_7-13-20.docx, Category: IRB Protocol;

The IRB determined that the proposed activity is not research involving human subjects as defined by DHHS and FDA regulations. To arrive at this determination, the IRB used "WORKSHEET: Human Research (HRP-310)." If you have any questions about this determination, please review that Worksheet in the HRPP Toolkit Library and contact the IRB office if needed.

Ongoing IRB review and approval for this activity is not required; however, this determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether IRB review is required, please submit a Modification to the IRB for a determination.

Appendix D

LSN Focus Group Interview Questionnaire

Focus Group Semi-Structured Interview Guide

Thank you again for joining us today. [Brief introduction of study staff on the call including name, connection to nursing and/or the project, and the role they will be taking during the focus group]

We are here to learn from you about the role school nurses play in supporting youth involved in trading sex and how we, as researchers, can help school nurses in practice continue that work as effectively as possible. We are conducting four of these focus groups with school nurses in Minnesota and will use the information we learn to help create materials to educate healthcare and service providers and to inform our research.

We are recording our conversation and will be taking notes because we do not want to miss anything you say. All names and identifiers that might be used will be removed before we share any materials outside of our study group, so you can be sure everything you say will be kept anonymous.

If you have any questions, feel free to ask them out loud or direct them privately to [study team member 2] in the chat. [Review consent forms]

Interview questions

1. To get started, please introduce yourself and describe to the group the setting you work in and population you serve.
2. Next, we would like to get an overview of what your role in the school (or district) is and how the health office fits into the school's functioning. What might you expect a typical week to look like for you and for the health office?
    a. Probe: What are some tasks that you can expect to perform most regularly?
    b. Probe: What are some reasons students most frequently visit the school health office? What are the most frequent reasons for 'unscheduled' visits?
    c. Probe: Who are the types of students who most frequently visit the school health office? What kind of formal or informal support might they have or need?

3. We know that school nurses know a lot about the most vulnerable students in our schools. One of the groups we are interested in learning more about are students who are absent from school frequently. Can you describe who these students are? *(Operational definition of "frequent": 15 days or more per year)*
    a. Probe: Why might students be absent from school only part of the day on a frequent basis? What about frequently missing full days of school?
    b. Probe: What vulnerable groups are particularly at risk for missing part of the school day? Why might they be at greater risk? (Example of vulnerable group: Students with addiction concerns)
    c. Probe: What about those at risk for frequently missing full days? Are they different groups than those that miss partial days? (Example of vulnerable group: Students with addiction concerns)

Probe: What are the outcomes for students frequently missing partial days of school and how are they different than the outcomes for students who miss full days?

Now that we have gotten to know each other a little bit, we are going to break up into two groups. We have preassigned these groups. I will be going with [Participant Names] and <study team member 2> will be facilitating the conversation with [Participant Names]. We will be coming back together in a little bit.

4. (Small Group) **(Experience with CSE youth)** As you know we are interested in learning more about how LSNs can support youth who are involved in trading sex for something of value like money or a place to stay. Everyone in this group indicated that they have had experience working with students who have traded sex or at least suspected that a student they worked with might be involved. Can you each describe the situation or situations?
    a. Probe: What led you to suspect/know this student was involved?
    b. Probe: What actions did you take? Or wish you took?
    c. Probe: Who did you talk to for support or what resources did you utilize?
    d. Probe: Where is that student now?
    e. Probe: How (if at all) did the student’s family factor in?
    f. Probe: What barriers have you experienced or do you think you might experience?

Thank you for all of your participation so far. We have a few more questions to go now that we are back with the full group.

5. While we were in small groups we talked a bit about what our schools actually look like. Now I want to brainstorm a little about what we could do better. What are some resources you wish you had to support students who were involved in trading sex?
    a. Probe: What barriers exist that keep you from obtaining these resources?
    b. Probe: Are there any particular types of professional development that would be helpful?
6. What are some questions you would like answered about working with students who are involved in trading sex?
    a. Probe: What questions can researchers help answer?
    b. Probe: What questions can community service providers help answer?
7. What advice would you have for a school nurse just starting out who found out a student was involved in trading sex?
8. Is there anything else you would like to add or anything we missed?

Before you go, we do want to direct everyone to the demographic questions. Finishing that demographics form which should only take about 10 minutes is what triggers the release of your $50 gift card. Your responses will not be connected with your name. We use this information only in how we describe who all participated in our study. If you want to, you can answer them right now and your gift card will be in your inbox later today. The link is in your email now. Thank you for participating in our study. If you have any questions or follow-up, please feel free to contact us. Good-bye.

Appendix E

LSN Focus Group Interview Structural Codes

Factors Influencing Chronically Absent Students

Code	**Description**	**Example/Note**
Student's Family	Participant describes the role of family in an individual young person's involvement or non-involvement in being chronically absent.	
Student's Family: Intra-familial relationships	Participant describes how relationships between members in the family influence a youth's involvement or non-involvement in being chronically absent. (Include relationships influenced by abusive or toxic communication)	
Student's Family: Family behaviors	Participant describes how behaviors exhibited by individual members or groups of members in the family influence a youth's involvement or non-involvement in chronic absenteeism. (Include abusive behavior)	(e.g. A parent picks the kid up without any questions asked when they're in school.)
Student's Family: Family attitudes	Participant describes how family member attitudes (as opposed to the actual actions taken to express those attitudes) influence a youth's involvement or non-involvement in trading sex.	
Student's Family: Family structure	Participant describes the structure of a youth's home as influential in the youth's involvement or non-involvement in chronic absenteeism. Family structure: the organization of who in the family lives in a household. Different family structures include: two-parent, single-parent, stepfamily, foster-family, extended-family, single-child family, etc. (Adapted from Mosby's Dictionary)	(e.g. Single parent goes out early to work → nobody to wake the kid up → kids wake up late and don't go to school.)
Student's Family: Family mental health	Participant describes a family member's mental health as influential in the youth's involvement or non-involvement in chronic absenteeism.	

Factors Influencing Chronically Absent Students

Code	**Description**	**Example/Note**
Student Mental Health	Participant describes the role of the individual student's mental health as influencing if the young person is chronically absent. Mental Health: A state of well-being in which the young person realizes their own abilities, can cope with normal stress, can work productively, and contribute to larger community. (Adapted from WHO)	
Student Mental Health: Antecedent factors	Participant describes the youth's mental health as contributing to the youth becoming chronically absent in the first place. This can be cyclical such that mental health state leads to continued involvement or leads to discontinuation of chronic absenteeism. Antecedent factor: factors that precede and lead to involvement (or non-involvement) in an action. An antecedent factor is the "cause" in cause and effect. (Adapted from Mosby's Dictionary)	
Student Mental Health: Emotional response	Participant describes the youth's mental health state as a response to being chronically absent. This could be somewhat cyclical where emotional response influences antecedent factors.	
Social Determinants	Participant describes specific characteristics of an individual student that may place the student at higher or lower risk for being chronically absent, regardless of quantitative evidentiary support.	
Social Determinants: Special education involvement	Participant describes involvement or not with special education as a factor in youth being chronically absent.	(e.g. student has an IEP)
Social Determinants: Housing stability	Participant describes housing stability or instability as a factor in youth being chronically absent.	(e.g. student is homeless or highly mobile)

Factors Influencing Chronically Absent Students		
Code	**Description**	**Example/Note**
	Homeless: lack of a fixed, regular, and adequate nighttime residence. This includes (1) sharing a home due to socioeconomic or similar reasons (e.g. staying with a friend, motel, etc.); (2) staying somewhere not designed for regular sleeping accommodation; or (3) living in cars, outdoors, substandard housing, public spaces, or similar. Unaccompanied youth: youth regularly not staying with parent or guardian. (Adapted from US Department of Education)	
Social Determinants: Race or ethnicity	Participant describes race and/or ethnicity as a factor in youth involvement or non-involvement in chronic absenteeism.	
Social Determinants: Access to resources	Participant describes access, limited access, or lack of access to specific necessary resources as a factor in youth involvement or non-involvement in chronic absenteeism. Do not include illegal resources that may be perceived as necessary (e.g. drugs or alcohol) but if legal to own object is illegally obtained, include that.	(e.g. Transportation, alarm clock, cell phone)
Social Determinants: Sexual orientation or gender identity	Participant describes sexual orientation or gender identity as a factor in youth being chronically absent.	
Social Determinants → Social Connection	Participants describe a relationship with peers or non-family member adults either in person or online that contributes to a young person's involvement or non-involvement in chronic absenteeism.	(e.g. hanging out with the "bad kids;" dating someone they have only interacted with online)
Social Determinants → Abuse history	Participant describes a history of physical, psychological, or substance abuse as a factor in youth being chronically absent.	(e.g. history of substance use)
Other Factors Facilitating Involvement	Participant describes additional factors contributing to youth being	

Factors Influencing Chronically Absent Students		
Code	**Description**	**Example/Note**
	chronically absent that do not fall into one of the other categories.	
Student physical Health	Participant describes physical factors (e.g. sleep disorder, asthma, DM) contributing to youth being chronically absent.	(e.g. sleep problem, chronic illness)

LSN Role in Supporting Chronically Absent Children		
Code	**Description**	**Example/Note**
LSN Interventions	Participant describes LSN intervention related to supporting individual chronically absent young people. This can include effectiveness of an intervention and/or context around an intervention. *Intervention:* Any direct care treatment an LSN performs on behalf of an individual student. This includes collaborative intervention, independent intervention, and the action of referring a student to another care provider or service.	
LSN Interventions: LSN independent	Participant describes an LSN intervention that is performed independent of other care providers. This should include independent referrals to services outside of the health office.	
LSN Interventions: Collaborative	Participant describes an LSN intervention that is performed in collaboration with other care providers. This should include collaborative referrals to services outside of the health office.	
LSN Interventions: No intervention	Participant describes no action on their part to address concerns or knowledge of a student in chronic absenteeism.	
LSN Interventions: Intervention Wishlist	Participant describes an intervention they wish they had performed or could perform in the future.	
Primary Prevention	Participant describes primary prevention actions taken by LSNs to prevent youth from becoming chronically absent. *Primary prevention*: a program of activities directed to improving general	(e.g. teaching staff about signs to look for)

LSN Role in Supporting Chronically Absent Children		
Code	**Description**	**Example/Note**
	well-being while also involving specific protection for selected issues	
Primary Prevention: Specific intervention	Participant describes a specific primary prevention intervention utilized to prevent youth being chronically absent.	
Primary Prevention: Intervention Wishlist	Participant describes a specific primary prevention intervention they had performed or could perform in the future.	

Barriers and Facilitators to Supporting Chronically Absent Children		
Code	**Description**	**Example/Note**
Barriers to Providing Care	Participant describes something that obstructs an actor's ability to provide nursing (or other disciplinary) care to a chronically absent student.	
Barriers to Providing Care: Systemic barriers	Participant describes something that obstructs an actor's ability to provide nursing (or other disciplinary) care to a chronically absent student at the systemic level.	(e.g. Lack of resources when there's family-related problem; School having a problem with access to resources)
Barriers to Providing Care: School level barriers	Participant describes something that obstructs an actor's ability to provide nursing (or other disciplinary) care to a young chronically absent student at the school level.	(e.g. LSN role being devalued in a school; wander away after getting off the bus)
Barriers to Providing Care: Individual LSN barriers	Participant describes something that obstructs an actor's ability to provide nursing (or other disciplinary) care to a chronically absent student at the individual LSN level.	(e.g. LSN discomfort in reaching out to parents or student)
Barriers to Providing Care: Student or family barriers	Participant describes something that obstructs an actor's ability to provide nursing (or other disciplinary) care to a chronically absent student at the individual student or family level.	(e.g. student disinterest in behavior change)
Facilitators to Providing Care	Participant describes something that simplifies or supports an actor's ability to provide nursing (or other disciplinary) care to a chronically absent student.	
Facilitators to Providing Care: Systemic facilitators	Participant describes something that simplifies or supports an actor's ability to provide nursing (or other disciplinary) care to a chronically absent student at the systems level.	(e.g. system to track student movement in schools to improve continuity of care if they transfer)

Barriers and Facilitators to Supporting Chronically Absent Children		
Code	**Description**	**Example/Note**
Facilitators to Providing Care: School level facilitators	Participant describes something that simplifies or supports an actor's ability to provide nursing (or other disciplinary) care to a chronically absent student at the school level.	(e.g. school climate is collaborative)
Facilitators to Providing Care: Individual LSN facilitators	Participant describes something that simplifies or supports an actor's ability to provide nursing (or other disciplinary) care to a chronically absent student at the individual LSN level.	(e.g. professional development)
Facilitators to Providing Care: Student or family facilitators	Participant describes something that simplifies or supports an actor's ability to provide nursing (or other disciplinary) care to a chronically absent student at the individual student or family level.	(e.g. caring adult relationship(s) at school)