2021

Modeling Substance Use and Mental Disorder Comorbidity Using Latent Variable and Network Approaches

Courtney T. Blondino
Virginia Commonwealth University

Modeling Substance Use and Mental Disorder Comorbidity Using Latent Variable and Network Approaches

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

by

Courtney Taylor Blondino

Bachelor of Science, Human Nutrition, Foods, and Exercise, Virginia Tech, 2015
Master of Public Health, Epidemiology, University of Kentucky, 2017

Director:

Elizabeth Prom-Wormley, Ph.D.
Assistant Professor
Division of Epidemiology
Department of Family Medicine and Population Health

Virginia Commonwealth University
Richmond, Virginia
May 2021

Dissertation Committee

Elizabeth Prom-Wormley, PhD, MPH (Committee Chair)
Assistant Professor
Division of Epidemiology
Department of Family Medicine and Population Health
Virginia Commonwealth University, Richmond, VA
Committee Members:

Juan Lu, PhD, MPH, MD
Associate Professor
Director of Epidemiology Graduate Programs
Division of Epidemiology
Department of Family Medicine and Population Health
Virginia Commonwealth University, Richmond, VA

Robert Perera, PhD
Associate Professor
Department of Biostatistics
Virginia Commonwealth University, Richmond, VA

Roxann Roberson-Nay, PhD
Associate Professor
Virginia Institute for Psychiatric and Behavioral Genetics
Virginia Commonwealth University, Richmond, VA

Michael Neale, PhD
Professor
Virginia Institute for Psychiatric and Behavioral Genetics
Virginia Commonwealth University, Richmond, VA

ACKNOWLEDGEMENTS

First and foremost, I would like to acknowledge my mentor, Elizabeth Prom-Wormley. Thank you for this incredible journey. Your teaching and unwavering support encouraged me to be the best student and epidemiologist possible. It has been an honor and a privilege to work with you. I sincerely thank my dissertation committee for their support and feedback which helped to make this dissertation stronger. I also would like to acknowledge the faculty, staff, and fellow students, especially James Clifford, in the Division of Epidemiology for their commitment to my success. Finally, I think my family and friends. My mother, Lisa Blankinship, has supported me in every endeavor – thank you for always believing in me. My sister, Megan Blondino, has loved me unconditionally – thank you for never letting me give up. My husband, Alex Roberson, has stood by my side for the last six years of graduate school – thank you for always being there for me.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

GLOSSARY OF ABBREVIATIONS

95% CI	95% Confidence Interval
AIC	Akaike Information Criteria
ADHD	Attention-Deficit/Hyperactivity Disorder
AOR	Adjusted Odds Ratio
AUC	Area Under the Curve
BIC	Bayesian Information Criteria
CIG	Cigarette
ECIG	Electronic cigarette
DSM	Diagnostic and Statistical Manual of Mental Disorders
ICD	International Statistical Classification of Diseases and Related Health Problems
GAIN-SS	Global Appraisal of Individual Needs – Short Screener
LCA	Latent Class Analysis
LMRT	Lo-Mendell-Rubin Test
LRT	Likelihood Ratio Test
ND	Nicotine Dependence
OR	Odds Ratio
PATH	Population Assessment of Tobacco and Health
PDNP	Prescription Drugs Not Prescribed
PTSD	Post-Traumatic Stress Disorder
ROC	Receiver Operating Characteristic
SUD	Substance Use Disorder

ABSTRACT

Modeling Substance Use and Mental Disorder Comorbidity Using Latent Variable and Network Approaches

By Courtney Taylor Blondino, MPH, Ph.D.

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at Virginia Commonwealth University

Virginia Commonwealth University, 2021

Director: Elizabeth Prom-Wormley, Ph.D.
Assistant Professor
Division of Epidemiology
Department of Family Medicine and Population Health

<u>Introduction.</u> Substance use disorder (SUD) is a common condition that affects millions of Americans and represents a substantial burden to the U.S. healthcare system. Addressing SUD has been complicated by comorbid mental disorders and co-occurring substance use. Consequently, detailing and addressing SUD and comorbid SUD represent an important goal to improve the health of Americans.

<u>Objective.</u> The research goal of this dissertation was to characterize the comorbidity between substance use, including tobacco use, and mental disorder symptoms measured as negative affect and externalizing symptoms in a population-based sample using latent variable and network approaches. Specifically, this project aims to: preliminarily assess comorbidity using multinomial regression between lifetime negative affect severity, externalizing severity and nicotine dependence, and current use of tobacco (cigarettes and e-cigarettes) and alcohol (Chapter 2); identify latent classes of

comorbid substance use as well as negative affect and externalizing symptoms and their ability to predict SUD severity (Chapter 3); detail substance use, negative affect, and externalizing symptom networks and test for differences in the network structure and connectivity by gender (Chapter 4); and use pairwise comparisons from the LCA and network results to address stability or movement of comorbidity structures over three waves of data (Chapter 5).

Methods. Waves 1 – 3 from the Population Assessment of Tobacco and Health Study were used. Various statistical analyses were used to complete each project including multinomial and ordinal regression, latent class analysis, cumulative ROC curve analysis, and network analysis.

Results. The associations between psychopathology (negative affect vs. externalizing severity) varied by different substance use combinations. Results from the latent class analysis identified a four-class solution as most optimal in characterizing comorbidity: low symptom (N=23,571, 72.9%), negative affect (N=4,098, 12.7%), externalizing (N=2,691, 8.3%), and comorbid (N=1,960, 6.1%). Network analysis results showed similarities between men and women. The strongest substance use/mental health symptom connections estimated as edge-weights (EW) in the network were between marijuana with lying (EW = 0.60, 95% CI = 0.49; 0.70), marijuana with engaging in fights (EW = 0.54, 95% CI = 0.27; 0.81), prescription drugs not prescribed (PDNP) with having trouble sleeping (EW = 0.53, 95% CI = 0.40; 0.66), and alcohol and impulsivity (EW = 0.48, 95% CI = 0.42; 0.53). Both latent class analysis and network analysis results identified relationships between (1) exclusive cigarette, dual cigarette and e-cigarette, marijuana, and PDNP with negative affect symptoms, and (2) alcohol with externalizing symptoms. Similar latent profiles emerged across the three waves specifically where the low symptom class was largest (65.5% to 72.9%) and the comorbid class was smallest (6.1% to 8.2%). Network structure and connectivity did not significantly differ by wave; however, edge-weight comparisons identified some stronger connections among the substance use behaviors and mental disorder symptoms from preceding to subsequent waves.

<u>Conclusions.</u> The results from the four different studies fill extensive gaps in the comorbidity research. This dissertation identified specific combinations of substance use behaviors and mental disorder symptoms, determined which sociodemographic factors play a role in specific comorbidity profiles, and assessed the patterns of comorbidity among three waves of data. These results support the need to approach substance use and mental disorders from a more holistic perspective, taking comorbidity into account to better support the overall wellbeing of the individual. The results can inform robust and targeted prevention strategies to effectively mitigate the substantial burden and societal costs of comorbidity in the U.S. population.

CHAPTER 2: THE ASSOCIATION BETWEEN NEGATIVE AFFECT AND EXTERNALIZING SEVERITY WITH CURRENT USE OF CIGARETTES, E-CIGARETTES, AND ALCOHOL IN ADULTS: WAVE 1 OF THE POPULATION ASSESSMENT OF TOBACCO AND HEALTH (PATH) STUDY

Introduction. Concurrent tobacco/alcohol use is common in adults, and associated with the severity of symptoms experienced by those with mental health disorders. However, few studies have explored this relationship across different combinations of tobacco products [i.e., conventional cigarette (CIG) and electronic cigarette (ECIG)] and alcohol.
Objective. Examine the association of lifetime mental disorder symptom severity and past 30-day combinations of CIG, ECIG, and alcohol use.
Methods. Data from the Wave 1 (2013-2014) Population Assessment of Tobacco and Health study were used. A total of 15,947 adults aged 18 years or older with complete study information were included. Multinomial logistic regression analyses were performed to determine the relationship between lifetime negative affect/externalizing severity and past 30-day use of tobacco and alcohol, adjusting for nicotine dependence (ND), sex, age, race, education, and income.
Results. Negative affect severity was more strongly associated with CIG and alcohol use (moderate AOR= 1.47, 95% CI= 1.22-1.77; high AOR= 1.29, 95% CI= 1.03-1.61) as well as alcohol-exclusive use (moderate AOR= 1.58, 95% CI= 1.27-1.96; high AOR= 1.31, 95% CI= 1.05-1.64) while externalizing severity was more strongly associated with ECIG and alcohol use (high AOR= 2.97, 95% CI= 1.84-4.81, moderate AOR= 2.29, 95% CI= 1.53-3.43) when accounting for ND compared to none. The relationship between externalizing severity with ECIG use was dependent on alcohol being used with ECIG.
Conclusions. The associations between psychopathology (negative affect vs. externalizing severity) vary by different combinations of alcohol, CIG, and ECIG. Further, these relationships may be mediated through ND. Future investigations into the comorbidity between mental disorder symptoms with tobacco and alcohol use should consider use of specific substances and their combination.

CHAPTER 3: LATENT CLASSES OF COMORBID SUBSTANCE USE AND NEGATIVE AFFECT AND EXTERNALIZING SYMPTOMS AND THEIR ROLE IN ADULT SUBSTANCE USE DISORDER SEVERITY

Introduction. SUD poses a substantial burden on the United States' health system. Many prevention efforts exist to slow the progression or prevent SUD from occurring. Substance use and mental health comorbidity profiles could predict SUD severity, further informing prevention and intervention strategies.
Objective. Identify latent classes of comorbid substance use as well as negative affect and externalizing symptoms and assess their ability to predict SUD severity.
Methods. Latent class analysis of past-month endorsement of negative affect and externalizing symptoms and past 30-day substance use will be used for each wave separately. We tested the degree to which demographic and social factors influence the probability of class membership. The probability of comorbidity class membership will be included in regression models to test the predictive probability of SUD severity.
Results. A four-class solution was considered to best fit the data and were categorized: low symptom (N=23,571, 72.9%), negative affect (N=4,098, 12.7%), externalizing (N=2,691, 8.3%), and comorbid (N=1,960, 6.1%). Substance use varied across the mental disorder symptoms. Exclusive cigarette use, dual cigarette and e-cigarette use, marijuana use, and prescription drugs not described more commonly occurred in the negative affect class while exclusive e-cigarette and alcohol use more commonly occurred with the externalizing class. Women and people with low socioeconomic status had higher odds of membership in the comorbid and negative affect classes. Social satisfaction was a very strong factor associated with the comorbid and negative affect classes. Latent class membership predicting SUD severity performed similarly to a model where the symptoms were grouped separately (i.e., negative affect symptoms, externalizing symptoms, and substance use behaviors).
Conclusions. A four-class solution best described the comorbidity structure in a nationally representative sample of U.S. adults. Certain substance use behaviors were more commonly associated with specific mental disorder symptoms. Demographic factors and a potentially modifiable social factor were significantly associated with latent

class membership. Overall, prediction of SUD severity was poor for latent class membership as well as substance use behaviors and mental disorder symptoms group separately. These results identify the need for prevention efforts required to mitigate development of more severe course of illness. Future work should consider other methodological approaches (e.g., factor mixture modeling and network analysis) to further investigate the comorbidity structure of U.S. adults.

CHAPTER 4: A NETWORK APPROACH TO SUBSTANCE USE, NEGATIVE AFFECT, AND EXTERNALIZING COMORBIDITY IN U.S. ADULTS

Introduction. Use of conventional cigarettes (CIG), alcohol, marijuana, and sedatives [i.e., benzodiazepines and barbiturates]) commonly co-occur with negative affect and externalizing disorders. It is unclear how these relationships extend to electronic cigarettes (ECIG) and prescription drugs not prescribed (i.e., sedatives, tranquilizers, and painkillers [PDNP]), and whether they differ by gender.

Objective. Detail substance use, negative affect, and externalizing symptom networks, and compare by gender.

Methods. Data from Wave 1 of the adult PATH sample was used to test a network model of past 30-day substance use, negative affect symptoms, and externalizing symptoms. Global and local differences in men and women networks were tested through visual comparisons, global strength invariance, network structure invariance, and edge strength invariance.

Results. Overall, networks were consistent between men and women. The strongest substance use/mental health symptom connections estimated as edge-weights (EW) were between marijuana with lying (EW = 0.60, 95% CI = 0.49; 0.70), marijuana with engaging in fights (EW = 0.54, 95% CI = 0.27; 0.81), PDNP with having trouble sleeping (EW = 0.53, 95% CI = 0.40; 0.66), and alcohol and impulsivity (EW = 0.48, 95% CI = 0.42; 0.53).

Conclusions. There were many weak connections throughout the substance use and negative affect/externalizing network. A few important connections were identified and encourage future study. In particular, PDNP was most strongly associated with negative affect while marijuana, alcohol and PDNP use were most strongly associated with externalizing.

CHAPTER 5: PRELIMINARY PATTERNS OF SUBSTANCE USE AND MENTAL DISORDER SYMPTOM COMORBIDITY IN ADULTS OVER TIME

Introduction. Patterns of co-occurring substance use and mental health conditions are well-described in youth and young adult populations. It remains unclear whether these patterns continue into adulthood.

Objective. Perform a preliminary assessment to determine the stability of substance use and mental disorder comorbidity across three years of data (2013-2016) using both latent class and network analyses.

Methods. Latent class analyses were conducted cross-sectionally for each wave of data (Wave 1, Wave 2, Wave 3). Class probability parameters, item response probability parameters, transition patterns and results from the multinomial logistic regression were compared across the three waves. Network models were estimated, and three tests of network invariance were used to test significant differences in network models by wave.

Results. Four-class solutions generated from the latent class analyses were compared by wave. Similar latent profiles emerged across the three waves specifically where the low symptom class was largest (65.5% to 72.9%) and the comorbid class was smallest (6.1% to 8.2%). Overall, when individuals transitioned from one class to another, they typically transitioned into the low symptom class (62.3% to 66.8%) from preceding to subsequent wave. Network structure and connectivity did not significantly differ by wave; however, edge-weight comparisons identified some stronger connections among the substance use behaviors and mental disorder symptoms from preceding to subsequent waves.

Conclusions. The comorbidity structure is consistent across waves. The connections between these behaviors and symptoms may become stronger at each wave. Therefore, investment of time, money, and other resources are encouraged to support those experiencing comorbidity as they are unlikely to change in adulthood

CHAPTER 1: INTRODUCTION

Substance use disorder

Substance use disorder (SUD) develops as a result of prolonged use of any psychoactive substance at high doses and/or frequencies, and is defined as the continued use of alcohol and/or drugs despite clinically significant impairment, including health problems, disability, and failure to meet major responsibilities at work, school, or home.[1–3] The essential feature of a SUD is a cluster of cognitive, behavioral, and physiological symptoms showing that the individual continues substance use despite significant substance-related problems.[2] Diagnosis of SUD is based on a pathological pattern of behaviors related to the use of a substance.[2]

SUD represents a significant public health burden because of the life-years lost due to disability, impaired quality of life, disruption of work and family relationships, and death from accidents or overdose.[4] In 2018, approximately 19.3 million American adults met diagnostic criteria for a past-year SUD,[5] and drug abuse and addiction cost society an estimated $600 billion every year.[6]

Substance use disorder and mental health comorbidity

SUD commonly co-occurs across substances and with mental disorders. Approximately 6% of American adults are affected with SUD.[2,7,8] Of those affected with SUD, about 50% have a co-occurring or comorbid mental illness such as negative affect (i.e., behaviors such as depression or anxiety where the distress of an affected individual is expressed inward) or externalizing disorders (i.e., behaviors such as attention-deficit hyperactivity disorder where the distress of an affected individual is expressed outward).[5] Further, many people affected with SUD also engage in use of other substances. For example, alcohol use disorder and nicotine dependence are commonly reported in approximately 25-50% of those with marijuana, cocaine, prescription opioid, and heroin use disorders.[9] Some common mental disorders that have been associated with SUDs (e.g., tobacco, alcohol, marijuana, cocaine, sedatives) include anxiety disorders, depressive disorders, conduct disorder, ADHD, and antisocial personality disorder.[9–17]

Comorbid substance use and mental disorders represent a substantial burden to the American health care system. Of the approximately 20 million adults in the United States (U.S.) who experience a SUD, half also have a co-occurring mental illness.[5] People with comorbid substance use and mental disorders suffer from more severe health outcomes compared to those who experience one disorder.[18] Substance use and mental disorders are the leading cause of disease burden in the U.S. This has increased from 2779 DALYs (age standardized disability adjusted life years) in 1990 to 3355 DALYs in 2015.[19] Additionally, the U.S. has the highest rate of death due to substance use and mental disorders together at an age standardized death rate = 12.0 per 100,000 compared to an average of 4.9 per 100,000 in similarly wealthy countries (e.g., France = 6.5 per 100,000; Canada = 5.8 per 100,000; United Kingdom = 5.2 per 100,000; Netherlands = 2.5 per 100,000).[19] The economic burden of substance use and mental disorder comorbidity, due to treatment spending from all public and private sources, is expected to increase from $171.7 billion in 2009 to $280.5 billion in 2020.[20]

Co-occurring mental disorders, without SUD, are also common. Negative affect disorders like depression and anxiety are frequently associated with one another.[21,22]

Examples of negative affect symptoms include feeling depressed, feeling anxious, having sleep trouble, or becoming distressed or upset about the past. Negative affect symptoms are commonly reported in those with externalizing disorders.[23] Externalizing disorders reflect distress expressed outward which is commonly diagnosed as attention deficit hyperactivity disorder (ADHD), oppositional defiant disorder, conduct disorder, antisocial personality disorder, and sometimes SUD.[22] Examples of externalizing symptoms include having a hard time paying attention or listening, feeling restless, acting impulsively, lying or conning, threatening people, and starting physical fights with people.

SUD is a common condition that affects millions of Americans and represents a substantial burden to the U.S. healthcare system. Addressing SUD has been complicated by comorbid mental disorders and co-occurring substance use. Consequently, detailing and addressing SUD and comorbid SUD represent an important goal to improve the health of Americans.

Current state of SUD measurement

The American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-V) is a classification of mental disorders, including SUD, with associated criteria designed to facilitate more reliable diagnoses of these disorders.[2] To date, the DSM-V is the standard reference for clinical practice and is considered the best description of how mental disorders are expressed.[2] In the DSM-V, substance-related disorders encompass ten separate classes of drugs: alcohol; caffeine; cannabis; hallucinogens (with separate categories for phencyclidine [or similarly acting arylcyclohexylamines] and other hallucinogens); inhalants; opioids; sedatives, hypnotics, and anxiolytics; stimulants (amphetamine-type substances, cocaine, and other stimulants); tobacco; and other (or unknown) substances. For a twelve-month period for diagnoses, two of the eleven criteria are required: (1) take substance in larger amounts or over longer period than intended, (2) express consistent desire to cut down or regulate use, (3) spent a great deal of time obtaining or using the substance, or recovering from its effects, (4) intense desire or urge for the substance (craving), (5) use results in failure to fulfill major role obligations, (6) continues use despite persistent social or interpersonal problem, (7) reduced involvement in activities because of use, (8) risky use in situations which are physically hazardous, (9) continued use despite physical or psychological problems, (10) requiring an increased dose of substance to achieve desired effect (tolerance), and (11) withdrawal symptoms. Diagnostic criteria allow a severity rating along with diagnosis: mild SUD defined as two or three symptoms, moderate SUD defined as four or five symptoms, and severe SUD defined as six or more symptoms.

The International Classification of Disease, Tenth Revision, Clinical Modification (ICD-10-CM) is the other main diagnostic classification systems for SUD in the United States.[24] The World Health Organization produces the ICD-10-CM with the primary focus for mental and behavioral disorder classification to help countries reduce the disease burden of mental disorders.[24] It is the standard transaction code set for diagnosis under the Health Insurance Portability and Accountability Act (HIPAA) and is used to track disease burden, mortality statistics, and to ensure appropriate billing.[25]

The ICD-10-CM lists the same ten substances as the DSM followed by a list of specifiers under larger categories including abuse, dependence, and use.

There are other instruments used in clinical and non-clinical settings to identify and measure SUD and SUD severity, overall or by specific substance. Common evidence-based instruments include the Addiction Severity Index (ASI), Alcohol Use Disorders Identification Test (AUDIT), Drug Abuse Screen Test (DAST), Fagerstrom Test for Nicotine Dependence (FTND), and the Global Appraisal for Individual Needs (GAIN).[26–28] These tools are useful to assess SUD and SUD severity in settings such as epidemiologic research of large populations where it is impractical to establish diagnoses.[27,28] This is important because substance use that does not result in a diagnosis of SUD remain pervasive throughout American society.[8] For example, 85.6% of American adults engage in alcohol use.[29] Of these, 25.8% engage in patterns of use that would not necessarily lead to a diagnosis of alcohol use disorder such as binge drinking (i.e., consuming 5 or more alcoholic drinks for males or 4 or more alcoholic drinks for females on the same occasion) in the past-month.[29] Additionally, 8.3% of Americans 12 years of age and older reported past-month marijuana use with intensity of use increasing (i.g.,11.1% of heavy use in 1992 to 35.4% in 2014).[30] Further, there is increasing evidence that environmental stressors such as the current COVID-19 pandemic can influence sub-threshold use towards problematic use.[31,32] Nevertheless, there are few effective strategies that address sub-threshold use in order to address population-level mental health issues in people with who do not meet criteria for SUD. Consequently, the use of other tools that can evaluate substance use beyond disordered substance use is particularly useful to detail population-level substance use that measure sub-threshold SUD. Nevertheless, a common limitation of current SUD instruments is that none of these tools address SUD comorbidity. Therefore, it has been challenging to characterize SUD comorbidity with currently available measurement tools. Instead, research has focused on modeling SUD comorbidity across measures of substance use and mental disorders.

Conceptual models of SUD-mental disorder comorbidity

Epidemiological research of substance use, negative affect, and externalizing disorders has typically studied comorbidity from three major perspectives. These models attempt to either identify or confirm an association between symptoms or disorders, test the causal relationship between comorbid disorders, or describe the patterns of overlap across disorders or symptoms.[12,22,33–41] To date, these models have concluded: (1) substance use behaviors and disorders co-occur,[34–36,42] (2) negative affect and externalizing symptoms and disorders co-occur,[22,38–40] and (3) substance use behaviors/disorders co-occur with negative affect and externalizing symptoms/disorders.[12,33,41] However, the use of each model has often been completed in isolation and this approach produces gaps in our understanding of comorbidity. The section below reviews the models that have been used to study comorbidity, their strengths and weaknesses, and identifies needs to expand insights that could be gained from these models.

<u>*Common cause model*</u>. To date, the major psychiatric classification systems (i.e., DSM-V and ICD-10) measure SUD, negative affect, and externalizing disorders as single latent constructs based on observable symptoms.[43] Therefore, SUD represents a

latent or unobservable construct of disease (denoted as a circle) that causes the observable symptoms (denoted as squares) which are measured to diagnose SUD (e.g., substance use taken in large amounts, cravings, social problems, and cessation attempts) (Figure 1.1). Based on the DSM's approach to diagnosis, at least two of the eleven symptoms listed in the criteria above are needed to result in a diagnosis of SUD. This model is called the common cause model. It has also been referred to as the medical model and has also been applied to physical conditions.[44,45]

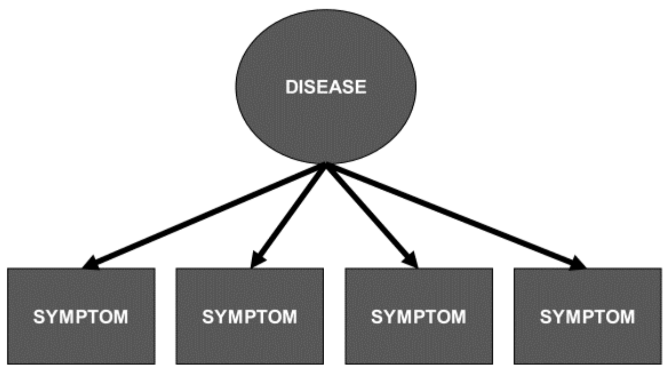


Figure 1.1: Common Cause Model of Disease

The common cause model assumes that the disease has a common pathogenic pathway or an etiology in which the mechanism is fully understood.[45] However, common pathogenic pathways for mental disorders, including SUDs, have not been identified.[45–48] Additionally, the common cause model is unidimensional and does not account for comorbidity. These models become complicated to interpret when we take multiple disorders, and their overlapping symptomatology, into account. This is problematic because there is a high degree of comorbidity that is not accounted for through our current diagnostic classification systems resulting in potential misclassification of disorder diagnosis. The internalizing-externalizing model presents an extension to the common cause model to account for comorbidity between internalizing and externalizing mental disorders.

Internalizing-Externalizing Model. The internalizing-externalizing model is a two-factor model of internalizing and externalizing factors that explain the interrelationships of psychiatric disorders, seen in Figure 1.2.[22,38] In this model, internalizing disorders like mood (major depressive disorder and dysthymia) and anxiety (generalized anxiety disorder, separation anxiety disorder, phobias, obsessive-compulsive disorder) disorders reflect a similar construct, and are associated with or explains the variance of the internalizing factor. Externalizing disorders like ADHD, oppositional defiant disorder, conduct disorder, antisocial personality disorder, and sometimes SUDs reflect a separate construct, an externalizing factor. The internalizing and externalizing factors can also be correlated in this model, indicating that internalizing and externalizing disorders are comorbid with each other.[22] This model has received considerable attention for understanding co-occurring psychiatric disorders, including SUDs, as latent factors in adults.[22,38–40]

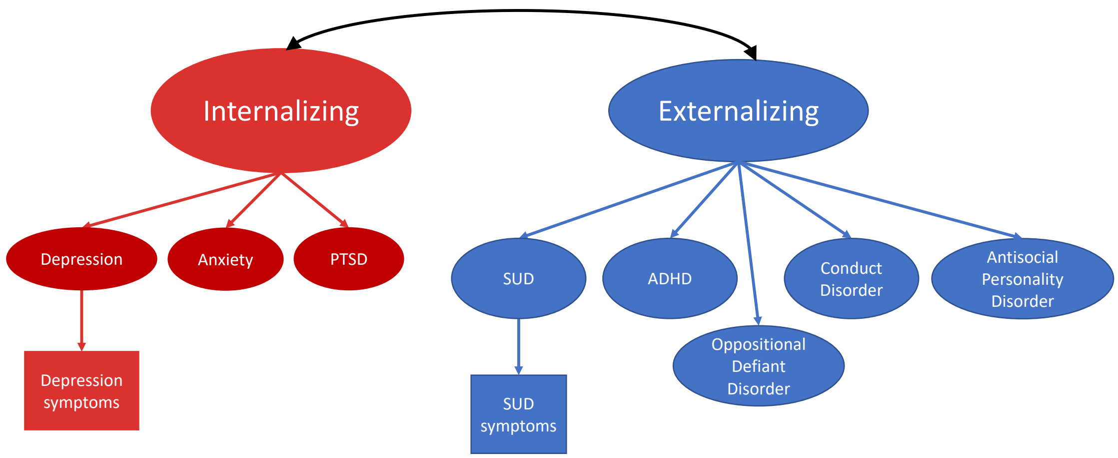


Figure 1.2: Internalizing-Externalizing Model. In the internalizing-externalizing model, the observable symptoms (in boxes) are caused by an unobserved internalizing and externalizing latent variables (in circles). This model allows there to be some correlation between the internalizing and externalizing latent variables.

The internalizing-externalizing model assumes that negative affect disorders represent a latent negative affect construct with externalizing disorders, including SUD represent a latent externalizing construct. Statistically, these latent factors represent the proportion of variance shared between the observed disorder (i.e., depression) and the latent construct (i.e., internalizing). This model extends the common cause model because it accounts for the comorbidity between internalizing and externalizing disorders. This model predominantly focuses on mental disorders as latent constructs and does not include a robust set of substance use behaviors. This method explains the relationships between the observed disorders or symptoms that explain the latent construct by calculating the model implied covariance. However, this calculation does not describe the unobserved heterogeneity in the population to identify different comorbidity patterns in a population. This is a limitation of the negative internalizing-externalizing model.

Network Model. The network model is a relatively new psychometric approach that can reduce the lack of clarity on the relationships between the observed disorders or symptoms that explain the latent construct and address the associations between the observed disorders or symptoms. A network model is likely to support a deeper understanding of comorbidity because it conceptualizes symptoms as mutually interacting, often reciprocally reinforcing elements of a complex network.[46] The network approach is based on the idea that comorbidities arise from shared symptoms between disorders which can capture complexity and individual variation in psychopathology.[49] The network approach naturally accommodates comorbidities as a central part of its theory.[50] In the network approach, comorbidity represents causal relationships between symptoms in which pathways can bridge symptoms that are part of multiple disorders.[46] Using a network model, symptoms, rather than disorders, are considered within the network structure. Rather than the disorder acting as the underlying cause of all symptoms, it is the symptoms that mutually interact and set a person into a disordered mental health state.

An example of the use of a network model is detailed in Figure 1.3 to summarize comorbidity of symptoms for SUD and depression. Symptoms found in depression and SUD include insomnia and weight loss. Within a network model, the symptoms make up a comorbid network structure of several symptoms that is specific to the person. This model conceptualizes how symptoms of different disorders function together specifically to produce a comorbid disordered state. The network approach explains the co-occurrence of mental disorder symptoms, including substance use behaviors, as resulting from direct interactions between these symptoms.[50] In network analysis, the term *interaction* is used to explain the reciprocal action or influence of symptoms. In the context of network analysis, *interaction* is not used to test whether an effect can be greater than (positive interaction, synergism) or less than what we would expect (negative interaction, antagonism).[51]

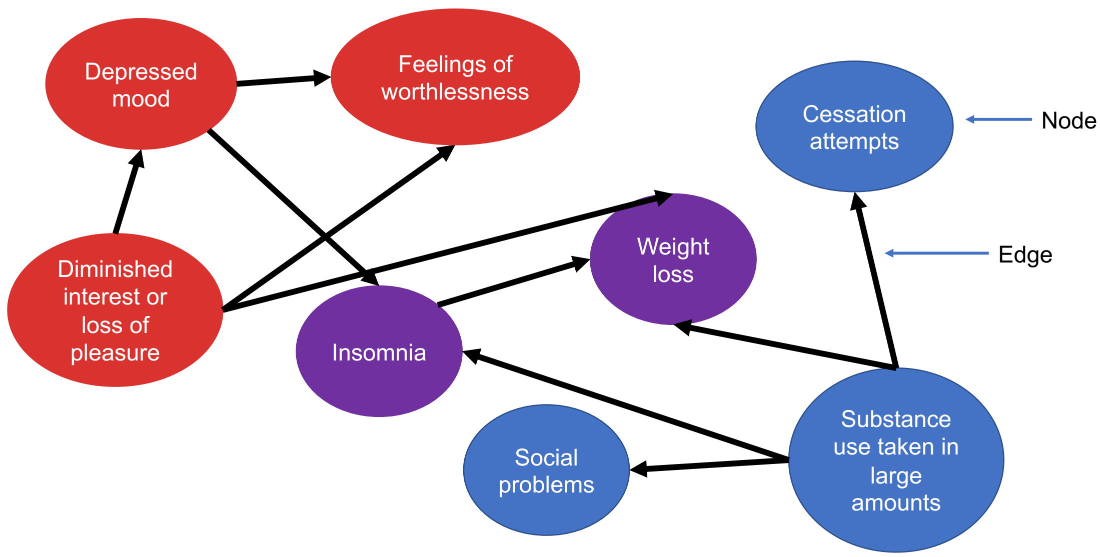


Figure 1.3: Network Model of Depression and SUD symptoms. The network model of depression and SUD is made of nodes (circles) and edges (lines connecting nodes). This is a directed network (arrows are directed from source to target node) where one symptom can lead to the activation of another. The depression symptoms in red are clustered together to the left of the network. The SUD symptoms in blue are clustered together to the right of the network. Insomnia and weight loss (in purple) are symptoms that occur in both depression and SUD and act as bridges between the disorders. The positioning and the distance between the symptoms/nodes within the network have implications for the comorbidity structure of depression and SUD.

Patterns of symptom-symptom or symptom-behavior interactions can be encoded in a network structure.[45] Measured symptoms and behaviors are represented as nodes. Nodes are connected by edges (seen in Figure 1.3). Edges represent the interactions between the nodes. Nodes that directly activate each other (i.e., demonstrate an association) are connected while nodes that do not directly activate each other are not. Changes that occur outside the network, external forces, can influence the symptoms and the interactions between the symptoms.[45]

Principles underlying the network approach imply the etiology of mental disorders as a process of spreading activation in a symptom network.[45] For example, if a symptom arises (for any reason), it may influence the probability that a connected symptom will activate as well.[45] A mental disorder will arise when a group of tightly coupled symptoms activate, and the cluster becomes self-sustaining.[45] Although symptom interactions may be most active within symptoms sets that are associated with a given mental disorder, these interactions do not stop at diagnostic boundaries.[45]

In network theory, diagnosis is conceptualized as a process where the presence of symptoms is identified by clinicians and any symptom-symptom interactions that sustain themselves.[45] An example of sustained symptom-symptom interactions could be a phenomenon in which one's depressive mood results in a lack of restful sleep which could lead to greater fatigue which may ultimately sustain their depressed mood, rendering the person to be diagnosed with a depressive disorder. Treatment could then evolve to intervene on the symptom interactions (i.e., directly change the state of one or more symptoms), the external field (i.e., remove triggering causes or add a protective layer to mitigate the symptom activation), or the network (i.e., modify the symptom-symptom connections).

Longitudinal trends for SUD and comorbidity

An inherent limitation of the models previously described is that the analysis is done at a single point in time. Previous studies have described that substance use behaviors, including past month substance use, can change over time. The gateway hypothesis of substance use posits that single and extensive use of alcohol and tobacco products can function as an entrance to polysubstance use, the use of at least two different psychotropic substances.[52–54] There is also evidence that certain mental health conditions can increase the risk of developing future mental health conditions, sometimes more severe. For example, chronicity of depressive symptoms increases the likelihood of anxiety and substance use disorders.[2,55,56] Studies of adolescents have reported that (1) externalizing problems (i.e., ADHD, ODD, CD) in youth precede substance use in both boys and girls whereas (2) substance use (i.e., alcohol and marijuana) in youth predict negative affect disorders in adulthood specifically for women.[57–61] Less is known about how these trends continue in adulthood. Therefore, greater investigation into substance use behaviors over time with mental health conditions are necessary to further develop the literature around longitudinal trends for SUD and comorbidity.

Common knowledge gaps across all chapters

The gold-standard diagnostic classification systems (i.e., DSM and ICD-10-CM) in the United States describe disorders as single latent constructs or single dimensions rather than considering disorders as multidimensional. Nevertheless, the American Psychiatric Association recognizes that mental disorders do not always fit completely within the boundaries of a single disorder.[2] This approach to diagnosis may be limited and could benefit from additional insight because (1) comorbidity is common and (2) the current tools do not consider comorbidity within the context of diagnosis. Additionally, in order to receive a diagnosis, a person must have the appropriate number of criteria to reach a diagnostic threshold and access to a physician or person qualified to diagnose. Using a threshold approach in current classification systems may underestimate the number of people who experience substance use and mental disorders, especially those that present as comorbidities. Consequently, current SUD research suffers from the unidimensional approach that does not account for comorbidity. Further, addressing SUD comorbidity could benefit from knowledge of the patterns of the symptoms underlying an SUD diagnosis. Such a symptoms approach to measuring comorbidity (e.g., past-month substance use or endorsement of mental disorder symptom) may be better in estimate the prevalence of comorbidity correctly. Furthermore, better prediction of additional health outcomes and more targeted prevention and intervention strategies are likely to result in a more accurate representation of comorbidity prevalence.

Patterns of comorbidity are not the same although current knowledge is based on homogeneous samples. People present with different combinations of substance use behaviors and mental health conditions due to biological, social, and environmental reasons.[2,12,16,62–66] Further, much of the comorbidity research so far has been conducted in clinical samples rather than population-based samples. Therefore, a robust set of substance use behaviors and mental disorder symptoms in a large sample of nationally representative adults are required to close this knowledge gap and appropriately characterize comorbidity. This assessment of comorbidity in a larger sample of U.S. adults will shed light on the comorbidity profiles that exist in the general population, expanding the current literature of clinical samples. Furthermore, factors that influence these associations must also be considered. Appropriate characterizations are important to target and personalize treatment and result in greater success in prognosis for people experiencing comorbidity.

Although SUD comorbidity is persistent across the life course, it is unclear whether patterns of comorbidity remain stable or change over time. Some studies report that comorbidity does not readily change, while others explain shifts in substance use and mental health conditions.[67–69] Consequently, assessment of comorbidity over time is needed to better understand the stability and/or continuity of comorbidity, and what factors may be associated with these trends. These studies will help to better understand the progression or regression of symptoms or behaviors in adults, and identify how to better support individuals experiencing comorbidity.

The goal of the dissertation

The research goal of this dissertation is to address the aforementioned knowledge gaps (current SUD research suffers from the unidimensional approach that does not account for comorbidity; patterns of comorbidity are not the same, although current knowledge is based on homogeneous samples; and it is unclear whether patterns of comorbidity remain stable or changes over time) by characterizing the comorbidity between substance use, including tobacco use, and mental disorder symptoms measured as negative affect and externalizing symptoms in a population-based sample. This is characterized in Table 1.1. These characterizations are needed to better support people experiencing substance use and mental disorder comorbidity.

Table 1.1: Assumptions, strengths, and limitations of conceptual models				
	Assumptions	Strengths	Limitations	Addressed in Dissertation
Common cause model	Disease has a common pathogenic pathway or an etiology in which the mechanism is fully understood	Used by DSM, ICD-10-CM	Etiology for SUD/mental disorders not fully understood Does not account for comorbidity	All chapters address the limitation of the common cause model by accounting for comorbidity
Internalizing-externalizing model	Internalizing disorders represent a latent internalizing construct with externalizing disorders, including SUD represent a latent externalizing construct Internalizing and externalizing factors can also be correlated	Extension of common cause model Accounts for high level comorbidity between internalizing and externalizing disorders	Rarely accounts for substance use behaviors Does not describe the unobserved heterogeneity in the population to identify different comorbidity patterns in a population	Chapter 2 and 5 (Aims 1 and 3)
Network model	Comorbidities arise from shared symptoms between disorders which can capture complexity and individual variation in psychopathology	Can use symptom level data Naturally accommodates comorbidities as a central part of its theory	New methodological approach Does not follow DSM or ICS-10-CM approach to diagnosis	Chapter 4 and 5 (Aims 2 and 3)

Chapter 2 is a preliminary assessment of comorbidity. Multinomial logistic regression analyses will be used to determine the association between lifetime negative affect severity, externalizing severity and nicotine dependence (ND) and current use of tobacco (cigarettes and e-cigarettes) and alcohol, adjusting for sex, age, race, education and income. Two adjusted multinomial regression models are considered. The first model only includes negative affect and externalizing severity, adjusting for the correlation between the two factors. The second model builds on the first model by including ND to determine if ND explained more of the association between substance use.

Chapter 3 addresses the knowledge gap of diagnostic classifications, comorbidity profiles, and addresses factors associated with the comorbidity profiles by (1) estimating latent classes of comorbid substance use as well as negative affect and externalizing symptoms and (2) assessing their ability to predict SUD severity. The latent class approach is a type of mixture modeling used to identify unobserved heterogeneity in a population and find meaningful groups of people that are similar based on their responses to measured items.[70] This approach follows the common cause model in that the measured items (i.e., substance use behaviors and mental disorder symptoms) give rise to the latent unobservable disorder or in this case, comorbidity, and extends the internalizing-externalizing model by accounting for one overall latent class. Latent class analysis goes beyond the multinomial regression and allows for the consideration of multiple substance use and mental disorder symptom combinations. Analyses related to this chapter also move beyond the descriptiveness of the latent class approach and use the predictive probabilities generated from the latent class analysis to predict a health outcome, SUD severity. The prediction analyses allow us to determine whether comorbidity versus a single construct (i.e., substance use, negative affect or externalizing separately) is important in predicting a distal health outcome.

Chapter 4 builds on the knowledge developed in Chapter 3 by detailing substance use, negative affect, and externalizing symptom networks. Network analysis is a complement to the latent class approach. Network analysis does not follow the common cause model yet it posits that the substance use behaviors and mental disorder symptoms mutually interact and comorbidities arise from shared symptoms between disorders which can capture complexity and individual variation in psychopathology. The network approach allows for us to determine if symptoms cross over diagnostic boundaries. We also extend past the network approach and test whether there are differences in the network structure and connectivity by gender.

Chapter 5 addresses the knowledge gap of stability or movement of comorbidity structures by assessing the structure over time using the results from the latent class and network analyses. Pairwise comparisons occur in two separate approaches. First, using results from latent class analyses of three waves of adult data, we (1) compare the class probabilities across the waves, (2) assess the item response patterns for each class by wave, and (3) identify the transition patterns to determine the stability or movement among the classes. Second, network comparison tests including global strength, network structure, and edge strength will be tested to determine if there are differences in the comorbidity networks by wave.

Setting and measures

All chapters use adult data from the Population Assessment of Tobacco and Health (PATH) Study.[71] Chapters 2, 3, and 4 use data from the first wave of participants. Chapter 5 uses data from waves 1, 2, and 3 to assess the comorbidity patterns in adults across time. Information regarding the sample size, dates of data collection, and the weighted response rate among participants is provided in Table 1.2.

Table 1.2: PATH Information for Waves 1, 2, and 3

	Sample Size	Data Collection	Weighted Response Rate*
Wave 1	32,320	September 2013 – December 2014	74.0%
Wave 2	28,362	October 2014 – October 2015	83.2%
Wave 3	28,148	October 2015 – October 2016	78.4%

*Weighted response rate among participants is conditional on Wave 1 participation.

The PATH study launched in 2011 in response to the Family Smoking Prevention and Tobacco Control Act in order to inform the Food and Drug Administration's regulatory activities.[71] This study is a collaborative effort among the National Institute on Drug Abuse, National Institutes of Health, and the Center for Tobacco Products at the Food and Drug Administration. There are eight primary objectives for the PATH study:

1. "Identify and explain between-person differences and within-person changes in tobacco-use patterns, including the rate and length of use by specific product type and brand, product/brand switching over time, uptake of new products, and dual- and poly-use of tobacco products (i.e., use of multiple products within the same time period and switching between multiple products).
2. Identify between-person differences and within-person changes in risk perceptions regarding harmful and potentially harmful constituents, new and emerging tobacco products, filters and other design features of tobacco products,

packaging, and labeling; and identify other factors that may affect use, such as social influences and individual preferences.

3. Characterize the natural history of tobacco dependence, cessation, and relapse, including readiness and self-efficacy to quit, motivations for quitting, the number and length of quit attempts, and the length of abstinence related to various tobacco products.
4. Update the comprehensive baseline and subsequent waves of data on tobacco-use behaviors and related health conditions, including markers of exposure and tobacco-related disease processes identified from the collection and analysis of biospecimens, to assess between-person differences and within-person changes over time in health conditions potentially related to tobacco use, particularly with use of new and different tobacco products, including modified-risk tobacco products.
5. Assess associations between TCA-specific actions and tobacco-product use, risk perceptions and attitudes, use patterns, cessation outcomes, and tobacco-related intermediate endpoints (e.g., biomarkers of exposure and biomarkers related to disease). Analyses will attempt to account for other potential factors, such as demographics, local tobacco-control policies, and social, familial, and economic factors, that may influence the observed patterns.
6. Assess between-person differences and within-person changes over time in attitudes, behaviors, exposure to tobacco products, and related biomarkers among and within population sub-groups identified by such characteristics as

race-ethnicity, gender, and/or age, or by risk factors, such as pregnancy or co-occurring substance use or mental health disorders.

7. To the extent to which sample sizes are sufficient, assess and compare samples of former and never users of tobacco products for between-person differences and within-person changes in relapse and uptake, risk perceptions, and indicators of tobacco exposure and disease processes.
8. Use data from the PATH Study's baseline and follow-up waves on tobacco-use behaviors, attitudes, and related health conditions, including potential markers of exposure and related disease processes identified from the analysis of biospecimens, to screen and subsample respondents for participation in formative and/or nested studies conducted during and after the PATH Study's waves of data and biospecimen collection"[71]

PATH is a nationally representative longitudinal cohort study of the civilian, non-institutionalized household population of the U.S. aged 9 and older at Wave 1, and participants engaged in all levels of tobacco use ranging from never using tobacco to frequent use. Participants were selected through a four-stage stratified area probability sample design, with a two-phase design for sampling adults at the final stage[71]:

1. Selected stratified sample of geographical primary sampling units (PSU) (i.e., county or group of counties)
2. Within each PSU, smaller geographical segments were formed and a sample of these segments was drawn
3. Sampling frame consisted of residential addresses located in the segments

4. Selected adults and youth from the sampled households identified at these addresses (with varying sampling rates of adults by age, race, and tobacco status).
    a. Adults were sampled in two phases:
        i. Sampling used information provided in the household screener
        ii. Sampling used information provided by the adult

Study domains include tobacco use behavior, attitudes and beliefs, and tobacco-related health outcomes. Specific topics are included in Table 1.3.

Table 1.3: Topics assessed in PATH Study

Tobacco Products	Measures/Topics Associated with Tobacco Products	Additional Topics
Cigarettes E-cigarettes/Electronic nicotine products Cigars (traditional, cigarillos, filtered) Pipe tobacco Hookah Smokeless tobacco (snus pouches and other forms of smokeless tobacco) Dissolvable tobacco Bidis and kreteks (youth only)	Ever use Recency of use Frequency of use Amount of use Brands used Purchase details Use of flavored products Harm and addictiveness Reasons for use	Poly use Nicotine dependence Packaging and health warnings Risk and harm perceptions Secondhand smoke exposure Marketing and advertising Media use Demographics Health Psychosocial and mental health Substance use Cessation Peer and family influences

Participants responded to tobacco-specific items including tobacco-use patterns, risk perceptions and attitudes towards current and newly emerging tobacco products, tobacco initiation, cessation, relapse behaviors, and health outcomes.[72] Participants also responded to non-tobacco items (e.g., media use, peer and family influences, health effect outcomes, and industry advertising and promotion).[72] A detailed list of the measures used in the dissertation are provided below.

Past Month Tobacco and Substance Use. Six substance use categories were used in this dissertation: exclusive cigarette, exclusive e-cigarette, dual cigarette and e-cigarette, alcohol, marijuana, and prescription drugs not prescribed (PDNP) including painkillers, sedatives, tranquilizers. Current cigarette use was endorsed if the respondent indicated ever smoking a cigarette (even one or two puffs), has smoked at least 100 or more cigarettes in his or her entire life, and now smokes cigarettes every day or some days, while also excluding the current use of e-cigarettes. Current e-cigarette use was endorsed if the respondent indicated ever using an e-cigarette (even one or two puffs), ever smoked e-cigarettes fairly regularly, and now uses e-cigarettes every day or some days, while also excluding the current use of cigarettes. Current dual cigarette and e-cigarette use was identified if the respondent indicated that they were a current cigarette and current e-cigarette user. Current alcohol, marijuana, and PDNP was endorsed if the respondent indicated ever using the substance and has used the substance within the past 30 days. Only past month or current use of the substances was considered (coded as 1, else = 0) to reduce the potential for recall bias and ensure for accurate overlap with negative affect and externalizing symptoms occurring in the same time frame.

Negative Affect and Externalizing Severity and Symptoms. Negative affect and externalizing symptoms were measured using the Global Appraisal of Individual Needs—Short Screener (GAIN-SS).[73] The GAIN-SS refers to negative affect as “internalizing symptoms”; however, these symptoms provided within the GAIN-SS are better represented as “negative affect”. Negative affect refers to the experience of negative drive states such as depression, anxiety, and stress,[74] and therefore, is a more appropriate term for these symptoms compared to “internalizing”. Negative affect will be the term used for the rest of the dissertation.

The GAIN-SS is derived from the full GAIN instrument and identifies individuals at risk for mental health disorders using a continuous measure of severity.[73] The full GAIN assessment is a validated, standardized biopsychosocial assessment and recommended for use in epidemiologic samples.[28,73] Four questions were used to measure negative affect symptoms that asked the last time you had significant problems with:

(1) “feeling trapped, lonely, sad, blue, depressed, or hopeless about the future,”

(2) “sleep trouble- such as bad dreams, sleeping restlessly or falling asleep during the day,”

(3) “feeling very anxious, nervous, tense, scared, panicked or something bad was going to happen,” and

(4) “becoming very distressed and upset when something reminded you of the past.”

Externalizing symptoms were also measured using the GAIN-SS.[28,73] Seven questions were used to assess externalizing symptoms. Items asked the last time you did the following two or more times:

(1) “lied or conned to get things you wanted or to avoid having to do something,”

(2) “had a hard time paying attention at school, work or home,”

(3) “had a hard time listening to instructions at school, work or home,”

(4) “were a bully or threatened other people,”

(5) “started physical fights with other people,”

(6) “felt restless or the need to run around or climb on things” and

(7) “gave answers before the other person finished asking the question.”

The items selected to identify negative affect and externalizing symptoms from the GAIN-SS instrument are ordinal and measures people across four times periods: past month, 2 to 12 months, over a year ago, and never.

Ethical considerations

This research uses publicly available secondary data, where information is recorded by the investigator in a manner that subjects cannot be identified (either directly or through identifiers). Therefore, this is considered exempt by the Virginia Commonwealth University School of Medicine Institutional Review Board.

CHAPTER 2: THE ASSOCIATION BETWEEN NEGATIVE AFFECT AND EXTERNALIZING SEVERITY WITH CURRENT USE OF CIGARETTES, E-CIGARETTES, AND ALCOHOL IN ADULTS: WAVE 1 OF THE POPULATION ASSESSMENT OF TOBACCO AND HEALTH (PATH) STUDY[1]

INTRODUCTION

Tobacco and alcohol are two of the most common substances used in the United States (U.S.).[75,76] In 2018, approximately 20.9% of U.S. adults were current conventional cigarette (CIG) smokers and 55.3% reported drinking alcohol in the past month.[77–79] Among individuals with alcohol use disorder, 23.8% also had nicotine dependence and 12.9% of individuals with nicotine dependence also had alcohol use disorder.[9] Concurrent use of CIG and alcohol represents a major public health concern because they have been associated with more negative health outcomes such as increased risk of cardiovascular disease, cirrhosis, head and neck cancers, liver cancer, pancreatitis, and psychiatric comorbidity than the exclusive use of either substance.[80–82] To date, it is unclear whether the factors associated with co-occurring tobacco and alcohol use are specific to CIG or extend to electronic cigarettes (ECIG).

Although dual use of ECIG and CIG is common and increasing in the U.S.,[83] the trends related to this form of tobacco use with alcohol remain unclear. In 2018, 57.3% and 25.2% of former CIG users were engaged in ever-use and current-use of ECIGs, respectively.[84] Approximately 9.7% of current ECIG users also engaged in CIG use.[84] In 2014, about 16% of current smokers were also current ECIG users.[85] Recent studies have reported that current ECIG users are at an increased risk of harmful alcohol use

[1] This chapter has been modified from the original manuscript accepted for publication in Addictive Behaviors: https://doi.org/10.1016/j.addbeh.2021.106890

compared to ECIG non-users,[86,87] with dual CIG and ECIG use resulting in more past-month total drinks compared to exclusive-ECIG users.[88] However, compared to studies of CIG use and alcohol, there is far less knowledge regarding the co-occurring use of ECIG and alcohol. Consequently, there is a need to examine the use of ECIG, CIG, and alcohol, which may be associated with more severe or different risk factors than dual or exclusive use of any of these three substances.

Negative affect (e.g., depression and anxiety) and externalizing [e.g., attention-deficit hyperactivity disorder (ADHD) and conduct disorder] psychopathology[2,16,73,89–91] are important mental health factors that have been consistently associated with exclusive use of either CIG or alcohol. A meta-analysis reported that current CIG smokers had a two-fold increased risk of depression relative to never and former CIG users.[92] Further, adults with depression are more likely to smoke and are less likely to be successful at quitting than adults without depression.[93] Whether this bidirectional association is maintained among ECIG users is unclear. The relationship between the use of alcohol, CIG, and ECIG, and negative affect and externalizing psychopathology is currently undetermined. Prior studies of the relationship between psychopathology and tobacco products, specifically ECIG, as well as alcohol typically focus on youth and young adults. These results indicate ECIGs are commonly used with other substances (i.e., CIG, alcohol, marijuana and opiates) and associated with mental health symptomatology (i.e., diagnosis of ADHD, PTSD, anxiety, and substance use disorders).[94–97] However, it is unclear if these associations are specific to youth and young adults, or if they also occur across adulthood.

This study addresses the aforementioned knowledge gaps by examining the association of lifetime mental disorder symptom severity and past 30-day combinations of CIG, ECIG, and alcohol use. We asked the following questions: (1) is there an association between negative affect/externalizing severity across combinations of CIG, ECIG, and alcohol use in US adults, and (2) is there a difference in severity based on tobacco product type (CIG vs. ECIG)? We expect (1) a significant, positive association between negative affect/externalizing severity across all combinations of CIG, ECIG, and alcohol use. For exploratory aim (2), we expect that this association varies with type and number of tobacco products used (i.e., CIG associated with negative affect; ECIG associated with externalizing/negative affect; CIG + ECIG associated with negative affect/externalizing).

METHODS

Study material and participants

Data from 32,320 adults aged 18 years and older participating in the first wave (2013-2014) of the Population Assessment of Tobacco and Health (PATH) study were used.[71] PATH is a nationally representative longitudinal cohort study of the civilian, non-institutionalized adult household population of the U.S., and participants engaged in all levels of tobacco use.[72] The household screener response rate was 54%.[71] The weighted response rate among participants was 74%.[73]

Study representativeness

Participants with missing data on tobacco and alcohol measures, mental health symptoms, or covariates were not included in the analysis (N=16,373). Survey respondents of the analytic sample endorsed greater substance use overall, negative affect/externalizing severity, and nicotine dependence (ND) than those not included in the analytic sample. The participants in the analytic sample were more likely to be men, aged 25-54 with lower levels of education and lower annual household income than those who were missing.

Measures

Current tobacco and alcohol use. Current tobacco and alcohol use was measured as an aggregate variable indicating the degree of past-month use of CIG, ECIG, and alcohol, and was developed from individual current-use items defined according to the National Health Interview Survey (2017) and listed in Table 2.1.[98]

Table 2.1: Individual criteria used to define current-use of alcohol, cigarettes, or e-cigarettes.

Code	Current Alcohol Use	Current Cigarette (CC) Use	Current E-Cigarette (EC) Use
1	• Ever used alcohol in past 30 days	• Ever smoked a CC (even 1-2 puffs) • Smoked ≥ 100 CC in lifetime • Smoke CC every day or some days	• Ever used an EC (even 1-2 times) • Ever smoked EC fairly regularly • Now use EC every day or some days
0	• Never used alcohol • Ever used alcohol but not in past 30 days	• Never smoked a CC • Smoked ≤ 99 CC in lifetime • Do not smoke CC now	• Never used an EC • Do not smoke EC regularly • Do not use EC now

Depending on responses, subjects were classified as current users (coded as 1), or non-current users (coded as 0).

The outcome variable was developed as an eight-level categorical variable: (1) alcohol-exclusive; (2) CIG-exclusive; (3) ECIG-exclusive; (4) CIG and alcohol; (5) ECIG and alcohol; (6) CIG and ECIG; (7) alcohol, CIG, and ECIG; and (8) non-use. This variable allowed us to evaluate the relationships between all combinations of alcohol, CIG, and ECIG use and negative affect/externalizing severity, with non-users as a reference group.

Negative affect/ externalizing severity. Negative affect and externalizing severity were measured in PATH using the Global Appraisal of Individual Needs—Short Screener (GAIN-SS).[73] The GAIN-SS is derived from the full GAIN instrument assessing individuals at risk for mental disorders using a continuous measure of severity. The full GAIN assessment is a reliable and validated biopsychosocial assessment recommended for use in epidemiologic samples.[28,73,99] There was good internal consistency among the negative affect (Cronbach's α=0.85) and externalizing (Cronbach's α=0.80) items in the analytic sample.

Items used to measure negative affect/externalizing symptoms are listed in Table 2.2. Responses were measured across four time periods: past month, 2-12 months, over a year ago, and never. Lifetime negative affect/externalizing items were coded as past month, 2-12 months, or over a year ago = 1 vs never = 0. The binary responses were summed to reflect a scale score of the number of lifetime symptoms. The score ranging from 0-4 negative affect symptoms and 0-7 externalizing symptoms were categorized into low (0), moderate (1-2), and high (3+) severity). These cut points were previously recommended based on validation analyses of the dimensional measures and have high predictive validity in other samples.[28,73,99] Higher scores indicate increased severity, a greater likelihood for diagnosis with a mental health disorder, and increased need for services.[73] Negative affect/externalizing severity were highly correlated with one another (r=0.68, ASE=0.0051, p<0.001).

Table 2.2 Items used to measure negative affect, externalizing, and nicotine dependence.	
Negative affect Symptoms*	Last time respondent had significant problem with: • feeling trapped, lonely, sad, blue, depressed, or hopeless about the future • sleep trouble - such as bad dreams, sleeping restlessly or falling asleep during the day • feeling very anxious, nervous, tense, scared, panicked, or that something bad was going to happen • becoming very distressed and upset when something reminded you of the past
Externalizing Symptoms*	Last time respondent engaged in the following behaviors 2-3 times: • lied or conned to get things you wanted or to avoid having to do something • had a hard time paying attention at school, work or home • had a hard time listening to instructions at school, work, or home • were a bully or threatened other people • started physical fights with other people • felt restless or the need to run around or climb on things • gave answers before the other person finished asking the question
Nicotine Dependence ǂ	WISDM: Primary • I find myself reaching for [product] without thinking about. • I frequently crave [product]. • My urges keep getting stronger if I don't use [product]. • Tobacco products control me. • My [product] use is out of control. • I usually want to use [product] right after I wake up. • I can only go a couple of hours without using [product]. • I frequently find myself almost using [product] without thinking about it. WISDM: Secondary • Using [product] would really help me feel better if I've been feeling down. • Using [product] helps me think better. • I [would] feel alone without my [product]. NDSS • I would find it really hard to stop using [product]. • I would find it hard to stop using [product]. • After not using [product] for a while, I need to use [product] in order to feel less restless and irritable. • After not using [product] for a while, I need to use [product] in order to keep myself from experiencing discomfort. DSM: Impaired Control • In the past 12 months, did you find it difficult to keep from using [product] in places where it was prohibited?

*Responses were measured across four time periods: past month, 2-12 months, over a year ago, and never. Participants indicating that they experienced a symptom at any time were coded as 1. Participants indicating that they never experienced an item were coded as 0. Binary responses were summed to reflect a scale score with a range of 0-4 symptoms for negative affect and 0-7 symptoms for externalizing. The scores were categorized into low (0), moderate (1-2), and high (3+) severity based on the recommended cut points.
ǂ Responses for WIDSM: Primary, WIDSM: Secondary, and NDSS were measured based on level of agreement from 1 = not true of me at all to 5 = extremely true to me. Response option for DSM: Impaired Control was 1 = Yes and 0 = No. These were summed to reflect a scale score with a range of 0-76 with higher values indicating greater ND.

Covariates. The role of nicotine dependence (ND) was included as a potential confounder. Adults with mental health disorders may have higher levels of ND as a result of tobacco product use.[100,101] Similarly, there is a strong association between ND and all levels of alcohol use.[102] People who engage in ECIG and CIG dual use have greater ND than exclusive use of either ECIG or CIG.[103,104] Sixteen items [8 from Wisconsin Inventory of Smoking Dependence Motives (WIDSM): Primary, 3 from WISDM: Secondary, 4 from Nicotine Dependence Syndrome Scale (NDSS), 1 from Diagnostic and Statistical Manual of Mental Disorders (DSM): Impaired Control] were used to measure ND and are listed in Table 2.2. These 16 items were recommended to use as a common instrument to assess ND across different tobacco product users from a differential item function analysis.[27] The items were summed into one continuous variable ranging from 0-76, with higher values indicating greater ND.

Sex, age, race/ethnicity, education, and annual household income were also included as covariates because they are consistently associated with mental health, and tobacco and alcohol use.[61,73,105–114]

Age, measured in PATH as a seven-level categorical variable, was re-categorized to have a uniform distribution with six levels (18-24, 25-34, 35-44, 45-54, 55-64, and 65 years or older). Education, measured in PATH as a six-level categorical variable, was re-categorized as a five-level categorical variable with a uniform distribution [less than high school, GED/high school graduate, some college (no degree) or Associate's degree, Bachelor's degree, and Advanced degree]. Race/ethnicity was measured as a four-level categorical race variable and included information from a separate variable that accounted for Hispanic ethnicity (Non-Hispanic White, Non-Hispanic Black, Non-Hispanic Other, and Hispanic Multicultural). The significance of the association between these variables and tobacco and alcohol use was tested as a series of unadjusted multinomial logistic regressions (Table 2.4).

Statistical analyses

Unadjusted multinomial logistic regression was used to test the association between tobacco and alcohol use and negative affect/externalizing severity. Tests were repeated after adjustment for sex, age, race, education, and annual household income. Two adjusted multinomial regression models were considered: the first model included only negative affect/externalizing severity, adjusting for the correlation between the two factors, while the second model also included ND to determine the degree to which ND explained the association between mental health severity and substance use. Odds ratios (OR) or adjusted odds ratios (AOR) and 95% confidence intervals (95% CI), profiled from estimates of standard error, are reported. All analyses were performed in SAS software, Version 9.4 (SAS Institute Inc, Cary, NC) and accounted for complex survey design and sampling weight using PROC SURVEYFREQ and PROC SURVEYLOGISTIC. Fay's method, a variant of balanced repeated replication method, was used to form replication weights in variance estimation in all analyses.

RESULTS

Descriptive statistics

Data from 15,947 participants with complete information were analyzed. Almost one quarter of the population engaged in alcohol-exclusive use (24.0%), 22.4% in CIG-exclusive use, and 1.3% in ECIG-exclusive use (Table 2.3). Across the different combinations of tobacco and alcohol use, 33.3% engaged in CIG and alcohol use, 1.7% engaged in ECIG and alcohol use, 2.0% engaged in CIG and ECIG, and 3.2% engaged in alcohol, CIG, and ECIG use. Almost half of the sample endorsed high negative affect (47.9%) and high externalizing (44.7%) severity. The mean ND was 37.0 (range=1-76, standard deviation=0.23) for the sample (Table 2.3).

Table 2.3: Overall Frequencies of the Analytic Sample (n = 15,947, Weighted N = 61,482,491)—PATH Wave 1 (2013-2014)

	n (Weighted %)		n (Weighted %)
Sex*		**Negative Affect Severity***	
Male	9039 (59.6)	Low	4310 (28.1)
Female	6908 (40.4)	Moderate	3731 (24.0)
Age*		High	7906 (47.9)
18-24 years old	4304 (17.7)	**Externalizing Severity***	
25-34 years old	3580 (24.3)	Low	4058 (26.8)
35-44 years old	2696 (18.3)	Moderate	4436 (28.5)
45-54 years old	2579 (18.5)	High	7453 (44.7)
55-64 years old	1871 (14.1)	**Nicotine Dependence**	37.0 (0.23)[a]
65 years or older	917 (7.1)		
Race*		**Tobacco and Alcohol Use***	
Non-Hispanic White	10257 (68.2)	Alcohol only	3603 (24.0)
Non-Hispanic Black	2305 (13.8)	CC only	3678 (22.4)
Non-Hispanic Other	1218 (6.4)	EC only	219 (1.3)
Hispanic Multiracial	2167 (11.7)	CC and Alcohol	5387 (33.3)
Education*		EC and Alcohol	288 (1.7)
Less than high school	2304 (13.4)	CC and EC	336 (2.0)
GED/High school graduate	5385 (35.5)	Alcohol, CC, and EC	558 (3.2)
Some college (no degree) or Associate's degree	5931 (34.9)	None	1878 (12.2)
Bachelor's degree	1685 (11.9)		
Advanced degree	642 (4.3)		
Annual Household Income*			
Less than $10,000	3532 (19.5)		
$10,000 to $24,999	4120 (24.8)		
$25,000 to $49,999	3746 (24.2)		
$50,000 to $99,999	2974 (20.2)		
$100,000 or more	1575 (11.4)		

* Indicates a significant difference at $p < 0.05$.

[a] Indicates mean and standard deviation (95% CL for the mean = 36.6-37.5)

Unadjusted multinomial logistic regression analysis

Compared to subjects with low negative affect severity, those with high negative affect severity were significantly more likely to engage in alcohol, CIG, and ECIG use (OR=3.42, 95% CI=2.48-4.72), CIG and ECIG use (OR=2.24, 95% CI=1.63-3.08), ECIG and alcohol use (OR=2.20, 95% CI=1.57-3.09), CIG and alcohol use (OR=2.28, 95% CI=1.97-2.65), CIG-exclusive use (OR=1.69, 95% CI=1.42-2.02), and alcohol-exclusive use (OR=1.42, 95% CI=1.20-1.69) than no use. Relative to those with low externalizing severity, subjects with high externalizing severity were more likely than not to engage in every level of tobacco and alcohol use except ECIG use, especially alcohol, CIG, and ECIG use (OR=4.56, 95% CI=3.31-6.30) and ECIG and alcohol use (OR=4.23, 95% CI=2.84-6.29). There were significant, positive associations between ND and alcohol, CIG, and ECIG use (OR=1.05, 95% CI=1.05-1.06), CIG and ECIG use (OR=1.08, 95% CI=1.07-1.08), CIG and alcohol use (OR=1.05, 95% CI=1.04-1.05), and CIG-exclusive use (OR=1.06, 95% CI=1.06-1.07). Females, relative to males, had significantly increased odds for CIG and ECIG use (OR=1.74, 95% CI=1.35-2.25), CIG and alcohol use (OR=1.21, 95% CI=1.06-1.38), ECIG-exclusive use (OR=1.99, 95% CI=1.45-2.74), and CIG-exclusive use (OR=1.65, 95% CI=1.43-1.90), except for alcohol-exclusive use (OR=0.71, 95% CI=0.61-0.84). There were significant associations by age, race, education, and annual household income (Table 2.4).

Table 2.4: Summary of Unadjusted Bivariate Associations by Level of Current Tobacco and Alcohol Use (n = 15,947, Weighted N = 61,482,491)

Variable	Alcohol, Cigarette, and E-cigarette OR (95% CI)	Cigarette and E-cigarette OR (95% CI)	E-cigarette and Alcohol OR (95% CI)	Cigarette and Alcohol OR (95% CI)	E-cigarette Only OR (95% CI)	Cigarette Only OR (95% CI)	Alcohol Only OR (95% CI)
Negative Affect Severity							
Low	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Moderate	**1.71 (1.14-2.56)**	1.50 (0.96-2.36)	1.36 (0.89-2.08)	**1.79 (1.52-2.11)**	1.12 (0.71-1.77)	**1.37 (1.10-1.71)**	**1.66 (1.36-2.01)**
High	**3.42 (2.48-4.72)**	**2.24 (1.63-3.08)**	**2.20 (1.57-3.09)**	**2.28 (1.97-2.65)**	1.30 (0.89-1.88)	**1.69 (1.42-2.02)**	**1.42 (1.20-1.69)**
Externalizing Severity							
Low	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Moderate	**2.32 (1.59-3.39)**	**1.67 (1.15-2.42)**	**2.84 (1.91-4.23)**	**2.00 (1.68-2.38)**	1.01 (0.67-1.52)	**1.21 (1.01-1.43)**	**2.04 (1.70-2.44)**
High	**4.56 (3.31-6.30)**	**2.22 (1.61-3.05)**	**4.23 (2.84-6.29)**	**2.58 (2.19-3.03)**	1.31 (0.89-1.94)	**1.31 (1.11-1.55)**	**2.29 (1.91-2.76)**
Nicotine Dependence							
	1.05 (1.05-1.06)	**1.08 (1.07-1.08)**	1.00 (0.99-1.01)	**1.05 (1.04-1.05)**	1.00 (0.99-1.01)	**1.06 (1.06-1.07)**	**0.97 (0.96-0.97)**
Sex							
Male	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Female	1.24 (0.98-1.56)	**1.74 (1.35-2.25)**	1.26 (0.94-1.69)	**1.21 (1.06-1.38)**	**1.99 (1.45-2.74)**	**1.65 (1.43-1.90)**	**0.71 (0.61-0.84)**
Age							
18-24 years old	Reference	Reference	Reference	Reference	Reference	Reference	Reference
25-34 years old	**1.75 (1.37-2.25)**	**3.62 (2.53-5.20)**	**1.69 (1.21-2.35)**	**2.42 (2.02-2.89)**	**1.86 (1.08-3.20)**	**2.53 (2.04-3.12)**	**1.40 (1.17-1.67)**
35-44 years old	1.28 (0.87-1.88)	**2.93 (2.00-4.31)**	0.84 (0.55-1.30)	**1.95 (1.57-2.42)**	1.49 (0.89-2.51)	**2.27 (1.81-2.86)**	0.90 (0.72-1.12)
45-54 years old	**0.57 (0.41-0.79)**	**1.78 (1.15-2.75)**	0.69 (0.44-1.10)	**1.55 (1.28-1.87)**	**2.01 (1.26-3.21)**	**2.25 (1.89-2.68)**	**0.76 (0.63-0.93)**
55-64 years old	**0.46 (0.32-0.66)**	**1.99 (1.31-3.02)**	0.64 (0.38-1.07)	1.21 (0.98-1.49)	1.04 (0.60-1.80)	**2.02 (1.64-2.49)**	**0.63 (0.48-0.83)**
65 years or older	**0.16 (0.07-0.35)**	1.42 (0.70-2.88)	0.31 (0.14-0.66)	**0.60 (0.47-0.76)**	1.03 (0.51-2.11)	**1.90 (1.47-2.47)**	**0.47 (0.36-0.61)**
Race							
White	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Non-Hispanic Black	**0.23 (0.16-0.35)**	**0.35 (0.23-0.55)**	**0.45 (0.27-0.73)**	**0.61 (0.51-0.73)**	**0.37 (0.20-0.67)**	**0.50 (0.42-0.60)**	**0.63 (0.51-0.77)**
Non-Hispanic Other	0.70 (0.44-1.12)	**0.60 (0.38-0.94)**	0.98 (0.56-1.72)	**0.72 (0.56-0.91)**	1.00 (0.57-1.76)	**0.66 (0.53-0.82)**	0.81 (0.62-1.06)
Hispanic Multicultural	**0.32 (0.23-0.46)**	**0.32 (0.20-0.52)**	**0.35 (0.22-0.57)**	**0.48 (0.40-0.57)**	**0.58 (0.40-0.86)**	**0.50 (0.41-0.61)**	**0.67 (0.53-0.83)**
Education							
Less than high school	Reference	Reference	Reference	Reference	Reference	Reference	Reference
GED/High school	**2.30 (1.61-3.28)**	1.15 (0.75-1.78)	**2.22 (1.16-4.26)**	**1.50 (1.24-1.81)**	1.51 (0.90-2.53)	0.96 (0.80-1.17)	**1.74 (1.39-2.17)**
Some college	**4.79 (3.28-6.99)**	**1.72 (1.16-2.56)**	**4.75 (2.60-8.67)**	**2.13 (1.78-2.56)**	**2.31 (1.46-3.64)**	0.96 (0.80-1.15)	**3.58 (2.89-4.43)**
Bachelor's degree	**4.91 (3.26-7.39)**	1.68 (0.99-2.86)	**5.04 (2.47-10.32)**	**2.34 (1.85-2.97)**	**1.86 (1.00-3.46)**	**0.65 (0.48-0.88)**	**7.16 (5.41-9.49)**
Advanced degree	**3.67 (1.98-6.82)**	0.69 (0.28-1.69)	2.32 (0.62-8.63)	**1.50 (1.07-2.10)**	1.85 (0.75-4.56)	**0.40 (0.27-0.59)**	**6.73 (4.80-9.44)**
Income							
< $10,000	Reference	Reference	Reference	Reference	Reference	Reference	Reference
$10,000-24,999	**1.94 (1.42-2.64)**	**1.49 (1.05-2.12)**	1.31 (0.82-2.09)	**1.33 (1.09-1.62)**	1.07 (0.62-1.83)	**1.26 (1.03-1.54)**	**1.30 (1.07-1.59)**
$25,000-49,000	**2.59 (1.85-3.62)**	1.41 (0.98-2.02)	**2.00 (1.41-2.83)**	**1.63 (1.36-1.97)**	1.21 (0.77-1.92)	1.08 (0.89-1.31)	**1.66 (1.35-2.05)**
$50,000-99,999	**2.95 (2.01-4.32)**	1.33 (0.86-2.05)	**3.03 (2.06-4.46)**	**1.98 (1.58-2.48)**	**1.82 (1.07-3.08)**	0.90 (0.71-1.14)	**3.04 (2.42-3.82)**
>=$100,000	**2.88 (1.83-4.53)**	1.06 (0.59-1.91)	**3.37 (2.16-5.24)**	**1.72 (1.38-2.14)**	1.33 (0.68-2.62)	**0.53 (0.38-0.72)**	**5.07 (3.91-6.57)**

Bolded values indicate estimate significant a $p < 0.05$

The “none” category is used in reference for the tobacco and alcohol use outcome.

Adjusted multinomial logistic regression analysis

Model 1: Negative affect/externalizing severity

Compared to subjects with low negative affect severity, those with high negative affect severity were significantly more likely to engage in alcohol, CIG, and ECIG use (AOR=2.01, 95% CI=1.30-3.09), CIG and alcohol use (AOR=1.61, 95% CI=1.30-2.00), and CIG-exclusive use (AOR=1.42, 95% CI=1.13-1.79) than none (Table 2.5). Participants with moderate negative affect severity, compared to low, were significantly more likely to engage in CIG and alcohol use (AOR=1.52, 95% CI=1.27-1.81), CIG-exclusive use (AOR=1.26, 95% CI=1.01-1.58), and alcohol-exclusive use (AOR=1.53, 95% CI=1.24-1.90) than none. Participants with high externalizing severity, compared to low, had 113% greater odds of alcohol, CIG, and ECIG use (AOR=2.13, 95% CI=1.36-3.34), 54% greater odds of CIG and ECIG use (AOR=1.54, 95% CI=1.04-2.28), 196% greater odds of ECIG and alcohol use (AOR=2.96, 95% CI=1.82-4.80), 74% greater odds of CIG and alcohol use (AOR=1.74, 95% CI=1.38-2.20), and 69% greater odds of alcohol-exclusive use (AOR=1.69, 95% CI=1.33-2.14) than no use. Participants with moderate externalizing severity, compared to low, were significantly more likely to engage in alcohol, CIG, and ECIG use (AOR=1.56, 95% CI=1.02-2.40), ECIG and alcohol use (AOR=2.32, 95% CI=1.55-3.46), CIG and alcohol use (AOR=1.54, 95% CI=1.26-1.88), alcohol-exclusive use (AOR=1.60, 95% CI=1.32-1.94) than no use.

Table 2.5: Model 1 - Multinomial Logistic Regression for Level of Current Tobacco and Alcohol Use (n = 15,947, Weighted N = 61,482,491)

Variable	Alcohol, Cigarette, and E-cigarette AOR (95% CI)	Cigarette and E-cigarette AOR (95% CI)	E-cigarette and Alcohol AOR (95% CI)	Cigarette and Alcohol AOR (95% CI)	E-cigarette Only AOR (95% CI)	Cigarette Only AOR (95% CI)	Alcohol Only AOR (95% CI)
Negative Affect Severity							
Low	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Moderate	1.37 (0.88-2.13)	1.23 (0.77-1.98)	0.98 (0.61-1.56)	**1.52 (1.27-1.81)**	1.03 (0.64-1.67)	**1.26 (1.01-1.58)**	**1.53 (1.24-1.90)**
High	**2.01 (1.30-3.09)**	1.46 (0.99-2.14)	1.20 (0.78-1.83)	**1.61 (1.30-2.00)**	1.00 (0.62-1.61)	**1.42 (1.13-1.79)**	1.21 (0.97-1.51)
Externalizing Severity							
Low	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Moderate	**1.56 (1.02-2.40)**	1.33 (0.90-1.97)	**2.32 (1.55-3.46)**	**1.54 (1.26-1.88)**	0.88 (0.57-1.37)	1.04 (0.86-1.25)	**1.60 (1.32-1.94)**
High	**2.13 (1.36-3.34)**	**1.54 (1.04-2.28)**	**2.96 (1.82-4.80)**	**1.74 (1.38-2.20)**	1.15 (0.69-1.93)	1.04 (0.82-1.32)	**1.69 (1.33-2.14)**
Sex							
Male	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Female	1.11 (0.87-1.40)	**1.57 (1.19-2.06)**	1.23 (0.91-1.66)	1.12 (0.98-1.29)	**2.01 (1.45-2.79)**	**1.52 (1.30-1.78)**	0.72 (0.61-0.84)
Age							
18-24 years old	Reference	Reference	Reference	Reference	Reference	Reference	Reference
25-34 years old	**1.72 (1.33-2.23)**	**3.95 (2.73-5.71)**	**1.70 (1.22-2.39)**	**2.44 (2.03-2.93)**	**1.97 (1.12-3.47)**	**2.74 (2.21-3.40)**	1.21 (0.99-1.47)
35-44 years old	1.22 (0.85-1.76)	**3.25 (2.17-4.87)**	0.80 (0.52-1.23)	**1.93 (1.55-2.42)**	1.54 (0.90-2.64)	**2.53 (2.00-3.20)**	**0.70 (0.55-0.89)**
45-54 years old	**0.57 (0.41-0.80)**	**1.90 (1.20-3.00)**	0.70 (0.43-1.12)	**1.54 (1.27-1.87)**	**2.08 (1.28-3.39)**	**2.32 (1.92-2.80)**	**0.65 (0.53-0.80)**
55-64 years old	**0.48 (0.33-0.70)**	**2.13 (1.39-3.28)**	0.68 (0.40-1.17)	1.23 (0.98-1.54)	1.11 (0.63-1.94)	**2.07 (1.65-2.60)**	**0.55 (0.41-0.73)**
65 years or older	**0.19 (0.09-0.40)**	1.58 (0.76-3.30)	**0.38 (0.17-0.85)**	**0.66 (0.50-0.88)**	1.13 (0.53-2.42)	**1.89 (1.41-2.54)**	**0.43 (0.32-0.59)**
Race							
White	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Non-Hispanic Black	**0.32 (0.21-0.48)**	**0.36 (0.23-0.57)**	0.64 (0.38-1.08)	**0.73 (0.60-0.88)**	**0.40 (0.22-0.73)**	**0.44 (0.37-0.53)**	1.00 (0.81-1.25)
Non-Hispanic Other	**0.59 (0.38-0.93)**	**0.60 (0.38-0.94)**	0.89 (0.51-1.56)	**0.69 (0.55-0.88)**	1.06 (0.62-1.81)	**0.72 (0.58-0.89)**	**0.64 (0.49-0.84)**
Hispanic Multicultural	**0.36 (0.25-0.52)**	**0.33 (0.20-0.54)**	**0.43 (0.26-0.73)**	**0.54 (0.44-0.66)**	**0.66 (0.44-0.99)**	**0.47 (0.38-0.59)**	0.90 (0.71-1.14)
Education							
Less than high school	Reference	Reference	Reference	Reference	Reference	Reference	Reference
GED/High school	**1.68 (1.15-2.44)**	0.98 (0.62-1.53)	1.63 (0.83-3.22)	**1.23 (1.01-1.50)**	1.32 (0.79-2.22)	0.90 (0.73-1.10)	**1.43 (1.12-1.82)**
Some college	**2.65 (1.77-3.97)**	1.32 (0.85-2.05)	**2.62 (1.39-4.97)**	**1.51 (1.23-1.86)**	**1.79 (1.12-2.86)**	0.89 (0.72-1.09)	**2.46 (1.94-3.13)**
Bachelor's degree	**2.48 (1.63-3.77)**	1.25 (0.69-2.28)	**2.25 (1.07-4.72)**	**1.50 (1.15-1.96)**	1.31 (0.72-2.40)	**0.63 (0.46-0.87)**	**4.13 (3.04-5.62)**
Advanced degree	**2.05 (1.07-3.90)**	0.53 (0.21-1.35)	1.06 (0.29-3.95)	0.97 (0.68-1.40)	1.29 (0.53-3.15)	**0.41 (0.28-0.59)**	**3.82 (2.59-5.63)**
Income							
< $10,000	Reference	Reference	Reference	Reference	Reference	Reference	Reference
$10,000-24,999	**1.78 (1.28-2.48)**	1.28 (0.88-1.85)	1.22 (0.76-1.97)	**1.25 (1.01-1.56)**	0.95 (0.56-1.61)	1.12 (0.90-1.40)	**1.26 (1.03-1.54)**
$25,000-49,000	**2.05 (1.44-2.93)**	1.05 (0.72-1.53)	**1.70 (1.17-2.48)**	**1.38 (1.12-1.71)**	0.98 (0.61-1.57)	0.92 (0.74-1.14)	**1.39 (1.12-1.73)**
$50,000-99,999	**2.17 (1.46-3.23)**	0.94 (0.59-1.48)	**2.47 (1.62-3.77)**	**1.61 (1.24-2.07)**	1.37 (0.81-2.32)	**0.76 (0.59-0.99)**	**2.22 (1.74-2.82)**
>=$100,000	**2.04 (1.26-3.29)**	0.80 (0.43-1.50)	**2.76 (1.72-4.41)**	**1.43 (1.11-1.85)**	1.03 (0.51-2.06)	**0.51 (0.36-0.72)**	**3.13 (2.33-4.22)**

Bolded values indicate estimate significant a $p < 0.05$
The “none” category is used in reference for the tobacco and alcohol use outcome.

Participants with high negative affect severity, compared to low, had the greatest odds for alcohol, CIG, and ECIG use rather than no use while adjusting for externalizing severity, sex, age, race, education, and annual household income. Participants with high externalizing severity, compared to low, had the greatest odds for ECIG and alcohol use rather than no use while adjusting for negative affect severity, sex, age, race, education, and annual household income.

Model 2: Negative affect, externalizing, and ND

Compared to subjects with low negative affect severity, those with high negative affect severity were more likely to engage in CIG and alcohol use (AOR=1.29, 95% CI=1.03-1.61) and alcohol-exclusive use (AOR=1.31, 95% CI=1.05-1.64) than no use (Table 2.6). Similar associations were found between moderate negative affect severity, relative to low, and CIG and alcohol use (AOR=1.47, 95% CI=1.22-1.77) and alcohol-exclusive use (AOR=1.58, 95% CI=1.27-1.96) than no use. Participants with high externalizing severity, compared to low, had 79% greater odds for alcohol, CIG, and ECIG use (AOR=1.79, 95% CI=1.15-2.78), 197% greater odds of ECIG and alcohol use (AOR=2.97, 95% CI=1.84-4.81), 53% greater odds of CIG and alcohol use (AOR=1.53, 95% CI=1.21-1.92), and 75% greater odds of alcohol-exclusive use (AOR=1.75, 95% CI=1.38-2.22) than no use. Subjects with moderate externalizing severity, compared to low, were more likely to engage in ECIG and alcohol use (AOR=2.29, 95% CI=1.53-3.43), CIG and alcohol use (AOR=1.41, 95% CI=1.16-1.72), and alcohol-exclusive use (AOR=1.62, 95% CI=1.33-1.97) than no use when adjusting for ND. ND was significantly associated with all combinations of tobacco and alcohol use, compared to none, except for ECIG and alcohol use (AOR=1.00, 95% CI=0.99-1.01) and ECIG-exclusive use (AOR=1.00, 95% CI=0.99-1.01).

Table 2.6: Model 2 - Multinomial Logistic Regression for Level of Current Tobacco and Alcohol Use (Including Nicotine Dependence) (n = 15,947, Weighted N = 61,482,491)

Variable	Alcohol, Cigarette, and E-cigarette AOR (95% CI)	Cigarette and E-cigarette AOR (95% CI)	E-cigarette and Alcohol AOR (95% CI)	Cigarette and Alcohol AOR (95% CI)	E-cigarette Only AOR (95% CI)	Cigarette Only AOR (95% CI)	Alcohol Only AOR (95% CI)
Negative Affect Severity							
Low	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Moderate	1.33 (0.85-2.07)	1.21 (0.74-1.97)	0.97 (0.61-1.56)	**1.47 (1.22-1.77)**	1.04 (0.64-1.68)	1.22 (0.96-1.55)	**1.58 (1.27-1.96)**
High	1.53 (1.00-2.36)	1.02 (0.68-1.53)	1.19 (0.78-1.81)	**1.29 (1.03-1.61)**	1.00 (0.62-1.60)	1.08 (0.85-1.38)	**1.31 (1.05-1.64)**
Externalizing Severity							
Low	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Moderate	1.42 (0.93-2.17)	1.16 (0.77-1.74)	**2.29 (1.53-3.43)**	**1.41 (1.16-1.72)**	0.87 (0.56-1.37)	0.92 (0.76-1.13)	**1.62 (1.33-1.97)**
High	**1.79 (1.15-2.78)**	1.23 (0.82-1.85)	**2.97 (1.84-4.81)**	**1.53 (1.21-1.92)**	1.16 (0.69-1.95)	0.88 (0.70-1.11)	**1.75 (1.38-2.22)**
Nicotine Dependence							
	1.06 (1.05-1.07)	**1.08 (1.07-1.09)**	1.00 (0.99-1.01)	**1.05 (1.04-1.05)**	1.00 (0.99-1.01)	**1.06 (1.05-1.06)**	**0.97 (0.97-0.98)**
Sex							
Male	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Female	1.17 (0.92-1.48)	**1.65 (1.26-2.16)**	1.22 (0.91-1.65)	**1.17 (1.01-1.35)**	**2.02 (1.46-2.80)**	**1.58 (1.34-1.86)**	**0.71 (0.60-0.83)**
Age							
18-24 years old	Reference	Reference	Reference	Reference	Reference	Reference	Reference
25-34 years old	**1.30 (1.01-1.68)**	**2.76 (1.92-3.97)**	**1.74 (1.23-2.46)**	**1.96 (1.63-2.37)**	**2.01 (1.12-3.60)**	**2.11 (1.69-2.63)**	**1.32 (1.09-1.61)**
35-44 years old	0.76 (0.53-1.09)	**1.74 (1.15-2.63)**	0.82 (0.53-1.29)	**1.34 (1.07-1.68)**	1.58 (0.90-2.76)	**1.62 (1.28-2.06)**	0.82 (0.64-1.04)
45-54 years old	**0.33 (0.24-0.45)**	0.92 (0.60-1.42)	0.72 (0.44-1.16)	1.00 (0.83-1.20)	**2.12 (1.31-3.43)**	**1.38 (1.13-1.68)**	**0.77 (0.62-0.95)**
55-64 years old	**0.28 (0.19-0.41)**	1.06 (0.67-1.68)	0.70 (0.41-1.20)	0.81 (0.64-1.03)	1.12 (0.64-1.98)	1.26 (0.99-1.60)	**0.65 (0.48-0.86)**
65 years or older	**0.12 (0.06-0.27)**	0.90 (0.43-1.92)	**0.39 (0.17-0.89)**	**0.48 (0.36-0.64)**	1.16 (0.53-2.53)	1.31 (0.96-1.77)	**0.51 (0.37-0.70)**
Race							
White	Reference	Reference	Reference	Reference	Reference	Reference	Reference
Non-Hispanic Black	**0.40 (0.26-0.61)**	**0.48 (0.30-0.77)**	0.64 (0.38-1.08)	0.87 (0.71-1.06)	**0.40 (0.22-0.73)**	**0.55 (0.45-0.67)**	0.95 (0.76-1.19)
Non-Hispanic Other	0.69 (0.44-1.09)	0.75 (0.48-1.18)	0.89 (0.51-1.57)	0.79 (0.62-1.00)	1.06 (0.62-1.83)	0.85 (0.67-1.08)	**0.63 (0.48-0.83)**
Hispanic Multicultural	**0.55 (0.38-0.81)**	**0.57 (0.35-0.94)**	**0.43 (0.25-0.73)**	**0.77 (0.62-0.95)**	**0.65 (0.43-1.00)**	**0.73 (0.58-0.91)**	0.79 (0.62-1.01)
Education							
Less than high school	Reference	Reference	Reference	Reference	Reference	Reference	Reference
GED/High school	**1.82 (1.24-2.66)**	1.10 (0.70-1.71)	1.60 (0.82-3.15)	**1.31 (1.07-1.60)**	1.31 (0.78-2.20)	0.97 (0.78-1.20)	**1.39 (1.09-1.78)**
Some college	**3.41 (2.26-5.14)**	**1.77 (1.13-2.76)**	**2.55 (1.33-4.87)**	**1.82 (1.46-2.25)**	**1.79 (1.12-2.85)**	1.09 (0.87-1.37)	**2.23 (1.73-2.85)**
Bachelor's degree	**4.40 (2.83-6.85)**	**2.48 (1.37-4.52)**	2.16 (1.00-4.67)	**2.32 (1.76-3.06)**	1.32 (0.72-2.40)	1.04 (0.75-1.46)	**3.33 (2.43-4.56)**
Advanced degree	**3.98 (2.09-7.59)**	1.16 (0.46-2.92)	1.02 (0.27-3.84)	**1.60 (1.11-2.31)**	1.29 (0.52-3.19)	0.70 (0.48-1.03)	**3.02 (2.02-4.52)**
Income							
< $10,000	Reference	Reference	Reference	Reference	Reference	Reference	Reference
$10,000-24,999	**1.86 (1.34-2.58)**	1.39 (0.96-2.01)	1.22 (0.75-1.97)	**1.30 (1.05-1.61)**	0.94 (0.56-1.59)	1.18 (0.95-1.46)	**1.26 (1.02-1.55)**
$25,000-49,000	**2.32 (1.62-3.32)**	1.27 (0.86-1.88)	**1.69 (1.16-2.47)**	**1.52 (1.24-1.87)**	0.96 (0.60-1.54)	1.04 (0.85-1.29)	**1.37 (1.10-1.70)**
$50,000-99,999	**2.45 (1.64-3.67)**	1.12 (0.71-1.78)	**2.44 (1.59-3.74)**	**1.76 (1.38-2.25)**	1.35 (0.80-2.28)	0.86 (0.67-1.10)	**2.15 (1.68-2.76)**
>=$100,000	**2.57 (1.56-4.25)**	1.09 (0.57-2.08)	**2.74 (1.70-4.42)**	**1.74 (1.34-2.26)**	1.02 (0.51-2.03)	**0.65 (0.45-0.92)**	**2.87 (2.12-3.89)**

Bolded values indicate estimate significant a $p < 0.05$
The “none” category is used in reference for the tobacco and alcohol use outcome.

Participants with high negative affect severity, compared to low, had the greatest odds for alcohol-exclusive use rather than no use while adjusting for externalizing severity, ND, sex, age, race, education, and annual household income. Participants with high externalizing severity, compared to low, had the greatest odds for ECIG and alcohol use rather than no use while adjusting for negative affect severity, ND, sex, age, race, education, and annual household income.

Additional models compared results across all categories of reference groups to establish differences for each category of tobacco/alcohol use (Appendix A, Supplemental Table 2.1). All significant associations between negative affect/externalizing severity and tobacco and alcohol combinations were significantly lower when referencing alcohol, CIG, and ECIG as well as ECIG and alcohol use. Conversely, significant positive associations were found between negative affect/externalizing severity and tobacco and alcohol combinations when referencing CIG and ECIG use, ECIG-exclusive, and CIG-exclusive. Results were mixed when referencing CIG and alcohol use, and alcohol-exclusive use.

DISCUSSION

Our study is one of the first to examine the relationships between negative affect/externalizing severity and combinations of CIG, ECIG, and alcohol use across adulthood. There were three major results. First, strong, positive associations with negative affect/externalizing severity at various levels of CIG, ECIG, and alcohol use were detected. Overall, negative affect severity was more strongly associated with CIG and alcohol use as well as alcohol-exclusive use while externalizing severity was more strongly associated with ECIG and alcohol use when accounting for ND. Second, ECIG may represent a new and underappreciated substance related to externalizing psychopathology. Alcohol was significantly associated with psychopathology when ECIG was included. Third, ND may mediate the relationship between negative affect/externalizing severity and various levels of CIG, ECIG, and alcohol use.

Patterns of tobacco and alcohol use vary by negative affect/externalizing severity

We detected specific patterns of association between tobacco and alcohol use with negative affect/externalizing severity. Specifically, high negative affect severity had a higher magnitude of association with CIG and alcohol use as well as alcohol-exclusive use. In contrast, externalizing severity was more strongly associated with ECIG and alcohol use. These results expand on recent positive associations that were detected between mental disorder symptoms and exclusive use of tobacco products in adults.[73] Specifically, multiple mental disorder symptoms (i.e., higher severity) was generally associated with use of more than one substance, except for alcohol. To date, individuals with co-occurring mental health disorders have been reported to have a more severe course of illness, health and social consequences, more difficulties when seeking and in treatment, or worse treatment outcomes than people with a single disorder.[18] Additionally, tobacco use has been reported to be higher among people with mental health problems (e.g., major depressive disorder, generalized anxiety, schizophrenia, and/or antisocial personality/conduct disorder).[90,115,116] These results suggest that patterns, rather than a dose-response, of tobacco and alcohol use are associated with negative affect/externalizing severity. Further investigation of these comorbidity patterns, including tobacco use, is required.

ECIG use associated with externalizing severity with co-occurring alcohol use

Negative affect/externalizing severity were not significantly associated with ECIG-exclusive use. This is inconsistent with previous work, perhaps due to differences in defining ECIG use.[73] Specifically, we expanded our study of “ECIG use” to include a commonly occurring form of tobacco use- dual use of ECIG and CIG. Our results provide a more detailed and nuanced description of the relationship between negative affect/externalizing psychopathology and ECIG use by parsing out co-occurring CIG and alcohol use from ECIG.

Concurrent ECIG and alcohol use, however, was significantly associated with externalizing severity. Further, compared to low externalizing severity, high and moderate externalizing severity showed stronger association with alcohol use of any kind (i.e., alcohol, CIG, and ECIG use; ECIG and alcohol use; CIG and alcohol use; and alcohol-exclusive use). This association between externalizing and alcohol is consistent with prior studies,[38,117,118] and this association remains when ECIGs are used with alcohol. This finding builds upon previous work that has established more harmful alcohol use with ECIG use in that externalizing symptoms are associated with this pattern of use. More research is needed to better understand the relationship between different combinations of tobacco and alcohol, including ECIG, and psychopathology.

ND may mediate the relationship between negative affect/externalizing severity and current tobacco and alcohol use in adults

The magnitude of the associations between negative affect/externalizing severity and levels of tobacco/alcohol use were reduced, although generally remained significant, when ND was included. The associations between negative affect severity and alcohol, CIG, and ECIG use and CIG-exclusive use as well as externalizing severity and CIG and ECIG use were no longer statistically significant. ND may explain more of the relationship between negative affect severity and alcohol, CIG, and ECIG use as well as CIG-exclusive use. Previous work has indicated that externalizing behaviors act as a precursor or factor involved in substance use, especially alcohol use.[38,117,118] Therefore, the relationship between externalizing severity and alcohol use in adults, whether exclusive or with tobacco, is expected to be mediated or have an indirect effect through ND. In an ad hoc mediation analysis,[119] ND was determined to be a significant mediator between negative affect/externalizing and tobacco and alcohol use. This is consistent with prior work that has identified ND as a mediator between mental conditions.[120] We also included a test for SUD severity (GAIN-SS) as a mediator in models including ND since it measures broader substance use behavior, including alcohol. However, no significant direct or indirect effect of SUD was detected. As mediation is inherently a causal hypothesis, we recommend future researchers to confirm this with a longitudinal analysis to accurately model a mediation pathway in the context of the transactional effect between tobacco initiation and ND development.

Strengths and limitations

These results should be interpreted while considering the following points. First, these data were collected in 2013-2014, so these analyses do not capture more recent ECIG products (i.e., pod-mods). Consequently, these results may not be generalizable to the current generation of ECIG devices. Second, the analytic sample size was reduced from the Wave 1 sample after removing participants with missing data. Many participants (N=13,865) were removed due to a skip pattern identified for the ND items used to calculate the composite ND item. If a participant was not a current tobacco user, a former 12-month tobacco user, or a current experimental tobacco user, they were not asked the ND items. ND is contingent upon tobacco initiation[121]; therefore, it was inappropriate to code these missing observations as 0. Therefore, there is systematic bias introduced by the missingness; however, results from sensitivity analyses in which all missing observations on the ND items were coded as 0 did not demonstrate differences that would alter the overall study conclusions. Third, use of self-reported data has the potential to introduce misclassification bias, which may underestimate the magnitude of associations. However, this would lead to an attenuation of effect sizes, rather than an overestimation. Fourth, the GAIN-SS measures negative affect/externalizing symptom severity rather than psychiatric diagnoses. We recognize use of symptom data as a strength, as we are more likely to capture true rates of mental health disorders without relying on disease-specific diagnoses. There is growing support for the use of subthreshold or transdiagnostic symptoms over traditional diagnoses to better explain the high rates of comorbidity among common mental disorders, particularly when characterizing population-based samples.[122] Therefore, these results represent the full distribution of severity across several mental health domains. Future investigations should test these associations with negative affect and externalizing symptoms rather than a composite severity score to further detail comorbidity patterns. Fifth, to answer our research questions, this study focused on current CIG, ECIG, and alcohol use, and ND. We could not determine if ND was due to the CIG or ECIG use or another tobacco product that was not included in these analyses. Future studies are encouraged to explore direct associations with other tobacco products and ND. Sixth, by using only data from Wave 1, direction of causation cannot be determined and future longitudinal studies are needed.

Conclusions

Negative affect and externalizing severity were strongly associated with multiple levels of CIG, ECIG, and alcohol use in this study. The magnitude of association varied by the tobacco product used. Overall, negative affect severity was more strongly associated with CIG and alcohol use as well as alcohol-exclusive use while externalizing severity was more strongly associated with ECIG and alcohol use when accounting for ND. ECIG may represent a new and underappreciated substance related to externalizing psychopathology. The magnitudes of these associations were reduced when ND was included in the model, indicating that ND may mediate the association between negative affect/externalizing severity and current tobacco and alcohol use. Future work is encouraged to investigate the different patterns of tobacco and alcohol use (i.e., using latent variable and network approaches) since our results suggest patterns of use rather than a dose-response relationship between tobacco and alcohol use and negative affect/externalizing severity. Longitudinal studies could provide deeper insight into the stability of these patterns over time.

CHAPTER 3: LATENT CLASSES OF COMORBID SUBSTANCE USE AND NEGATIVE AFFECT AND EXTERNALIZING SYMPTOMS AND THEIR ROLE IN ADULT SUBSTANCE USE DISORDER SEVERITY

INTRODUCTION

Substance use disorder (SUD) results from the prolonged use of any psychoactive substance at high doses and/or frequencies, and is defined as continued use despite associated health and social problems.[2,123] SUD poses a substantial burden on the United States' health system, with almost 20 million American adults meeting diagnostic criteria for a past-year SUD in 2018.[5] Of the 20 million adults in the United States (U.S.) who experience a SUD, half also have a co-occurring mental disorder.[5] The co-occurrence of mental disorders is common among people who use substances or engage in polysubstance use, consuming more than one substance over a defined period.[10,12,35] People with co-occurring tobacco use, substance use, and mental disorders have more severe courses of mental illness, more severe health and social consequences, more difficulties seeking and receiving treatment, and worse treatment outcomes compared to people with a singular disorder.[18] Therefore, people who experience greater comorbidity may be at a greater risk for SUD severity.

Study of Comorbidity Characterizes Common Patterns of Co-Occurring Disorders

Historically, comorbidity research involving substance use or symptoms underlying mental disorders has generally focused on the identification and description of either polysubstance use or mental disorder classes, separately. These studies have identified several specific subcategories of either SUD or mental disorders. For example, a recent study examined past-month polysubstance use among a small sample of psychiatric inpatients with co-occurring mental disorder and SUDs. This study identified three polysubstance use profiles: cannabis and alcohol (35.1%), alcohol only (49.3%), and polysubstance use including cocaine plus alcohol and marijuana (15.7%).[35] Other latent class analyses of substance use focus on a sample of specific substance users. For example, a five-class solution was most optimal in describing polysubstance use in a small sample of lifetime cocaine users: past 30-day tobacco use only (45%), past 30-day alcohol, marijuana and tobacco use (31%), past 30-day tobacco, prescription opioid and sedative use (13%), past 30-day cocaine, alcohol, marijuana and tobacco use (9%), and past 30-day cocaine and multiple polysubstance use (2%).[37] Another study of opioid-dependent patients identified a two-class solution: severe comorbidity with high rates of other SUDs specifically amphetamine and sedative (10%), and a less severe comorbidity class with moderate rates of nicotine, alcohol, cannabis, and cocaine disorders (90%).[36] Overall, the comorbidity characteristics are different depending on the population being studied.

In another study assessing only mental disorder comorbidity, a four-class solution best described the sample: low psychopathology (84.0%), internalizing (9.9%), externalizing (4.5%), and high psychopathology (1.6%).[41] The high psychopathology class was more strongly associated with lifetime suicide attempt, compared to the internalizing and externalizing classes.[41] The internalizing and externalizing classes had overall higher odds of lifetime affective and substance use disorders, respectively.[41] These four class solutions are commonly identified in mental disorder only comorbidity.[41,124]

Detailing comorbidity between substance use and mental disorders has only recently been addressed and confirms the strength of association between these two conditions. For example, a more recent study evaluated the class structure of substance use and mental disorders together and identified a four-class solution: low disorder (73.6%); mental health and low SUD (10.6%); alcohol, cannabis, and low mental health disorder (12.2%); and polysubstance use and moderate mental health disorder (3.5%).[12] Mental disorders were more likely to occur with polysubstance use disorders in young adults.[12] Another study accounting for both mental and SUDs also identified a four-class solution: little psychopathology (62.5%), internalizing disorders (16.9%), externalizing disorders (16.4%), and both internalizing/externalizing disorders (4.2%).[124] The most severe class, internalizing/externalizing disorder, demonstrated elevated rates for both mental disorders as well as alcohol (85.4%), cannabis (76.2%), and hard drug use disorders (61.1%) while the internalizing class had moderate alcohol use disorder endorsement (34.2%) and the externalizing class had greater endorsement of alcohol (84.2%), cannabis (82.0%), and hard drug disorders (53.7%).[124] As a whole, these studies confirm and demonstrate a consistent comorbidity structure that likely takes form in four classes. These studies also identify a larger, low psychopathology group and a smaller yet more severe comorbidity group with more moderate comorbidity somewhere in between. The severe comorbidity group has higher endorsement of mental disorders and substance use behaviors. Understanding and identifying these comorbidity profiles have the potential to better support people with comorbidity through risk assessment and interventions.[41]

Limitations in Current Approach to Studying Comorbidity

Despite recent advancements in the study of comorbid SUD, several gaps in knowledge remain. First, the current comorbidity literature usually considers diagnosis to measure substance use and mental disorders. However, the overlap between symptoms from multiple disorders and the simultaneous co-occurrence of multiple disorders is very common. The current classification systems are limited in their ability to characterize comorbidity and may neglect to appropriately account for these overlaps. Therefore, using symptom-level data and a more recent measure of substance use behaviors (i.e., past-month substance use) will not only account for these overlaps, but these measurements could also indicate severity. For example, people who engaged in past-month use compared to people who used in the past-year may be more likely to have higher levels of SUD severity. Furthermore, there may be an underestimation of the population level burden of comorbidity due to the measurement of substance use and mental disorder through diagnostic classifications only. People who use substances or experience mental health problems that are not severe enough to receive a diagnosis or have access to a diagnosis are not included.[117] Most substance-related health and social problems occur among individuals who are not addicted or have a SUD diagnosis.[8] Consequently, using subthreshold measures may better address this issue of underreporting while also accounting for the overlap between substance use behaviors and mental health conditions.

Second, tobacco use is rarely considered in comorbidity research, despite consistent literature supporting the association of tobacco use and mental disorders.[12,41,92] Nicotine dependence occurs in a substantial proportion of individuals with alcohol use disorder (23.8%), marijuana use disorder (32.6%), cocaine use disorder (47.7%), prescription opioid use disorder (45.4%), and heroin use disorder (66.3%).[9] Negative affect (e.g., depression and anxiety) and externalizing [e.g., attention-deficit hyperactivity disorder (ADHD) and conduct disorder] psychopathology[2,16,73,89–91] are important mental health factors that have been consistently associated with exclusive use of conventional cigarettes (CIG). A meta-analysis reported that current CIG users had a two-fold increased risk of depression relative to never and former CIG users.[92] Further, adults with depression are more likely to smoke and are less likely to be successful at quitting than adults without depression.[93] With the increase in access and use of electronic cigarettes (ECIG) alone and in CIG users,[84,125] it is increasingly important to consider this alternative method of nicotine delivery in comorbidity research alongside CIG use.

Third, comorbidity is typically studied using smaller samples of more severely disordered participants,[35] which does not provide a sense of the etiology of comorbidity for people who are affected with lower levels of severity. This is unfortunate, because people with lower levels of comorbid severity are expected to make up at least 2/3 of the U.S. population.[12,41,124] Consequently, there is a gap in our understanding of comorbidity across the population. Specifically, it is unclear whether the latent class structure of SUD comorbidity from clinical samples will be detected in a population-based sample where severity of SUD is generally lower. Nevertheless, population-based study samples can measure SUD severity using screeners like the Addiction Severity Index (ASI), Alcohol Use Disorders Identification Test (AUDIT), Drug Abuse Screen Test (DAST), Fagerstrom Test for Nicotine Dependence (FTND), and the Global

Appraisal for Individual Needs (GAIN).[26–28] These instruments are useful to assess SUD and SUD severity in settings that cannot specifically assess diagnosis. Use of these instruments can therefore capture people who (1) have subthreshold levels of impairment, and (2) may not have access to a physician to receive a diagnosis. Therefore, it is possible to model the relationship among a robust set of substance use behaviors, including tobacco, and mental disorder symptoms in a large, nationally representative sample to detail the comorbidity between SUDs and mental disorders on a population-level.

Use of Comorbidity Details to Predict SUD Severity

In addition to the aforementioned limitations, the detailing of comorbidity between co-occurring substance use and mental disorder has primarily been descriptive (i.e., developing comorbidity profiles) with some tests of association between comorbidity profiles and other factors (i.e., demographic characteristics, early life factors [i.e., parental factors], psychiatric diagnoses, suicide attempts, self-efficacy in abstinence, and treatment involvement).[12,35,41] However, there have been few advancements in improving treatment outcomes and reducing the prevalence of mental disorders and/or SUDs.[9] Consequently, there has been a stated need for comorbidity research due to the "insufficient information" that exists today.[126] Recent studies have suggested the value in accounting for patterns of substance use and comorbidity to categorize SUD. Such detail is expected to improve the identification of individuals at risk for high SUD severity.[41] Consequently, it is possible that establishing a "comorbidity profile" may be useful to screen individuals for SUD risk in order to appropriately address additional risk factors.

Comorbidity profiles are also expected to be associated with sociodemographic characteristics like sex, age, race, education, annual household income, and perceived social connectedness. Being female is a significant predictor of membership for a mental health disorder class,[12] internalizing or negative affect class and high psychopathology class.[41] People who are older in age (65 years and older) and black, non-Hispanic are significantly protected from membership in internalizing or negative affect, externalizing and high psychopathology classes.[41] Higher income level is also indicative of a protective relationship from membership in internalizing or negative affect, externalizing, and high psychopathology classes.[41] Perceived social connectedness and support have been found to protect adults against substance use behaviors and mental disorder symptoms.[127,128] Therefore, these sociodemographic characteristics must be considered when developing comorbidity profiles to predict SUD severity.

Purpose, research questions, and hypotheses

Prediction modeling using the probability of class membership is an extension to previous LCA work that can further inform prevention and potential intervention strategies among polysubstance users with varying levels of mental disorder symptoms. The three goals for this study are: (1) identify latent classes of comorbid substance use behaviors and mental disorder symptoms using LCA, (2) determine if there are differences in comorbidity by demographic and social factors, and (3) predict SUD severity using the probability of comorbidity class membership severity in the first wave of adult data from the Population Assessment of Tobacco and Health study. Based on previous literature, we expect substance use to vary across the mental disorder symptoms. More substance use behaviors will cluster with highly comorbid mental disorder classes. Certain demographic factors will increase the risk of highly comorbid class membership, while social satisfaction will decrease risk. Higher probabilities of class membership in the highly comorbid and externalizing classes will predict greater SUD severity compared to probability of class membership in the low comorbidity and negative affect classes. This research will support population-level strategies to prevent or treat SUDs, including nicotine dependence, and the development of more severe mental health problems.

METHODS

Setting

Wave 1 adult data (N=32,320) from the Population Assessment of Tobacco and Health (PATH) study was used.[71] These data are cross-sectional and were collected between September 2013 and December 2014. PATH is a nationally representative longitudinal cohort study of the civilian, non-institutionalized household population of the U.S., and participants engaged in all levels of tobacco use ranging from never using tobacco to frequent use.

The weighted response rate among participants was 74.0% for Wave 1.[73] Participants responded to tobacco-specific items including tobacco-use patterns, risk perceptions and attitudes towards current and newly emerging tobacco products, tobacco initiation, cessation, relapse behaviors, and health outcomes.[72] Participants also responded to non-tobacco items (e.g., media use, peer and family influences, health effect outcomes, and industry advertising and promotion).[72]

Study representativeness

Participants with missing data on the substance use, negative affect, and externalizing measures were not included in the analysis (N= 2,109). Survey respondents of the analytic sample endorsed significantly greater substance use overall, negative affect symptoms, and externalizing symptoms (except for fighting) compared to those not included in the analytic sample. The participants in the analytic sample were more likely to be Non-Hispanic white, men, aged 25-54 with higher levels of education and annual household income than those who were missing.

Measures

Seventeen variables were studied: six substance use variables, four negative affect variables, and seven externalizing variables.

<u>Past Month Tobacco and Substance Use.</u> Six substance use categories were used in this study: exclusive cigarette, exclusive e-cigarette, dual cigarette and e-cigarette, alcohol, marijuana, and prescription drugs not prescribed (PDNP) including painkillers, sedatives, tranquilizers. Only past month or current use of the substances was considered (coded as 1, else = 0) to reduce the potential for recall bias and ensure for accurate overlap with negative affect and externalizing symptoms occurring in the same time frame.

Past Month Negative Affect and Externalizing Symptoms. Negative affect and externalizing symptoms were measured using the Global Appraisal of Individual Needs—Short Screener (GAIN-SS).[73] The items selected to identify negative affect and externalizing symptoms from the GAIN-SS instrument are ordinal and measures people across four times periods: past month, 2 to 12 months, over a year ago, and never. Participants indicating that they experienced a symptom within the past month were coded as 1. Participants indicating that they experienced the symptom 2 to 12 months ago, over a year ago, and never were coded as 0. Only past month or current negative affect and externalizing symptoms were considered reducing the potential for recall bias and ensure accurate overlap with substance use occurring in the same time frame.

Past Month Substance Use Disorder Symptoms. Substance use disorder (SUD) severity was measured using the Global Appraisal of Individual Needs—Short Screener (GAIN-SS).[73] Seven questions were used to measure SUD symptoms that asked the last time:

(1) “used alcohol/drugs weekly or more often,”

(2) “spent a lot of time getting alcohol/drugs,”

(3) “spent a lot of time using or recovering from alcohol or other drugs,”

(4) “kept using alcohol or other drugs even though it was causing social problems, leading to fights, or getting you into trouble with other people,”

(5) “your use of alcohol or other drugs reduced your involvement in activities at work, school, home or social events,”

(6) “had withdrawal problems such as shaky hands, throwing up, having trouble sitting still or sleeping,” and

(7) “use of alcohol/drugs to avoid withdrawal.”

These items are ordinal and measures people across four time periods: past month, 2 to 12 months, over a year ago, and never. Participants indicating that they experienced a symptom within the past month were coded as 1. Participants indicating that they experienced the symptom 2 to 12 months ago, over a year ago, and never were coded as 0. Only past month or current SUD symptoms was considered to reduce the potential for recall bias and ensure accurate overlap with substance use, negative affect symptoms, and externalizing symptoms occurring in the same time frame. The binary responses were summed to reflect a scale score of the number of current SUD symptoms. The score ranging from 0 to 7 SUD symptoms will be categorized into low (0), moderate (1-2), and high (3+) severity). This is the final measure used as the SUD severity outcome. These are the recommended cut points based on validation analyses of the dimensional measure.[28] The ordinal severity categories were informed by other studies, showing concurrent and high predictive validity in other samples.[28,73,99] A higher score indicates increased severity, a greater likelihood for diagnosis with a SUD, and increased need for services.[73] There is substantial overlap with the symptoms identified in the GAIN-SS and the symptoms identified in the DSM-5 as the GAIN-SS uses symptoms from the DSM to generate dimensional symptom count measures (Table 3.1).[28,129]

Table 3.1: Overlap between SUD Symptoms for DSM-5 Diagnostic Criteria and GAIN-SS

DSM Criterion	**DSM Symptom Description**	**GAIN-SS Symptom Description**
Impaired Control	Taking substance in larger amounts or over a longer period than initially intended.	Used alcohol/drugs weekly or more often
	Expressing a persistent desire to cut down or regulate use and may report unsuccessful attempts to do so.	
	Spending a great deal of time obtaining the substance, using the substance, or recovering from its effects.	Spent a lot of time getting alcohol/drugs
		Spent a lot of time using or recovering from alcohol or other drugs
	Craving manifested by an intense desire or urge for the drug that may occur at any time but is more likely in an environment in which drug use has previously occurred.	
Social Impairment	Recurrent substance use results in failure to fulfill major role obligations at work, school, or home.	Your use of alcohol or other drugs reduced your involvement in activities at work, school, home or social events
	Continuing substance use despite persistent social/interpersonal problems exacerbated by the effects of the substance.	Kept using alcohol or other drugs even though it was causing social problems, leading to fights, or getting you into trouble with other people
	Giving up or reducing important social, occupational, or recreational activities because of substance use.	Your use of alcohol or other drugs reduced your involvement in activities at work, school, home or social events
Risky Use	Recurrent substance use in physically hazardous situations.	
	Continuing use despite knowledge that persistent physical or psychological problems are	

<table>
<tr><td></td><td>exacerbated or caused by substance use.</td><td></td></tr>
<tr><td rowspan="3">Pharmacology</td><td>Markedly increased dose of the substance required to achieve desired effect.</td><td></td></tr>
<tr><td rowspan="2">Withdrawal symptoms specific to a drug class.</td><td>Had withdrawal problems such as shaky hands, throwing up, having trouble sitting still or sleeping</td></tr>
<tr><td>Use of alcohol/drugs to avoid withdrawal</td></tr>
<tr><td colspan="3">Note: SUD is diagnosed with the occurrence of two or more symptoms. 2 – 3 symptoms = mild presentation, 4 – 5 symptoms = moderate presentation and 6 or more symptoms = severe presentation. GAIN-SS symptom severity is categorized as 0 symptoms = low severity, 1-2 symptoms = moderate severity, and 3 or more symptoms = high severity.</td></tr>
</table>

Covariates. Sociodemographic variables such as sex, age, race, education, annual household income, and social satisfaction were included as auxiliary variables to predict the probability of class membership. These factors were chosen because they have been identified by previous studies to significantly predict latent class membership.[12,41,127,128] Sex was a binary variable with one level representing male and the other level representing female. Age, measured in PATH as a seven-level categorical variable, was re-categorized to have a uniform distribution with six levels (18-24, 25-34, 35-44, 45-54, 55-64, and 65 years or older). Race/ethnicity was measured as a four-level categorical race variable and included information from a separate variable that accounted for Hispanic ethnicity (Non-Hispanic White, Non-Hispanic Black, Non-Hispanic Other, and Hispanic Multicultural). Education, measured in PATH as a six-level categorical variable, was re-categorized as a five-level categorical variable with a uniform distribution [less than high school, GED/high school graduate, some college (no degree) or Associate's degree, Bachelor's degree, and

Advanced degree]. Annual household income was measured as a five-level categorical variable: less than $10,000, $10,000 to $24,999, $25,000 to $49,999, $50,000 to $99,999, and $100,000 or more. Level of satisfaction with social activities and relationships was measured as a five-level categorical variable: extremely satisfied, very satisfied, moderately satisfied, a little satisfied, and not at all satisfied.

Statistical analysis

Latent Class Analysis. Latent Class Analysis (LCA) was used to assign participants into classes of substance use and/or mental disorders using their responses to self-report measures of substance use, negative affect, and externalizing symptoms while accounting for demographic and environmental factors as predictors of class membership. The general goal of an LCA is to define an unobserved, latent variable (i.e., comorbidity) as a set of classes where the observed variables or items (i.e., substance use and negative affect-externalizing symptoms) are locally independent.[130] Local independence or conditional independence is a condition in which the observed items are independent of one another, condition on the level of the latent variable.[130] This means that the relationship or correlation between the observed items represents a distinct domain that is fully explained by the level of the latent class and that there is no residual correlation between the items.[70,131] LCA accounts for the observed covariation between substance use and mental disorder symptoms and offers objective indices of class classification accuracy that are not available in traditional cluster analysis methods.[132]

LCA models are produced by estimating item probability parameters and class probability parameters.[70] Item probability parameters represent the probability of endorsing an item conditional on latent class membership. It can also be referred to as the item response probabilities or conditional item probabilities.[70] An LCA estimates the probability of being in a latent class conditional on the probability of endorsing a measured item. The combination of these two parameters is used to estimate the probability of being in a latent class with the marginal item probability for item $u_j = 1\ (j = 1,2, \ldots r)$ given by:

$$P(u_j = 1) = \sum_{k=1}^{K} P(C = k)P(u_j = 1|C = k) \tag{1}$$

where u is a categorical latent variable C with K classes $(C = k;\ k = 1, 2, \ldots, K)$ across r binary items.[70] Specifically, class probability parameter $P(C = k)$ reflects the probability that a person in a given latent class has of endorsing the specific item.[70] The class probability parameter specifies the prevalence of each class in the population or the relative frequency of class membership.[70] Further, the item probability parameter or conditional item probability for a given class is defined by the logistic regression:

$$P(u_j = 1|C = k) = \frac{1}{1 + \exp(-v_{jk})} \tag{2}$$

where the v_{jk} is the logit for each of the u_j's for each of the latent classes, k.[70] The class probability parameter or the prevalence of each class in the population is $\delta_k = P(C = k)$.[70]

Model Selection. Five LCA models, ranging from 2-6 classes were tested and included the seventeen variables of interest. Models were evaluated using measures of model fit, model parsimony, and entropy.

Model fit quantifies how well a model explains the data. Model fit is assessed as the -2 log-likelihood which and reflects how much unexplained variation there is in the estimated model. Model fit comparisons generally test which model best explained the data by using a likelihood ratio test (LRT) which compares the relative fit of two models that differ by a set of parameter restrictions.[133] The traditional LRT is a hypothesis test that compares two nested models:

$$LRT = -2 \log e \frac{L_s(\hat{\theta})}{L_g(\hat{\theta})} \tag{3}$$

where s represents the likelihood for the null model and g represents the likelihood of the nested model. The estimated values of a traditional LRT follow a chi-square distribution with the number of degrees of freedom equal to the difference in parameter numbers between the two models that will be compared.[133,134]

Model fit comparisons for LCA do not meet certain regularity assumptions that must be satisfied[133,134] and as such the use of classic LRT for model comparison is not appropriate. Specifically, LCA violates the assumption that the additive property of the likelihood ratio statistic can assess the statistical difference between pairs of hierarchically related models, meaning that one model is a constrained form of the other.[135] Consequently, model fit for the LCA was assessed using a Lo-Mendell-Rubin adjusted likelihood ratio test (LMRT) rather than the traditional LRT.[133,134] The LMRT approximates the likelihood ratio test distribution which can be used for comparing nested latent class models.[70] LMRT uses the adjusted asymptotic distribution of the likelihood ratio statistic and compares the improvement in fit between neighboring class models (i.e., comparing $k-1$ and the k class models) with a p-value that can be used to determine if there is a statistically significant improvement in the fit for the inclusion of one more class.[70]

Model parsimony was used to assess how well the model explains the data while accounting for differences in the number of model parameters estimated in models with different. Parsimony was evaluated using Akaike information criteria (AIC), Bayesian information criteria (BIC), and sample-size adjusted BIC. AIC was estimated as:

$$AIC = -2\log L + 2p \tag{4}$$

where L is the likelihood function, p is the number of free model parameters.[70,136] The benefit of using AIC is that it measures the closeness of the estimated model to the true model among the competing models.[136] The smallest value of the AIC is selected as the more parsimonious model.[136]

Bayesian information criteria (BIC)[137] was estimated as:

$$BIC = -2\log L + p\log(n) \tag{5}$$

where $-2\log L$ is -2 times the log-likelihood of the estimated model and n is the number of observations. BIC aids model selection by penalizing the number of factors in a model.[138] The smaller BIC indicates a more parsimonious model.

Sample-size adjusted BIC[139] builds on BIC by replacing the same size n in the BIC equation above with n^*. It was estimated as:

$$n^* = (n+2)/24 \quad (6)$$

$$adjusted\ BIC = -2\log L + p\log[(n+2)/24] \quad (7)$$

The benefit of sample-adjusted BIC is that it accounts for the sample size. Simulation studies have confirmed that BIC and sample-size adjusted BIC are better indicators of the number of latent classes than AIC.[70]

Entropy is commonly used as a model selection criterion that indicates the model's ability to classify a person in a latent class (i.e., level of separation).[133] It measures aggregated classification uncertainty and reflects accuracy of class membership assignment.[133] Classification uncertainty is assessed at the individual level by posterior probabilities from the estimated model.[140] Therefore, entropy identifies the estimated model's ability to classify an individual into a class based on their poster probability of having endorsed a specific item. There is little distinction between classes when posterior probabilities across the classes are very similar. Entropy ranges from 0 to 1 and a higher entropy represents a better fit; values > 0.80 indicate the latent classes highly discriminate.[133]

Joint evaluation of parsimony, entropy, and model fit was used to identify the LCA model that best explained the data. Consequently, models with lower values for AIC and BIC were preferred. Models with a larger entropy value were preferred because they strongly discriminate between classes. Statistically significant results from the LMRT were used to determine the model with the lowest number of classes that best fit the data. Interpretability and average latent class probabilities were also considered in determining optimal class solution.[70,141]

The best fitting model was also used to assign class membership to each participant. This assignment of class membership was then used in the multinomial logistic regression and prediction analyses as detailed below.

Multinomial Logistic Regression. Multinomial regression was used to determine whether any covariates were significantly associated with membership of a latent class. [142] Multinomial regression was conducted using the three-step method (R3STEP) via the AUXILIARY statement in Mplus. This approach was used to identify the variables to use as covariates in the third step multinomial logistic regression. A multinomial regression tests the association between any set of categorical or continuous predictors with a categorical outcome as:

$$P(C = 1|X) = \frac{1}{1 + exp^{(\alpha+\beta X)}} \tag{8}$$

where *p* has a categorical latent variable C with X as the covariate of interest (i.e., sex, age, race, education, annual household income, and level of satisfaction with social activities and relationships). The intercept is denoted as α and the regression coefficient is β. This approach was used in order for the latent class model and the latent class predictor model to be obtained automatically[142] rather than introducing potential bias by performing a multinomial regression after the latent class models were selected.

Prediction Analysis. Cumulative receiver operator characteristic (ROC) curves (developed from the cumRoc3 MACRO[143]) were used to: (1) estimate the ability of the class membership probabilities created from the LCA to predict SUD severity; (2)

estimate ability of the substance use variables, negative affect symptoms, and externalizing symptoms to each separately predict SUD severity; and (3) compare the predictive ability of the two approaches. This comparison was addressed in order to determine if establishing latent class membership performed better compared to the use of separate variables for predicting SUD severity.

The classic ROC curve is computed by comparing a binary outcome Y with a continuous measure X where each observed level of X is evaluated as a candidate cut point discriminating observed Y = 1 (positive) from Y = 2 (negative).[143] Traditionally, ROC curves have been used to establish the value of a diagnostic test measured as a binary outcome. Results from a ROC curve analysis provides results that support in the identification of the threshold that distinguishes a positive test from a negative test.[144] The correct classifications among positive outcomes are the true positives (TP). The correct classifications of the negative outcomes are the true negatives (TN). The incorrect classifications among negative outcomes are the false positives and the among the positive outcomes the false negatives. These classifications are used to compute the sensitivity (i.e., probability that an observation with a positive outcome is correctly classified as positive [sensitivity = TP/(TP + FN)]) and specificity (i.e., probability that an observation with a negative outcome is correctly classified as negative [specificity = TN/(TN + FP)]) of a test. The coordinates of a ROC curve are computed where the x-axis is the false positive rate (i.e., 1 – specificity) and the y-axis is the sensitivity or true positive rate.[143]

A cumulative ROC curve analysis extends the classical empirical ROC curve by discriminating three or more ordinal outcome levels on a shared continuous scale.[143]

The cumulative ROC calculates the area under the curve (AUC) which explains the ability of a continuous measurement to discriminate between ordinal outcome levels (i.e., 3-level SUD severity outcome). In this case, the AUC is the probability that an observation with a higher severity SUD outcome will have a higher continuous measurement (i.e., higher class probability) than an observation with a lower severity SUD outcome.[143] An AUC of 0.5 represents no discriminating ability (i.e., no better than chance) versus an AUC of 1.0 represents perfect discrimination between the groups.[145]

The probability of the SUD severity (P_{SUD}) for class membership (x_1) was estimated as follows:

$$\mathrm{P}_{SUD} = \frac{\exp\,[\beta_0 + \ \beta_1 x_1]}{(1 + \exp[\beta_0 + \ \beta_1 x_1])} \tag{9}$$

where β_0 is the intercept, and β_1 is the estimated regression coefficient for the probability of latent class membership (x_1).[146]

Predictive probabilities were generated from four ordinal logistic regression models. The main model of interest estimated the probability of class membership (generated from the LCA) and SUD severity. Three additional models were run using predictive probabilities from regressing substance use variables, negative affect symptoms, and externalizing symptoms separately on SUD severity.

<u>Statistical Programs, Handling Missingness and Complex Sampling Design</u>. Data management, summary statistics, and the prediction analysis were performed in SAS 9.4. All LCA was conducted in MPlus.[147] Participants with missing data were not included in the latent class analysis (N = 2,109, missing data patterns = 256). Complex

sampling design was accounted for in SAS 9.4 using PROC SURVEYFREQ (to generate summary statistics) and PROC SURVEYLOGISTIC (to generate the predictive probabilities from the logistic regression models), and in MPlus using the WEIGHT option.

RESULTS

Summary statistics

The sample was 51.9% female and 66% Non-Hispanic White. Age was evenly distributed among the sample. Most of the sample had at least a GED or high school education (88.4%), had an annual household income of more than $25,000 (65.9%), and were very (46.1%) or extremely (22.3%) satisfied with their social activities and relationships. Current alcohol use was most frequently reported (52.4%), followed by exclusive cigarette use (16.6%), marijuana use (7.1%) and PDNP (5.1%). Sleep trouble was the most common past-month negative affect symptom reported (26.7%), followed by feeling very anxious (16.1%), feeling depressed (13.4%), and becoming distressed about the past (12.5%). The most frequently endorsed past-month externalizing symptom was giving answers before the other person finished asking the question (32.0%), followed by having a hard time paying attention (14.6%) and having a hard time listening to instructions (10.4%). Most of the sample indicated low past month SUD severity (63.2%, Table 3.2).

Table 3.2: Wave 1 Summary Statistics	
	Wave 1 (N=32320)
	N (Weighted %)
Sex	
Male	16306 (48.1)
Female	15980 (51.9)
Age	
18-24	9110 (13.0)
25-34	6337 (17.7)
35-44	4930 (16.5)
45-54	4846 (17.9)
55-64	3971 (16.6)
65+	3110 (18.2)
Race	
Non-Hispanic White	19295 (66.0)
Non-Hispanic Black	4496 (11.2)
Non-Hispanic Other	2429 (7.5)
Hispanic Multiracial	4817 (13.3)
Education	
Less than high school	4233 (11.6)
GED/High school graduate	9765 (29.5)
Some college (no degree)	11300 (31.0)
Bachelor's degree	4498 (17.8)
Advanced degree	2311 (10.1)
Annual household income	
Less than $10,000	5668 (13.7)
$10,000- $24,999	6768 (20.4)
$25,000- $49,999	6670 (23.0)
$50,000- $99,999	6140 (24.9)
$100,000 or more	3914 (18.0)
Satisfaction with social activities and relationships	
Extremely satisfied	6942 (22.3)
Very satisfied	13742 (46.1)
Moderately satisfied	8157 (23.7)
A little satisfied	2376 (5.6)
Not at all satisfied	1001 (2.3)
Past month tobacco and substance use	
Exclusive CIG	10381 (16.6)
Exclusive ECIG	578 (0.9)
Dual CIG + ECIG	996 (1.5)
Alcohol	17787 (52.4)
Marijuana	4392 (7.1)
PDNP	1950 (5.1)
Past month negative affect symptoms	
Depressed	5692 (13.4)

Sleeping	9564 (26.7)
Anxious	6864 (16.1)
Distressed/Past	5605 (12.5)
Past month externalizing symptoms	
Lied	3245 (7.1)
Attention	5831 (14.6)
Listening	4128 (10.4)
Bully	737 (1.7)
Fights	404 (0.7)
Restless	2953 (6.2)
Answered	11399 (32.0)
Past month SUD severity	
Low	16481 (63.2)
Moderate	8985 (31.7)
High	2156 (5.1)

Class membership and item-response probabilities

A four-class model was identified as best fitting the data and was selected to conduct additional analyses (Table 3.3). Classes from the 4-class model were labeled based on the highest conditional probabilities that characterized the class. The characteristics and patterns for each class are detailed below. Figure 3.1 displays the probability of being categorized within one of the four classes (i.e., the class membership probabilities) given the specific patterns of past-month substance use and endorsement of negative affect and externalizing symptoms in the past month (i.e., item response probabilities or conditional probability).

Table 3.3: Wave 1 LCA Model Fit and Parsimony													
	AIC	BIC	Sample-Size Adjusted BIC	Entropy	H_o LL	LMRT	p-value	LC 1	LC 2	LC 3	LC 4	LC 5	LC 6
2 class	311322.5	311616	311504.7	0.864	-176279	41085.86	<0.05	7339 (22.7%)	24981 (77.3%)				
3 class	307151.6	307596	307427.5	0.773	-155626	4184.516	<0.05	2637 (8.2%)	7375 (22.8%)	22308 (69.0%)			
4 class	303521.4	304116.6	303891	0.844	-153520	3640.512	<0.05	1960 (6.1%)	2691 (8.3%)	23571 (72.9%)	4098 (12.7%)		
5 class	302018.6	302764.7	302481.9	0.695	-151679	1508.737	0.6834	1854 (5.7%)	3800 (11.8%)	2594 (8.0%)	5497 (17.0%)	18575 (57.5%)	
6 class	301484.8	302381.8	302041.7	0.718	-150917	559.354	0.7626	734 (2.3%)	1383 (4.3%)	5385 (16.7%)	2500 (7.7%)	3658 (11.3%)	18659 (57.7%)
NOTE: AIC = Akaike information criteria, BIC = Bayesian information criteria, LL = log likelihood, LMRT = Lo Mendell Rubin Test, LC = latent class													

Low-Symptom Class. Most participants were categorized as being in the low symptom class (N=23,571, 72.9%). This class, overall, had lower conditional probabilities for endorsing all items compared to the other classes. Consequently, participants in this class had a low probability of endorsing most substance use and negative affect/externalizing items. However, the conditional probability of exclusive cigarette use was marginally higher for the low symptom class compared to the externalizing class (13.6% vs. 12.9%) meaning that a person in the low symptom class had a 13.6% probability for endorsing exclusive cigarette use in the past month.

Negative Affect Class. The negative affect class (N=4,098, 12.7%) had higher conditional probabilities for the four negative affect symptoms compared to the low symptom class, and externalizing class. A person was more likely to endorse exclusive CIG, dual CIG and ECIG, marijuana, and PDNP use if they were in the negative affect class compared to the low symptom, and externalizing classes. This class represents a population of people who more commonly endorse the four negative affect symptoms along with past-month substance use, excluding ECIG.

Externalizing Class. The externalizing class (N=2,691, 8.3%) had higher conditional probabilities for all seven externalizing symptoms compared to low symptom, and negative affect classes except for the “start physical fights with other people” (Externalizing class = 0.7%, Negative Affect class = 2.00%). The conditional probability for exclusive ECIG and alcohol use was greater for those in the externalizing class compared to the low symptom, and negative affect classes. Therefore, this class represents adults who experience higher levels of externalizing symptoms along with exclusive ECIG and alcohol use.

Comorbid Class. Approximately 6% of participants were categorized as being in the comorbid class (N=1,960, 6.1%). Compared to the other classes, this class had higher conditional probabilities for all items except for alcohol use (Comorbid class = 57.8%, Externalizing class = 64.4%). This class represents a small population of people who, overall, have high endorsement of all seventeen items and, therefore, may indicate more severe presentation of substance use and mental disorder symptom severity.

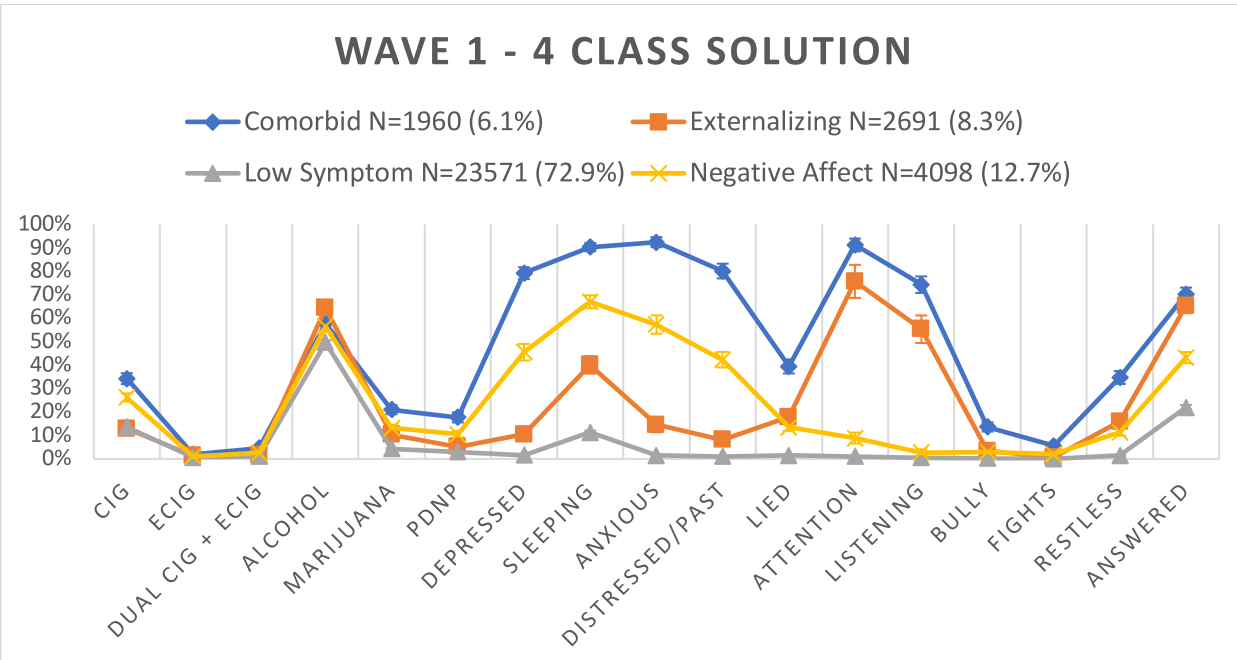


Figure 3.1: Four class solution of substance use behaviors and mental disorder symptoms.

Associations between Sociodemographic Factors and Substance Use/Mental Disorder Classes

Males were significantly less likely than females to be classified in the comorbid (OR = 0.72, 95% CI = 0.63-0.82, $p < 0.05$) and negative affect (OR = 0.74, 95% CI = 0.66-0.83, $p < 0.05$) classes relative to the low symptom class (Table 3.4).

As age increased, the odds of class membership decreased for all classes relative to the low symptom class. Therefore, the youngest age group (18-24 years) had the highest odds of class membership compared to the oldest age group (65 years and older), relative to the low symptom class (Comorbid Class OR = 10.02, 95% CI = 7.06-14.24, $p < 0.05$; Negative Affect Class OR = 3.88, 95% CI = 3.11-4.83, $p < 0.05$; Externalizing Class OR = 3.39, 95% CI = 2.69-4.28, $p < 0.05$). Respondents who identified as Non-Hispanic Black, Hispanic Multicultural, and Non-Hispanic Other were significantly less likely than respondents who identified as Non-Hispanic White to be classified in any of the classes relative to the low symptom class.

Table 3.4 Wave 1 - Association Between Demographic and Social Variables with Probability of Latent Class Membership			
	Comorbid Class OR (95% CI)	**Negative Affect OR (95% CI)**	**Externalizing OR (95% CI)**
Sex			
Female	REF	REF	REF
Male	0.72 (0.63-0.82)*	0.74 (0.66-0.83)*	1.00 (0.87-1.14)
Age			
18-24 years	10.02 (7.06-14.24)*	3.88 (3.11-4.83)*	3.39 (2.69-4.28)*
25-34 years	6.00 (4.17-8.64)*	2.39 (1.90-3.01)*	1.81 (1.40-2.33)*
35-44 years	4.10 (2.83-5.94)*	1.69 (1.32-2.16)*	1.46 (1.11-1.91)*
45-54 years	3.77 (2.60-5.47)*	1.52 (1.19-1.93)*	1.13 (0.86-1.50)
55-64 years	2.27 (1.54-3.36)*	1.44 (1.12-1.85)*	0.98 (0.73-1.32)
65 years and older	REF	REF	REF
Race			
Non-Hispanic White	REF	REF	REF
Non-Hispanic Black	0.48 (0.40-0.59)*	0.76 (0.64-0.89)*	0.62 (0.51-0.77)*
Non-Hispanic Other	0.73 (0.57-0.94)*	0.72 (0.56-0.91)*	0.69 (0.53-0.89)*
Hispanic Multiracial	0.50 (0.41-0.61)*	0.78 (0.66-0.93)*	0.65 (0.52-0.80)*
Education			
Less than high school	1.53 (1.05-2.21)*	1.62 (1.22-2.15)*	0.79 (0.57-1.09)
GED/High school graduate	1.37 (0.98-1.92)	1.42 (1.10-1.83)*	0.80 (0.61-1.05)
Some college (no degree)	1.79 (1.29-2.47)*	1.36 (1.06-1.75)*	1.08 (0.85-1.37)
Bachelor’s degree	1.35 (0.94-1.93)	1.16 (0.88-1.53)	1.10 (0.85-1.42)
Advanced degree	REF	REF	REF
Income			
Less than $10,000	2.54 (2.03-3.18)*	1.61 (1.32-1.95)*	0.83 (0.65-1.06)
$10,000- $24,999	2.02 (1.62-2.51)*	1.51 (1.25-1.83)*	0.55 (0.44-0.68)
$25,000- $49,999	1.45 (1.16-1.81)*	1.11 (0.92-1.33)	0.84 (0.68-1.04)
$50,000- $99,999	1.00 (0.78-1.27)	0.99 (0.82-1.20)	1.21 (1.00-1.46)*
$100,000 or more	REF	REF	REF
Level of satisfaction with social activities and relationships			
Extremely Satisfied	REF	REF	REF
Very satisfied	1.65 (1.31-2.07)*	1.55 (1.31-1.85)*	1.42 (1.18-1.70)*
Moderately satisfied	8.15 (6.54-10.15)*	4.53 (3.78-5.43)*	2.66 (2.18-3.26)*
A little satisfied	34.19 (26.40-44.29)*	11.09 (8.70-14.14)*	3.08 (2.17-4.39)*
Not at all satisfied	95.87 (66.32-138.58)*	22.62 (15.44-33.16)*	3.67 (1.62-8.31)*

Note: Low symptom class was the reference level for the outcome.
* Indicates a p-value < 0.05

Compared to those with an advanced degree, participants with lower education levels were significantly more likely to be in the comorbid (Less than High School OR =

1.53; Some College/No Degree OR = 1.79) and negative affect (Less than High School OR = 1.62; GED/High School Graduate OR = 1.42; Some College/No Degree OR = 1.36) classes relative to the low symptom class. The associations between education level and the externalizing class were not statistically significant. Compared to an income of $100,000 or more, those with annual household incomes of $99,999 and below were significantly more likely to be in the comorbid (Less than $10,000 OR = 2.54; $10,000-$24,999 OR = 2.02; $25,000-$49,999 OR = 1.45) and negative affect (Less than $10,000 OR = 1.61; $10,000-$24,999 OR = 1.51) classes, compared to the low symptom class (Table 3.4).

A reduction in social satisfaction was associated with membership in comorbid, negative affect, and externalizing classes. For example, compared to being extremely satisfied, as social satisfaction decreased, the likelihood of being in the comorbid class increased (Not at all satisfied OR = 95.87, 95% CI = 66.32-138.58, $p < 0.05$). Similarly, compared to participants who were extremely satisfied, participants who were not at all satisfied were about 23 times more likely to be categorized in the negative affect class (OR = 22.62, 95% CI = 15.44-33.16, $p < 0.05$) and almost four time more likely to be in the externalizing class (OR = 3.67, 95% CI = 1.62-8.31, $p = 0.002$). This relationship was not detected for the alcohol class.

Prediction modeling

Data generated from the LCA model (i.e., class membership and probability of class membership) were exported from Mplus and imported into SAS to determine the predictive ability of the latent class on SUD severity. Class membership significantly predicted SUD severity (Table 3.5).

Table 3.5: Prediction of the probability of class membership and SUD severity	
Latent Class	SUD Severity Odds Ratio (95% Confidence Interval)
Low Symptom Class	REF
Comorbid Class	2.31 (2.07-2.59)*
Externalizing Class	1.54 (1.38-1.72)*
Negative Affect Class	1.51 (1.38-1.65)*
* Indicates a p-value < 0.0001.	

Relative to the low symptom class, membership in the comorbid class increased the odds of SUD severity by 2.31 times (OR = 2.31, 95% CI = 2.07-2.59, $p < 0.0001$). The externalizing and negative affect classes had similar relationships with SUD severity (Externalizing OR = 1.54, 95% CI = 1.38-1.72, $p < 0.0001$; Negative Affect OR = 1.51, 95% CI = 1.38-1.65, $p < 0.0001$). These estimates were unadjusted since the sociodemographic covariates were accounted for in the development of the latent classes.

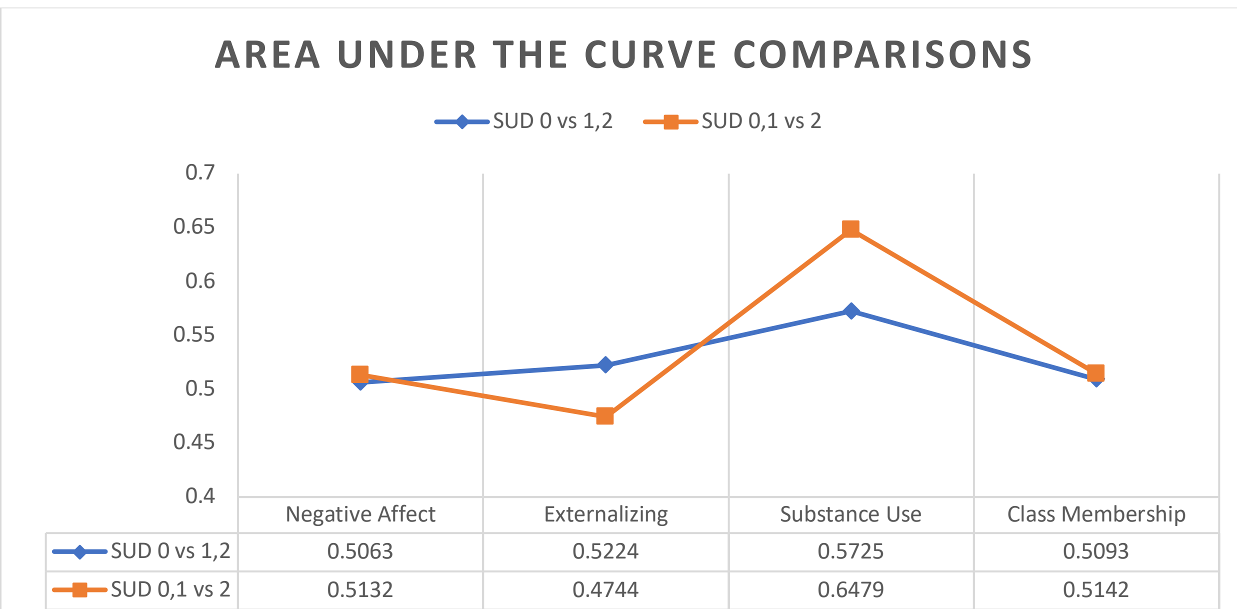


Figure 3.2: Area under the curve comparisons generated from the cumulative ROC Curves

Cumulative ROC curves (see Appendix B, Supplemental Figures 3.1-3.8) were generated to determine if latent class membership (combining substance use behaviors, negative affect, and externalizing symptoms) was a better predictor of specific levels of SUD severity compared to any of the indicators that make up the latent class separately. Past-month substance use behaviors best predicted SUD severity. For example, the area under the curve (AUC; i.e., the degree of separability) for past-month substance use to predict low SUD severity versus moderate/high SUD severity was 0.57 (Figure 3.2). This means that the substance use variables would only be correct in predicting SUD severity about 57% of the time. Further, the AUC improved when predicting low/moderate SUD severity versus high SUD severity (AUC = 0.6479). Therefore, at this threshold (low/moderate vs high SUD severity), the predictive ability increased from 57% to 65%. The ability of the latent class to predict SUD severity was marginally better than the negative affect and externalizing indicators for the low/moderate versus high SUD severity (Class Membership AUC = 0.51; Negative Affect AUC = 0.51; Externalizing AUC = 0.47). Therefore, substance use behaviors measured alone performed better at predicting SUD severity compared to comorbid substance use as reflected in comorbid latent class membership.

Overall, AUCs ranged from 0.47 (low/moderate SUD severity vs high SUD severity level for externalizing) to 0.65 (low/moderate SUD severity vs high SUD severity for substance use behaviors) meaning that predictions of SUD severity were only correct about 47-65% of the time.

DISCUSSION

To our knowledge, this is one of the first studies to use a latent class approach to describe comorbidity between substance use behaviors and mental disorder symptoms together in a large nationally representative sample of U.S. adults. This study also used information from the latent class analysis to predict a health outcome, SUD severity. There are three major results from this study. First, a four-class solution (i.e., low symptom class, negative affect class, externalizing class, and comorbid class) best described the data. These classes also allowed us to understand what symptoms and substance use behaviors commonly occur together in this sample, confirming our hypothesis that substance use would vary across mental disorder symptoms. For example, exclusive cigarette, dual cigarette and e-cigarette, marijuana, and PDNP use more commonly occurred in the negative affect class while exclusive e-cigarette and alcohol use more commonly occurred with the externalizing class. Second, sociodemographic factors were significantly associated with latent class membership and social satisfaction was a strong factor associated with the comorbid and negative affect classes. Third, latent class membership predicting SUD severity performed similarly to a model where the symptoms were grouped separately (i.e., negative affect symptoms, externalizing symptoms, and substance use behaviors).

Class prevalences and underestimation

The four-class solution was determined to best fit the data for this sample. This is consistent with prior latent class results that have identified a four-class solution to be most optimal. Salom et al. identified a four-class solution of comorbid polysubstance use and mental health disorders in young adults: low disorder (73.6%); mental health and low SUD (10.6%); alcohol, cannabis, and low mental health disorder (12.2%); and polysubstance use and moderate mental health disorder (3.5%).[12] Other studies have also found that four-class solutions are most optimal in their samples with the low psychopathology class being the largest group (62.5% to 84.0%) followed by an internalizing or negative affect class with some substance use endorsement (i.e., alcohol use disorder), an externalizing class with high endorsement of substance use problems (i.e., nicotine dependence, alcohol use disorder, and drug use disorder), and a comorbid or both internalizing or negative affect/externalizing with high endorsement of substance use problems class as the smallest group (1.1% to 4.2%).[41,124] Therefore, our results confirm and support a four-class solution to best explain substance use and mental disorder comorbidity in U.S. adults.

Similar to other latent class findings, the low symptom class was most common in this sample (72.3%, N=23,571). This suggests that most American adults may engage in some substance use (i.e., current CIG or alcohol use) while also experiencing some mental disorder symptoms, specifically impulsivity or sleep problems, but otherwise have low endorsement of other substance use behaviors and mental disorder symptoms. Almost thirty percent (27.7%) of the sample, however, were categorized in the other three remaining classes based on higher probabilities of endorsing substance use or mental disorder symptom items. These people have the potential to be underestimated or not accounted for based on the current classification systems due to their subthreshold levels of possible impairment. This presents a missed opportunity to identify the comorbid substance use and mental disorder symptoms, and potentially prevent the comorbidity from becoming progressively worse. Therefore, this part of the population could benefit from refined detection and possible intervention (e.g., access to support or educational materials). Early detection could result in better intervention outcomes and better overall mental health outcomes.

Patterns of substance use varies by negative affect and externalizing classes

Exclusive cigarette, dual cigarette and e-cigarette, marijuana, and PDNP use had higher endorsement in the negative affect class. In comparison, exclusive e-cigarette and alcohol use had higher endorsement in the externalizing class. These patterns may be helpful in identifying people at risk for development of more severe comorbidity in public health spaces. The implications of these results are considered below.

Negative affect class

Previous work reports that people who engage in conventional cigarette use are at an increased risk of negative affect disorders like depression and anxiety.[90,92] The relationship between negative affect symptoms and dual cigarette and e-cigarette use is not as well understood. Our prior analysis (see Chapter 2) showed that the combined use of alcohol, cigarette and e-cigarette was significantly associated with high negative affect severity while dual cigarette and e-cigarette use was not significantly associated with any level of negative affect severity after adjusting for demographic covariates.[148] The LCA results add to our understanding of psychopathology, confirming the association between dual cigarette and e-cigarette use with negative affect symptoms, especially as this dual use is currently increasing in adults, specifically young adults.[83]

There was higher endorsement for past-month marijuana use in the negative affect class compared to the externalizing class. There is a vast literature demonstrating the association between marijuana use and negative affect symptoms such as anxiety and depression.[149–152] Additionally, marijuana use has been increasing at a greater rate in women, who are more likely to endorse negative affect symptoms, compared to men over the last decade (40% increase for men, 53% increase for women from 2006 to 2016).[153] Therefore, the reason there is greater endorsement of marijuana in the negative affect versus the externalizing class may be because women are more likely to make up the negative affect class.

The finding of high endorsement of PDNP in the negative affect class is consistent with the literature. Evidence suggests that opioid use, a substance measured within PDNP, is associated with PTSD symptoms.[154] Other studies have also identified that people in SUD treatment for nonmedical use of prescription painkillers like opioids, almost half (43%) have a diagnosis or symptoms of anxiety and depression.[155]

Externalizing class

The item response probability for exclusive e-cigarette use was higher for the externalizing class compared to the low symptom and negative affect classes. This is not supported by our prior analysis (see Chapter 2) where exclusive e-cigarette use was not significantly associated with negative affect or externalizing severity using a multinomial logistic regression.[148] However, e-cigarette and alcohol use together were significantly associated with externalizing severity.[148] Given that alcohol use is widely accepted to be associated with externalizing behaviors,[38,117,118] we hypothesized that the alcohol use may drive the relationship between e-cigarette and externalizing severity when alcohol was used with e-cigarettes (i.e., dual use). Nevertheless, the results from this LCA study support that exclusive e-cigarette use may be more related to externalizing symptoms than negative affect. As e-cigarettes continue to increase in use,[125] it is important to understand how this electronic nicotine delivery system differs from conventional cigarette use. Some studies have identified similarities in that both deliver nicotine and result in poor health outcomes specifically related to the lungs.[125] However, regarding comorbidity with mental disorder symptoms, e-cigarettes may present differently than conventional cigarettes. Alcohol use was also endorsed at a greater probability in the externalizing class compared to the negative affect and low symptom classes. This is consistent with prior studies that suggest that alcohol use contributes to a latent factor of externalizing behaviors.[38,117,118]

Sociodemographic characteristics and latent class membership

As age increased, the odds of class membership decreased. This is consistent with previous work, where younger people are at greater risk for mental health and substance use problems compared to people in older age categories.[41] This could indicate an increase in substance use initiation which is typical in younger age categories.[156] This also matches with the age of onset for most mental disorders as roughly 50% to 75% of all lifetime mental disorders start by the mid-teens and mid-20s, respectively.[2] Therefore, broad prevention strategies that address all latent class profiles could be helpful in supporting younger people with comorbidity.

Compared to Non-Hispanic White participants, those in all other race categories were less likely to be in any of the latent classes. Another study has also identified the association between race and comorbidity, whereas those who identify as Non-Hispanic White are at increased risk of latent class membership compared to individuals in other racial/ethnic groups, specifically Non-Hispanic Black and Hispanic Other.[41] There are many potential reasons why this occurs. First, Non-Hispanic White populations are overrepresented in psychopathology and comorbidity research.[123] It is also likely that diagnoses of comorbidity are optimized for the Non-Hispanic White population rather than across all groups. Therefore, it has been more difficult to draw associations with other racial/ethnic groups. Second, due to the historical distrust in the U.S. healthcare system, people of other racial/ethnic groups may be less likely to participate in research and indicate that they participate in substance use behaviors or experience mental disorder symptoms.[157] However, a likely conclusion that is not an artifact of study sampling or potential misclassification could be the strong levels of resiliency in other racial/ethnic groups, specifically seen in African Americans or those who identify as Black.[158] This is known as the black-white mental health paradox and explains that Black Americans tend to experience similar or relatively low rates of psychiatric disorders compared to Whites despite higher stress exposure, greater material hardship, and worse physical health.[158] While it is important to support those who identify as Non-Hispanic White, it remains important to continue being inclusive of all racial/ethnic categories in comorbidity research to develop more consistent results and provide the appropriate level of support and targeted prevention efforts.

Women had higher odds of membership in the comorbid and negative affect classes. This is consistent with previous LCA work that has shown that women are more likely to be in a comorbid or internalizing/negative affect class.[12,41,124] Men, however, were not significantly associated with membership in the externalizing class. Men typically have higher rates of alcohol use and endorse more externalizing symptoms and disorders. This may due to the robust set of other items included in the generation of the latent classes.

Compared to those at higher levels of socioeconomic status, people at lower levels of socioeconomic status had a greater risk of membership in the comorbid and negative affect classes. This is consistent with previous studies. For example, a longitudinal study of 34,653 noninstitutionalized U.S. adults identified that low levels of household income were associated with several lifetime mental disorders and a reduction in household income was associated with increased risk of incident mood, anxiety or substance use disorders compared to respondents with no change in income.[159] Additionally, prior work has identified that higher income and education levels represent a protective relationship from membership in internalizing or negative affect, externalizing, and high psychopathology classes.[41]

The magnitudes of the association between social satisfaction and the comorbid as well as negative affect classes were very large. Social satisfaction was also associated with the externalizing class. Those who were less satisfied with their social lives had greater odds of externalizing class membership. Although not incredibly precise, the association between social satisfaction and class membership may be a place to intervene, as social satisfaction is (1) a more easily modifiable factor compared to other demographic characteristics and (2) demonstrates a protective association. For example, as people became less satisfied with their social activities and relationships, their odds of membership in the comorbid, negative affect, or externalizing classes significantly increased. Therefore, if social satisfaction can improve (i.e., becoming more content with one's activities and relationships with the support of a psychological therapist or counselor), there could be a decrease in the risk of belonging to the comorbid, negative affect, or externalizing classes. This buffering effect through social support has been demonstrated previously[160] and could be an opportunity to intervene or prevent further development of comorbid or negative affect psychopathology.

Limited ability to predict SUD severity

The comorbid class had the strongest association with SUD severity when predicting SUD severity using class membership in a multinomial logistic regression. The comorbid class had greater endorsement of past month substance use behaviors, except for alcohol use, and mental disorder symptoms compared to the other classes. This finding indicates that people with endorsement of more items at greater rates are associated with greater SUD severity. Therefore, we assumed that by grouping symptoms together and describing them as they occur using a latent modeling approach, we would be better able to predict health outcomes than assessing these symptoms separately. However, the cumulative ROC analyses showed that the ability of latent class membership to predict SUD severity was no better than the symptoms grouped separately (i.e., negative affect symptoms, externalizing symptoms, and substance use behaviors). Further, predictions of SUD severity were only correct about 47-65% of the time.

The poor ability of latent class membership to predict SUD severity may be due to the incongruency between measurement tool used and population assessed. The outcome measure, SUD severity, was created to approximate SUD diagnosis. This tool was originally validated in populations which were oversampled with SUD in order to appropriately distinguish between SUD and no SUD.[28] Therefore, the tool used to measure SUD may not perform as well in a sample of people who experience subthreshold levels of SUD or other mental disorders. Classes generated from other methods like factor mixture modeling[131] that can account for heterogeneous groups (i.e., SUD and no SUD) may be better in predicting SUD severity. Another reason could be due to misclassification bias introduced by the measurement used for negative affect and externalizing symptoms, and SUD severity. We could be misclassifying individuals by collapsing 2 to 12 months, over a year ago, and never into one reference category to compare to the past-month endorsement. Additionally, the negative affect and externalizing symptoms were correlated, which may also affect the ability of either items to predict SUD severity. Future studies should consider using a factor mixture modeling approach to determine comorbidity's predictive ability of SUD severity. Until then, it may be that disorders do better at predicting SUD severity compared to subthreshold or symptom-level measures.

Strengths and limitations

There are several strengths and limitations to this study. First, this study used data from a large, nationally representative sample of U.S. adults allowing for the generalizability of these results to the adult, noninstitutionalized population in the U.S. However, participants included in this study differed significantly from participants with missing data as those included had greater endorsement of substance use, negative affect symptoms, and externalizing symptoms. Therefore, these participants may not represent the U.S. adult population. Additionally, these data are cross-sectional and this study cannot resolve causal inference. As more waves of data are collected, we will be able to assess the stability of this class structure along with changes in class membership using longitudinal methods (e.g., latent transition analysis) especially considering the more recent changes in substance use over time (e.g., the increase in e-cigarette and marijuana use).

Second, the substance use and mental disorder symptoms measure comorbidity within the same time frame (i.e., past month endorsement of substance use as well as the negative affect-externalizing symptoms and SUD severity). There is potential misclassification due to how the measure was developed by collapsing the 2 to 12 months, over a year ago, and never response options into one group (coded as 0 vs past-month coded as 1). People who endorsed a symptom in the last 2 to 12 months or over a year ago differ from people who never endorsed a symptom. Future work could consider (1) developing a three-level categorical variable that separates those who never endorsed a symptom from those who endorsed a symptom in the last 2 to 12 months and over a year ago to compare with the past-month level, or (2) maintain the original four levels of the item to avoid losing information through dichotomizing or re-categorizing the variables. Our binary measurement, however, allowed us to model current comorbid polysubstance use and negative affect and externalizing symptoms while also predicting SUD severity within the same time period. Additionally, by using symptom measures, we accommodate and provide a better understanding of comorbidity compared to a diagnostic classification system.[117] Additionally, it is important to note that although the cut points for SUD severity have acceptable reliability, validity, and overlap with DSM-5 SUD criterion in this sample, we may not be appropriately treating these variances as a continuous probability. Future research is encouraged to evaluate the SUD severity items and consider measurement techniques such as the use of quantiles[161] to confirm that the categorization of the SUD severity variable is appropriate for the population being studied (i.e., based on the distribution of population's responses).

Third, the assessment of factors associated with class membership is limited to the demographic and social factors included in this study. There could be other factors associated with comorbidity that were not included and could result in residual confounding. Future work should investigate the association of other environmental factors on class membership to better understand the influence of additional social determinants of health on comorbidity.

Fourth, we ran an ordinal regression model for the prediction analysis. However, the model violated the proportional odds assumption (chi-square = 439.2, p-value < 0.0001). This means that the relationship between any two pairs of the outcome groups (i.e., low vs moderate/high SUD severity and low/moderate vs high SUD severity) was not statistically the same. We also ran a multinomial regression model because of the assumption violation. However, results from a multinomial regression were consistent with the ordinal regression results and, therefore, we presented the ordinal regression results in order to be synonymous with results presented from the cumulative ROC curves. Nevertheless, predicting SUD severity is a strength. It is an extension to the primarily descriptive ability of using an LCA. Further, by utilizing the ROC curves, we were able to identify the ability of latent class membership to predict SUD severity and compare that to the substance use behaviors and mental disorder symptoms separately.

Fifth, a LCA model was used to assign participants into comorbidity classes using their responses to self-report measures of substance use, negative affect, and externalizing symptoms. This approach is considered to be important to discover classes based on observed data and characterize participants based on latent class membership. LCA was selected to compare results with previous studies that assessed for comorbidity and be used in clinical and research settings specifically for risk assessment and treatment.[41] Further, the interpretability of LCA results (i.e., classifications and assigning individuals to groups based on their item endorsement) can be easily translated for use in clinical settings by identifying individuals at potential risk for increased comorbidity severity based on their current substance use and mental disorder symptoms. It is possible that messaging could be developed based on the latent classes identified in this study. Results could be shared with policy makers so they may allocate more resources toward developing comorbid support in clinical spaces.

Nevertheless, there are several limitations related to the use of the LCA model. Specifically, by using categorical data to assign individuals into discrete classes, there may be a loss of information that would emerge from a model that accounts for a continuous distribution. For example, we observed parallel trends of the item response probabilities across the classes. This observation suggests that there could be a continuous distribution to these data and that a dimensional presentation, rather than discrete, may be more appropriate in characterizing the comorbidity within the population. Therefore, the discrete latent classes generated from the LCA may not represent the actual types of individuals in the population.[162] Additionally, the conditional independence assumption of the LCA model can also be seen as a limitation. Conditional independence simplifies the presentation of underlying classes in a population based on consistent patterns in the data (i.e., item response probabilities) yet it may be an over-simplification or biased representation of the true heterogeneity in the population.[163] It may not be true that the latent class fully explains the relationship between the observed variables. A possible solution to address both major LCA limitations is the use of a factor mixture model. The factor mixture model uses a hybrid of latent class and factor analysis where the latent variable allows for the classification of individuals into groups while the factor models the heterogeneity of the disorder within the latent class, relaxing the conditional independence assumption.[131] This is useful because comorbidity class membership and the range of severity within and across classes can be modeled concurrently.[164] A factor mixture model estimates a factor score for each individual which quantifies the heterogeneity within a class; however, there is no model-based classification of individuals because individuals are assumed to be from the same homogeneous population.[164] Network analysis can also assess comorbidity structure without the assumption of conditional independence, and is an approach used in Chapter 4.

Sixth, we could not use the bootstrap likelihood ratio test as an additional examination of model fit when accounting for complex sampling design during the LCA. LCA includes a bootstrap likelihood ratio test to test for model fit across models with various classes. However, it could not be performed when using weighted data to account for complex sampling design. Consequently, we relied on the Lo-Mendell-Rubin adjusted likelihood ratio test along with other parsimony metrics to decide on the optimal class solution. This is considered a promising and appropriate approach when determining the number of classes from an LCA model.[165]

Seventh, this is a sample of mainly healthy people and as such, the dimensionality of comorbidity may be different here compared to a sample of people diagnosed with psychiatric conditions (e.g., those who are institutionalized). This approach should be replicated in other samples to confirm or refute the dimensionality of comorbidity. Nevertheless, the results from this sample detail the patterns of mental disorder symptoms and substance use behaviors in a broader population in order to appropriately characterize comorbidity at a population level. This is important because substance use and/or mental disorder symptoms that do not result in a diagnosis remain pervasive throughout American society.[8] Undiagnosed individuals may go untreated and untreated mental illness, including SUD, represents $300 billion due to losses in productivity annually.[166] Therefore, it is important to identify and detail patterns of substance use and mental health outcomes throughout the full population in addition to those at highest risk for disorders or those who are affected.[117131]

Conclusions

In a nationally representative sample of U.S. adults, four latent classes were most optimal at describing mental disorder symptom and substance use comorbidity. Negative affect symptoms were commonly seen with exclusive CIG use, dual CIG and

EIG use, marijuana, and PDNP use. Externalizing symptoms were commonly seen with exclusive ECIG use and alcohol. Social satisfaction may be an important factor to consider when intervening on comorbidity. Comorbidity of latent class membership was similar to negative affect, externalizing, and substance use behaviors, separately, in predicting SUD severity. This may suggest that network psychometrics may be a better approach to understanding the predictive ability of comorbidity for other health outcomes. Future research may benefit from using a network approach to better understand comorbidity.

CHAPTER 4: A NETWORK APPROACH TO SUBSTANCE USE, NEGATIVE AFFECT, AND EXTERNALIZING COMORBIDITY IN U.S. ADULTS[2]

INTRODUCTION

Substance use (e.g., tobacco, alcohol, marijuana, and sedatives [i.e., benzodiazepines and barbiturates]) commonly co-occurs with negative affect disorders (i.e., behavioral problems that manifest and are maintained within the individual[167]) such as depression and anxiety, and externalizing disorders (i.e., behavioral problems that manifest as negative outward behavior acting on the external environment[168]) like attention-deficit hyperactivity disorder (ADHD). These comorbidities are summarized in Table 4.1.[11,89,91,149–152,169–183]

Many of the most consistent results regarding comorbidity have focused mainly on disorder, as seen in Table 4.1. However, other papers have used different substance use measures (e.g., initiation, recency of use, quantity of use) and they too have seen comorbidity with mental disorders and mental disorder symptoms.[151,152,169,178,179,181] Given the consistency of results in disorder and other use measures, it is worth the effort to focus on lower levels of symptomatology and explore the etiology of comorbidity below the diagnostic threshold.

[2] A modified version of this chapter was submitted for publication to Addictive Behaviors, Special Issue on Networks, Complexity and Addictive Behaviors.

Table 4.1: Summary of Previously Reported Comorbidity by Substance

Substance Use Disorder	Co-Occurring Substance use	Comorbid Mental Health
Tobacco	Alcohol Marijuana	Major depressive disorder Generalized anxiety disorder ADHD
Alcohol	Marijuana Opioids	Anxiety disorders Depressive disorders ADHD Conduct disorder
Marijuana	Alcohol use disorder Tobacco use disorder Substance use disorders	Depression Anxiety Conduct disorder
Sedatives	Tobacco use disorder Alcohol use disorder Opioids General illicit drug use	Depressive disorders Anxiety disorders
Opioids	Tobacco Alcohol Marijuana Sedatives	Depression Anxiety PTSD Conduct disorder

To date, it is unclear whether the same patterns of comorbidity identified with substances such as conventional cigarettes (CIG), alcohol, and marijuana extend to relatively new products including electronic cigarettes (ECIG) and use of prescription drugs in a manner not previously prescribed (e.g., sedatives, tranquilizers, and painkillers) which have increased in popularity over the past several years. For example, the prevalence of ECIG use has increased from 2.4% in 2012 to 7.6% in 2018.[125] Further, the patterns of some substance use in ECIG users have been reported to be similar to that of CIG users. Specifically, ECIG use frequently occurs with alcohol use and other substances.[87,94,95,97] Similar to ECIG, the prevalence of PDNP (i.e., prescription drug use not prescribed, nonmedical use of a prescription drug including recreational use) has also been increasing for the last fifteen years with overdose deaths involving prescription opioids being four times greater in 2018 than in 1999.[184]

People with SUDs and mental disorders are at a higher risk for nonmedical use of prescription opioids.[185] Of those in SUD treatment for nonmedical use of prescription painkillers, almost half (43%) have a diagnosis or symptoms of anxiety and depression.[155] Additionally, opioids and sedatives are sometimes combined for recreational use, resulting in a higher risk for comorbid mental conditions as well as nonfatal and fatal overdoses.[186,187] Therefore, as the prevalence of these substances increase there has also been increasing evidence for their comorbidity with negative affect/externalizing behaviors and other substances.

Patterns of mental health comorbidity focus on diagnoses rather than symptoms

Most research on the patterns of comorbidity between substance use, externalizing, and negative affect behaviors has focused on diagnoses of disorders rather than the symptoms underlying these diagnoses.[22,39] However, this approach neglects the inclusion of people who experience subthreshold levels of impairment and results in a potential loss of information when summing symptoms to reach diagnosis.[33] Most substance-related health and social problems occur among individuals who are not addicted or have a SUD diagnosis.[8] Additionally, many of these symptoms cross over diagnostic boundaries.[33,188] Consequently, there is a substantial gap in understanding the overlap between substance use and mental disorder symptoms. By using another measure like past-month substance use, it is possible to capture people who use substances with and without a diagnosis of a SUD, allowing for a more robust assessment of comorbidity patterns and accounting for overlaps between substance use behaviors and mental disorder symptoms. If the focus were to remain solely on diagnosis, there would be no evidence to support people who experience subthreshold use and comorbid mental disorder symptoms. Therefore, past-month substance use offers an opportunity to study a larger population of people who use substances (i.e., those with and without disorder) as well as identify and intervene at the subthreshold level to better support individuals experiencing comorbid substance use.

Gender differences in the comorbidity of substance use and mental health

Much of the substance use and SUD research has largely been conducted in men. However, the prevalence of substance use in the U.S. has been increasing in women.[123] Further, negative affect/externalizing symptoms present differently in men and women.[2,123] For example, men are more likely to experience externalizing disorders, while women are more likely to report negative affect disorders.[123,183,189,190] Furthermore, comorbid psychiatric conditions occur more frequently in women with SUDs compared to men.[190,191] Consequently, the comorbidity between substance use and psychopathology may also vary by gender.

Study goals and hypotheses

Network analyses of substance use or SUDs have yet to account for comorbid mental disorders.[34,192] Therefore, the primary goal of this study is to detail a network system of past-month substance use as well as a wide range of negative affect/externalizing symptoms, and quantify how well a given node can be predicted by all other nodes it is connected to in the network using nodewise predictability. A secondary goal of this study is to determine whether there are gender differences in the comorbidity network structure. Based on prior literature, we hypothesize that tobacco, alcohol, marijuana, and PDNP will connect with negative affect, specifically depression and anxiety symptoms, and tobacco, alcohol, and marijuana will connect with externalizing symptoms, specifically impulsivity and conduct disorder symptoms. We also expect differences in network structure by gender, with men experiencing greater connection between substance use and externalizing symptoms and women with greater connection between substance use and negative affect symptoms.

METHODS

Setting

Wave 1 adult data (N=32,320) from the Population Assessment of Tobacco and Health (PATH) study were used.[71] These data are cross-sectional and were collected between September 2013 and December 2014. PATH is a nationally representative longitudinal cohort study of the civilian, non-institutionalized household population of the U.S., and participants engaged in all levels of tobacco use ranging from never using tobacco to frequent use.

The weighted response rate among participants was 74.0% for Wave 1.[73] Participants responded to tobacco-specific items including tobacco-use patterns, risk perceptions and attitudes towards current and newly emerging tobacco products, tobacco initiation, cessation, relapse behaviors, and health outcomes.[72] Participants also responded to non-tobacco items (e.g., media use, peer and family influences, health effect outcomes, and industry advertising and promotion).[72]

Study representativeness

Participants with missing data on the substance use, negative affect, and externalizing measures were not included in the analysis (N= 2,109). Survey respondents of the analytic sample endorsed significantly greater substance use overall, negative affect symptoms, and externalizing symptoms (except for fighting) compared to those not included in the analytic sample. The participants in the analytic sample were more likely to be Non-Hispanic white, men, aged 25-54 with higher levels of education and annual household income than those who were missing.

Measures

Past Month Tobacco and Substance Use. Six substance use categories were used in this study: exclusive cigarette, exclusive e-cigarette, dual cigarette and e-cigarette, alcohol, marijuana, and prescription drugs not prescribed (PDNP) including painkillers, sedatives, tranquilizers. The summary of past-month substance use is described in Table 4.2. Current dual cigarette and e-cigarette use were identified if the respondent indicated they were a current cigarette and current e-cigarette user. Current alcohol, marijuana, and PDNP was endorsed if the respondent indicated ever using the substance and has used the substance within the past 30 days. Only past month or current use of the substances was considered (coded as 1, else = 0) to reduce the potential for recall bias and ensure for accurate overlap with negative affect and externalizing symptoms occurring in the same time frame. These substance use variables were nodes in the networks.

Table 4.2: Summary of Past-Month Substance Use and Symptoms of Negative affect and Externalizing Disorders

Past-Month Tobacco and Substance Use	**Variable Definition**
Exclusive CIG (or CIG)	Ever smoking a cigarette (even one or two puffs), has smoked at least 100 or more cigarettes in his or her entire life, and now smokes cigarettes every day or some days, while also excluding the current use of e-cigarettes
Exclusive ECIG (or ECIG)	Ever using an e-cigarette (even one or two puffs), ever smoked e-cigarettes fairly regularly, and now uses e-cigarettes every day or some days, while also excluding the current use of cigarettes
Dual CIG + ECIG	That they were a current cigarette and current e-cigarette user
Alcohol	Ever using alcohol and has used alcohol within the past 30 days
Marijuana	Ever using marijuana and has used marijuana within the past 30 days
PDNP	Ever using prescription drugs not prescribed (PDNP) (i.e., painkillers, sedatives, and tranquilizers) and has used PDNP within the past 30 days
Past-Month Negative affect Symptoms*	**The last time you had significant problems with:**
Depressed	Feeling trapped, lonely, sad, blue, depressed, or hopeless about the future
Sleeping	Sleep trouble such as bad dreams, sleeping restlessly or falling asleep during the day
Anxious	Feeling very anxious, nervous, tense, scared, panicked or something bad was going to happen
Distressed/Past	Becoming very distressed and upset when something reminded you of the past
Past-Month Externalizing Symptoms*	**The last time you did the following two or more times:**
Lied	Lied or conned to get things you wanted or to avoid having to do something
Attention	Had a hard time paying attention at school, work or home
Listening	Had a hard time listening to instructions at school, work or home

Bully	Were a bully or threatened other people
Fights	Started physical fights with other people
Restless	Felt restless or the need to run around or climb on things
Answered	Gave answers before the other person finished asking the question
* The items selected to identify negative affect and externalizing symptoms from the GAIN-SS instrument are ordinal and measures people across four times periods: past month, 2 to 12 months, over a year ago, and never. Participants indicating that they experienced a symptom within the past month were coded as 1. Participants indicating that they experienced the symptom 2 to 12 months ago, over a year ago, and never were coded as 0.	

Past Month Negative Affect and Externalizing Symptoms. Negative affect and externalizing symptoms were measured using the Global Appraisal of Individual Needs—Short Screener (GAIN-SS).[73] The summary of past-month negative affect and externalizing symptoms are described in Table 4.2. The items selected to identify negative affect and externalizing symptoms from the GAIN-SS instrument are ordinal and measures people across four times periods: past month, 2 to 12 months, over a year ago, and never. Participants indicating that they experienced a symptom within the past month were coded as 1. Participants indicating that they experienced the symptom 2 to 12 months ago, over a year ago, and never were coded as 0. Only past month or current negative affect and externalizing symptoms were considered reducing the potential for recall bias and ensure accurate overlap with substance use occurring in the same time frame. The negative affect and externalizing symptoms, along with the substance use variables, were nodes in the networks.

Covariates. Networks were stratified by gender to confirm significant differences in comorbidity networks by gender. Men and women experience substance use and mental disorders differently; therefore, it is important to test these differences by stratifying the networks. Previous work shows that women are more likely to experience negative affect symptoms while men are more likely to experience externalizing symptoms.[41,123] Men also tend to participate in substance use more regularly than women and experience substance use problems at twice the rate as females.[123] Therefore, it is important to determine how these comorbidities present by gender.

Statistical Analysis

Summary Statistics. Summary statistics were generated for the sample using PROC SURVEYFREQ in SAS 9.4 to account for complex sampling design.

Network Analysis. A network model can support a deeper understanding of comorbidity because it conceptualizes symptoms as mutually interacting, often reciprocally reinforcing elements of a complex network.[46] The network approach is based on the idea that comorbidities arise from shared symptoms between disorders which can capture complexity and individual variation in psychopathology.[49] The network approach naturally accommodates comorbidities as a central part of its theory.[50] In the network approach, comorbidity represents causal relationships between symptoms in which pathways can bridge symptoms that are part of multiple disorders.[46] Using a network model, symptoms, rather than disorders, are considered within the network structure. Rather than the disorder acting as the underlying cause of all symptoms, it is the symptoms that mutually interact and set a person into a disordered mental health state.

Within a network model, the symptoms make up a comorbid network structure of several symptoms that is specific to the person. This model conceptualizes how symptoms of different disorders function together specifically to produce a comorbid disordered state. The network approach explains the co-occurrence of mental disorder symptoms, including substance use behaviors, as resulting from direct interactions between these symptoms.[50] In network analysis, the term *interaction* is used to explain the reciprocal action or influence of symptoms. In the context of network analysis, *interaction* is not used to test whether an effect can be greater than (positive interaction, synergism) or less than what we would expect (negative interaction, antagonism).[51] Patterns of symptom-symptom or symptom-behavior interactions can be explained using a network structure.[45]

An example of the use of a network model is detailed in Figure 4.1 to summarize comorbidity of symptoms for SUD and depression. The network model of SUD and depression is made of symptoms denoted as nodes (circles) and the associations between the symptoms denoted as edges (lines connecting nodes). Every node in a network is connected by edges. Edges represent the interactions between the nodes. Nodes that directly activate each other (i.e., are associated with one another) are connected with an estimated edge, while nodes that do not directly activate each other are not. This figure details a directed network where arrows are directed from one node to another, indicating that one symptom can lead to the activation of another. Depression symptoms (red) are clustered together to the left of the network. SUD symptoms (blue) are clustered together to the right of the network. Insomnia and weight loss (in purple) are symptoms that occur in both depression and SUD and act as bridges between the disorders. The positioning and the distance between the symptoms/nodes within the network have implications for the comorbidity structure of depression and SUD.

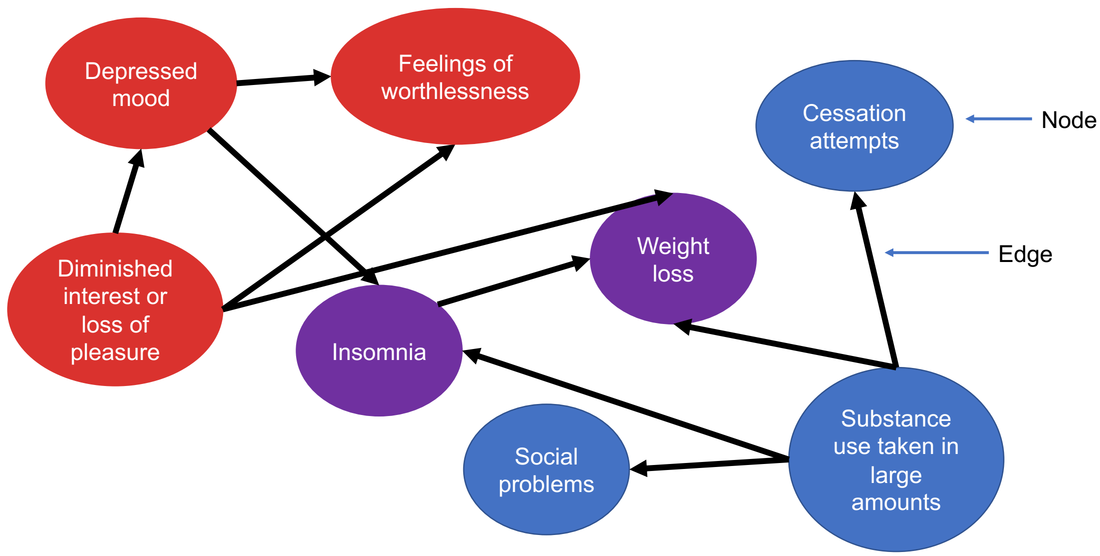


Figure 4.1: Network Model of Depression and SUD symptoms.

Overall Network Modeling Strategy. Two sets of network analyses were completed to evaluate the connections between substance use behaviors and negative affect/externalizing symptoms. The first analysis estimated network models in the entire sample and also tested for the consistency of network connections by gender (i.e., network comparison tests). The second set of network analyses consisted of model evaluation to establish the nodewise metrics (i.e., centrality and predictability), and accuracy/stability of the network models as detailed below. This second set of analyses is common to all networks, regardless of the presence or absence of gender differences.

Network Model Estimation. All networks were estimated using an Ising Model in R 3.6.0 using the *IsingFit* package.[193] Ising model selection uses the Extended Bayesian Information Criteria (EBIC) to measure model parsimony for moderate sample sizes and for a high number of variables by accounting for the number of unknown parameters and the complexity of the model space.[194–196] Models determined to best explain the data using EBIC were interpreted for relevant relationships.[193,197] Edges between two nodes were estimated at most pairwise, after adjusting for all other substance use, negative affect, and externalizing variables.[196] Edges were compared against each other to determine strength. Networks were visualized using the *qgraph* package.[198] Blue edges illustrate positive partial correlations; red edges illustrate negative partial correlations. The wider the edge, the stronger the correlation.

The Ising model contains two node-specific parameters: an interaction parameter and a node parameter. The interaction parameter, $\beta_{jk,}$ represents the strength of the interaction between variables j and k. The node parameter, τ_j, represents the autonomous disposition of the variable to take the value of one, regardless of neighboring variables. The model estimates these parameters with logistic regressions, iteratively, (i.e., one variable is regressed on all others).[196] The conditional probability of Xj given all other nodes $X\backslash j\ is$ given by:

$$\mathrm{P}\theta\left(\mathrm{x_j}\middle|\mathrm{x_{Jy}}\right) = \frac{\exp\left[\tau_j x_j + x_j \ \sum_{k\in V\backslash j}\beta_{jk}x_k\right]}{1+\exp\left[\tau_j + \ \sum_{k\in V\backslash j}\beta_{jk}x_k\right]} \tag{10}$$

where $x = (x_1, x_2, \ldots, x_n)$ and $x_i = 0 \; or \; 1$. The node parameter or intercept is τ_j , the threshold of the variable. The interaction parameter or slope is $\beta_{jk,}$ the connection strength between the relevant nodes.[196]

A network approach was conducted to estimate the edges, or connections between the nodes (i.e., denoted as lines between the nodes and are called edges) as partial correlations among a set of binary items (i.e., current substance use behaviors [exclusive cigarette, exclusive e-cigarette, dual cigarette and e-cigarette, alcohol, marijuana, and PDNP], four negative affect symptoms, and seven externalizing symptoms).[34,196,199]

<u>Network Comparisons to Test for Gender Differences.</u> Gender differences across networks were evaluated using two approaches. First, visual comparisons using an average layout established differences in the magnitude and direction (i.e., positive or negative) of edge-weights between nodes. Second, three tests of network invariance were used to test significant differences in network models by gender: global strength invariance, network structure invariance, and edge strength invariance.

Global strength invariance. The global strength invariance hypothesis tested whether the overall level of connectivity in a network was identical between men and women. The global strength invariance hypothesis tests the weighted absolute sum of all edges in the networks (i.e., S or the sum of the unique variance in the network).[200] The null hypothesis specifies that the connectivity for a network in men (ω_{ij}^1) and women (ω_{ij}^2) are equal: $H_0\colon \sum_{i=1}^{p}\sum_{j>i}\left|\omega_{ij}^1\right| = \sum_{i=1}^{p}\sum_{j>i}\left|\omega_{ij}^2\right|$, where ω_{ij}^G is the edge-weights between nodes i and j of network G.

For all $i < j$, the distance S between two networks if defined as:

$$S\left(\omega_{ij}^{1}, \omega_{ij}^{2}\right) = \left|\sum_{i=1}^{p}\sum_{j>i}\left(\left|\omega_{ij}^{1}\right| - \left|\omega_{ij}^{2}\right|\right)\right| \qquad (11)$$

The test of global strength invariance was accomplished through permutation testing procedure as implemented in the *NetworkComparisonTest* package[201] to statistically assess the difference in global strength by gender. Briefly, permutation testing was conducted by repeatedly rearranging the data and calculating the test statistic of each permutation.

Network structure invariance. A test of the network structure invariance hypothesis was conducted to determine whether network structures were identical by gender. This test was conducted by comparing the maximum differences in the edge weights between all nodes in the networks.[200] The null hypothesis that specifies all edges are equal is: $H_0\colon \Omega^1 = \Omega^2$, where Ω^G denotes a symmetric $p\ x\ p$ matrix that contains the edge weights of graphical model G. Therefore, gender difference in network structure would be detected if any of the edge weights between the nodes are determined to be different by gender.

The test of network structure invariance computed the maximum difference (M) in network edge-weights (ω_{ij}^{G} between nodes i and j of network G) by gender. The maximum statistic provides the largest value among each element of a vector which contains the differences in unique edge weights of networks in men (ω_{ij}^{1}) and women (ω_{ij}^{2}). This is defined as:

$$M\left(\omega_{ij}^{1} - \omega_{ij}^{2}\right) = max_{ij}\left|\omega_{ij}^{1} - \omega_{ij}^{2}\right| \qquad (12)$$

for $i < j$ (i.e., the upper triangle of the matrix Ω). This metric functioned as the test statistic, and followed the same permutation procedure used to test the global strength invariance.

Edge strength invariance. A test of the edge strength invariance hypothesis was conducted to determine if a specific edge between two nodes was equally strong by gender. Edge strength is also referred to as the edge weight, quantified as the magnitude of an edge. This is the magnitude of association between two nodes. The null hypothesis for this test is $H_0: \omega_{ij}^1 = \omega_{ij}^2$. This was assessed by taking the absolute difference in edge strength (ω_{ij}) between two nodes ($i \text{ and } j, for\ i < j$) of interest then testing differences between nodes across all other node combinations in the networks[200]:

$$E\left(\omega_{ij}^1, \omega_{ij}^2\right) = |\omega_{ij}^1 - \omega_{ij}^2| \tag{13}$$

<u>Network Model Evaluation</u>. Once network models were produced, the network structure was detailed across four categories: centrality (i.e., the influence of a node in a network), nodewise predictability (i.e., how well a given node can be predicted by all other nodes it is connected to in the network), model accuracy (i.e., the degree to which the model correctly describes the data), and model stability (i.e., the degree to which network estimates are expected in other samples). The aggregate evaluation of these edge-related metrics provides additional detail regarding how nodes within a network connect with one another and the degree to which a given network model is expected to consistently explain the data.

Centrality. Three centrality metrics (closeness, betweenness, and node strength) were computed for all three networks (full sample, men only, women only) in order to detail how nodes (i.e., substance use behaviors and negative affect-externalizing symptoms) interact with one another within a network.

Closeness quantifies how well a node is indirectly connected to other nodes.[197] Closeness is a measure of reach or importance of an individual node, based on the number of connections of that node, localized to that node. It considers the indirect ties to other nodes in addition to immediate connections. The closeness of a node is the reciprocal of the sum of the shortest path distances from the node to all n-1 other nodes. The higher closeness centrality, the shorter reach to other symptoms in the network meaning the more connected the symptom is to other symptoms in the network.

Betweenness refers to how critical a node is to a network as a bridging node to all other nodes in the network. It quantifies the number of times a node act as a bridge between the shortest path of two other nodes.[197] Betweenness is a measure of centrality based on the shortest path length connecting any two nodes. For a given node, betweenness is the sum of the fraction of all possible shortest paths that pass through that node. The more of these shortest paths that go through a node, the higher their betweenness centrality. Betweenness identifies bridges or go-betweens to identify other symptoms that may be key players in the comorbidity network.

The **node strength** also known as degree of a node is the number of edges that touches that node. It quantifies how well a node is directly connected to other nodes.[197]

Nodes that have many edges would be considered to have a high node strength because it indicates more connectedness to other nodes.

All centrality estimates were standardized using z-scores in order to compare the metrics across the networks. Additional details for each centrality metric calculated from the example network model visualized in Figure 4.2 is summarized in Table 4.3.

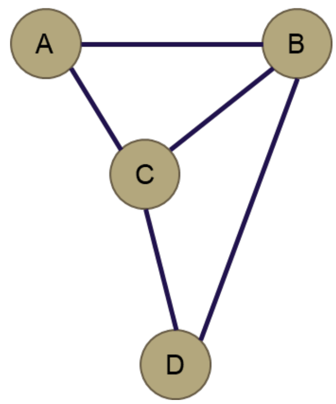


Figure 4.2: Example Network Model

Table 4.3: Calculations, Examples and Interpretations of Node Centrality Metrics Derived from Example Network Model

Metric	**Calculation**	**Figure Example**	**Interpretation**
Closeness	Reciprocal of the sum of the shortest path distances from the given node to all n-1 other nodes. Calculated as: 1/(sum of shortest path distances from given node/3).	Values for closeness centrality for each node: A = 0.75, B = 1.0, C = 1.0, and D = 0.75. Sum of *shortest* paths for node A: (A—B) = 1 (A—C) = 1 (A—D) = 2 Total = 4 Closeness for node A is calculated: 1/[4/3] = 0.75.	Nodes B and C have the highest closeness value because they only have a shortest path of 1 to all other nodes in the graph. Nodes A and D have a shortest path of 2 to one another.
Betweenness	For a given node, the sum of the *fraction* of all possible shortest paths that pass	Betweenness centrality is as follows: A = 0.0, B = 0.166, C = 0.166, and D = 0.	Nodes B and C have higher betweenness centrality compared to nodes A and D. Nodes B and C are

	through that node.	All possible shortest paths: (A—B), (A—C), (A—C—D), (A—B—D), (C—D), (B—D), and (B—C). Only the A—D path that includes C is counted in the calculation of for betweenness. Betweenness for node C is = 1/6 = 0.166.	between nodes A and D. One must go through B and C to connect to nodes A and D.
Strength	Count of the number of edges that touches a given node.	Strength A = 2 B = 3 C = 3 D = 2	Nodes B and C have the highest strength and have greater connectedness compared to nodes A and D.

Nodewise Predictability. In addition to evaluating network structure as summarized above, it is important to also analyze nodewise predictability. The concept of predictability complements the interpretation of network structures. Specifically, nodewise predictability quantifies how well a given node can be predicted by all the other adjacent nodes it is connected to in the network.[202] Estimating predictability is crucial for three reasons. First, it considers how much of the variance at a given node is explained by the edges connected to it. Consequently, an edge that explains 50% of the variance of a node will be considered more important than an edge that explains 0.5% of the variance of the node. Second, the predictability at one node can provide an expectation regarding the extent to which a specific node is influencing another node. Therefore, nodewise predictability can produce expectations regarding whether a node can be influenced by intervening on the nodes that are connected to it. Third, estimates of predictability across nodes indicates whether a network (or portion of the network) is influenced by itself through strong interactions between nodes (i.e., high predictability) or whether it is mostly determined by other factors that are not included in the network (i.e., low predictability).[202] Consequently, interpretation of nodewise predictability can yield important insight about the whole network in addition to those related to network structure (i.e., centrality).

Estimation of Predictions in Network Model. Nodewise predictability is estimated by computing the mean of the conditional distribution of a specific node given all its neighboring nodes.

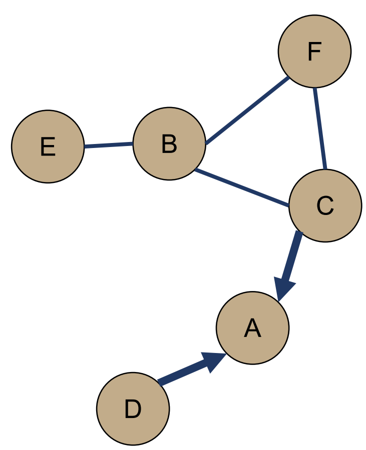


Figure 4.3: Six-node network to determine nodewise predictability of Node A

For example, in a six-node network (Figure 4.3) consisting of nodes A, B, C, D, E and F, an estimated network is produced and the probability of observing a node (A) within the network given the nodes that are connected to it (C and D) is estimated as:

$$P(A = k|C, D) = \frac{\exp\{\mu_k\}}{\sum_{l=1}^{K} \exp\{\mu_k\}} \tag{14}$$

where *A* is a node measured as a binary variable and *k* indicates the category, *K* is the number of categories that the node (for a binary node, *K* = 2). *C* and *D* represent the nodes adjacent to *A*. μ_k represents the mean of the conditional distribution at node *A* and is estimated as:

$$\mu_k = \beta_{0k} + \beta_{Ck}C + \beta_{Dk}D \tag{15}$$

where $\beta_{Ck}C$ and $\beta_{Dk}D$ represent the edge weights of nodes C and D on node A. Therefore, the probability of observing a specific value at node A depends on the influence of nodes C and D on node A.

Quantifying Predictability Using Categorical Data. The estimation of nodewise predictability for categorical data establishes how close estimated predictabilities at each node compared to the observed values in the data. The predictability of a network that uses continuous data is estimated as a proportion of the variance for the network model that is explained by the predictability measure, measured as:

$$R_A^2 = 1 - \frac{var(\hat{A} - A)}{var(A)} \tag{16}$$

where var is the variance, $\hat{A}$ is a vector of predictions for A as defined in equation 14, and A is the vector of observed values in the data. All variables are centered to reflect a mean of zero in order to remove the possibility that an intercept from a given node can affect the predictability measure. When R^2 = 1, a node can be perfectly predicted by its neighboring nodes. In comparison, when R^2 = 0, a node cannot be predicted by all its neighboring nodes in the network.

The estimation of nodewise predictability for categorical data differs from that of continuous data. In particular, the use of categorical data necessitates the estimation of a value of “normalized accuracy” which parallels the estimation of nodewise predictability by centering the mean to be equal to zero. Normalized accuracy is estimated by removing the marginal effects at each node (i.e., probabilities of the categories when ignoring all other variables) to determine how well a given node was predicted by all other nodes in the network. The utility of normalized accuracy can be exemplified using a hypothetical sample with 100 observations, where ten observations have a score of zero and 90 observations with a score of one. The marginal probabilities for the node are $p0$ = 0.1 and $p1$ = 0.9. Further, if all other nodes in the network do not contribute to predicting whether a node has a value of 1 or 0, it is possible to predict a value of one for all cases. Subsequently, a 90% correct classification would be estimated. However, this is misleading and results in an inflated estimate of predictability because nothing can be predicted by all the other nodes. Normalized accuracy is estimated to remove the accuracy that occurs from the “trivial” prediction from other nodes using marginal of the variable ($p1$ = 0.9) alone. Therefore, normalized accuracy is estimated as the ratio between the additional accuracy due to the remaining nodes in the network and one minus the accuracy of the node alone:

$$A_{norm} = \frac{A - max\{p_0, p_1, \dots, p_m\}}{1 - \max\{p_0, p_1, \dots, p_m\}} \quad (17)$$

where $A = \frac{1}{n}\sum_{i=1}^{n} \mathbb{I}\,(y_i = \hat{y}_i)$, which reflects the proportion of correct predictions (accuracy/correct classification). Additionally, p_0, p_1,… p_m represent the marginal probabilities of the categories for a node where $\mathbb{I}$ is the indicator function for the event $F_i = \hat{F}_i$. For binary variables, the marginal probabilities are defined as p_0 and $p_1 = 1 - p_0$. Therefore, A_{norm} indicates how much the node of interest can be predicted by all other nodes in the network, beyond what is trivially predicted by the marginal distribution. When $A_{norm} = 0$, none of the other nodes contribute to the marginal in predicting the node of interest. When $A_{norm} = 1$, all other nodes perfectly predict the node of interest.[202]

Interpreting and Visualizing Predictability Using Categorical Data. It is valuable to interpret estimates of A and A_{norm} via network model visualization as a multi-colored ring surrounding a node.

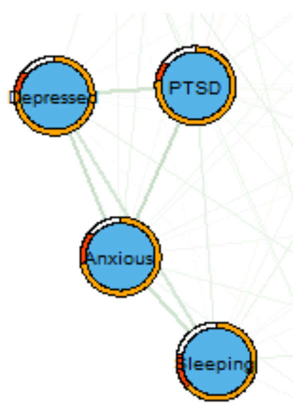


Figure 4.4: Visualizing the Nodewise Predictability of Categorical Data

In Figure 4.4, the accuracy of a node from an intercept-only model (i.e., a model estimating node predictability without the influence of the other nodes) is represented by the orange part of the ring. The red part of the ring reflects the additional accuracy of the node that is achieved by all the remaining nodes. The sum of the red and orange sections represents A, or the accuracy of the full model. A_{norm} is the normalized accuracy that is estimated as the ratio between the additional accuracy due to the remaining nodes in the network and one minus the accuracy of the node alone (white + red sections).

Model estimation to compute nodewise predictability was completed using the *mgm* package.[203] The estimated models for the nodewise prediction were the same as those estimated using the *IsingFit* package because the approach was the same: neighborhood selection-based method to estimate the binary-valued Ising model.[202] The *predict* ( ) function was used to compute the predictability for each node in the network, specifying accuracy/correct classification, normalized accuracy, and the accuracy of the intercept (marginal) model to visualize the decomposition of total accuracy.

<u>Accuracy and Stability.</u> Typical parameter estimates in a model provide some indication of the degree of uncertainty around the estimate (e.g., standard error or confidence intervals). However, such estimates are not automatically generated from a network model. Consequently, a preliminary network model cannot provide insight regarding the uncertainty of the parameters estimated. Additional calculations of accuracy and stability related to the network parameter estimates can be produced to establish confidence in the model's ability to estimate the true value from the data.

Tests of accuracy and stability established confidence in the network model's ability to generate the accurate estimate, allowing for appropriate interpretation of results. Therefore, network accuracy and stability tested the inferences about the network structure and centrality indices. Accuracy and stability were calculated in three steps: (1) estimating the accuracy of estimated edge-weights, by drawing bootstrapped confidence intervals, (2) investigating the stability of centrality indices, and (3) testing whether edge-weights and centrality estimates for different variables differ from each other using a bootstrapped difference test.[197]

Edge-weight accuracy (i.e., the accuracy of estimated network connections) was assessed by obtaining confidence intervals around the estimated edge-weights using non-parametric bootstrapping (*bootnet* package[204]).[197] Confidence intervals generated around the estimated edge-weights identifies the precision of the edge-weight and whether the confidence intervals overlap with the bootstrapped confidence intervals of other edge-weights. An edge-weight with high accuracy has a narrow confidence interval that does not overlap with the confidence interval of other edges. The non-parametric bootstrapping evaluated whether edge-weights for the variables in the network differed from each other in three steps: (1) estimate the difference between the bootstrap value of two edge-weights using non-parametric bootstrap, (2) construct bootstrapped confidence interval around difference scores, and (3) test the model with estimated connections against a model of the null hypothesis to establish whether edge-weights differed from one another by checking if zero was in the bootstrapped confidence interval.[197] Results from the edge-weight accuracy test are visualized as a plot.

Centrality stability refers to the degree to which an estimate of a centrality metric (i.e., closeness, betweenness, or strength) is consistent after re-estimating the network in other samples with characteristics similar to the study sample. Centrality stability can be estimated for each metric separately, and answers the question: “Does the order of the centrality indices remain the same after re-estimating the network with a smaller sample?”

The stability of the centrality indices is quantified as a correlation stability coefficient (i.e., CS-coefficient). The CS-coefficient represents the maximum proportion of observations that can be dropped from the original sample. The higher the CS-coefficient the greater the stability of the centrality indices. A CS-coefficient should not be below 0.25 and preferably above 0.5 for appropriate interpretation of the results. In step 2, centrality stability was investigated by using a case-dropping subset bootstrap procedure where a centrality metric was obtained for the dataset. Then, networks were re-estimated after subsetting the sample to determine if the CS-coefficient for the centrality indices retained a correlation of 0.5 in at least 95% of the samples.[197]

The estimation of a bootstrapped difference test (nonparametric bootstrap) was used to test the degree to which edge and centrality estimates differ from each other across variables. A bootstrapped difference test uses the difference between the bootstrapped value of one edge weight/centrality and another edge weight/centrality using non-parametric bootstrap and constructs a bootstrapped confidence interval around difference scores. The bootstrapped difference test identifies whether (1) a specific edge (e.g., A—B) is significantly larger than another edge (e.g., A—C) and (2) the centrality of node A is significantly larger than the centrality of node B. If the confidence interval generated from the bootstrapped difference test includes zero than the two edges or two centrality metrics of interest are considered to not differ significantly from each other. This bootstrapped difference test was done for the estimated edge-weights and node strength.

Missingness and Complex Sampling Design. Missing data were removed using listwise deletion (N = 2,109). Complex sampling design was not accounted for in the estimation of the network models.

RESULTS

Summary statistics

32,320 participants were included in the overall sample, and 30,211 participants had complete data for all nodes. The sample was 51.9% female and 66.0% Non-Hispanic White. Age was evenly distributed across the sample. Most of the sample had at least a GED or high school education (88.4%) and an annual household income of more than $25,000 (65.9%). Past-month alcohol use was most frequently reported (52.4%), followed by CIG use (16.6%), and marijuana use (7.1%, Table 4.4). Sleep trouble was the most common past-month negative affect symptom reported (26.7%), followed by feeling very anxious (16.1%), feeling depressed (13.4%), and becoming distressed about the past (12.5%) (Table 4.4). Giving answers before the other person finished asking the question was the most common past-month externalizing symptom (32.0%), followed by having a hard time paying attention (14.6%) and listening to instructions (10.4%).

Table 4.4: Demographic Characteristics of the Overall Sample by Gender

	Men (N=16306, 48.1%) N (Weighted %)	**Women (N=15980, 51.9%)** N (Weighted %)	**Overall (N=32320, 100%)** N (Weighted %)
Age*			
18-24	4609 (13.6)	4495 (12.5)	9110 (13.0)
25-34	3232 (18.6)	3103 (16.9)	6337 (17.7)
35-44	2448 (16.4)	2478 (16.7)	4930 (16.5)
45-54	2428 (18.0)	2409 (17.9)	4846 (17.9)
55-64	2039 (16.7)	1929 (16.5)	3971 (16.6)
65+	1547 (16.8)	1558 (19.5)	3110 (18.2)
Race*			
Non-Hispanic White	9815 (66.8)	9467 (65.4)	19295 (66.0)
Non-Hispanic Black	2129 (11.0)	2364 (11.3)	4496 (11.2)
Non-Hispanic Other	1266 (7.8)	1162 (7.3)	2429 (7.5)
Hispanic Multiracial	2383 (12.4)	2429 (14.1)	4817 (13.3)
Education*			
Less than high school	2287 (12.1)	1938 (11.1)	4233 (11.6)
GED/High school graduate	5187 (30.6)	4570 (28.5)	9765 (29.5)
Some college (no degree)	5353 (29.6)	5942 (32.4)	11300 (31.0)
Bachelor's degree	2237 (17.6)	2260 (18.1)	4498 (17.8)
Advanced degree*	1132 (10.1)	1176 (10.0)	2311 (10.1)
Annual household income*			
Less than $10,000	2519 (11.9)	3144 (15.3)	5668 (13.7)
$10,000- $24,999	3287 (19.4)	3477 (21.4)	6768 (20.4)
$25,000- $49,999	3453 (23.3)	3214 (22.8)	6670 (23.0)
$50,000- $99,999	3338 (25.8)	2797 (24.0)	6140 (24.9)
$100,000 or more	2220 (19.6)	1692 (16.5)	3914 (18.0)
Past month tobacco and substance use			
CIG*	5435 (19.0)	4942 (14.3)	10381 (16.6)
ECIG*	299 (1.0)	278 (0.8)	578 (0.9)
Dual CIG + ECIG*	533 (1.7)	463 (1.3)	996 (1.5)
Alcohol*	9550 (56.3)	8231 (48.8)	17787 (52.4)
Marijuana*	2611 (9.1)	1780 (5.3)	4392 (7.1)
PDNP *	914 (4.7)	1035 (5.4)	1950 (5.1)
Past month negative affect symptoms			
Depressed*	2513 (12.1)	3178 (14.6)	5692 (13.4)
Sleeping*	4313 (24.6)	5249 (28.7)	9564 (26.7)
Anxious*	2931 (14.0)	3931 (18.0)	6864 (16.1)
Distressed/Past*	2459 (11.2)	3143 (13.6)	5605 (12.5)
Past month externalizing symptoms			
Lied*	1763 (8.2)	1480 (6.0)	3245 (7.1)
Attention*	2712 (13.9)	3114 (15.2)	5831 (14.6)
Listening	1976 (10.2)	2148 (10.6)	4128 (10.4)
Bully	384 (1.8)	352 (1.6)	737 (1.7)
Fights*	258 (0.9)	146 (0.5)	404 (0.7)
Restless*	1661 (7.3)	1292 (5.3)	2953 (6.2)
Answered*	5459 (30.5)	5937 (33.4)	11399 (32.0)

*Significantly different between men and women at p = 0.05 level.

Network Comparisons by Gender

The range of the magnitudes of the tetrachoric correlations between were similar for men and women: substance use and negative affect symptoms r_{Men} = 0.001-0.33, r_{Women} = 0.07-0.30, substance use and externalizing symptoms r_{Men} = 0.01-0.31, r_{Women}= -0.03-0.36, and negative affect and externalizing symptoms r_{Men} = 0.34-0.61, r_{Women}= 0.32-0.61 (Table 4.5). This suggested few gender differences between comorbid substance use and negative affect/externalizing symptoms.

Table 4.5: Tetrachoric correlations for Men (N =16,306) and Women (N =15,980)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. CIG	1	-0.94*	-0.999*	0.03*	0.26*	0.20*	0.20*	0.18*	0.21*	0.21*	0.10*	0.09*	0.12*	0.16*	0.17*	0.07*	0.06*
2. ECIG	-0.94*	1	-0.958*	-0.02	0.01*	0.10*	0.08*	0.10*	0.10*	0.08*	0.03	0.09*	0.14*	0.09	-0.03	0.04	0.02
3. Dual CIG + ECIG	-0.999*	-0.92*	1	0.06*	0.21*	0.21*	0.12*	0.15*	0.19*	0.15*	0.03	0.17*	0.14*	0.05	0.09	0.09*	0.09*
4. Alcohol	0.04*	0.01	0.06*	1	0.38*	0.04	0.06*	0.07*	0.07*	0.07*	0.12*	0.12*	0.08*	0.07*	0.08*	0.15*	0.23*
5. Marijuana	0.27*	0.10*	0.18*	0.32*	1	0.28*	0.30*	0.20*	0.29*	0.30*	0.35*	0.25*	0.22*	0.27*	0.36*	0.26*	0.19*
6. PDNP	0.21*	0.06	0.21*	0.06*	0.30*	1	0.32*	0.33*	0.34*	0.34*	0.26*	0.25*	0.26*	0.19*	0.28*	0.17*	0.18*
7. Depressed	0.16*	0.001	0.16*	0.07*	0.25*	0.28*	1	0.70*	0.79*	0.75*	0.49*	0.58*	0.54*	0.45*	0.44*	0.46*	0.32*
8. Sleeping	0.12*	0.05	0.16*	0.10*	0.17*	0.32*	0.71*	1	0.73*	0.66*	0.42*	0.56*	0.54*	0.35*	0.36*	0.46*	0.35*
9. Anxious	0.16*	0.01	0.20*	0.06*	0.22*	0.33*	0.79*	0.74*	1	0.79*	0.49*	0.61*	0.58*	0.45*	0.47*	0.50*	0.37*
10. Distressed/Past	0.18*	0.05	0.22*	0.07*	0.26*	0.32*	0.77*	0.69*	0.81*	1	0.54*	0.57*	0.55*	0.46*	0.47*	0.49*	0.35*
11. Lied	0.07*	0.11*	0.12*	0.19*	0.30*	0.26*	0.50*	0.43*	0.53*	0.56*	1	0.53*	0.52*	0.52*	0.51*	0.40*	0.38*
12. Attention	0.03*	0.07*	0.17*	0.15*	0.23*	0.22*	0.58*	0.55*	0.61*	0.59*	0.57*	1	0.91*	0.47*	0.34*	0.52*	0.48*
13. Listening	0.07*	0.01	0.16*	0.07*	0.21*	0.23*	0.57*	0.55*	0.61*	0.59*	0.53*	0.91*	1	0.48*	0.38*	0.51*	0.47*
14. Bully	0.17*	0.07	0.14*	0.12*	0.27*	0.26*	0.45*	0.43*	0.53*	0.53*	0.54*	0.48*	0.50*	1	0.66*	0.38*	0.33*
15. Fights	0.17*	0.07	0.11*	0.05	0.31*	0.30*	0.36*	0.35*	0.42*	0.46*	0.46*	0.36*	0.40*	0.73*	1	0.46*	0.28*
16. Restless	0.06*	0.03	0.15*	0.11*	0.26*	0.16*	0.45*	0.43*	0.49*	0.47*	0.45*	0.53*	0.51*	0.45*	0.46*	1	0.49*
17. Answered	0.01	0.05	0.13*	0.22*	0.20*	0.14*	0.34*	0.36*	0.40*	0.39*	0.44*	0.49*	0.45*	0.42*	0.31*	0.52*	1

* Significant association at $p < 0.05$ level.
Note: Correlations for women are on the top diagonal. Correlations for men are on the bottom diagonal.

The average layouts between networks for men and women did not indicate substantial differences by gender (Figure 4.5). The tobacco cluster was also quite similar for both men and women. The edge weight between “Bully” and “Fights” was thicker (i.e., greater) in the male network compared to the female network. Some nodes had more or fewer edges, depending on the network. The following nodes had more edges in women: PDNP, feeling depressed, feeling anxious, attention difficulties, fighting, and restlessness. The following nodes had more edges in men: CIG, alcohol, feeling distressed about the past, lying, listening difficulties, and giving answers before person finished asking the question.

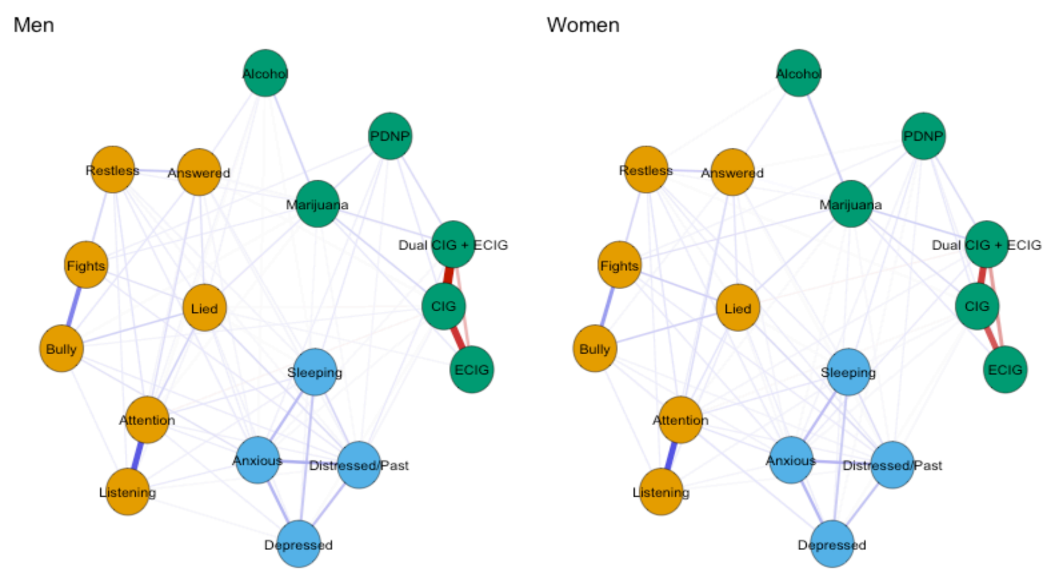


Figure 4.5 Network Structure by Gender

The edge-weights (EW) were significantly different ($p < 0.05$) between men and women for eight edges: (1) alcohol—marijuana ($EW_{Men} = 0.87$, $EW_{Women} = 1.08$), (2) alcohol—sleeping problems ($EW_{Men} = 0.08$, $EW_{Women} = 0$), (3) marijuana—feeling anxious (EW_{Men} = 0, EW_{Women} = 0.15), (4) ECIG—lying (EW_{Men} = 0.25, EW_{Women} = 0), (5) alcohol—lying (EW_{Men} = 0.31, EW_{Women} = 0), (6) alcohol—attention difficulties (EW_{Men} = 0.25, EW_{Women} = 0), (7) lying—attention difficulties (EW_{Men} = 0.86, EW_{Women} = 0.56), alcohol—listening difficulties (EW_{Men} = -0.11, EW_{Women} = 0).

Despite some node-specific relationships that differed by gender, the overall structure of the networks (maximum difference = 1.33, p-value = 0.32) and the global strength (Men = 53.4, Women = 50.9, p-value = 0.46) did not significantly differ between men and women. Therefore, the overall structure and connectivity was not different across men and women, and we focus detailing overall network results for men and women together first, then we subsequently provide information for men and women separately to further detail these networks.

Overall Network

The overall network consisted of 17 nodes (Figure 4.6). The network had 94 non-zero edges out of 136 possible edges (density=0.691), indicating that 69.1% of possible connections were identified in the network. The figure below shows the estimated network structure of 6 substance use behaviors (in green), 4 negative affect symptoms (in blue), and 7 externalizing symptoms (in yellow). The network structure is an Ising model, which is a network of partial correlation coefficients. Especially strong connections emerged between the tobacco use nodes, between “Attention” and “Listening”, and “Fights” and “Bully”. The negative affect symptoms were positioned between the substance use behaviors and externalizing symptoms, with many of the nodes lying on the periphery of the network.

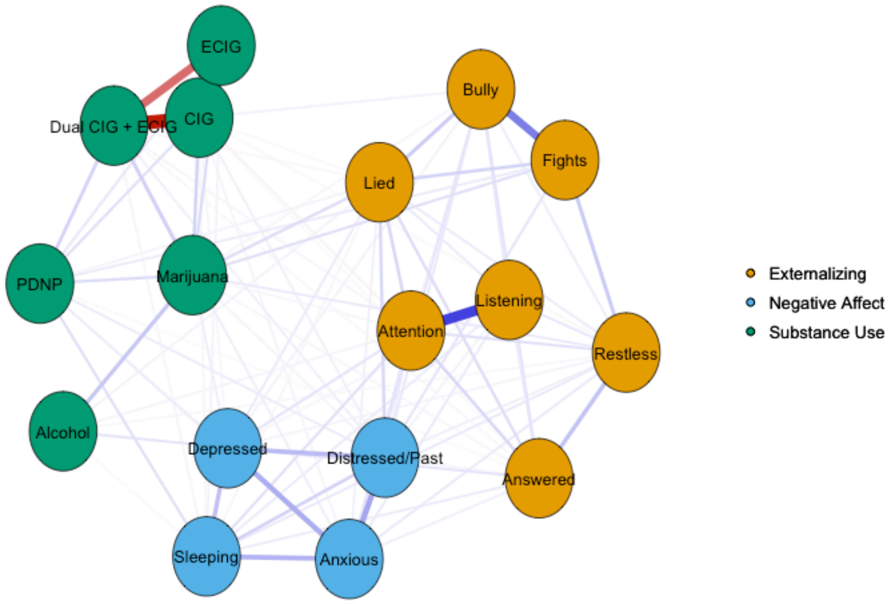


Figure 4.6: Overall Network of Substance Use, Negative Affect and Externalizing Comorbidity

The assessment of the accuracy of estimated network connections demonstrated that many edge-weights significantly differ from one-another (Appendix C, Supplemental Figure 4.1). Results from the edge-weights significant difference test for the overall sample network can be found in Appendix C, Supplemental Figure 4.2. Tobacco products were negatively associated with one another (CIG—ECIG = -4.74 [95% CI = -5.50; -3.98], dual CIG + ECIG—CIG = -4.60 [95% CI = -5.82; -3.39], dual CIG + ECIG—ECIG = -2.66 [95% CI = -4.11; -1.21]) (Appendix C, Supplemental Table 4.1).

Externalizing symptoms also demonstrated strong connections with one another. For example, attention difficulties had the strongest connection with listening difficulties

(EW = 3.47, 95% CI = 3.36; 3.58). Bullying was positively associated with fighting (EW = 2.40, 95% CI = 2.10; 2.70).

The connections between PDNP were strongest with negative affect symptoms. Specifically, the connections with the largest magnitudes were PDNP—sleeping problems (EW = 0.53, 95% CI = 0.40; 0.66), PDNP—feeling anxious (EW=0.31, 95% CI = 0.17; 0.46), and PDNP—feeling distressed about the past (EW = 0.31, 95% CI = 0.16; 0.45).

The connections between marijuana, alcohol, and PDNP use were strongest with externalizing symptoms. Specifically, the connections with the largest magnitudes were marijuana—lying (EW = 0.60, 95% CI = 0.49; 0.70), marijuana—fighting (EW = 0.54, 95% CI = 0.27; 0.81), alcohol—answered (EW = 0.48, 95% CI = 0.42; 0.53), marijuana—restlessness (EW = 0.37, 95% CI = 0.26; 0.49) and PDNP—fighting (EW = 0.36, 95% CI = 0.001; 0.72).

The investigation of the stability of centrality indices demonstrated that closeness (CS coefficient = 0.517) and strength (CS coefficient = 0.594) were stable enough for interpretation. The betweenness CS coefficient (0.206), however, was too low to interpret for the overall network (Appendix C, Supplemental Figure 4.3). Significant differences between node strength were also tested and are displayed in Appendix C, Supplemental Figure 4.4.

The centrality metrics are provided in the Table 4.6, and also depicted as z-scores in Appendix C, Supplemental Figure 4.5. “CIG” has the greatest strength in the network (strength = 2.39), followed by “Dual CIG + ECIG” (strength = 1.47), “ECIG” (strength = 0.92), and “Anxious” (strength = 0.38). Nodes with the greatest closeness centrality include “Lied” (closeness = 2.34), “Marijuana” (closeness = 1.46), “Fights” (closeness = 1.24), and “Bully” (closeness = 0.75). Alcohol was lowest for strength, and ECIG was lowest for closeness. Both are seen on the periphery of the network.

Table 4.6. Node Centrality Indices for the Overall Sample		
	Strength	Closeness
CIG	2.39	-1.00
ECIG	0.92	-1.51
Dual CIG + ECIG	1.47	-0.59
Alcohol	-2.01	-1.07
Marijuana	-0.18	1.46
PDNP	-1.02	-0.41
Depressed	-0.35	-0.59
Sleeping	-0.21	0.05
Anxious	0.38	-0.08
Distressed/Past	0.27	0.13
Lied	-0.11	2.34
Attention	0.70	0.49
Listening	-0.25	-0.68
Bully	-0.18	0.75
Fights	-0.51	1.24
Restless	-0.54	-0.003
Answered	-0.78	-0.53

Men-Only Network

The men-only network consisted of 17 nodes (N = 15,268) visualized in Figure 4.7. The network had 85 non-zero edges out of 136 possible edges. Figure 4.7 shows the estimated network structure of 6 substance use behaviors (in green), 4 negative affect symptoms (in blue), and 7 externalizing symptoms (in yellow). The network structure is an Ising model, which is a network of partial correlation coefficients. Similar connections emerged between the tobacco use nodes, between “Attention” and “Listening”, and “Fights” and “Bully”.

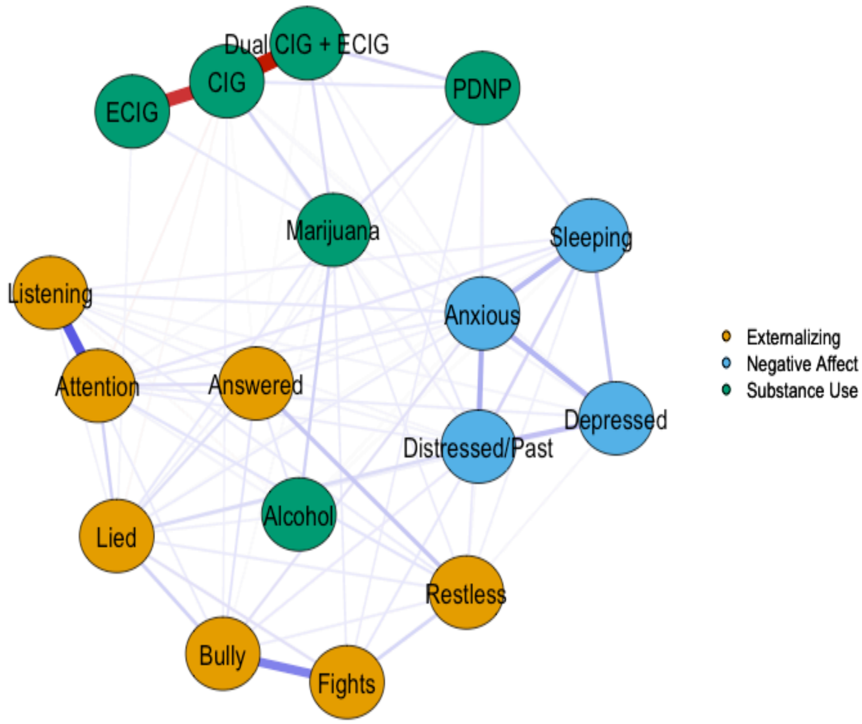


Figure 4.7: Men-Only Network of Substance Use, Negative Affect and Externalizing Comorbidity

The assessment of the accuracy of estimated network connections demonstrated that many edge-weights significantly differed from one-another (Appendix C, Supplemental Figure 4.6). Results from the edge-weights significant difference test for the overall sample network can be found in Appendix C, Supplemental Figure 4.7. Especially strong connections emerged among “Dual CIG + ECIG” and “CIG” (edge weight = -5.31), “CIG” and “ECIG” (edge weight = -4.16), “Dual CIG + ECIG” and “ECIG” (edge weight = -1.55), “Attention” and “Listening” (edge weight = 3.45), and “Bully” and

"Fights" (edge weight = 2.59). Other connections were absent like "Exclusive CIG" and "Sleeping" (edge weight = 0) (Appendix C, Supplemental Table 4.2).

The investigation of the stability of centrality indices demonstrated that strength (CS coefficient = 0.517) was stable enough for interpretation. Closeness (CS coefficient = 0.361) was lower than the preferred 0.50, but higher than 0.25. Closeness metrics should be interpreted with caution. The betweenness CS coefficient (0.128), however, was too low to interpret for the men-only network (Appendix C, Supplemental Figure 4.8). Significant differences between node strength were also tested and are displayed in Appendix C, Supplemental Figure 4.9.

The centrality metrics are provided in the Table 4.7, and also depicted as z-scores in Appendix C, Supplemental Figure 4.10. "CIG" had the greatest strength in the network (strength = 2.62), followed by "Dual CIG + ECIG" (strength = 1.16), "Anxious" (strength = 0.71), and "Attention" (strength = 0.67). Nodes with the greatest closeness centrality included "Distressed/Past" (closeness = 1.62), "Lied" (closeness = 1.57), and "Bully" (closeness = 1.15). Alcohol was lowest for all centrality metrics.

Table 4.7: Node Centrality Indices for the Men-Only Sample

	Strength	Closeness
CIG	2.62	-0.27
ECIG	0.07	-1.77
Dual CIG + ECIG	1.26	0.0004
Alcohol	-1.87	-1.91
Marijuana	-0.32	0.46
PDNP	-1.19	-0.55
Depressed	-0.39	-0.31
Sleeping	-0.37	0.27
Anxious	0.71	0.59
Distressed/Past	0.56	1.62
Lied	0.11	1.57
Attention	0.67	-0.01

Listening	-0.09	-0.76
Bully	-0.08	1.15
Fights	-0.49	0.65
Restless	-0.55	0.09
Answered	-0.66	-0.81

Women-Only Network

The women-only network consisted of 17 nodes (N = 14,925) visualized in Figure 4.8. The network had 84 non-zero edges out of 136 possible edges. Figure 4.8 shows the estimated network structure of 6 substance use behaviors (in green), 4 negative affect symptoms (in blue), and 7 externalizing symptoms (in yellow). The network structure is an Ising model, which is a network of partial correlation coefficients. Similar connections emerged between the tobacco use nodes, between “Attention” and “Listening.”

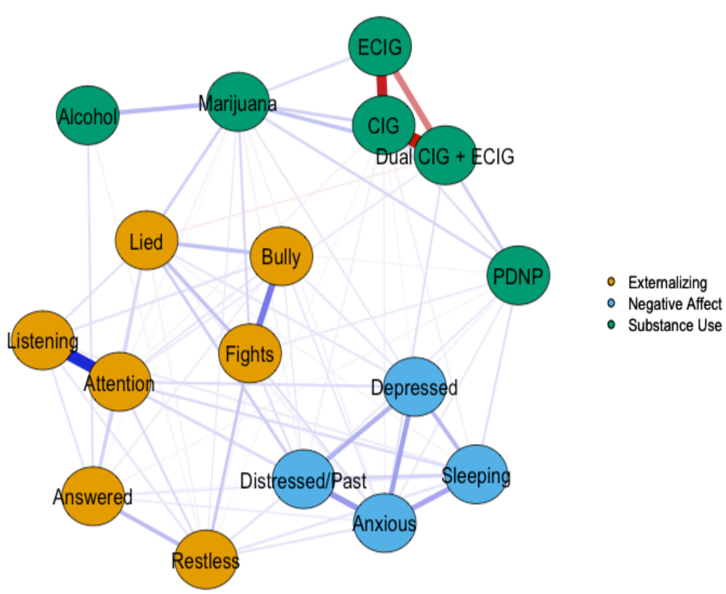


Figure 4.8: Women-Only Network of Substance Use, Negative Affect and Externalizing Comorbidity

The assessment of the accuracy of estimated network connections suggested that many edge-weights significantly differed from one-another (Appendix C, Supplemental Figure 4.11). Results from the edge-weights significant difference test for the overall sample network can be found in Appendix C, Supplemental Figure 4.12. Especially strong connections emerged among “Dual CIG + ECIG” and “CIG” (edge weight = -3.98), “CIG” and “ECIG” (edge weight = -3.48), “Dual CIG + ECIG” and “ECIG” (edge weight = -1.96), “Attention” and “Listening” (edge weight = 3.49), and “Bully” and “Fights” (edge weight = 2.04). Other connections were absent like “CIG” and “Alcohol” (edge weight = 0) (Appendix C, Supplemental Table 4.3).

Closeness (CS coefficient = 0.439) and strength (CS coefficient = 0.361) were lower than the preferred 0.50, but higher than 0.25. Closeness and strength metrics should be interpreted with caution. The betweenness CS coefficient (0.128), however, was too low to interpret for the women-only network (Appendix C, Supplemental Figure 4.13). Significant differences between node strength were also tested and are displayed in Appendix C, Supplemental Figure 4.14.

The centrality metrics are provided in Table 4.8, and also depicted as z-scores in Appendix C, Supplemental Figure 4.15. “CIG” had the greatest strength in the network (strength = 1.91), followed by “Dual CIG + ECIG” (strength = 1.35), “Attention” (strength = 1.13), and “Anxious” (strength = 1.03). Nodes with the greatest closeness centrality included “Lied” (closeness = 2.25), “Anxious” (closeness = 1.13), and “Marijuana” (closeness = 0.95). Alcohol was lowest for strength, and PDNP was lowest for closeness. Both are seen on the periphery of the network.

Table 4.8: Node Centrality Indices for the Women-Only Sample		
	Strength	Closeness
CIG	1.91	-0.34
ECIG	0.17	-1.34
Dual CIG + ECIG	1.35	0.38
Alcohol	-2.27	-0.94
Marijuana	0.08	0.95
PDNP	-1.11	-1.42
Depressed	-0.13	0.41
Sleeping	-0.29	0.01
Anxious	1.03	1.13
Distressed/Past	0.37	0.63
Lied	0.02	2.25
Attention	1.13	0.31
Listening	-0.05	-0.56
Bully	-0.48	-0.12
Fights	-0.23	0.76
Restless	-0.57	-0.97
Answered	-0.95	-1.15

Nodewise Predictability

Nodewise predictability results are summarized in Table 4.9 and are graphically depicted in Figure 4.9. The predictability measures accuracy/correct classification and normalized accuracy. The accuracy of the intercept (marginal) model was also used to estimate the decomposition of the total accuracy in the intercept model (Table 4.9, Accuracy of Intercept column, orange) and the contribution of other variables (Table 4.9, Correct Classification – Accuracy of Intercept Model column). Figure 4.9 visualizes the results from Table 4.9. The colors in the ring around the node indicate the accuracy of the intercept model (orange) and the total accuracy (orange plus red). The normalized accuracy is the ratio red/ (red + white).

Table 4.9. Nodewise Predictability Values

Variable	Total Accuracy (Accuracy/Correct Classification)	Normalized Accuracy	Accuracy of Intercept (Marginal) Model	Correct Classification – Accuracy of Intercept Model (Contribution from other variables)
CIG	0.687	0.037	0.675	0.012
ECIG	0.982	0.000	0.982	0
Dual CIG + ECIG	0.969	0.000	0.969	0
Alcohol	0.590	0.041	0.572	0.018
Marijuana	0.861	0.011	0.859	0.002
PDNP	0.938	0.001	0.938	0
Depressed	0.878	0.312	0.823	0.055
Sleeping	0.808	0.357	0.701	0.107
Anxious	0.876	0.422	0.785	0.091
Distressed/Past	0.883	0.330	0.825	0.058
Lied	0.904	0.048	0.899	0.005
Attention	0.908	0.498	0.817	0.091
Listening	0.914	0.332	0.871	0.043
Bully	0.977	0.024	0.976	0.001
Fights	0.987	-0.027	0.987	0
Restless	0.909	0.009	0.908	0.001
Answered	0.707	0.186	0.640	0.067

Results from the CIG node are detailed as an example by which to interpret results. The normalized accuracy (i.e., estimate of nodewise predictability for use with categorical variables) was 0.037. The normalized accuracy was computed by taking the ratio of the contribution from other variables (0.012) to the contribution from other variables (0.012) plus one minus the total accuracy: 0.012/0.012 + 0.313 = 0.037. Therefore, 3.7% of the CIG node could be predicted by all other nodes in the network. Further, the total accuracy of the CIG node was 68.7% (0.675 + 0.012 = 0.687). Therefore, most of accuracy of the CIG node (67.5%) was due to contributions of this node specifically. Since the other variables do not strongly contribute to the predictability of CIG, it is expected that successful intervention on past-month CIG use specifically could potentially address use by 68.7%. In contrast, intervention for other behaviors related to past-month CIG use is likely to influence this behavior by 3.7%.

Results for the CIG node are compared against the Anxious node, where a greater proportion of the predictability was due to other nodes. The normalized accuracy of the Anxious node was 0.422 meaning that 42.2% of the Anxious node could be predicted by all other nodes in the network. Furthermore, the total accuracy of the Anxious node was 87.6%, meaning that the majority of accuracy of the Anxious node (42.2%) was due to contributions of other nodes in the network, not the Anxious node specifically (45.4%). Unlike the CIG node, successful intervention on other nodes connected to the Anxious node could potentially address this symptom by 87.6%.

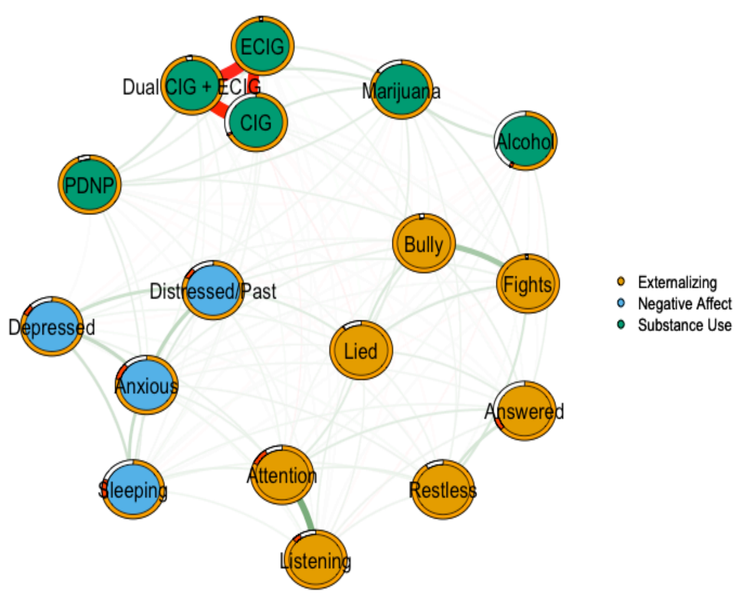


Figure 4.9: Mixed Graphical Model Estimated on the Data. Green edges indicate positive relationships and red edges indicate negative relationships. The orange part of the ring indicates the accuracy of the intercept model. The red part of the ring is the additional accuracy achieved by all remaining variables. The sum of both orange and red is the accuracy of the full model *A*. The normalized accuracy Anorm is the ratio between the additional accuracy due to the remaining variables (red) and one minus the accuracy of the intercept model (white + red).

Nodes with greater strength and greater magnitudes in edge-weights were predicted better (e.g., Dual CIG + ECIG and Attention) than nodes with fewer or weaker edges (e.g., Alcohol). Alcohol (0.590), CIG (0.687), and Answered (0.707), and had the lowest total accuracy in the network whereas Fights (0.987), ECIG (0.982), and Bully (0.977) had the highest total accuracy. Interestingly, the predictability of fights, ECIG, and Bully had no contribution from the other variables (correct classification = 0 or near 0). Other nodes also had a correct classification of 0 (i.e., Dual CIG + ECIG and PDNP) meaning that other nodes in the network did not predict the node at hand beyond the intercept model. Additional accuracy due to the remaining variables contributed to the predictive ability of the negative affect items as well as Attention, Listening, and Answered externalizing items (denoted by the red portion of the ring).

The average predictability as estimated across the accuracy/correct classification column for all nodes was 0.85, indicating that 85% of the variance of the network was explained by the nodes in the network. Therefore, the network was largely determined by itself through strong mutual interactions between nodes. Intervention on any of these nodes would likely result in a decrease of a neighboring symptom, especially for nodes with higher contribution from other variables in the network.

DISCUSSION

To our knowledge, this is one of the first studies to investigate a comorbidity network including substance use behaviors and a wide range of mental disorder symptoms in a large sample of U.S. adults. There were two major results from this study. First, networks for men and women did not significantly differ in structure and connectivity, although there were significant differences by gender between specific nodes. Second, the overall network structure and edge-weights confirmed the connections of substance use behaviors, negative affect symptoms, and externalizing symptoms in the network. Yet, there were edges that crossed the construct boundaries (i.e., substance use, negative affect, externalizing), identifying connections across these three constructs. Furthermore, high predictability of all nodes indicated the network was largely determined by itself through strong mutual interactions between nodes. These results suggest that symptom connections, including substance use, (1) do not significantly differ between men and women, and (2) connect based on construct with some overlap.

No gender differences between overall networks but some gender-specific differences between nodes

There were no significant gender differences for the overall comorbidity networks. This was inconsistent with our hypothesis based on expectations developed in the prior literature.[123,190] In general, these studies report a higher prevalence of alcohol, tobacco, illicit substance use, and externalizing problems in men and greater negative affect symptoms and comorbidity in women.[123,183,189,190] Differences in the network structures may be due to the measurement of past-month substance use and mental health symptom endorsement rather than diagnosis. There were gender differences in higher severity due to use of aggregate sum scores. However, this difference was not present in a comorbidity network of people experiencing subthreshold levels of use or symptoms.[205,206]

Some gender-specific differences were discovered between nodes. These are detailed below:

(1) **Alcohol—marijuana** (EW_{Men} = 0.87, EW_{Women} = 1.08): One would likely expect that this connection between alcohol and marijuana would have a larger magnitude for men, especially given that men experience more substance use and externalizing behaviors.[123,183,189,190] However, as men have a higher prevalence of marijuana use, women have demonstrated a greater increase in use over time.[153] From 2006-2016, the rate of marijuana use increased 40% for men (from 8.1% to 11.3%), and 63% for women (from 4.1% to 6.7%).[153] Data also suggest that women have a faster development of cannabis use disorder with poorer outcomes compared to men.[153] Therefore, the alcohol—marijuana connection may be a greater connection of interest in women, but should also not be discounted in men.

(2) **Alcohol—sleeping problems** (EW_{Men} = 0.08, EW_{Women} = 0): There was a small positive connection between alcohol use and sleeping problems for men where there was no edge present for women, indicating that alcohol use was associated with sleeping problems in men. The relationship between alcohol use and sleep disturbances is well established.[207] A recent study conducted in the United Kingdom identified that men who maintained a heavy volume of drinking over three decades, had unstable consumption patterns, and sustained hazardous drinking had worse sleep profiles compared to men without these problems while

results for women were mixed.[208] This relationship is consistent with the literature and should continue to be a relationship of interest in comorbidity research.

(3) **Marijuana—feeling anxious** ($EW_{Men} = 0$, $EW_{Women} = 0.15$): Women have demonstrated a greater increase in marijuana use over time compared to men, although men have a higher prevalence of marijuana use.[153] Women with cannabis use disorder are more likely to experience anxiety and depression compared to men with cannabis use disorder.[153] This relationship is consistent with prior work and underscores the importance of marijuana use and anxiety problems in women.

(4) **ECIG—lying** ($EW_{Men} = 0.25$, $EW_{Women} = 0$): Men reported more ECIG use and lying behaviors compared to women. This is consistent with research that has established women reporting less substance use, including e-cigarettes, and externalizing behaviors compared to men.[123,183,189,190,209] Additionally, this significantly different edge demonstrates that this connection is not present in women where other positive connections exist for the lying node (e.g., attention difficulties—lying). This relationship may be of more importance in men versus women.

(5) **Alcohol—lying** ($EW_{Men} = 0.31$, $EW_{Women} = 0$) *and* **Alcohol—attention difficulties** ($EW_{Men} = 0.25$, $EW_{Women} = 0$): Men consistently drink more alcohol and have a higher likelihood of alcohol use disorders compared to women.[210–212] Similarly, men experience more externalizing symptoms than women.[123,183,189,190] The relationship between with alcohol and externalizing symptoms is in line with

prior research and these results confirm the importance of this relationship, especially with lying and attention difficulties, in men.

(6) **Lying—attention difficulties** (EW_{Men} = 0.86, EW_{Women} = 0.56): Lying and attention difficulties are externalizing behaviors that are associated with each other in both men and women. The association between attention difficulties, commonly seen with ADHD, and emotional dysregulation is well recognized.[213] Emotional dysregulation encompasses emotional expressions and experiences that are context-inappropriate, which is clinically expressed as irritability.[213] Those with irritability can react to the external stimuli in ways that are overly angry and aggressive[214]; however, the connection between attention difficulties and lying is not well understood. This connection should be investigated further, especially in men as we see a stronger magnitude of association compared to women.

(7) **Alcohol—listening difficulties** (EW_{Men} = -0.11, EW_{Women} = 0): The relationship between alcohol and listening difficulties is not as well understood compared to other externalizing symptoms identified with alcohol use. Listening difficulties can be classified as an inattention ADHD symptom.[2] The comorbidity between ADHD and alcohol has been identified[14,15,215] yet our results reflect a negative relationship between listening difficulties and alcohol for men and no relationship for women. This may be because the inattention ADHD symptom may be less likely to be associated with alcohol use compared to the hyperactivity/impulsivity ADHD symptoms.[216] However, results are mixed.[216–219] Further work is needed to better understand this relationship by gender.

Overall network connections with negative affect and externalizing symptoms varies by substance

The largest edge-weight between any tobacco product and a negative affect or externalizing symptom was found between dual use of CIG and ECIG and feeling anxious. Furthermore, out of all potential connections with negative affect and externalizing symptoms, ECIG was only connected to sleeping problems. Previous studies of adolescents and young adults found ECIG use was associated with ADHD, PTSD, anxiety and other SUDs.[94,95,97] More research is needed to confirm the relationship between ECIG and mental health symptoms in adults. It is possible that these differences in results are due to the measurement of tobacco use. It is rare for studies to exclude ECIG use from CIG use and vice versa, which misclassifies the relationships between CIG and ECIG use to be more strongly associated than reality. Consequently, the study of this association may be important to consider in future research.

PDNP was connected to negative affect and externalizing symptoms with the strongest connection found with sleeping problems, followed by fighting, feeling anxious, and becoming distressed about the past. These results confirm previous work which identified nonmedical prescription drug disorders with externalizing behaviors[220] as well as negative affect behaviors, specifically opioids with PTSD symptoms.[154]

Marijuana had a relatively strong connection between two externalizing symptoms (i.e., lying and fighting). These connections with conduct disorder specific symptoms confirm and reinforce the association with marijuana use as stronger than with previously identified negative affect symptoms (i.e., anxiety and depression).[149–152] Although these connections were not as strong as those within the construct, they exist and demonstrate the overlapping nature among these symptoms and behaviors.

Alcohol was connected to the impulsivity item which is consistent with previous alcohol focused literature demonstrating an association between highly impulsive behaviors and alcohol use.[221] A bidirectional relationship has been identified in that impulsivity significantly increases the risk for initiation, continuation, and excessive alcohol use and can also result from acute intoxication and long-term alcohol abuse. [221] These results confirm this association specifically with a past month measure of alcohol use and through the externalizing symptom of giving an answer before a question is finished being asked.

Finally, several weaker and negative connections remain across substances and negative affect/externalizing symptoms although there were some strong connections across constructs, which is supported by the prior literature.[89,149,169–173] These broad, though weaker connections emphasize the complexity of the comorbidity across substance use and negative affect/externalizing disorders.

Node Centrality and Predictability

Strength and closeness were the only centrality metrics stable enough to interpret for the overall, men, and women networks. Exclusive CIG use and dual CIG and ECIG use were the two nodes with the largest strength in all three networks. This means that these two types of tobacco use had the most connections to other nodes in these networks. This may be due to the oversampling of tobacco users in the PATH data or due to tobacco use being common across the other behaviors.

Lying, acting as a bully, physical fighting, feeling anxious, feeling distressed about the past, and marijuana use were the nodes with the highest closeness centrality for the three networks. Based on these results, we could consider lying, acting as a bully, and physical fighting as more important externalizing nodes in the comorbidity structure compared to the others given their stronger measure of indirect ties to other nodes in the network. Similarly, we could assume that feeling anxious and feeling distressed about the past may be a more important negative affect node in the comorbidity structure as it has a closer measure of reach compared to the other negative affect symptoms. Marijuana use could also be considered an important substance use node in the comorbidity structure based on its high closeness centrality.

Nodes with more and stronger edges had higher node predictability compared to other nodes in the network with lower strength. The overall high predictability of all nodes in the network, however, has implications for potential intervention. Since, on average, 85% of the variance of a node was explained by its neighbors, then one could intervene on one of these symptoms which could affect the entire network. The negative affect variables had the highest predictability contributed from other nodes in the network. Therefore, if we wanted to reduce anxiousness, the network model suggests intervening on the variables that are closely connected to the anxious node: sleep problems, feeling depressed, and feeling distressed about the past. Nodewise predictability tells us we might reduce anxiousness by approximately 87.6% (total accuracy) if we were to intervene on sleep problems, feeling depressed, and feeling distressed about the past.

The node with the overall lowest centrality in all networks was past month alcohol use. This is consistently low across all networks, perhaps because in the way alcohol was measured in this study. Past month alcohol use is not indicative of severity or problematic alcohol use. Therefore, people who indicated past month alcohol use were not likely to also indicate other substance use behaviors and mental disorder symptoms, as demonstrated by the low strength and closeness. More severe measures of problematic alcohol use, however, may perform differently in a comorbidity network. Future work should consider other measures of alcohol use in determining comorbidity structure of substance use behaviors and mental disorder symptoms.

Strengths and limitations

These results should be evaluated in light of the following limitations. First, conclusions that are drawn from this study are not indicative of severe psychopathology or SUD because it uses a population-based sample and data from subthreshold behaviors. Therefore, we cannot draw conclusions about disorders. Nevertheless, the purpose of this project was to better understand how the wide range of behaviors and symptoms interact in a typical sample of adults. Second, the items included in the networks were dichotomized either from combining multiple measures as seen in the substance use items or from collapsing the ordinal negative affect/externalizing items. This strategy results in a loss of information, but allows for easier interpretation of the results, especially since all items overlap regarding time (i.e., past-month endorsement).

Additionally, the items included in the network analysis represent three separate constructs (i.e., substance use, negative affect problems, and externalizing problems) and this presents a potential limitation regarding nodewise predictability because these constructs are correlated. This could result in edge-weights indicating how similar the variables are and do not necessarily reflect mutual influence. Therefore, further investigation in the nodewise predictability of these items is warranted. Future analyses could include three categorical items that represent the constructs and then test for nodewise predictability. Third, these models did not adjust for the influence of other sociodemographic variables. Therefore, there may be some residual confounding. Fourth, these data are cross-sectional. We cannot draw any causal inferences from these networks. Future research is encouraged to study these networks over time. Fifth, networks can only be estimated with complete data. Approximately 2,109 participants were missing data on all seventeen items and not included in the estimation of the overall network. There may be potential for social desirability or misclassification biases in that the people with missing information on these items did not want to endorse their substance use or mental disorder symptoms.

A strength of this project includes the use of accuracy, stability and comparison tests to ensure that the inferences made by these study results were appropriate.

Conclusions

Results emphasized many weak connections throughout the substance use and negative affect/externalizing network and identified a few important connections for future study. In particular, PDNP was most strongly associated with negative affect while marijuana, alcohol and PDNP use were most strongly associated with externalizing. Future work should replicate these analyses in other large samples, including additional nodes of importance and/or sociodemographic factors that may play a role in the comorbidity structure and assess the stability of these networks over time.

CHAPTER 5: PRELIMINARY PATTERNS OF SUBSTANCE USE AND MENTAL DISORDER SYMPTOM COMORBIDITY IN ADULTS OVER TIME

INTRODUCTION

Comorbidity of substance use and mental health problems are more common than exclusive substance use or mental health problems only.[67] Comorbidity has been detected consistently across age groups, and studies within a specific age group have demonstrated different results. For example, younger age groups are at a greater risk of experiencing substance use and mental health comorbidity compared to older age groups.[41,222] It is unclear, however, whether patterns of comorbidity in adults change over time.

Substance use comorbidity over time

Several longitudinal studies identifying co-occurring substance use over time have focused primarily on youth and young adults. These studies have discovered a similar result in that substance use behaviors are relatively stable over time; however, if there is a change in behavior, it usually moves from less severe to more severe (e.g., alcohol only to using multiple substances simultaneously [polysubstance use]). Generally, early substance use (e.g., alcohol and illicit substances) in adolescents is strongly associated with later substance use disorder (SUD) development.[223] Longitudinal association studies in youth have demonstrated that ever tobacco use predicts subsequent substance use[224] while others have identified heavy alcohol use predicting marijuana use during college.[225,226] A latent transition analysis of adolescents identified three substance use groups (mild alcohol use, alcohol and moderate marijuana use, and polysubstance use). Overall, adolescents generally remained in the same group over time; however, when they did transition, adolescents were most likely to move to a more harmful substance use status.[227] Another longitudinal analysis of vocational students (16 to 20 years) in Germany found a similar result. Approximately, 10% of alcohol users at baseline transitioned to polysubstance use of alcohol, cigarettes, and marijuana at 18-month follow up.[52] It is unclear whether this stability and potential transition to more harmful substance use continues in adulthood.

Mental health comorbidity over time

Compared to co-occurring substance use, more is understood about co-occurring mental health conditions in adults. There is evidence that less severe mental disorders precede more severe disorders.[2] Certain mental health conditions can increase the risk of development of future mental health conditions. For example, chronicity of depressive symptoms increases the likelihood of anxiety and substance use disorders.[2,55,56] Epidemiologic studies have provided evidence for both continuity and change of mental disorder comorbidity.[67–69] Overall, the highest stability rates are documented in low psychopathology and externalizing classes, whereas the internalizing or negative affect and highly comorbid classes are moderately stable. Furthermore, results from a latent transition analysis of a nationally representative sample demonstrated that internalizing or negative affect presentations progressed toward remission, while comorbid and externalizing presentations was notably symptomatic across time.[67]

Substance use and mental health comorbidity over time

Less, however, is understood about the stability of substance use and mental disorder comorbidity in adults over time. To date, comorbidity studies have reported that (1) externalizing problems (e.g., ADHD, ODD, CD) in youth precede substance use in both boys and girls whereas (2) substance use (e.g., alcohol and marijuana) in youth predict internalizing or negative affect disorders in adulthood specifically for women.[57–61] A systematic review and meta-analysis of prospective cohort studies suggests a positive association between anxiety and later alcohol use disorders.[228] More longitudinal approaches to assessing the comorbidity structure over time is needed. These studies will help to better understand the progression or regression of symptoms or behaviors in adults, and identify how to better support individuals experiencing comorbidity.

Study goals and hypotheses

The primary goal of this study is to perform a preliminary assessment of the substance use and mental disorder symptom comorbidity patterns across three years of data (2013-2016) using both latent class and network analyses. The secondary goal of this study is to describe the similarities and differences in the patterns of comorbidity across LCA and network analyses. Based on the current understanding of substance use and mental disorder symptoms over time, we hypothesize that overall adult comorbidity patterns will not significantly differ across time. However, we expect minor changes in the relationship between a few behaviors and substances across waves.

METHODS

Setting

The Population Assessment of Tobacco and Health (PATH) study is a nationally representative longitudinal cohort study of the civilian, non-institutionalized household population of the U.S., and participants engaged in all levels of tobacco use ranging from never using tobacco to frequent use.[72] Three waves of data were included in this study.

<u>Wave 1.</u> Wave 1 adult data (N=32,320) are cross-sectional and were collected between September 2013 and December 2014. The weighted response rate among participants was 74.0% for Wave 1.[73] This interview rate is conditional on completion of the Wave 1 screener.

<u>Wave 2</u>. Wave 2 adult data (N=28,362) were collected between October 2014 and October 2015. The weighted response rate among participants was 83.2% and is conditional on Wave 1 participation.[71]

<u>Wave 3</u>. Wave 3 adult data (N=28,148) were collected between October 2015 and October 2016. The weighted response rate among participants was 78.4% and is conditional on Wave 1 participation.[71]

<u>Study Representativeness</u>. New participants introduced at Waves 2 or 3 were excluded. This includes youth that "aged up" into the adult questionnaires. Therefore, we included the same adults from Wave 1 through Wave 3 resulting in Wave 2 N = 26,444 and Wave 3 N = 26,239.

Participants with missing data on the substance use, negative affect, and externalizing measures were not included in the analysis (N_{Wave1}= 2,109, N_{Wave2}= 852, N_{Wave3}= 880). Survey respondents of the analytic sample endorsed significantly greater substance use overall, negative affect symptoms, and externalizing symptoms (except for fighting in Waves 1-3 and bullying in Waves 2-3) compared to those not included in the analytic sample. The participants in the analytic sample were more likely to be Non-Hispanic white, men, aged 25-54 with higher levels of education and annual household income than those who were missing.

Measures

Past Month Tobacco and Substance Use. Six substance use categories were used in this study: exclusive cigarette, exclusive e-cigarette, dual cigarette and e-cigarette, alcohol, marijuana, and prescription drugs not prescribed (PDNP) including painkillers, sedatives, tranquilizers. Only past month or current use of the substances was considered (coded as 1, else = 0) to reduce the potential for recall bias and ensure for accurate overlap with negative affect and externalizing symptoms occurring in the same time frame.

Past Month Negative Affect and Externalizing Symptoms. Negative affect and externalizing symptoms were measured using the Global Appraisal of Individual Needs—Short Screener (GAIN-SS).[73] Four questions were used to measure negative affect symptoms that asked the last time you had significant problems with:

(1) “feeling trapped, lonely, sad, blue, depressed, or hopeless about the future,”

(2) "sleep trouble- such as bad dreams, sleeping restlessly or falling asleep during the day,"

(3) "feeling very anxious, nervous, tense, scared, panicked or something bad was going to happen," and

(4) "becoming very distressed and upset when something reminded you of the past."

Externalizing symptoms were also measured using the GAIN-SS. Seven questions were used to assess externalizing symptoms. Items asked the last time you did the following two or more times:

(1) "lied or conned to get things you wanted or to avoid having to do something,"

(2) "had a hard time paying attention at school, work or home,"

(3) "had a hard time listening to instructions at school, work or home,"

(4) "were a bully or threatened other people,"

(5) "started physical fights with other people,"

(6) "felt restless or the need to run around or climb on things" and

(7) "gave answers before the other person finished asking the question."

The items selected to identify negative affect and externalizing symptoms from the GAIN-SS instrument are ordinal and measures people across four times periods: past month, 2 to 12 months, over a year ago, and never. Participants indicating that they experienced a symptom within the past month were coded as 1. Participants indicating that they experienced the symptom 2 to 12 months ago, over a year ago, and never were coded as 0. Only past month or current negative affect and externalizing symptoms were considered reducing the potential for recall bias and ensure accurate overlap with substance use occurring in the same time frame.

Covariates. Sex, age, race, education, annual household income, and level of satisfaction with social activities and relationships were included as covariates in the generation of latent class models for Waves 1, 2, and 3. Sex was a binary variable with one level representing male and the other level representing female. Age, measured in PATH as a seven-level categorical variable, was re-categorized to have a uniform distribution with six levels (18-24, 25-34, 35-44, 45-54, 55-64, and 65 years or older). Race/ethnicity was measured as a four-level categorical race variable and included information from a separate variable that accounted for Hispanic ethnicity (Non-Hispanic White, Non-Hispanic Black, Non-Hispanic Other, and Hispanic Multicultural). Education, measured in PATH as a six-level categorical variable, was re-categorized as a five-level categorical variable with a uniform distribution [less than high school, GED/high school graduate, some college (no degree) or Associate’s degree, Bachelor’s degree, and Advanced degree]. Annual household income was measured as a five-level categorical variable: less than $10,000, $10,000 to $24,999, $25,000 to $49,999, $50,000 to $99,999, and $100,000 or more.

Level of satisfaction with social activities and relationships was measured as a five-level categorical variable: extremely satisfied, very satisfied, moderately satisfied, a little satisfied, and not at all satisfied.

These covariates were included as auxiliary variables to predict the probability of class membership. Covariates were not included or adjusted for in the development of the networks for Waves 1, 2, and 3.

Statistical analysis

Summary Statistics. Data management and summary statistics for the three waves were done in SAS 9.4. Data were then exported from SAS and imported into Mplus to conduct the LCA. Results from the LCA were then imported back into SAS to evaluate the pairwise comparisons. The original data generated in SAS were also imported into R to estimate the network structures.

Latent Class Comparisons

Latent Class Analysis. Latent class analysis (LCA) is a type of mixture modeling used to identify unobserved heterogeneity in a population and find meaningful groups of people that are similar based on their responses to measured items.[70,141] The observed items (i.e., six substance use behaviors, four negative affect symptoms, and seven externalizing symptoms) are independent of each other given an individual's response on the latent variable meaning that the latent variable (i.e., comorbidity class) explains why the observed items are related to one another.[130] LCA accounts for the observed covariation between substance use and mental disorder symptoms and offers objective indices of class classification accuracy that are not available in traditional cluster analysis methods.[132]

Two parameters are estimated in the LCA model: item probability parameters and class probability parameters. Item probability parameters represent the probability of endorsing an item conditional on latent class membership. It can also be referred to as the item response probabilities or conditional item probabilities. Class probability parameters reflect the probability that a person in a given latent class has of endorsing the specific item. The class probability parameter specifies the prevalence of each class in the population or the relative frequency of class membership. Therefore, a LCA estimates the probability of being in a latent class conditional on the probability of endorsing a measured item.[70] More detail on LCA is provided in the methods section of Chapter 3.

Model Selection. A four-class solution was determined to be most optimal in the LCA for Wave 1 (Chapter 3). Therefore, only a four-class solution was generated for Waves 2 and 3 for comparison. Lo-Mendell-Rubin adjusted likelihood ratio test (LMRT), Akaike information criteria (AIC), Bayesian information criteria (BIC), sample-size adjusted BIC, and entropy were tested to show model fit and parsimony. More detail on these fit and parsimony tests are provided in the methods section of Chapter 3. A smaller AIC and BIC, a larger entropy, and statistically significant results from the LMRT are conditions that determine a more optimal class solution.

Multinomial Logistic Regression. Multinomial regression was used to determine whether any covariates were significantly associated with membership of a latent class.[142] Multinomial regression was conducted using the three-step method (R3STEP) via the AUXILIARY statement in Mplus. This approach was used in order for the latent class model and the latent class predictor model to be obtained automatically[142] rather than introducing potential bias by performing a multinomial regression after the latent class models were selected. More detail of the multinomial logistic regression procedure is provided in the methods section of Chapter 3.

Latent Class Analysis Comparisons. Class probability parameters, item response probability parameters, transition patterns and results from the multinomial logistic regression were compared across the three waves. Differences in class and item response probability parameters were compared. Transition patterns were identified to determine the stability of movement among the classes across the waves. Odds ratios from the multinomial logistic regression were discussed.

Handling Missingness and Complex Sampling Design. Data management, summary statistics, and transition tables for latent class comparisons were performed in SAS 9.4. All LCA was conducted in MPlus. Missing data were removed ($N_{\text{Wave 1}} = 256$, $N_{\text{Wave 2}} = 166$, $N_{\text{Wave 3}} = 198$). Complex sampling design was accounted for in SAS 9.4 using PROC SURVEYFREQ (to generate summary statistics), and in Mplus using the WEIGHT option.

Network Comparisons

Network Analysis. Patterns of associations or interactions between substance use behaviors and mental disorder symptoms can be encoded in a network structure.[45] Measured symptoms and behaviors (i.e., substance use behaviors and negative affect and externalizing symptoms) are represented as nodes. Nodes are connected by edges. Edges represent the interactions between the nodes. Nodes that directly activate each other (i.e., demonstrate an association) are connected while nodes that do not directly activate each other are not. Three networks were generated for comparison, one for each wave. The resulting networks produced patterns of symptom-symptom or symptom-substance use interactions.[45]

Network Model Estimation. All networks were estimated using an Ising Model in R 3.6.0 (*IsingFit* package[193]) to estimate the associations between the nodes (i.e., edges) as partial correlations among a set of binary items (i.e., current substance use behaviors [exclusive cigarette, exclusive e-cigarette, dual cigarette and e-cigarette, alcohol, marijuana, and PDNP], four negative affect symptoms, and seven externalizing symptoms).[34,196,199] Ising model selection uses the Extended Bayesian Information Criteria (EBIC) to measure model parsimony for moderate sample sizes and for a high number of variables by accounting for the number of unknown parameters and the complexity of the model space.[194–196] Models determined to best explain the data using EBIC were interpreted for relevant relationships.[193,197] Edges between two nodes were estimated at most pairwise, after adjusting for all other substance use, negative affect, and externalizing variables.[196] Edges were compared against each other to determine strength. Networks were visualized using the *qgraph* R package.[198] Blue edges illustrate positive partial correlations; red edges illustrate negative partial correlations. The wider the edge, the stronger the correlation.

The Ising model contains two node-specific parameters: the interaction parameter and the node parameter. Details on how these parameters were calculated are provided in the methods section of Chapter 4.

Network Comparisons to Test for Differences by Wave. Differences by wave were evaluated using two approaches. First, visual comparisons using an average layout established differences in the magnitude and direction (i.e., positive or negative) of edge-weights between nodes. Second, three tests of network invariance were used to test significant differences in network models by wave. Greater detail of the three tests of network invariance is provided in the methods section of Chapter 4. These tests were done in a pairwise fashion in the following order: Wave 1 was compared to Wave 2, Wave 2 was compared to Wave 3, and Wave 1 was compared to Wave 3.

The *global strength invariance hypothesis* tested whether the overall level of connectivity in a network was identical between the waves. The global strength invariance hypothesis tests the weighted absolute sum of all edges in the networks or the sum of the unique variance in the network.[200]

A test of the *network structure invariance hypothesis* determined whether network structures were identical by wave by comparing the maximum differences in the edge-weights between all nodes in the networks.[200]

Edge strength invariance hypothesis was tested to determine if a specific edge between two nodes was equally strong between the waves. Edge strength is also referred to as the edge weight, quantified as the magnitude of an edge. This is the magnitude of association between two nodes.[200]

Handling Missingness and Complex Sampling Design. Participants with missing data were removed using listwise deletion ($N_{\text{Wave 1}}$ = 2,109, $N_{\text{Wave 2}}$ = 852, $N_{\text{Wave 3}}$ = 880). Complex sampling design was not accounted for in the estimation of the network models.

RESULTS

Summary statistics

The overall sample size decreased from Wave 1 (N=32,320) to Wave 3 (N=26,239) as shown in Table 5.1. Women (51.9 to 52.1%) and those who identified as

Non-Hispanic White (65.8 to 66.0%) made up the majority of the samples across the waves. Age was evenly distributed. Most of the samples had at least a GED or high school education, an annual household income of more than $25,000, and were at least moderately satisfied with their social activities and relationships.

Endorsement of past month substance use and mental disorder symptoms remained stable across the waves. Alcohol was the most commonly reported past month substance used (52.4% to 54.4%), followed by CIG (16.6%) and marijuana (7.1% to 9.7%). Sleep problems were the most commonly reported past month negative affect symptom (25.4% to 27.0%) followed by feeling anxious (16.0% to 16.5%). Giving answers before the other person finished asking the question was the most common past month externalizing symptom (28.5% to 32.0%), followed by having a hard time paying attention (14.6% to 15.4%) and listening to instructions (10.4% to 11.3%).

Table 5.1: Characteristics of the Samples by Wave

	Wave 1 (N=32,320) N (Weighted %)	**Wave 2 (N=26,444)** N (Weighted %)	**Wave 3 (N=26,239)** N (Weighted %)
Sex			
Male	16306 (48.1)	13067 (47.9)	12830 (47.9)
Female	15980 (51.9)	13354 (52.1)	13386 (52.1)
Age			
18-24	9110 (13.0)	6259 (11.1)	6546 (10.8)
25-34	6337 (17.7)	5674 (17.8)	5824 (17.8)
35-44	4930 (16.5)	4200 (16.8)	3971 (16.4)
45-54	4846 (17.9)	4030 (17.5)	3804 (17.5)
55-64	3971 (16.6)	3507 (17.3)	3389 (17.5)
65+	3110 (18.2)	2770 (19.5)	2703 (20.0)
Race			
Non-Hispanic White	19295 (66.0)	15757 (65.9)	15368 (65.8)
Non-Hispanic Black	4496 (11.2)	3774 (11.2)	3808 (11.2)
Non-Hispanic Other	2429 (7.5)	1948 (7.7)	1946 (7.6)
Hispanic Multiracial	4817 (13.3)	3949 (13.4)	4067 (13.5)
Education			
Less than high school	4233 (11.6)	3159 (10.9)	3101 (10.8)
GED/High school graduate	9765 (29.5)	7516 (27.8)	7591 (27.8)
Some college (no degree)	11300 (31.0)	9567 (32.3)	9416 (32.0)
Bachelor’s degree	4498 (17.8)	3971 (18.4)	3944 (18.7)
Advanced degree	2311 (10.1)	2106 (10.6)	2074 (10.8)
Annual household income			
Less than $10,000	5668 (13.7)	4358 (12.3)	4192 (11.4)
$10,000- $24,999	6768 (20.4)	5598 (19.9)	5384 (19.1)
$25,000- $49,999	6670 (23.0)	5665 (22.9)	5672 (22.9)
$50,000- $99,999	6140 (24.9)	5415 (26.2)	5546 (26.8)
$100,000 or more	3914 (18.0)	3519 (18.7)	3658 (19.8)
Satisfaction with social activities and relationships			
Extremely satisfied	6942 (22.3)	5285 (20.9)	5630 (21.4)
Very satisfied	13742 (46.1)	11295 (46.8)	10578 (44.6)
Moderately satisfied	8157 (23.7)	7015 (24.2)	6895 (24.6)
A little satisfied	2376 (5.6)	1975 (5.8)	2119 (6.6)
Not at all satisfied	1001 (2.3)	812 (2.3)	939 (2.7)
Past month tobacco and substance use			
CIG	10381 (16.6)	8373 (16.6)	7904 (16.6)
ECIG	578 (0.9)	593 (1.2)	703 (1.5)
Dual CIG + ECIG	996 (1.5)	1069 (2.0)	938 (1.8)
Alcohol	17787 (52.4)	15312 (54.4)	14749 (53.9)
Marijuana	4392 (7.1)	4363 (8.9)	4630 (9.7)
PDNP	1950 (5.1)	1707 (5.4)	1737 (5.8)
Past month negative affect symptoms			
Depressed	5692 (13.4)	4639 (13.6)	4421 (13.2)
Sleeping	9564 (26.7)	7745 (27.0)	7152 (25.4)
Anxious	6864 (16.1)	5602 (16.5)	5433 (16.0)
Distressed/Past	5605 (12.5)	4577 (13.1)	4493 (12.7)
Past month externalizing symptoms			
Lied	3245 (7.1)	2399 (6.6)	2360 (6.4)
Attention	5831 (14.6)	4818 (15.3)	4798 (15.4)
Listening	4128 (10.4)	3480 (11.3)	3478 (11.3)
Bully	737 (1.7)	635 (1.7)	641 (1.7)
Fights	404 (0.7)	331 (0.7)	336 (0.7)

Restless	2953 (6.2)	2125 (5.9)	2112 (5.6)
Answered	11399 (32.0)	8390 (29.8)	8033 (28.5)

Latent Class Comparisons

Four class solution

The four-class model was selected for interpretation in Wave 1 because (1) the AIC, BIC, and sample-size adjusted BIC were smallest for the four-class solution compared to the three- and two- class solutions, and (2) the LMRT was statistically significant, rejecting the five-class model when compared to the four-class model. A four-class solution was also selected for Waves 2 and 3 to compare latent classes across waves (Table 5.2).

Table 5.2: Model Parsimony and Fit Statistics for Five-Class Solution by Wave

	AIC	BIC	Sample-Size Adjusted BIC	Entropy	H_o LL	LMRT	p-value	LC 1	LC 2	LC 3	LC 4
Wave 1	303521	304117	303891	0.844	-153520	3641	<0.05	1960 (6.1%)	2691 (8.3%)	23571 (72.9%)	4098 (12.7%)
Wave 2	252434	253015	252789	0.847	-127721	3134	<0.05	1727 (6.5%)	2316 (8.8%)	3478 (13.2%)	18922 (71.6%)
Wave 3	249796	250377	250151	0.787	-125429	1196	0.5789	2140 (8.2%)	5400 (20.6%)	17176 (65.5%)	1524 (5.8%)

NOTE: AIC = Akaike information criteria, BIC = Bayesian information criteria, LL = log likelihood, LMRT = Lo Mendell Rubin Test, LC = latent class

Class probability

The four classes were labeled based on the characteristics of the item response probabilities of the specific class. The low symptom class was most common across the three waves and gradually decreased from Wave 1 to Wave 3 ($N_{Wave\ 1}$ = 23,571, 72.9%; $N_{Wave\ 2}$ = 18,922, 71.6%; $N_{Wave\ 3}$ = 17,176, 65.5%). The comorbid class was the least common across the three waves and gradually increased from Wave 1 to Wave 3 ($N_{Wave\ 1}$ = 1,960, 6.1%; $N_{Wave\ 2}$ = 1,727, 6.5%; $N_{Wave\ 3}$ = 2,140, 8.2%), seen in Table 5.3.

Table 5.3: Class Probability by Wave

	Comorbid Class	Low Comorbid Class	Externalizing Class	Low Symptom Class	Negative Affect Class	Substance Use Class
	N (Weighted %)	N (Weighted %)	N (Weighted %)	N (Weighted %)	N (Weighted %)	N (Weighted %)
Wave 1	1960 (6.1%)	--	2691 (8.3%)	23571 (72.9%)	4098 (12.7%)	--
Wave 2	1727 (6.5%)	--	2316 (8.8%)	18922 (71.6%)	3478 (13.2%)	--
Wave 3	2140 (8.2%)	5400 (20.6%)	--	17176 (65.5%)	--	1524 (5.8%)

The externalizing and negative affect classes remained stable at Waves 1 and 2. At Wave 3, however, the classes changed in composition. Rather than externalizing and negative affect classes, low comorbid and substance use classes emerged.

Item response probabilities

Figures 5.1 – 5.3 display the item-probability plots for the four-class solution for Waves 1, 2, and 3. Tables 5.4 – 5.6 presents the corresponding conditional probability or item response probability values for the four comorbidity classes for Waves 1, 2, and 3. There are seventeen items (six substance use, four negative affect, and seven externalizing items) along the x-axis of each plot. The y-axis represents the probability of endorsing a given item. The four lines, called profiles, correspond to the four classes in the LCA solution and the values are the conditional item probabilities for each of the seventeen items across the four classes.

The comorbid and the low symptom classes were the most extreme classes that consistently emerged across the three waves. The comorbid class, overall, had high probability of endorsing most items while the low symptom class generally had low probability of endorsing all items except for endorsing past 30-day alcohol use.

Similar profiles emerged across Waves 1 and 2, particularly for the negative affect and externalizing classes. The negative affect class had high item response probabilities on the four negative affect items: feeling depressed (40.2% to 45.5%), sleeping problems (63.5% to 67.0%), feeling anxious (57.3% to 58.3%), and feeling distressed about the past (37.7% to 42.3%). The externalizing class had high item response probabilities on the seven externalizing items: lying (13.8% to 17.9%), attention problems (75.7% to 78.0%), listening problems (55.3% to 56.9%), bullying (2.8% to 3.4%), getting into physical fights (0.7% to 1.0%), restlessness (15.9% to 16.9%), and giving an answer before a question is finished being asked (62.9% to 65.3%).

In Wave 3, a low comorbid and substance use class emerged. The low comorbid class resembled the comorbid class with similar peaks yet overall lower item endorsement compared to the comorbid class. This was specifically noticeable for the negative affect and externalizing symptoms where the item response probabilities were second highest relative to the comorbid class for all negative affect and externalizing items except for getting into physical fights. The substance use class had higher item response endorsement for cigarette (43.1%), alcohol (76.1%), and marijuana (69.9%)

use compared to all other classes. Item response probabilities were also high for e-cigarette use (2.9%), dual cigarette and e-cigarette use (3.4%), and PDNP (7.9%).

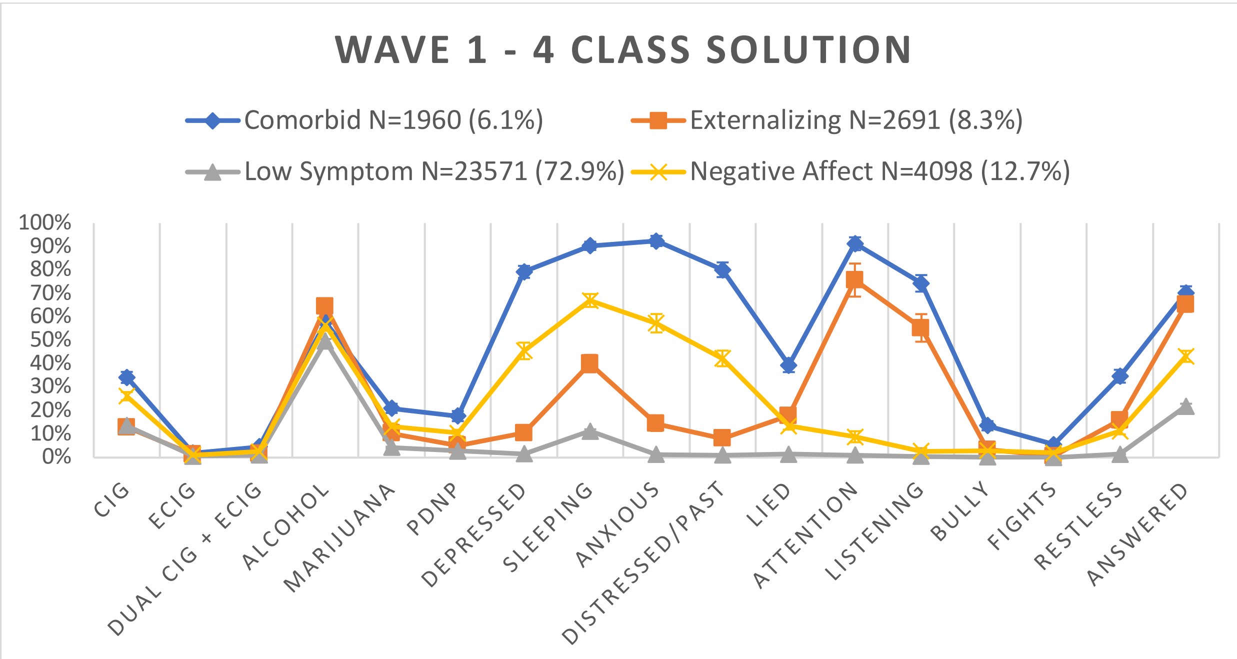


Figure 5.1: Wave 1 - Four-Class Solution of Substance Use Behaviors and Mental Disorder Symptoms

Table 5.4: Wave 1 Item Response Probability Values for Four Latent Classes				
	Comorbid N=1960 (6.1%)	Externalizing N=2691 (8.3%)	Low Symptom N=23571 (72.9%)	Negative Affect N=4098 (12.7%)
CIG	34.20%	12.90%	13.60%	26.20%
ECIG	1.90%	1.30%	0.70%	1.00%
Dual CIG + ECIG	4.60%	1.60%	1.00%	2.50%
Alcohol	57.80%	64.40%	49.60%	56.40%
Marijuana	21.00%	10.30%	4.30%	13.20%
PDNP	17.70%	5.20%	2.90%	10.50%
Depressed	79.20%	10.50%	1.60%	45.50%
Sleeping	90.30%	40.10%	11.30%	67.00%
Anxious	92.40%	14.50%	1.40%	57.30%
Distressed/Past	80.10%	8.20%	1.10%	42.30%
Lied	39.40%	17.90%	1.50%	13.50%
Attention	91.20%	75.70%	1.00%	8.90%
Listening	74.30%	55.30%	0.50%	2.70%
Bully	13.60%	3.40%	0.20%	3.00%
Fights	5.60%	0.70%	0.00%	2.00%
Restless	34.70%	15.90%	1.50%	11.30%
Answered	70.30%	65.30%	22.00%	43.20%

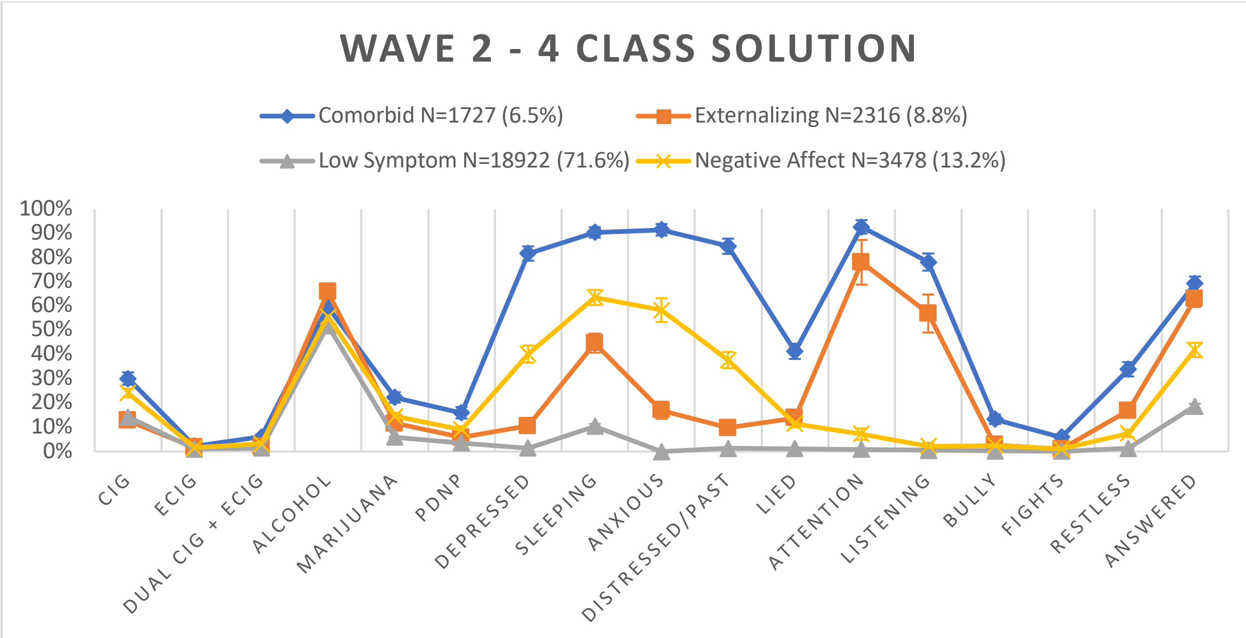


Figure 5.2: Wave 2 - Four-Class Solution of Substance Use Behaviors and Mental Disorder Symptoms

Table 5.5: Wave 2 Item Response Probability Values for Four Latent Classes

	Comorbid N=1727 (6.5%)	Externalizing N=2316 (8.8%)	Low Symptom N=18922 (71.6%)	Negative Affect N=3478 (13.2%)
CIG	30.10%	12.80%	14.20%	24.10%
ECIG	2.20%	1.70%	1.00%	1.60%
Dual CIG + ECIG	6.00%	2.00%	1.40%	3.10%
Alcohol	59.70%	65.80%	52.00%	55.70%
Marijuana	22.30%	11.60%	5.90%	14.50%
PDNP	16.00%	5.80%	3.50%	9.10%
Depressed	81.60%	10.60%	1.50%	40.20%
Sleeping	90.30%	44.70%	10.40%	63.50%
Anxious	91.40%	17.00%	0.00%	58.30%
Distressed/Past	84.60%	9.80%	1.30%	37.70%
Lied	41.30%	13.80%	1.20%	11.50%
Attention	92.60%	78.00%	0.90%	7.20%
Listening	78.10%	56.90%	0.50%	2.20%
Bully	13.30%	2.80%	0.20%	2.50%
Fights	5.90%	1.00%	0.10%	1.10%
Restless	33.90%	16.90%	1.30%	7.50%
Answered	69.20%	62.90%	18.70%	41.90%

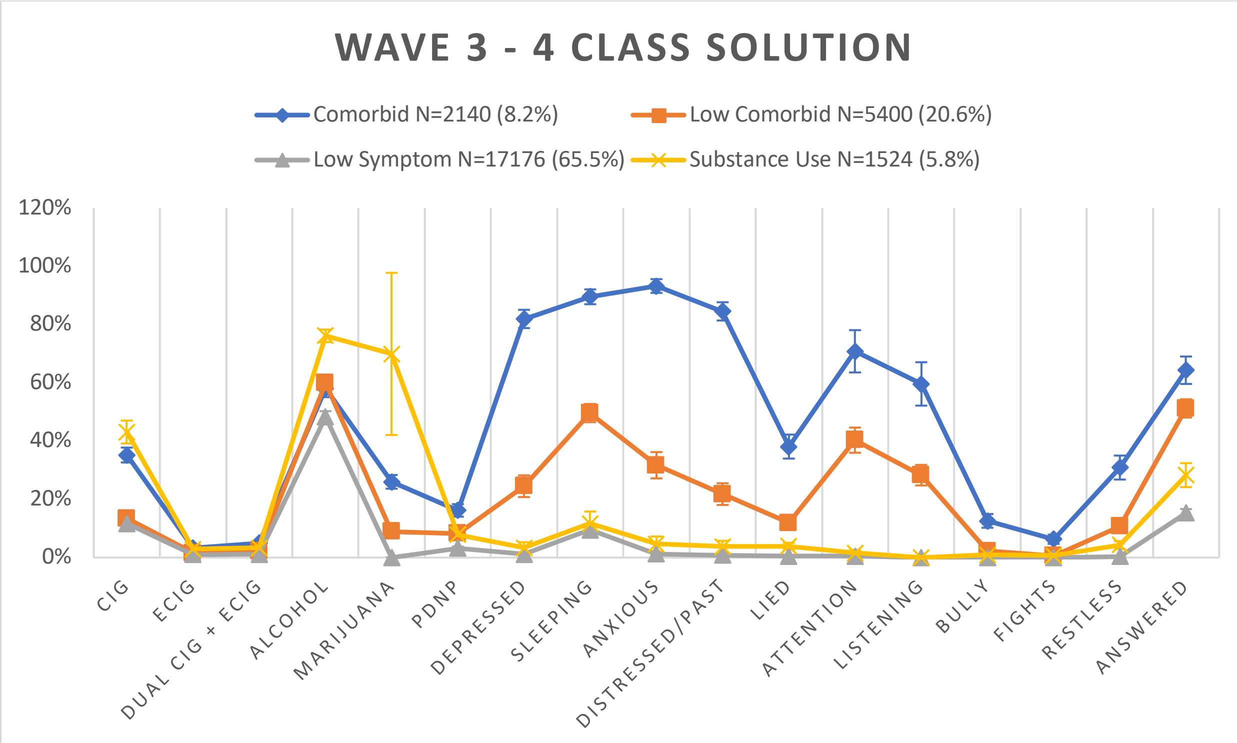


Figure 5.3: Wave 3 - Four-Class Solution of Substance Use Behaviors and Mental Disorder Symptoms

Table 5.6: Wave 3 Item Response Probability Values for Four Latent Classes

	Comorbid N=2140 (8.2%)	Low Comorbid N=5400 (20.6%)	Low Symptom N=17176 (65.5%)	Substance Use N=1524 (5.8%)
CIG	35.20%	13.50%	11.70%	43.10%
ECIG	3.30%	1.60%	1.00%	2.90%
Dual CIG + ECIG	5.00%	2.10%	1.10%	3.40%
Alcohol	58.00%	59.90%	48.30%	76.10%
Marijuana	26.00%	8.90%	0.00%	69.90%
PDNP	16.20%	8.30%	3.20%	7.90%
Depressed	81.90%	24.50%	1.10%	3.30%
Sleeping	89.50%	49.60%	9.50%	11.70%
Anxious	93.20%	31.70%	1.30%	4.70%
Distressed/Past	84.50%	21.80%	0.80%	3.80%
Lied	38.10%	11.90%	0.50%	3.90%
Attention	70.80%	40.30%	0.50%	1.60%
Listening	59.60%	28.30%	0.00%	0.00%
Bully	12.60%	2.30%	0.10%	0.90%
Fights	6.30%	0.60%	0.00%	0.70%
Restless	30.90%	10.70%	0.40%	4.30%
Answered	64.30%	51.20%	15.30%	28.30%

Transitions based on cross-sectional results

Individuals in the sample were assigned to one of the four classes based on the LCA posterior probabilities. This was done for Waves 1, 2, and 3 and class membership information was merged across the waves to create cross-classification tables. The tables were used to describe individual movement among the comorbidity classes over time. Table 5.7 includes cross tabulations for three transition points (i.e., Wave 1 to Wave 2, Wave 2 to Wave 3, and Wave 1 to Wave 3). To read the Table 5.7, start with the preceding wave first and look to where people move to the subsequent wave. Interpret the proportion forward whereas 9.5% of individuals in the negative affect class at Wave 1 transitioned to the comorbidity class at Wave 2. It is not appropriate, however, to interpret this proportion backwards (i.e., 9.5% of those in the comorbid class at Wave 2 were in the negative affect class at Wave 1).

Table 5.7: Transition Tables based on Cross-Sectional LCA Results (%)											
	W2					W3					
W1	COM	EXT	LS	NA		W2	COM	L COM	LS	SU	
COM	8.6%	9.2%	65.4%	16.8%	100%	COM	7.6%	21.5%	64.3%	6.6%	100%
EXT	9.5%	10.2%	64.3%	15.9%	100%	EXT	7.8%	20.2%	65.0%	6.1%	100%
LS	9.5%	8.9%	65.1%	16.2%	100%	LS	8.4%	20.5%	65.3%	5.8%	100%
NA	9.8%	10.0%	62.3%	15.0%	100%	NA	8.2%	20.7%	64.9%	6.1%	100%
	W3										
W1	COM	L COM	LS	SU							
COM	8.2%	21.0%	64.9%	6.0%	100%						
EXT	7.1%	20.2%	66.8%	5.9%	100%						
LS	8.7%	20.2%	65.0%	6.1%	100%						
NA	7.6%	22.6%	64.2%	5.6%	100%						
W1 = Wave 1, W2 = Wave 2, W3 = Wave 3, COM = comorbidity class, NA = negative affect class, EXT = externalizing class, LS = low symptom class, L COM = low comorbid, SU = substance use											

The values shaded in grey are values that describe stability in membership status. For example, 8.6% of individuals who were in the comorbidity class in Wave 1 remained in the comorbidity class in Wave 2. The values that are not shaded describe movement among the classes. For example, of the individuals who were in the comorbid class in Wave 1, 16.8% transitioned into the negative affect class in Wave 2. The low symptom class was the most stable class for all transitions (62.3% to 66.8%). From Wave 1 to Wave 2, the negative affect class was second to the low symptom class in stability (15.0%) followed by the externalizing class (10.2%).

As seen in all three transition tables, overall, when individuals transitioned, they typically to transitioned into the low symptom class (62.3% to 66.8%). The second most common transition was to the low comorbid class (20.2% to 22.6%) from Wave 2 to Wave 3 and Wave 1 to Wave 3.

Sociodemographic characteristics

In Waves 1 and 2, males were significantly less likely than females to be classified in the comorbid and negative affect classes relative to the low symptom class, seen in Tables 5.8 – 5.10. In Wave 3, this relationship between sex and probability of latent class membership extended to the low comorbid class where men were less likely to be classified in the low comorbid class (OR = 0.79, 95% CI = 0.70-0.89), Table 5.10. However, males were more likely than females to be classified in the substance use class relative to the low symptom class in Wave 3 (OR = 1.98, 95% CI = 1.73-2.27), Table 5.10.

For all waves, a trend emerged for age across all classes: as age increased, the odds of class membership decreased for all classes relative to the low symptom class. Therefore, the youngest age group (18-24 years) had the highest odds of class membership compared to the oldest age group (65 years and older), relative to the low symptom class (Tables 5.8 – 5.10). In Waves 1 and 2, the relationship between age and probability of latent class membership was largest in magnitude for the comorbid class. In Wave 3, the magnitude of association between age and the comorbid class (18-24 years OR_{Wave3} = 13.99, 95% CI = 9.83-19.90) was smaller than the substance use class (18-24 years OR_{Wave3} = 30.75, 95% CI = 18.99-49.81), Table 5.10.

Respondents who identified as Non-Hispanic Black, Hispanic Multicultural, and Non-Hispanic Other were significantly less likely than respondents who identified as Non-Hispanic White to be classified in any of the classes relative to the low symptom class across all waves except for the relationship between Non-Hispanic Black and the substance use class in Wave 3 (OR = 1.10, 95% CI = 0.93-1.29).

Generally, all education levels relative to having an advanced degree increased one's odds for membership in the comorbid classes for all waves, negative affect classes for Waves 1 and 2, and the substance use class for Wave 3, relative to the low symptom class across all three waves. The opposite relationship emerged between education and the externalizing classes for Waves 1 and 2, and the low comorbid class for Wave 3. Relative the having an advanced degree, generally all education levels demonstrated a protective effect for membership in the externalizing classes, especially in Wave 2 (Tables 5.8 – 5.9). The same relationship between education and latent class membership emerged between income and latent class membership except for the low comorbid class in Wave 3. Yet, these associations were not statistically significant (Less than \$10,000 OR_{Wave3} = 1.20, 95% CI = 0.94-1.53; \$10,000-\$24,999 OR_{Wave3} = 1.14, 95% CI = 0.93-1.39; \$25,000-\$49,999 OR_{Wave3} = 1.07, 95% CI = 0.89-1.28; \$50,000-\$99,999 OR_{Wave3} = 1.01, 95% CI = 0.85-1.19), Table 5.10.

Compared to being extremely satisfied, as social satisfaction decreased, the likelihood of being in the comorbid (Not at all satisfied OR_{Wave1} = 95.87, 95% CI = 66.32-138.58; Not at all satisfied OR_{Wave2} = 83.35, 95% CI = 54.05-128.53; Not at all satisfied OR_{Wave3} = 88.94, 95% CI = 58.93-134.24), negative affect (Not at all satisfied OR_{Wave1} = 22.62, 95% CI = 15.44-33.16; Not at all satisfied OR_{Wave2} = 18.34, 95% CI = 11.66-28.84), externalizing (Not at all satisfied OR_{Wave1} = 3.67, 95% CI = 1.62-8.31; Not at all satisfied OR_{Wave2} = 6.31, 95% CI = 3.12-12.78), low comorbid (Not at all satisfied OR_{Wave3} = 8.78, 95% CI = 5.47-14.12), and substance use (Not at all satisfied OR_{Wave3} = 4.21, 95% CI = 2.40-7.41) classes significantly increased. A dose-response relationship was identified with each level decrease of social satisfaction for every latent class across all waves.

Table 5.8: Wave 1 - Association Between Demographic and Social Variables on Probability of Latent Class Membership*

	Comorbid Class OR (95% CI)	Externalizing Class OR (95% CI)	Negative Affect Class OR (95% CI)
Sex			
Female	REF	REF	REF
Male	0.72 (0.63-0.82)	1.00 (0.87-1.14)	0.74 (0.66-0.83)
Age			
18-24 years	10.02 (7.06-14.24)	3.39 (2.69-4.28)	3.88 (3.11-4.83)
25-34 years	6.00 (4.17-8.64)	1.81 (1.40-2.33)	2.39 (1.90-3.01)
35-44 years	4.10 (2.83-5.94)	1.46 (1.11-1.91)	1.69 (1.32-2.16)
45-54 years	3.77 (2.60-5.47)	1.13 (0.86-1.50)	1.52 (1.19-1.93)
55-64 years	2.27 (1.54-3.36)	0.98 (0.73-1.32)	1.44 (1.12-1.85)
65 years +	REF	REF	REF
Race			
Non-Hispanic White	REF	REF	REF
Non-Hispanic Black	0.48 (0.40-0.59)	0.62 (0.51-0.77)	0.76 (0.64-0.89)
Non-Hispanic Other	0.73 (0.57-0.94)	0.69 (0.53-0.89)	0.72 (0.56-0.91)
Hispanic Multiracial	0.50 (0.41-0.61)	0.65 (0.52-0.80)	0.78 (0.66-0.93)
Education			
Less than high school	1.53 (1.05-2.21)	0.79 (0.57-1.09)	1.62 (1.22-2.15)
GED/High school graduate	1.37 (0.98-1.92)	0.80 (0.61-1.05)	1.42 (1.10-1.83)
Some college (no degree)	1.79 (1.29-2.47)	1.08 (0.85-1.37)	1.36 (1.06-1.75)
Bachelor's degree	1.35 (0.94-1.93)	1.10 (0.85-1.42)	1.16 (0.88-1.53)
Advanced degree	REF	REF	REF
Income			
Less than $10,000	2.54 (2.03-3.18)	0.83 (0.65-1.06)	1.61 (1.32-1.95)
$10,000- $24,999	2.02 (1.62-2.51)	0.55 (0.44-0.68)	1.51 (1.25-1.83)
$25,000- $49,999	1.45 (1.16-1.81)	0.84 (0.68-1.04)	1.11 (0.92-1.33)
$50,000- $99,999	1.00 (0.78-1.27)	1.21 (1.00-1.46)	0.99 (0.82-1.20)
$100,000 or more	REF	REF	REF
Level of satisfaction with social activities and relationships			
Extremely satisfied	REF	REF	REF
Very satisfied	1.65 (1.31-2.07)	1.42 (1.18-1.70)	1.55 (1.31-1.85)
Moderately satisfied	8.15 (6.54-10.15)	2.66 (2.18-3.26)	4.53 (3.78-5.43)
A little satisfied	34.19 (26.4-44.29)	3.08 (2.17-4.39)	11.09 (8.70-14.14)
Not at all satisfied	95.87 (66.32-138.58)	3.67 (1.62-8.31)	22.62 (15.44-33.16)

*Reference low symptom class

Table 5.9: Wave 2 - Association Between Demographic and Social Variables on Probability of Latent Class Membership*

	Comorbid Class OR (95% CI)	Externalizing Class OR (95% CI)	Negative Affect Class OR (95% CI)
Sex			
Female	REF	REF	REF
Male	0.57 (0.49-0.66)	1.01 (0.87-1.17)	0.70 (0.62-0.79)
Age			
18-24 years	8.17 (5.73-11.64)	2.75 (2.09-3.62)	2.59 (2.09-3.22)
25-34 years	5.48 (3.79-7.92)	2.29 (1.73-3.04)	1.88 (1.49-2.36)
35-44 years	4.39 (3.01-6.39)	1.97 (1.48-2.63)	1.40 (1.11-1.78)
45-54 years	3.60 (2.46-5.28)	1.30 (0.96-1.77)	1.44 (1.14-1.82)
55-64 years	2.33 (1.56-3.49)	1.17 (0.85-1.62)	1.31 (1.03-1.68)
65 years +	REF	REF	REF
Race			
Non-Hispanic White	REF	REF	REF
Non-Hispanic Black	0.43 (0.35-0.53)	0.51 (0.39-0.65)	0.81 (0.69-0.95)
Non-Hispanic Other	0.80 (0.59-1.09)	0.58 (0.43-0.77)	0.74 (0.55-0.99)
Hispanic Multiracial	0.62 (0.50-0.78)	0.53 (0.41-0.69)	0.73 (0.61-0.88)
Education			
Less than high school	1.28 (0.86-1.90)	0.51 (0.34-0.77)	1.42 (1.04-1.94)
GED/High school graduate	1.26 (0.89-1.78)	0.62 (0.47-0.82)	1.40 (1.07-1.85)
Some college (no degree)	1.58 (1.14-2.21)	0.94 (0.73-1.19)	1.48 (1.13-1.93)
Bachelor’s degree	1.08 (0.74-1.59)	0.91 (0.71-1.17)	1.06 (0.79-1.41)
Advanced degree	REF	REF	REF
Income			
Less than $10,000	2.14 (1.66-2.76)	0.78 (0.57-1.07)	1.63 (1.31-2.04)
$10,000- $24,999	1.87 (1.46-2.39)	0.89 (0.69-1.15)	1.52 (1.23-1.87)
$25,000- $49,999	1.29 (1.01-1.65)	0.98 (0.79-1.22)	1.18 (0.97-1.44)
$50,000- $99,999	1.01 (0.78-1.30)	0.90 (0.73-1.10)	0.90 (0.73-1.10)
$100,000 or more	REF	REF	REF
Level of satisfaction with social activities and relationships			
Extremely satisfied	REF	REF	REF
Very satisfied	1.79 (1.35-2.35)	1.60 (1.3-1.97)	1.69 (1.4-2.03)
Moderately satisfied	8.36 (6.4-10.91)	3.24 (2.59-4.05)	4.07 (3.35-4.94)
A little satisfied	33.21 (24.27-45.45)	4.47 (3.05-6.54)	10.9 (8.35-14.23)
Not at all satisfied	83.35 (54.05-128.53)	6.31 (3.12-12.78)	18.34 (11.66-28.84)

*Reference low symptom class

Table 5.10: Wave 3 - Association Between Demographic and Social Variables on Probability of Latent Class Membership*

	Comorbid Class OR (95% CI)	Low Comorbid Class OR (95% CI)	Substance Use Class OR (95% CI)
Sex			
Female	REF	REF	REF
Male	0.78 (0.68-0.91)	0.79 (0.70-0.89)	1.98 (1.73-2.27)
Age			
18-24 years	13.99 (9.83-19.90)	3.23 (2.66-3.94)	30.75 (18.99-49.81)
25-34 years	8.31 (5.81-11.90)	2.00 (1.64-2.45)	17.64 (10.87-28.62)
35-44 years	5.49 (3.80-7.94)	1.45 (1.18-1.80)	11.45 (6.99-18.76)
45-54 years	4.15 (2.87-6.00)	1.26 (1.02-1.56)	7.46 (4.53-12.30)
55-64 years	2.16 (1.48-3.15)	1.27 (1.03-1.57)	6.39 (3.81-10.69)
65 years +	REF	REF	REF
Race			
Non-Hispanic White	REF	REF	REF
Non-Hispanic Black	0.45 (0.37-0.56)	0.61 (0.52-0.73)	1.10 (0.93-1.29)
Non-Hispanic Other	0.76 (0.57-1.01)	0.68 (0.53-0.86)	0.61 (0.47-0.80)
Hispanic Multiracial	0.57 (0.46-0.71)	0.61 (0.51-0.73)	0.54 (0.45-0.66)
Education			
Less than high school	1.64 (1.12-2.40)	0.76 (0.57-0.99)	2.08 (1.38-3.13)
GED/High school graduate	1.34 (0.94-1.89)	0.81 (0.65-1.01)	2.34 (1.60-3.43)
Some college (no degree)	1.68 (1.20-2.34)	1.04 (0.85-1.27)	2.25 (1.55-3.26)
Bachelor’s degree	0.98 (0.68-1.42)	0.89 (0.72-1.11)	1.62 (1.10-2.38)
Advanced degree	REF	REF	REF
Income			
Less than $10,000	2.87 (2.20-3.76)	1.20 (0.94-1.53)	2.35 (1.84-2.99)
$10,000- $24,999	2.19 (1.70-2.81)	1.14 (0.93-1.39)	1.94 (1.55-2.43)
$25,000- $49,999	1.43 (1.12-1.83)	1.07 (0.89-1.28)	1.42 (1.14-1.76)
$50,000- $99,999	1.12 (0.86-1.45)	1.01 (0.85-1.19)	1.07 (0.85-1.34)
$100,000 or more	REF	REF	REF
Level of satisfaction with social activities and relationships			
Extremely satisfied	REF	REF	REF
Very satisfied	1.92 (1.51-2.46)	1.56 (1.32-1.85)	1.23 (1.04-1.44)
Moderately satisfied	7.52 (5.98-9.45)	3.50 (2.91-4.20)	1.92 (1.59-2.30)
A little satisfied	36.74 (27.49-49.11)	7.17 (5.43-9.47)	2.51 (1.86-3.39)
Not at all satisfied	88.94 (58.93-134.24)	8.78 (5.47-14.12)	4.21 (2.40-7.41)

*Reference low symptom class

Network Comparisons

Wave 1

The Wave 1 network consisted of 17 nodes (Figure 5.4). The network had 94 non-zero edges out of 136 possible edges (density=0.691), indicating that 69.1% of possible connections were identified in the network. The network structure is an Ising model, which is a network of partial correlation coefficients. Especially strong connections emerged between the tobacco use nodes, between “Attention” and “Listening”, and “Fights” and “Bully”. The negative affect symptoms were positioned between the substance use behaviors and externalizing symptoms, with many of the nodes lying on the periphery of the network. Edge-weights are shown in the Appendix D (Supplemental Table 5.1).

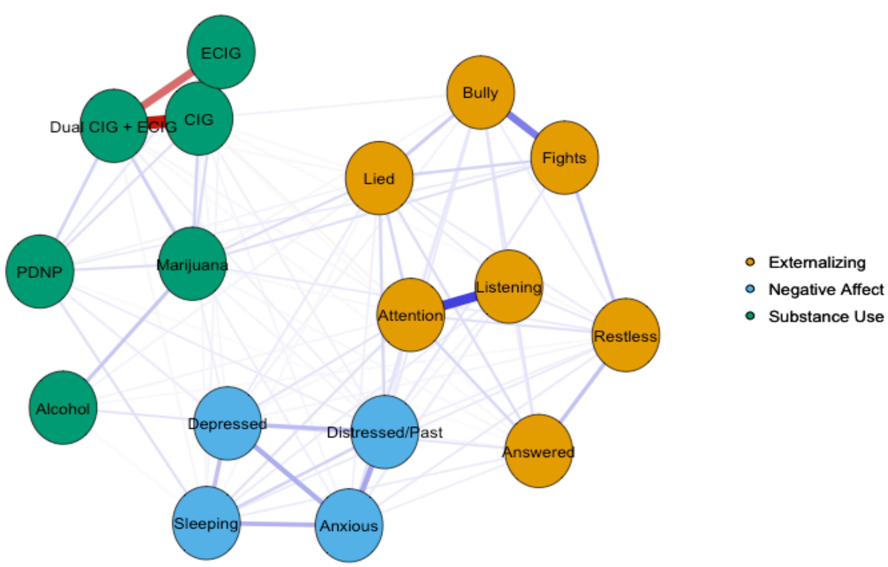


Figure 5.4: Visualization of Wave 1 Network

Wave 2

The Wave 2 network consisted of 17 nodes and had 84 non-zero edges out of 136 possible edges (density = 0.618), indicating that 61.8% of possible connections were identified in the network (Figure 5.5). The network structure is an Ising model, which is a network of partial correlation coefficients. Similar to Wave 1, strong connections emerged between the tobacco use nodes, between “Attention” and “Listening”, and “Fights” and “Bully”. The nodes were clustering based on their respective groups rather than the negative affect symptoms lying between the substance use behaviors and negative affect symptoms, as seen in the Wave 1 network. Edge-weights are shown in the Appendix D (Supplemental Table 5.2).

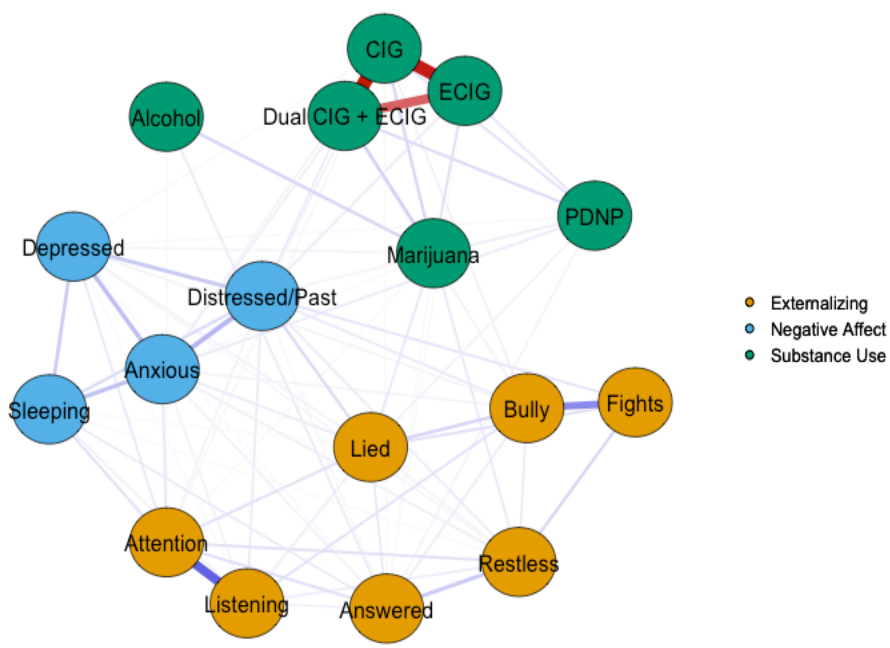


Figure 5.5: Visualization of Wave 2 Network

Wave 3

The Wave 3 network consisted of 17 nodes and had 95 non-zero edges out of 136 possible edges (density = 0.699), indicating that 69.9% of possible connections were identified in the network (Figure 5.6). The network structure is an Ising model, which is a network of partial correlation coefficients. Similar to Waves 1 and 2, strong connections emerged between the tobacco use nodes, between “Attention” and “Listening”, and “Fights” and “Bully”. Edge-weights are shown in the Appendix D (Supplemental Table 5.3).

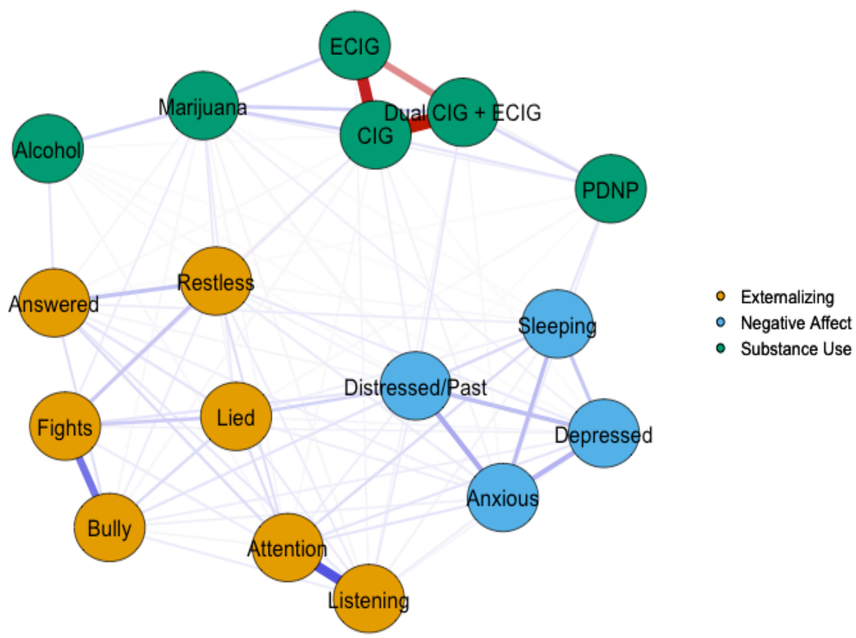


Figure 5.6: Visualization of Wave 3 Network

Wave 1 vs Wave 2 Comparison

There were not too many noticeable differences when visually comparing the Wave 1 and Wave 2 networks (Figure 5.7). The direction of the edges (e.g., positive or negative) was the same in both networks. The edge-weight between “Bully” and “Fights” appears larger in Wave 2 compared to Wave 1. Some nodes had more or fewer connections, depending on the network. For example, alcohol had six connections in the Wave 1 network versus only three connections in the Wave 2 network.

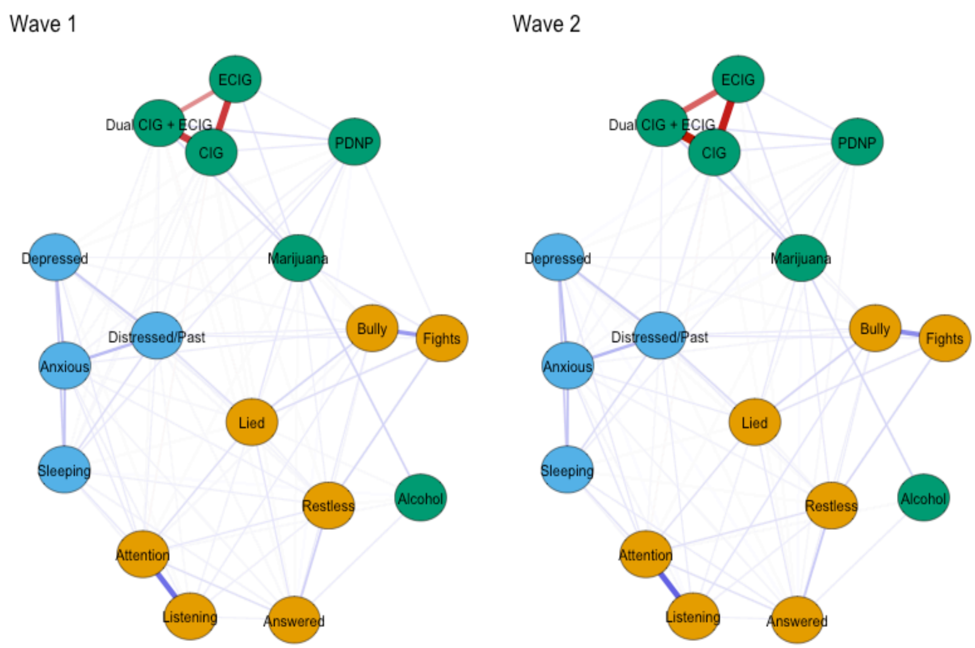


Figure 5.7: Visual Comparison of Wave 1 and Wave 2 Networks

Nine edges (edge-weights = EW) were significantly different ($p < 0.05$) between Wave 1 and Wave 2 (Table 5.11). Overall, these edges increased in magnitude from Wave 1 to Wave 2 except for the connections between sleeping problems—restless ($EW_{Wave\ 1}$= 0.37, $EW_{Wave\ 2}$ = 0.09), marijuana—PDNP ($EW_{Wave\ 1}$= 0.62, $EW_{Wave\ 2}$ = 0.37), and CIG—sleeping problems ($EW_{Wave\ 1}$= 0.16, $EW_{Wave\ 2}$ = 0). Some connections existed in one wave where it did not in another: CIG—sleeping ($EW_{Wave\ 1}$= 0.16, $EW_{Wave\ 2}$ = 0) and ECIG—distressed about the past ($EW_{Wave\ 1}$= 0, $EW_{Wave\ 2}$ = 0.42).

Table 5.11: Significant Edge Differences between Wave 1 and Wave 2

Node 1	Node 2	W1 Edge	W2 Edge	P-value
CIG	Marijuana	0.78	0.91	0.01
Marijuana	PDNP	0.62	0.37	0.01
CIG	Sleeping	0.16	0	0.04
ECIG	Distressed/Past	0	0.42	0.03
Anxious	Lied	0.25	0.46	0.03
Distressed/Past	Listening	0.20	0.46	0.01
Attention	Listening	3.47	3.72	0.02
Bully	Fights	2.40	2.80	0.04
Sleeping	Restless	0.37	0.09	0.01

Despite some node-specific relationships that differed by wave, the overall structure of the networks (maximum difference = 1.56, p-value = 0.23) and the global strength (Wave 1 = 56.0, Wave 2 = 59.3, p-value = 0.27) did not significantly differ between Wave 1 and Wave 2. Therefore, the overall structure and connectivity was not different between Wave 1 and Wave 2.

Wave 2 vs Wave 3 Comparison

There were fewer differences between Waves 2 and 3 versus Waves 1 and 2 when visually comparing the networks (Figure 5.8). The direction of the edges was the same in both networks. The magnitudes of the edge-weight appear very similar between Waves 2 and 3. Some nodes had more or fewer connections depending on the network. For example, marijuana had eight connections in the Wave 2 network versus twelve connections in the Wave 3 network.

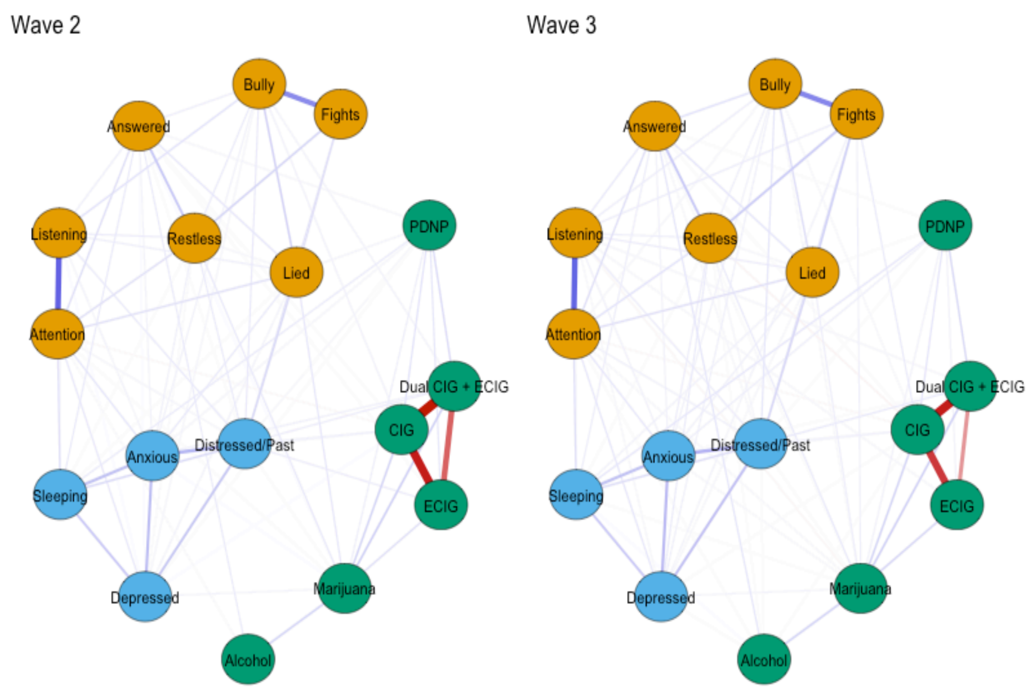


Figure 5.8: Visual Comparison of Wave 2 and Wave 3 Networks

Nine edges were significantly different ($p < 0.05$) between Wave 2 and Wave 3 (Table 5.12). Five edges increased in magnitude from Wave 2 to Wave 3: Dual CIG + ECIG—listening difficulties ($EW_{Wave\ 2}= 0$, $EW_{Wave\ 3} = 0.22$), listening difficulties—fighting ($EW_{Wave\ 2}= 0$, $EW_{Wave\ 3} = 0.40$), feeling depressed—restlessness ($EW_{Wave\ 2}= 0.14$, $EW_{Wave\ 3} = 0.44$), alcohol—answered ($EW_{Wave\ 2}= 0.44$, $EW_{Wave\ 3} = 0.55$), and bullying—answered ($EW_{Wave\ 2}= 0.38$, $EW_{Wave\ 3} = 0.65$). The remaining four edges decreased in magnitude from Wave 2 to Wave 3.

Table 5.12: Significant Edge Differences between Wave 2 and Wave 3

Node 1	Node 2	W2 Edge	W3 Edge	P-value
ECIG	Distressed/Past	0.42	0	0.03
Dual CIG + ECIG	Listening	0	0.22	0.02
Marijuana	Listening	0	-0.2	0.02
Listening	Fights	0	0.4	0.02
CIG	Restless	0	-0.29	0.02
Depressed	Restless	0.14	0.44	0.02
Distressed/Past	Restless	0.43	0.23	0.04
Alcohol	Answered	0.44	0.55	0.03
Bully	Answered	0.38	0.65	0.05

Despite the nine node-specific relationships that differed by wave, the overall structure of the networks (maximum difference = 1.32, p-value = 0.23) and the global strength (Wave 2 = 59.3, Wave 3 = 60.0, p-value = 0.75) did not significantly differ between Wave 2 and Wave 3. Therefore, the overall structure and connectivity was not different between Wave 2 and Wave 3.

Wave 1 vs Wave 3 Comparison

There were few noticeable differences when visually comparing the Wave 1 and Wave 3 networks (Figure 5.9). The direction of the edges (e.g., positive or negative) was the same in both networks. The edge-weight between “Bully” and “Fights” appears larger in Wave 3 compared to Wave 1.

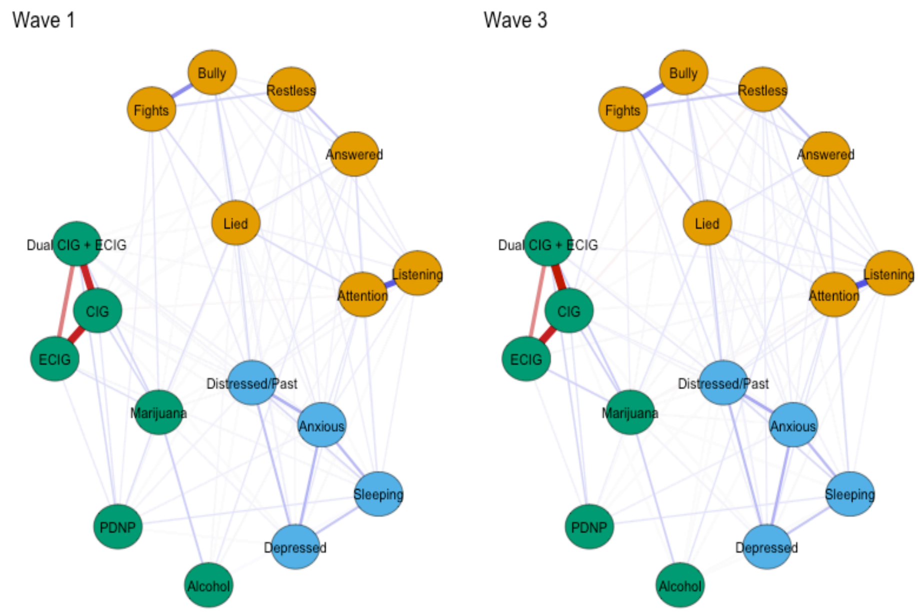


Figure 5.9: Visual Comparison of Wave 1 and Wave 3 Networks

Twenty-two edges were significantly different ($p < 0.05$) between Wave 1 and Wave 3 (Table 5.13). Half of these edges increased in magnitude, specifically CIG—marijuana ($EW_{Wave\ 1} = 0.78$, $EW_{Wave\ 3} = 0.93$), anxious—distressed about the past ($EW_{Wave\ 1} = 1.63$, $EW_{Wave\ 3} = 1.77$), and attention difficulties—listening difficulties ($EW_{Wave\ 1} = 3.47$, $EW_{Wave\ 3} = 3.66$).

Table 5.13: Significant Edge Differences between Wave 1 and Wave 3

Node 1	Node 2	W1 Edge	W3 Edge	P-value
CIG	Alcohol	0	-0.07	0.01
CIG	Marijuana	0.78	0.93	0.01
Marijuana	PDNP	0.62	0.32	0.01
Alcohol	Depressed	0	-0.09	0.03
CIG	Sleeping	0.16	0	0.04
ECIG	Sleeping	0.22	0	0.04
PDNP	Distressed/Past	0.31	0.08	0.04
Anxious	Distressed/Past	1.63	1.77	0.05
Alcohol	Lied	0.2	0	0.01
Marijuana	Lied	0.6	0.39	0.02
Sleeping	Lied	0.11	0.3	0.03
Sleeping	Attention	0.53	0.69	0.04
Dual CIG + ECIG	Listening	0	0.22	0.04
Marijuana	Listening	0	-0.2	0.02
Distressed/Past	Listening	0.2	0.4	0.05
Attention	Listening	3.47	3.66	0.03
Listening	Fights	0	0.4	0.03
Bully	Fights	2.40	2.88	0.03
CIG	Restless	-0.09	-0.29	0.04
Depressed	Restless	0.19	0.44	0.01
Sleeping	Restless	0.37	0.14	0.02
Listening	Answered	0.34	0.48	0.05

Global strength did not significantly differ between Wave 1 and Wave 3 (Wave 1 = 56.0, Wave 3 = 60.0, p-value = 0.24). There was not a significant difference in the maximum difference in edge weights between Waves 1 and 3 (maximum difference = 0.82, p-value = 0.60). Therefore, the overall structure and connectivity was not different between Wave 1 and Wave 3.

Wave 1, Wave 2, and Wave 3 Network

The network including Waves 1, 2, and 3 consisted of 51 nodes (Figure 5.10). The network had 233 non-zero edges out of 1275 possible edges (density=0.183), indicating that 18.3% of possible connections were identified in the network. The network structure is an Ising model, which is a network of partial correlation coefficients. Edge-weights within a respective wave reduced in magnitude. For example, the edge-weight between cigarette and e-cigarette for the Wave 1 only network was -4.74 and the edge-weight in the network with three waves is -2.28. Edge-weights are shown in the Appendix D (Supplemental Table 5.4-5.6).

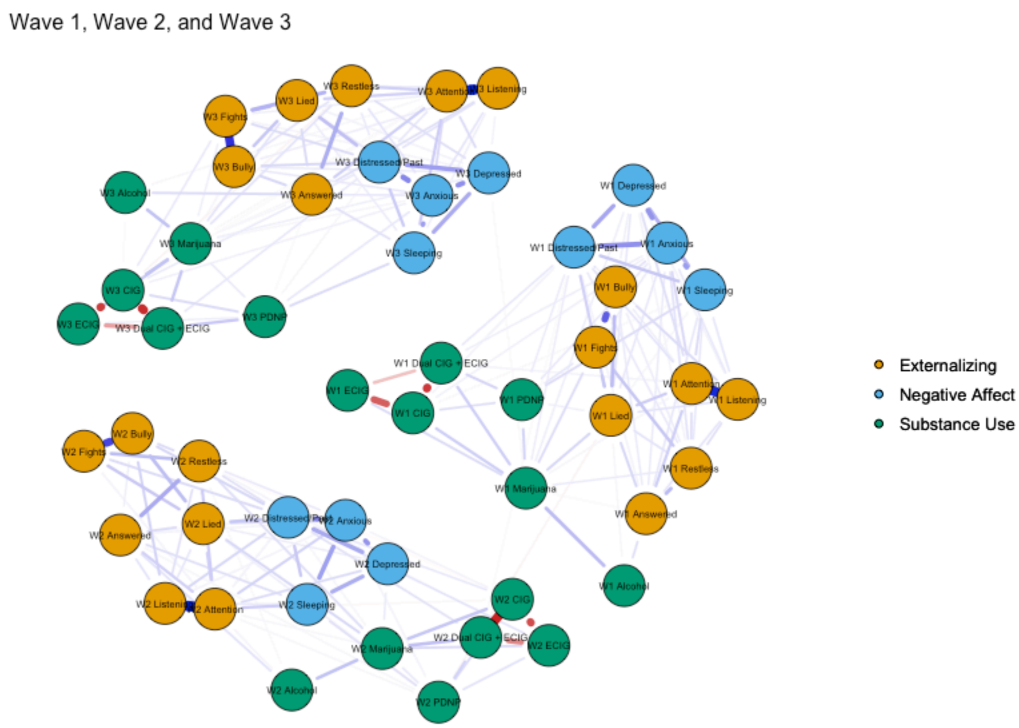


Figure 5.10: Visualization of the Wave 1, Wave 2, and Wave 3 Network

Nodes clustered by wave with very little overlap. Only two edges connected from Wave 1 to Wave 2: distressed about the past from Wave 1—dual cigarette and e-cigarette from Wave 2 (EW = 0.05, tetrachoric correlation = 0.02, p-value = 0.30), and lying from Wave 1—PDNP from Wave 2 (EW = -0.14, tetrachoric correlation = -0.02, p-value = 0.31). One edge connected from Wave 1 to Wave 3: marijuana from Wave 1—feeling depressed from Wave 3 (EW = -0.07, tetrachoric correlation = -0.03, p-value = 0.02). No edges connected from Wave 2 to Wave 3. All other tetrachoric correlations between waves were zero or near zero and were not statistically significant.

DISCUSSION

This is one of the first studies to use results from latent class and network analyses to preliminarily assess whether comorbidity between substance use behaviors and mental disorder symptoms changes over time in adults. There were three major results from this study. First, both latent class profiles and network analyses suggested that the comorbidity structure remained stable over time. Second, results from the latent class comparisons demonstrated that for people that did transition to another class, these transitions moved from a more severe class to a less severe or low symptom class. Third, the edge strength invariance test suggested stronger connections among the substance use behaviors and mental health symptoms from preceding to subsequent waves.

Overall stability in latent profiles with transitions to low symptom class

Similar latent profiles emerged across the three waves in the cross-sectional review of the four-class solution specifically where the low symptom class was largest (65.5% to 72.9%) and the comorbid class was smallest (6.1% to 8.2%). These consistencies also emerged in the multinomial regression analyses to determine which sociodemographic factors were significantly associated with class membership. These results support previous work that have established a stability in substance use and mental disorder comorbidity over time[67–69] and are generally consistent with youth and young adult studies.[227]

The largest difference was seen in Wave 3 where low comorbid and substance use classes emerged rather than the externalizing and negative affect classes from Waves 1 and 2. There are two potential reasons for this difference. First, it could be the case that a four-class solution is not most optimal in Wave 3. Results from the entropy and the Lo-Mendell-Rubin tests did not support the selection of a four-class solution as most optimal. Yet, a four-class solution was selected in order to easily compare the latent classes from Wave 3 to the latent classes from Waves 1 and 2. Another reason could be that the composition of the latent classes shifted meaning that the comorbidity profile changed from Wave 2 to Wave 3. However, this is not a probable reason because the same people are included at each wave and a shift in their comorbidity profiles is unlikely to occur over the course of a year.

The low symptom class was the most stable over the transition periods. Our results suggest that there was a greater transition from the original class (i.e., comorbidity, negative affect, externalizing, low comorbid or substance use) to the low symptom class compared to stability from the preceding to the subsequent waves. This is inconsistent with substance use comorbidity research in adolescents, as they have identified transitions from less to more severe substance use behaviors.[12,223,226,227] However, results are consistent with prior mental disorder comorbidity literature that explains both a continuity and a change.[67,227] This transition may indicate that the individual is receiving the resources and support necessary to remit to a less severe class. This could also result from characterizing symptom or behavior patterns rather than diagnoses. Past month endorsement of a symptom or behavior may be more flexible to transition compared to a diagnosis. A true longitudinal assessment like a latent transition analysis is needed to confirm these transitions over time.

No difference in network structure over time but an increase in association between symptoms and behaviors

There was no significant difference in overall structure and connectivity between any pairwise network comparison. This finding supports the stability discovered in the cross-sectional assessments of the latent classes. However, there were significant differences in edge weights between the waves. These differences (1) occurred within constructs (e.g., between two substance use behaviors) and across constructs (e.g., between a substance use behavior and negative affect symptom), and (2) generally demonstrated an increase in magnitude from the preceding wave to the subsequent wave. These discoveries enforce that the comorbidity structure was not dynamic, and that connections were becoming stronger across a three-year time period.

There are many reasons why connections may increase across time. This could be due to biological factors (i.e., onset of new disorder symptoms) or time-varying changes (i.e., age or an increase in education and income). Another important consideration, however, is the change in substance use and mental health conditions due to cultural or environmental shifts. For example, there is potential for greater access to and use of electronic cigarettes as new electronic nicotine delivery systems are developed. Marijuana is also becoming more widely available in the United States because of changes in policies either decriminalizing or legalizing recreational use. Other worldly events (e.g., the COVID-19 pandemic) may give rise to an increase in substance use behaviors and mental health problems.[32] The biological, time-varying, cultural and environmental shifts have the potential to increase the connections of substance use behaviors and mental disorder symptoms over time.

The network with all three waves of data was very sparse, and interesting connections were not found between nodes in different waves. The data were merged by an individual's identification number; however, time was not accounted for in the network model, meaning that this was not a true longitudinal analysis. Additionally, the nodes in these networks only capture past month endorsement of substance use behaviors and mental disorder symptoms. Therefore, it is unlikely that there would be connections detected between the waves because nodes within a respective wave represent a different time. Prior work has identified comorbid longitudinal relationships by using diagnostic level measures after one year[68] and three years[67] from data collected at baseline. Relationships have also been identified between depressive symptoms (measured as the frequency of depressive symptoms within one week using the Center for Epidemiologic Studies Depression Scale [CES-D]) and past month as well as past year major depressive disorder diagnosis over seven years of data collection.[229] However, these studies did not utilize a network approach. There are longitudinal network models for panel data in development that should be leveraged to more accurately account for the longitudinal nature of these data and explore possible connections over time.

Complementary latent class and network results

Both analyses support stability in the comorbidity structures over time. The preliminary assessments of comorbidity patterns of both approaches complement the major results that (1) the comorbidity structure exists and remains relatively stable over time, (2) if a transition occurs in the comorbidity profile, it is likely to move from more to less severe, and (3) connections among the substance use behaviors and mental disorder symptoms may be growing larger from preceding to subsequent wave.

Strengths and limitations

Results from this study should be interpreted with consideration of the following limitations. First, these data were collected between 2013-2016. Three years is likely not a long enough time frame to detect significant changes in comorbidity. Although there were considerable cultural and environmental changes during this time (like the increase and influx of e-cigarette availability and products, respectively, as well as changes in marijuana legislation across the U.S.), we recommend a more updated longitudinal assessment of the comorbidity structure especially in light of the COVID-19 pandemic. Second, the approaches used in this study were preliminary assessments of the comorbidity patterns in a nationally representative sample of adults. A latent transition analysis was not conducted. Additionally, we did not evaluate the differences in item response patterns due the overwhelming nature of possible combinations (i.e., seventeen items across three time periods). A true longitudinal assessment (i.e., latent transition analysis) is needed to ensure optimal class solution across the waves and confirm the transitions found in this study. Third, the ability to perform a longitudinal network analysis for panel data is currently limited. This method is in the early stages of development; therefore, we were limited to pairwise network comparisons. Researchers should consider using this new approach (i.e., cross-lagged network models) to investigate the comorbidity structure in future analyses. Fourth, PATH participants who did not start the study at Wave 1 were not included in our study. This decreased the sample size; however, we could account for any cohort effects by excluding them from the study. Furthermore, participants included in the analyses differed significantly from participants that were excluded due to missing data. Consequently, these results may not be generalizable to the U.S. adult population. Fifth, the network models did not adjust for the influence of other sociodemographic variables were not included in these analyses. Therefore, there may be some residual confounding. Sixth, accuracy and stability test for Waves 2 and 3 were not conducted; therefore, network results should be interpreted with caution.

Conclusions

This is the first study to use complementary statistical methods, latent class and network analysis, to evaluate substance use behavior and mental disorder symptom comorbidity patterns in adults over time. These results suggest that the comorbidity structure exists and remains stable. Furthermore, the connections between these behaviors and symptoms are possibly becoming stronger. Therefore, investment of time, money, and other resources are encouraged to support those experiencing comorbidity as they are unlikely to change in adulthood. It is important to target and maintain interventions based on comorbidity structures because the structure is not changing in composition, but is changing in strength. There is a need to support people based on the comorbidities that they present, not just one behavior or symptom at a time.

CHAPTER 6: DISCUSSION

To date, current approaches to detection and prevention of comorbid SUD have been limited by a focus on SUD exclusively even though substance use often co-occurs across multiple substances and is often comorbid with mental disorders. This dissertation sought to address the following knowledge gaps: (1) current SUD research suffers from the unidimensional approach that does not account for comorbidity; (2) patterns of comorbidity are not the same, although current knowledge is based on homogeneous samples; and (3) it is unclear whether patterns of comorbidity remain stable or changes. Therefore, the goal of this dissertation was to characterize the comorbidity between substance use, including tobacco use, and mental disorder symptoms measured as negative affect and externalizing symptoms in a population-based sample by: preliminarily assessing comorbidity using multinomial regression between lifetime negative affect severity, externalizing severity and nicotine dependence, and current use of tobacco (cigarettes and e-cigarettes) and alcohol (Chapter 2); identifying latent classes of comorbid substance use as well as negative affect and externalizing symptoms and their ability to predict SUD severity (Chapter 3); detailing substance use, negative affect, and externalizing symptom networks and testing for differences in the network structure and connectivity by gender (Chapter 4); and using pairwise comparisons from the LCA and network results to address stability or movement of comorbidity structures over three waves of data (Chapter 5).

Confirmation of the comorbidity structure in U.S. adults prompts a multidimensional approach to substance use and mental disorders

Results from Chapters 3 and 4 distinguished different comorbidity groups and identified how the substance use and mental disorder symptoms connected with each other, respectively. These results confirm a robust comorbidity structure in U.S. adults by characterizing seventeen substance use behaviors and mental disorder symptoms into groups and identifying specific connections. The LCA results suggest that approximately 21% of the sample made up the negative affect or externalizing classes at Waves 1 and 2, and about 26% made up the low comorbid or substance use classes at Wave 3. These individuals reflect a subpopulation with possible subthreshold levels of impairment that may not be identified with current classification systems of substance use and mental disorders. Network analysis results confirm that connections between behaviors and symptoms overlap and cut across constructs (i.e., diagnostic boundaries). Furthermore, these results identified comorbidity patterns, not singular disorder in the population. This provides reason to reconsider our current unidimensional approach to substance use and mental disorder comorbidity because the prevalence of potential subthreshold level comorbidity in the population is happening at a greater rate than the high comorbidity class (21-26% vs 6-8%). These results support inclusion and regular study of additional substance use behaviors and mental disorder symptoms at subthreshold levels. The robust comorbidity structure can provide insight into the overall wellbeing of an individual. Dissertation results have the potential to increase comorbidity awareness in clinicians and further help clinicians to better target comorbidity because specific aspects of substance use and mental disorder symptoms that are more likely to occur with each other were identified. Future studies could build upon these findings and explore how this comorbidity awareness can be applied to clinical settings. In the future, comorbidity research done in the clinical space could be translated to encourage increased communication about substance use behaviors and mental disorder symptoms with clinicians and their clients to consider comorbidity during health screenings, and support those affected with multiple conditions more efficiently.

Substance use varied by mental disorder symptoms suggesting different comorbidity profiles

Analysis of the Population Assessment of Tobacco and Health Study provided a unique opportunity to study patterns of comorbidity across multiple tobacco products (i.e., electronic- and conventional- cigarette use) in addition to substance use and mental health comorbidity. A relatively novel tobacco product, e-cigarettes, was included in the study while also accounting for the dual use of conventional cigarettes and e-cigarettes rather than simply classifying any tobacco use through a measure of nicotine dependence or considering conventional cigarette use only. Dual users represent a novel and distinct class of tobacco users that must be accounted for, especially when exploring comorbidity.[230] This approach (1) allowed for a more detailed investigation into how tobacco products present and connect with comorbid substance use behaviors and mental disorder symptoms, and (2) limited the potential misclassification bias introduced when dual users are not classified outside of conventional or e-cigarette use.

Results from Chapters 3 and 4 indicated that conventional cigarette use and dual use of conventional cigarettes and e-cigarettes were associated with negative affect symptoms, while the exclusive use of e-cigarettes was associated with externalizing symptoms. This result differs from a prior study from Conway et al. (2017) that reported e-cigarette use to have a larger magnitude of association with negative affect/internalizing severity compared to externalizing symptoms.[73] This difference is likely due to the classification of tobacco product use. The Conway et al. paper measured current e-cigarette use without excluding conventional cigarette use. Nevertheless, dual use of e-cigarette and cigarette use is increasing in the U.S.[83] and about 16% of current smokers were also current e-cigarette users in 2014.[85] Additionally, results in Chapter 2 identified patterns of association with mental disorder symptoms varied by dual use and exclusive use. The results from this dissertation encourage the study of three separate classes of tobacco products (i.e., exclusive conventional cigarette use, exclusive e-cigarette use, and dual use of conventional and e-cigarettes) in order to provide a clearer understanding of comorbidity profiles related to substance use behaviors and mental disorder symptoms in U.S. adults.

Chapter 2 determined that the associations between psychopathology (negative affect vs. externalizing severity) varied by different combinations of alcohol, conventional cigarette and e-cigarette use. Negative affect severity was associated with cigarette and alcohol use together as well as alcohol-exclusive use, while externalizing severity was associated with e-cigarette and alcohol use together. These results confirm that associations between negative affect and externalizing severity varies by different combinations of alcohol, cigarette, and e-cigarette use.

Chapter 3 results built on those of Chapter 2 by including a more robust set of substance use and mental disorder symptom variables (i.e., adding marijuana, prescription drugs not prescribed [PDNP], four negative affect symptoms, and seven externalizing symptoms) and extending past multinomial regression by using a latent class analysis approach to detail the patterns of specific substance use behaviors that have different relationships with mental disorder symptoms. Specifically, exclusive cigarette use, dual cigarette and e-cigarette use, marijuana use, and PDNP were associated with a negative affect class. In contrast, exclusive e-cigarette and alcohol use were associated with an externalizing class. Results from Chapter 3 confirmed the relationship between conventional cigarette use and negative affect identified in Chapter 2, and provided more clarity on the relationships between dual cigarette and e-cigarette use and negative affect as well as exclusive e-cigarette use and alcohol use with externalizing symptoms.

Chapter 4 complemented the results from Chapter 3 by identifying the magnitudes of specific connections between a unique pair of variables. PDNP, marijuana use, dual use of cigarette and e-cigarette, and conventional cigarette use had strong connections with negative affect symptoms. PDNP use was most strongly associated with negative affect symptoms. Marijuana and alcohol use were most strongly associated with externalizing symptoms. Results from the nodewise predictability analysis identified which nodes were most important in influencing the other nodes in the network, an important discovery regarding intervention.

It is often thought that interventions can be best developed using longitudinal data. However, the use and incorporation of marginal effects in models have been utilized more frequently in health systems research to establish expectations related to interventions, particularly for cross-sectional data.[231] An underappreciated result from network models are the estimates of nodewise predictability, which are produced using marginal effects. The nodewise predictability results discussed in Chapter 4 provided a quantitative understanding into how effective intervention could be as well as guidance on how to intervene on substance use behaviors and mental disorder symptoms (i.e., through a specific node of interest or neighboring nodes). Most of the nodes in the model had a small normalized accuracy, meaning that most of the accuracy or predictability of these nodes in the network were due to the contribution of the node in question specifically rather than through the contribution of other nodes. The negative affect symptoms, attention problems, listening problems, and impulsivity had larger normalized accuracy in that the accuracy of these symptoms had larger contributions by other nodes in the network. Therefore, intervention on any of the nodes would likely influence any other behavior or symptom in the network since the network was largely determined by itself through strong mutual interactions between nodes.

These results could help to inform future research in clinical spaces to target specific behaviors and symptom combinations. This type of research could identify a potential opportunity for clinicians and their patients/clients to have an open conversation about substance use behaviors that may influence their mental health and vice versa. A clinician could consider alternative approaches for someone with comorbidity versus someone affected with a single condition. For example, if a person were to present with co-occurring dual cigarette and e-cigarette use, a clinician could consider asking questions about the person’s co-occurring negative affect symptoms.

The alternative can happen as well: if a person were to present with negative affect symptoms (i.e., feeling depressed, feeling anxious, experiencing sleeping problems, and/or feeling distressed about the past), a clinician could ask questions about the person’s co-occurring tobacco use, specifically conventional cigarette use or the combination of cigarettes and e-cigarettes together.

Sociodemographic characteristics were associated with comorbidity

The results in Chapters 3, 4, and 5 were consistent with previous work, particularly as it applies to gender.[2,41,123] For instance, compared to men, women had greater odds of membership in the comorbid, low comorbid, and negative affect latent classes (Chapters 3 and 5). Compared to women, men had greater odds of membership in the substance use latent class (Chapter 5). Results from Chapter 4 expanded the gender difference literature related to comorbidity by identifying specific connections between comorbid substance use and mental disorder symptoms by gender. Specifically, alcohol use and sleeping problems, exclusive e-cigarette use and lying, alcohol use and lying, and alcohol use and attention difficulties were all stronger for men than they were for women.

Chapters 3 and 5 emphasized the importance of age on comorbidity. Participants of any age category (i.e., 18-24 years, 25-34 years, 35-44 years, 45-54 years, and 55-64 years) compared to those ages 65 and older, had greater odds of latent class membership for all classes. The magnitude of the association gradually decreased as age increased. This is consistent with previous work where younger people are at greater risk for mental health and substance use problems compared to people in older age categories.[41] This matches the age of substance use initiation which typically occurs in younger age categories, and age of onset for most mental disorders as roughly 50% to 75% of all lifetime mental disorders start by the mid-teens and mid-20s, respectively.[2,154] Therefore, strategies targeting younger ages, specifically those between the ages of 18-24 years, could be helpful in reducing comorbidity in younger ages and possibly prevent comorbidity as age increases.

Chapters 3 and 5 highlighted the role of race/ethnicity on comorbidity and encourage additional study in this area. For example, participants who described themselves as belonging to non-White racial categories (i.e., Non-Hispanic Black, Non-Hispanic Other, Hispanic Multiracial) were less likely to be in any of the following latent classes compared to those who categorized themselves as Non-Hispanic White: comorbid, externalizing, negative affect, low comorbid, and substance use classes. This result does support other findings typically identified in the Black-White mental health paradox.[158] This paradox has generally supported the idea that Black Americans experience similar or relatively low rates of psychiatric disorders compared to Whites despite higher stress exposure, greater material hardship, and worse physical health.[158] Previous work exploring the Black-White mental health paradox has focused on single psychiatric conditions[158] and these results identify that this paradox is also present for comorbidity.

Chapters 3 and 5 identified the role of education and income in comorbidity. In general, low education and income were positively associated with membership in the comorbid, negative affect, and substance use latent classes. However, a negative relationship was discovered with low education and income and the externalizing class in Waves 1 and 2, and the low comorbid class in Wave 3. Prior work has identified that higher education and income levels represent a protective relationship from membership in internalizing or negative affect, externalizing, and high psychopathology classes.[41] Therefore, our result of low education and income being less likely to occur with externalizing and low comorbid classes is different than what has previously been identified. Future research should continue to include socioeconomic status variables in the assessment of comorbidity to further clarify this association.

A social support variable was included in Chapters 3 and 5 because the relationship between social support and substance use behaviors/disorders (1) is well-established in youth, but results are mixed, and (2) may be a potential modifiable factor to use as part of intervention strategies to address substance use and mental disorder symptom comorbidity. This variable also provided insight into how an individual's interpersonal relationships were associated with comorbidity as previous research has only focused on a single outcome (e.g., substance use only).[232,233] The associations between social satisfaction and latent class membership reflected a potential dose-response relationship where a decrease in social satisfaction significantly increased odds of class membership. This represents a very interesting opportunity for potential intervention because social satisfaction is an easier factor to influence or change compared to the other demographic factors included in the analysis (i.e., sex, age, race, education, and income). Specifically, the probability of class membership in comorbid, negative affect, low comorbid, or substance use classes could decrease if social satisfaction can be improved by increasing satisfaction with activities and relationships. Epidemiologic and community-based participatory research studies have identified the benefit of improving social support and relationships to reduce the likelihood of developing mental health and substance use problems.[234–239] Consequently, social satisfaction should be considered and implemented for public health prevention strategies related to substance use and mental disorder symptom comorbidity, supporting and expanding community-wide efforts to develop and increase social satisfaction.

Comorbidity structure remained stable with transition to lower severity groups but identification of stronger connections across three data points

Results from Chapter 5 confirmed prior research regarding substance use and mental disorder symptom comorbidity[67,69]: the behaviors were stable across three years. Both the LCA and network analyses showed that the overall comorbidity profiles and network structures were consistent across waves. Further, evaluation of the possible latent class transitions among the waves identified that people more commonly transitioned from more severe class to a less severe class. However, stronger connections were discovered in subsequent waves when specifically testing for significant differences in edge-weights of substance use and mental disorder symptom connections between the waves. Consequently, the connections between these behaviors and symptoms may become stronger over time. Investment of time, money, and other resources early in adulthood are encouraged to support those experiencing comorbidity as the co-occurring behaviors and symptoms are likely to become more severe in adulthood.

Latent class analysis and network analysis produce complementary results

Although LCA and network approaches are different and follow different conceptual frameworks, results from both arrived at similar conclusions described in Table 6.1. Both LCA and network analysis identified relationships between (1) exclusive cigarette, dual cigarette and e-cigarette, marijuana, and PDNP with negative affect symptoms, and (2) alcohol with externalizing symptoms.

Table 6.1: Associations between substance use and mental disorder symptoms identified through LCA and/or network analysis

Past Month Substance Use	**Past Month Mental Disorder Symptom**	**LCA**	**Network Analysis**
Exclusive cigarette	Negative affect	Yes	Yes
Dual cigarette and e-cigarette	Negative affect	Yes	Yes
Marijuana	Negative affect	Yes	Yes
	Externalizing		Yes
PDNP	Negative affect	Yes	Yes
Exclusive e-cigarette	Externalizing	Yes	
Alcohol	Externalizing	Yes	Yes

Latent class analysis was best at distinguishing different comorbidity in the population while also accounting for the potential influence of sociodemographic factors compared to the network analysis. Although the latent class analysis was unsuccessful at using latent class membership to predict SUD severity, a strong relationship between class membership and SUD severity was detected. This confirms and underscores the importance of the relationship between comorbidity and SUD severity.

Network analysis was best at demonstrating the total number of connections between substance use behaviors and mental disorder symptoms compared to the LCA, while also showing which behaviors/symptoms were most influential in the comorbidity network. These results identify important comorbid substance use behaviors and mental disorder symptoms, informing a more targeted approach to comorbidity. There were no significant differences in network structure or connectivity by gender, but specific connections were different and these differences were consistent with other literature.[2,153,209,240–243]

These approaches complement each other because they fill the gaps of the other approach. Latent class results identified heterogeneous groups in the population which helped to inform which items were likely to happen with each other. Network analysis results provided information regarding the strength of associations between two nodes. For example, alcohol use had relatively high item response probabilities across all classes, but was highest in the externalizing class. Network analysis results identified that alcohol use was more strongly associated with the impulsivity externalizing symptom compared to other externalizing symptoms. In addition to identifying which substance use behaviors or mental disorder symptoms likely occur with one another, network analysis complements the latent class results by identifying the magnitude of the associations. Latent variable and network approaches should continue to be used in comorbidity studies to further explore the comorbidity structure in other populations including additional substance use behaviors and mental disorder symptoms.

Future considerations to address dissertation limitations

Symptom-level data (i.e., past month endorsement of substance use and experiencing negative affect/externalizing symptoms) were used in these analyses to address research gaps identified in Chapter 1. A strength of using symptom-level data was that it limited recall bias and accurately accounted for comorbidity overlap (i.e., comorbidity occurring within the same time frame). However, it did not identify problematic or severe comorbidity. This limitation was obvious in two places in the dissertation. First, the poor ability to predict SUD severity using latent class membership in Chapter 3 may have been because of the past-month measurement. Second, past-month alcohol use was not necessarily indicative of problematic or harmful alcohol use. This point was acknowledged specifically in Chapters 4 and 5 regarding why the alcohol node was not well centralized or connected to others in the network. These points should be considered when interpreting results from this dissertation.

Results estimated in this dissertation may be subject to bias due to missing data. The sample was large and missing data did not influence the statistical power of the models tested. However, the missing data may have represented a misclassification bias in two ways. First, participants with missing data were significantly different than those included in the analyses. Those included in the analyses were more likely to endorse substance use, negative affect symptoms, and externalizing symptoms compared to those who were missing. Also, the analytic sample were more likely to be Non-Hispanic white, men, aged 25-54 years with higher levels of education and annual household income than those who were missing. Consequently, these results may not be generalizable to the U.S. adult population. Second, there was an expectation for social desirability bias to play a role in the missing data, meaning that participants might be less likely to endorse their true substance use behaviors and negative affect/externalizing symptoms because of the stigmatization surrounding these measures. This effect is expected to underestimate the study results. Although those included in the analysis were more likely to endorse substance use behaviors and mental disorder symptoms, social desirability may still be at play and should be considered when interpreting results.

Other statistical approaches are encouraged as the comorbidity research continues to develop. First, the use of LCA was limited by the conditional independence assumption and its inability to account for heterogeneous groups within the population (i.e., SUD and no SUD). Consequently, factor mixture modeling (FMM)[131] is suggested to address this limitation. Unlike LCA, FMM does not operate under the conditional independence assumption, meaning that it is not the latent class only that truly defines why the classes emerge as they do. FMM may also better account for the people with and without SUD in the sample and, therefore, has the potential to create latent classes that better predict SUD severity.

Second, the comorbidity structure using LCA and network analysis was assessed separately without the ability to account for both the variance that is unique to pairs of variables (network approach) and the variance that is shared across all variables (LCA approach). Therefore, a hybrid latent class and network model, also referred to as residual network modeling[244], should be a method considered in future work. The hybrid latent class and network model allows for the estimation of structural equation modeling (like LCA) without the assumption of conditional independence, and the estimation of a network structure, while considering the fact that the covariance between items may be partly due to latent factors.[244] This approach may further detail the etiology of comorbidity.

Finally, preliminary assessments of the comorbidity patterns over time were done by assessing the latent class and network structure cross-sectionally at three separate time periods. However, a true longitudinal analysis to test whether the stability or changes were statistically significant over time was not performed. Latent transition analysis[70,226] is a necessary next analysis to confirm the suspected trends discovered in Chapter 5. A time-series network model[245] should also be considered to similarly estimate the comorbidity network structure over time. This method for panel data is in development, but early results suggest it could be the network equivalent to a latent transition analysis.[245]

Implications of dissertation results and final conclusions

In summary, there are three specific results from the dissertation that could apply to public health practice. First, identification of specific substance use and mental disorder symptom connections can be a useful starting point in discussing comorbidity. Past-month PDNP was consistently identified to be strongly associated with negative affect symptoms while alcohol use was consistently identified to be strongly associated with externalizing symptoms. Therefore, building awareness of co-occurring negative affect and externalizing symptoms in individuals who are engaged in these past month substance use behaviors is an appropriate strategy in approaching the comorbidity conversation and future comorbidity research particularly in clinical spaces.

Second, the nodewise predictability results showed strong mutual interactions between all nodes. This implies that interventions on any of the six substance use behaviors, four negative affect symptoms, and seven externalizing symptoms would likely result in a change in the comorbidity network. Nodes with a greater proportion of predictability due to other nodes (i.e., negative affect items, attention difficulties, listening difficulties, and impulsivity) may be important to target due to their influence of other nodes in the network. However, targeting one specific behavior or symptom may not be the most effective strategy specifically given the pairwise comparison results that show a stability in the comorbidity profile. Therefore, it is likely that interventions might be most effective when targeting multiple behaviors and symptoms together.

Last, sociodemographic variables can be helpful in identifying potential risk for specific comorbidity profiles. For example, a young woman between the ages of 18-24 years with a lower education level or income is at potential risk for membership in the comorbid or negative affect classes. This demographic information could be used in public health practice to offer services or programs to people who may likely fit into this risk profile. Studies have identified the use of individual characteristics to create risk profiles in machine learning algorithms to predict substance use disorder treatment success.[246–248] Risk profiles have been generated and used in community and clinical settings to effectively target interventions.[249,250] Additionally, the dose-response relationship identified with social satisfaction and comorbidity represents a unique opportunity to encourage overall social support and healthy interpersonal relationships, especially when providing mental health and substance use services. Some studies have identified that social support interventions (e.g., support group involvement and utilizing family/friend support in a community-based substance abuse program) resulted in reduced substance use.[251–253] Improving social satisfaction could result in reduced substance use and may be extended to reducing comorbidity.

Characterizations of the comorbidity structure provide more information on how to approach substance use and mental disorders. Using a large sample of U.S. adults, this study identified specific combinations of substance use behaviors and mental disorder symptoms, determined which sociodemographic factors play a role in specific comorbidity profiles, and assessed the patterns of comorbidity among three waves of data. These results support the need to approach substance use and mental disorders from a more holistic perspective, taking comorbidity into account to better support the overall wellbeing of the individual. The results can inform robust and targeted prevention strategies to effectively mitigate the substantial burden and societal costs of comorbidity in the U.S. population.

APPENDICES

APPENDIX A: CHAPTER 2

Supplemental Table 2.1: Model 2 - Presentation of Different Reference Levels for Current Substance Use Outcome (Including Nicotine Dependence) (n = 15,947, Weighted N = 61,482,491)

Variable	Alcohol, Cigarette, and E-cigarette AOR (95% CI)	Cigarette and E-cigarette AOR (95% CI)	E-cigarette and Alcohol AOR (95% CI)	Cigarette and Alcohol AOR (95% CI)	E-cigarette Only AOR (95% CI)	Cigarette Only AOR (95% CI)	Alcohol Only AOR (95% CI)	None AOR (95% CI)
Negative Affect Severity (ref=low)								
Moderate	REF	0.91 (0.51-1.63)	0.73 (0.39-1.38)	1.11 (0.74-1.67)	0.78 (0.45-1.38)	0.92 (0.61-1.40)	1.19 (0.81-1.75)	0.75 (0.48-1.18)
High	REF	0.67 (0.40-1.11)	0.78 (0.44-1.38)	0.84 (0.56-1.27)	0.65 (0.38-1.10)	0.71 (0.46-1.10)	0.86 (0.57-1.28)	0.65 (0.43-1.01)
Externalizing Severity (ref=low)								
Moderate	REF	0.82 (0.48-1.39)	1.61 (0.92-2.82)	0.99 (0.66-1.51)	0.62 (0.34-1.10)	**0.65 (0.44-0.97)**	1.14 (0.73-1.77)	0.70 (0.46-1.08)
High	REF	0.69 (0.41-1.17)	1.66 (0.94-2.94)	0.86 (0.57-1.28)	0.65 (0.35-1.18)	**0.49 (0.33-0.74)**	0.98 (0.64-1.49)	**0.56 (0.36-0.87)**
Nicotine Dependence	REF	**1.02 (1.01-1.03)**	**0.94 (0.93-0.95)**	**0.99 (0.98-1.00)**	**0.95 (0.93-0.96)**	1.00 (0.99-1.00)	**0.92 (0.91-0.93)**	**0.94 (0.94-0.95)**
Negative Affect Severity (ref=low)								
Moderate	1.10 (0.62-1.96)	REF	0.81 (0.46-1.43)	1.22 (0.81-1.82)	0.86 (0.44-1.69)	1.01 (0.68-1.50)	1.30 (0.87-1.96)	0.83 (0.51-1.35)
High	1.50 (0.90-2.48)	REF	1.17 (0.67-2.02)	1.26 (0.88-1.81)	0.97 (0.58-1.64)	1.06 (0.75-1.51)	1.28 (0.90-1.83)	0.98 (0.65-1.47)
Externalizing Severity (ref=low)								
Moderate	1.23 (0.72-2.08)	REF	**1.97 (1.11-3.50)**	1.22 (0.84-1.78)	0.75 (0.46-1.25)	0.80 (0.56-1.14)	1.39 (0.93-2.09)	0.86 (0.57-1.30)
High	1.46 (0.86-2.47)	REF	**2.42 (1.32-4.44)**	1.24 (0.87-1.78)	0.94 (0.53-1.68)	0.72 (0.50-1.03)	1.42 (0.95-2.12)	0.82 (0.54-1.23)
Nicotine Dependence	**0.98 (0.98-0.99)**	REF	**0.93 (0.92-0.94)**	**0.97 (0.97-0.98)**	**0.93 (0.92-0.94)**	**0.98 (0.98-0.99)**	**0.91 (0.90-0.91)**	**0.93 (0.92-0.94)**
Negative Affect Severity (ref=low)								
Moderate	1.36 (0.72-2.57)	1.24 (0.70-2.19)	REF	1.51 (0.97-2.35)	1.07 (0.56-2.06)	1.26 (0.79-1.99)	**1.62 (1.03-2.55)**	1.03 (0.64-1.64)
High	1.29 (0.72-2.28)	0.86 (0.50-1.49)	REF	1.08 (0.73-1.61)	0.84 (0.46-1.52)	0.91 (0.61-1.36)	1.10 (0.73-1.65)	0.84 (0.55-1.28)
Externalizing Severity (ref=low)								
Moderate	0.62 (0.35-1.09)	**0.51 (0.29-0.90)**	REF	**0.62 (0.42-0.91)**	**0.38 (0.23-0.63)**	**0.40 (0.28-0.59)**	0.71 (0.47-1.06)	**0.44 (0.29-0.66)**
High	0.60 (0.34-1.06)	**0.41 (0.23-0.76)**	REF	**0.51 (0.34-0.78)**	**0.39 (0.22-0.68)**	**0.30 (0.19-0.46)**	**0.59 (0.38-0.91)**	**0.34 (0.21-0.55)**
Nicotine Dependence	**1.06 (1.05-1.07)**	**1.08 (1.06-1.09)**	REF	**1.05 (1.04-1.06)**	1.00 (0.99-1.02)	**1.06 (1.05-1.07)**	**0.97 (0.96-0.99)**	1.00 (0.99-1.01)
Negative Affect Severity (ref=low)								
Moderate	0.90 (0.60-1.36)	0.82 (0.55-1.23)	0.66 (0.43-1.03)	REF	0.71 (0.44-1.15)	**0.83 (0.71-0.98)**	1.07 (0.92-1.26)	**0.68 (0.57-0.82)**
High	1.19 (0.79-1.80)	0.79 (0.55-1.14)	0.93 (0.62-1.38)	REF	0.77 (0.49-1.23)	**0.84 (0.73-0.97)**	1.02 (0.85-1.22)	**0.78 (0.62-0.97)**
Externalizing Severity (ref=low)								
Moderate	1.01 (0.66-1.53)	0.82 (0.56-1.20)	**1.62 (1.10-2.39)**	REF	**0.62 (0.41-0.94)**	**0.65 (0.58-0.74)**	1.14 (0.97-1.36)	**0.71 (0.58-0.86)**
High	1.17 (0.78-1.76)	0.80 (0.56-1.15)	**1.95 (1.28-2.95)**	REF	0.76 (0.47-1.22)	**0.58 (0.51-0.66)**	1.14 (0.94-1.39)	**0.66 (0.52-0.82)**
Nicotine Dependence	**1.01 (1.01-1.02)**	**1.03 (1.02-1.03)**	**0.95 (0.94-0.96)**	REF	**0.96 (0.95-0.97)**	**1.01 (1.01-1.01)**	**0.93 (0.93-0.93)**	**0.95 (0.95-0.96)**

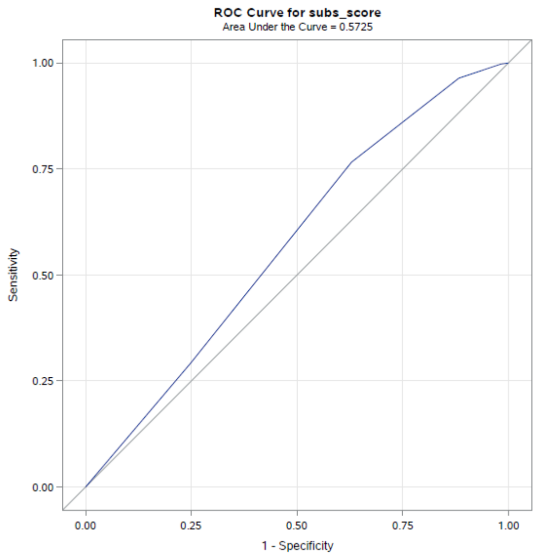
Bolded values indicate estimate significant a p < 0.05
Each model adjusts for sex, age, race, education, and annual household income.

Supplemental Table 2.1 CONTINUED: Model 2 - Presentation of Different Reference Levels for Current Substance Use Outcome (Including Nicotine Dependence) (n = 15,947, Weighted N = 61,482,491)

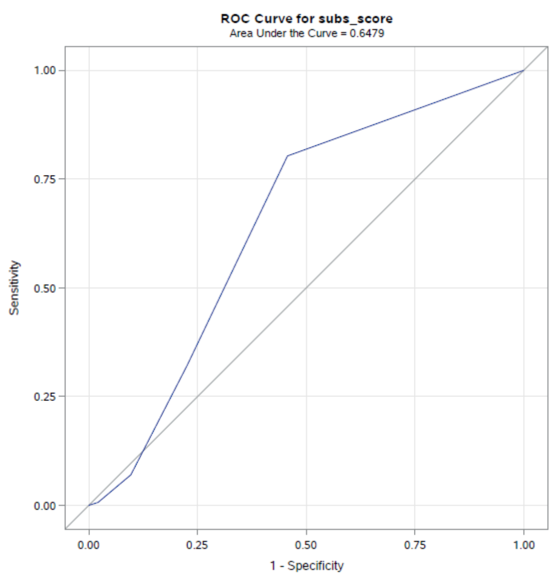
Variable	Alcohol, Cigarette, and E-cigarette AOR (95% CI)	Cigarette and E-cigarette AOR (95% CI)	E-cigarette and Alcohol AOR (95% CI)	Cigarette and Alcohol AOR (95% CI)	E-cigarette Only AOR (95% CI)	Cigarette Only AOR (95% CI)	Alcohol Only AOR (95% CI)	None AOR (95% CI)
Negative Affect Severity (ref=low)								
Moderate	1.28 (0.73-2.24)	1.16 (0.59-2.28)	0.94 (0.49-1.80)	1.41 (0.87-2.29)	REF	1.18 (0.72-1.91)	1.51 (0.95-2.42)	0.96 (0.60-1.55)
High	1.54 (0.91-2.61)	1.03 (0.61-1.72)	1.20 (0.66-2.18)	1.29 (0.81-2.06)	REF	1.09 (0.69-1.71)	1.32 (0.84-2.05)	1.00 (0.63-1.61)
Externalizing Severity (ref=low)								
Moderate	1.63 (0.91-2.91)	1.33 (0.80-2.19)	**2.62 (1.59-4.31)**	**1.62 (1.06-2.46)**	REF	1.06 (0.71-1.58)	**1.85 (1.21-2.83)**	1.14 (0.73-1.79)
High	1.55 (0.85-2.83)	1.06 (0.60-1.90)	**2.57 (1.47-4.51)**	1.32 (0.82-2.14)	REF	0.76 (0.47-1.24)	1.51 (0.93-2.47)	0.87 (0.51-1.46)
Nicotine Dependence								
	1.06 (1.05-1.07)	**1.08 (1.06-1.09)**	1.00 (0.99-1.01)	**1.05 (1.04-1.06)**	REF	**1.06 (1.05-1.07)**	**0.97 (0.96-0.98)**	1.00 (0.99-1.01)
Negative Affect Severity (ref=low)								
Moderate	1.09 (0.72-1.65)	0.99 (0.67-1.47)	0.80 (0.50-1.26)	**1.20 (1.03-1.41)**	0.85 (0.52-1.38)	REF	**1.29 (1.07-1.55)**	0.82 (0.65-1.04)
High	1.41 (0.91-2.19)	0.94 (0.66-1.34)	1.10 (0.74-1.64)	**1.19 (1.03-1.38)**	0.92 (0.58-1.45)	REF	**1.21 (1.00-1.46)**	0.92 (0.73-1.17)
Externalizing Severity (ref=low)								
Moderate	**1.54 (1.03-2.29)**	1.26 (0.88-1.80)	**2.47 (1.68-3.64)**	**1.53 (1.35-1.73)**	0.95 (0.63-1.41)	REF	**1.75 (1.45-2.10)**	1.08 (0.89-1.32)
High	**2.03 (1.36-3.02)**	1.39 (0.97-2.00)	**3.37 (2.18-5.21)**	**1.73 (1.53-1.97)**	1.31 (0.81-2.13)	REF	**1.98 (1.62-2.42)**	1.13 (0.90-1.43)
Nicotine Dependence								
	1.00 (1.00-1.01)	**1.02 (1.01-1.03)**	**0.95 (0.94-0.95)**	**0.99 (0.99-0.99)**	**0.95 (0.94-0.96)**	REF	**0.92 (0.92-0.93)**	**0.95 (0.94-0.95)**
Negative Affect Severity (ref=low)								
Moderate	0.84 (0.57-1.24)	0.77 (0.51-1.15)	**0.62 (0.39-0.98)**	0.93 (0.80-1.09)	0.66 (0.41-1.06)	**0.78 (0.64-0.94)**	REF	**0.64 (0.51-0.79)**
High	1.17 (0.78-1.75)	0.78 (0.55-1.11)	0.91 (0.61-1.36)	0.98 (0.82-1.18)	0.76 (0.49-1.19)	**0.83 (0.69-1.00)**	REF	**0.76 (0.61-0.96)**
Externalizing Severity (ref=low)								
Moderate	0.88 (0.56-1.37)	0.72 (0.48-1.08)	1.42 (0.94-2.13)	0.87 (0.74-1.03)	**0.54 (0.35-0.83)**	**0.57 (0.48-0.69)**	REF	**0.62 (0.51-0.75)**
High	1.02 (0.67-1.55)	0.70 (0.47-1.05)	**1.70 (1.10-2.63)**	0.88 (0.72-1.06)	0.66 (0.41-1.08)	**0.51 (0.41-0.62)**	REF	**0.57 (0.45-0.73)**
Nicotine Dependence								
	1.09 (1.08-1.10)	**1.11 (1.10-1.11)**	**1.03 (1.02-1.04)**	**1.08 (1.07-1.08)**	**1.03 (1.02-1.04)**	**1.09 (1.08-1.09)**	REF	**1.03 (1.02-1.03)**
Negative Affect Severity (ref=low)								
Moderate	1.33 (0.85-2.07)	1.21 (0.74-1.97)	0.97 (0.61-1.56)	**1.47 (1.22-1.77)**	1.04 (0.64-1.68)	1.22 (0.96-1.55)	**1.58 (1.27-1.96)**	REF
High	1.53 (1.00-2.36)	1.02 (0.68-1.53)	1.19 (0.78-1.81)	**1.29 (1.03-1.61)**	1.00 (0.62-1.60)	1.08 (0.85-1.38)	**1.31 (1.05-1.64)**	REF
Externalizing Severity (ref=low)								
Moderate	1.42 (0.93-2.17)	1.16 (0.77-1.74)	**2.29 (1.53-3.43)**	**1.41 (1.16-1.72)**	0.87 (0.56-1.37)	0.92 (0.76-1.13)	**1.62 (1.33-1.97)**	REF
High	**1.79 (1.15-2.78)**	1.23 (0.82-1.85)	**2.97 (1.84-4.81)**	**1.53 (1.21-1.92)**	1.16 (0.69-1.95)	0.88 (0.70-1.11)	**1.75 (1.38-2.22)**	REF
Nicotine Dependence								
	1.06 (1.05-1.07)	**1.08 (1.07-1.09)**	1.00 (0.99-1.01)	**1.05 (1.04-1.05)**	1.00 (0.99-1.01)	**1.06 (1.05-1.06)**	**0.97 (0.97-0.98)**	REF

Bolded values indicate estimate significant a p < 0.05
Each model adjusts for sex, age, race, education, and annual household income.

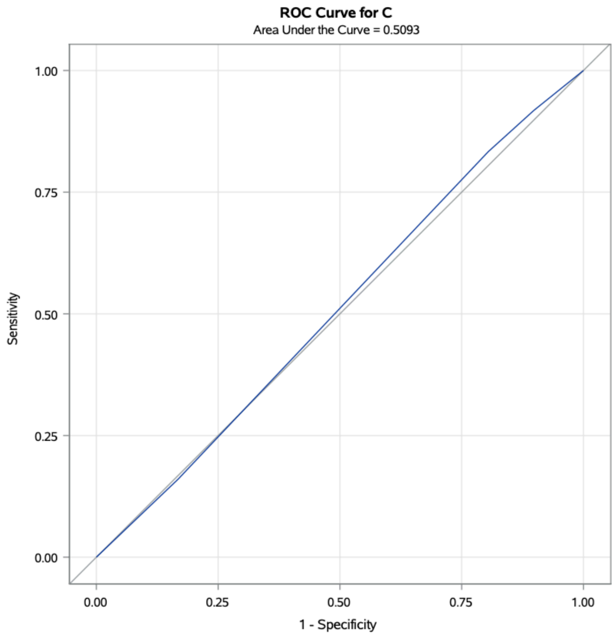
APPENDIX B: CHAPTER 3



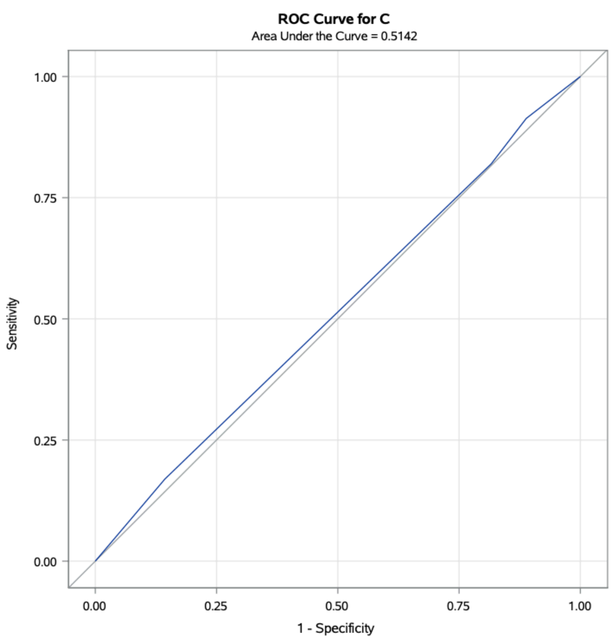
Supplemental Figure 3.1: Cumulative ROC Curve for Substance Use: SUD 0 vs 1,2



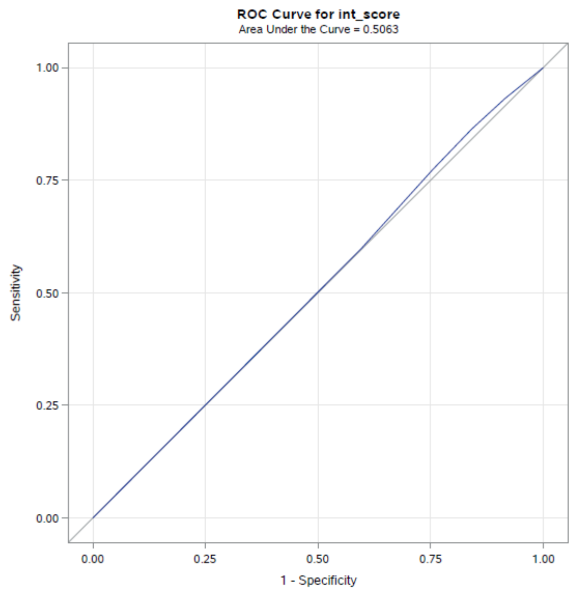
Supplemental Figure 3.2: Cumulative ROC Curve for Substance Use: SUD 0,1 vs 2



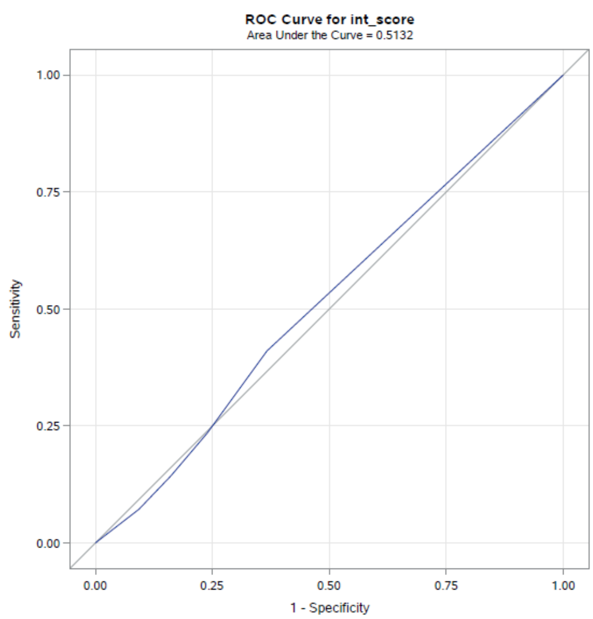
Supplemental Figure 3.3: Cumulative ROC Curve for Class Membership: SUD 0 vs 1,2



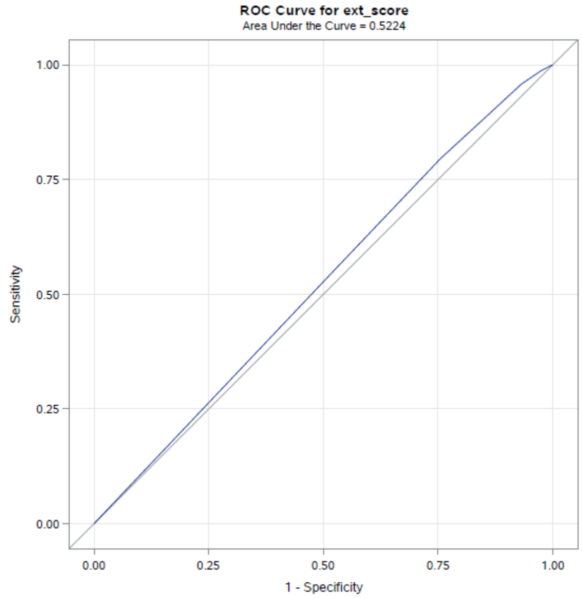
Supplemental Figure 3.4: Cumulative ROC Curve for Class Membership: SUD 0,1 vs 2



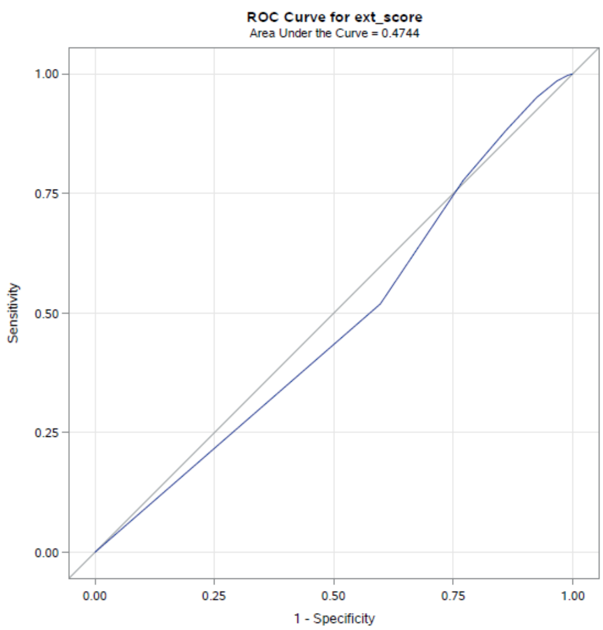
Supplemental Figure 3.5: Cumulative ROC Curve for Negative Affect: SUD 0 vs 1,2



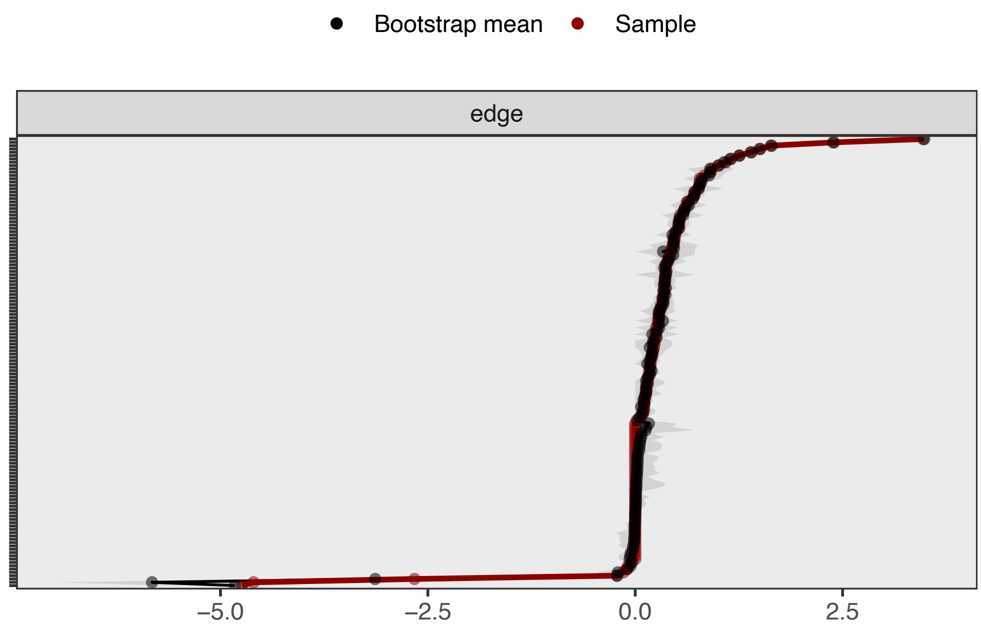
Supplemental Figure 3.6: Cumulative ROC Curve for Negative affect: SUD 0,1 vs 2



Supplemental Figure 3.7: Cumulative ROC Curve for Externalizing: SUD 0 vs 1,2

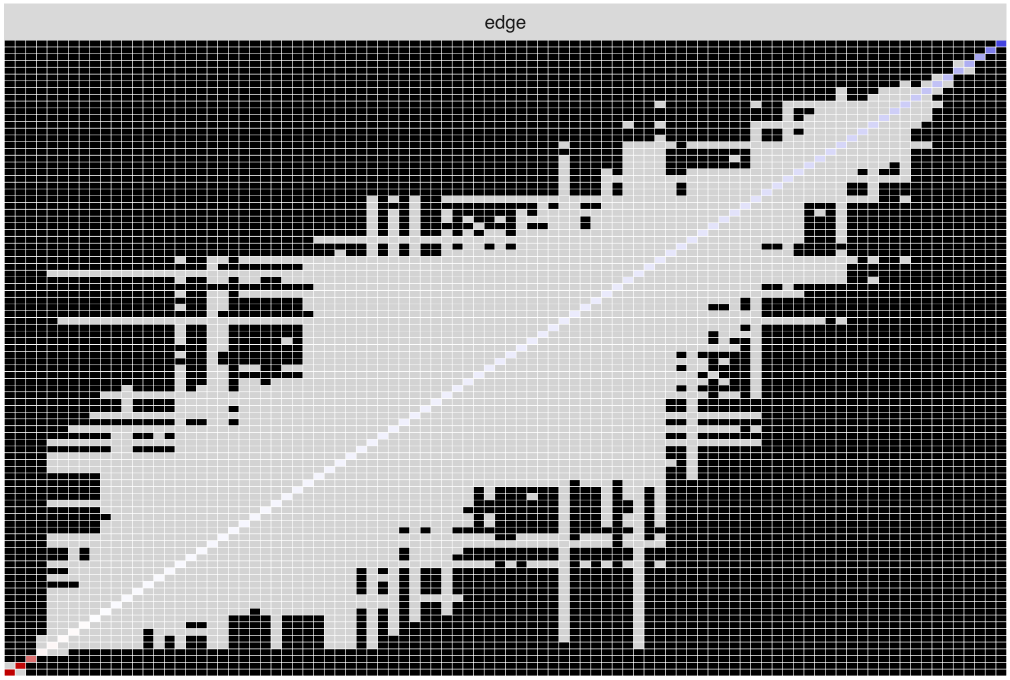


Supplemental Figure 3.8: Cumulative ROC Curve for Externalizing: SUD 0,1 vs 2



Supplemental Figure 4.1: Results from Edge-Weight Accuracy Test for the Overall Sample Network

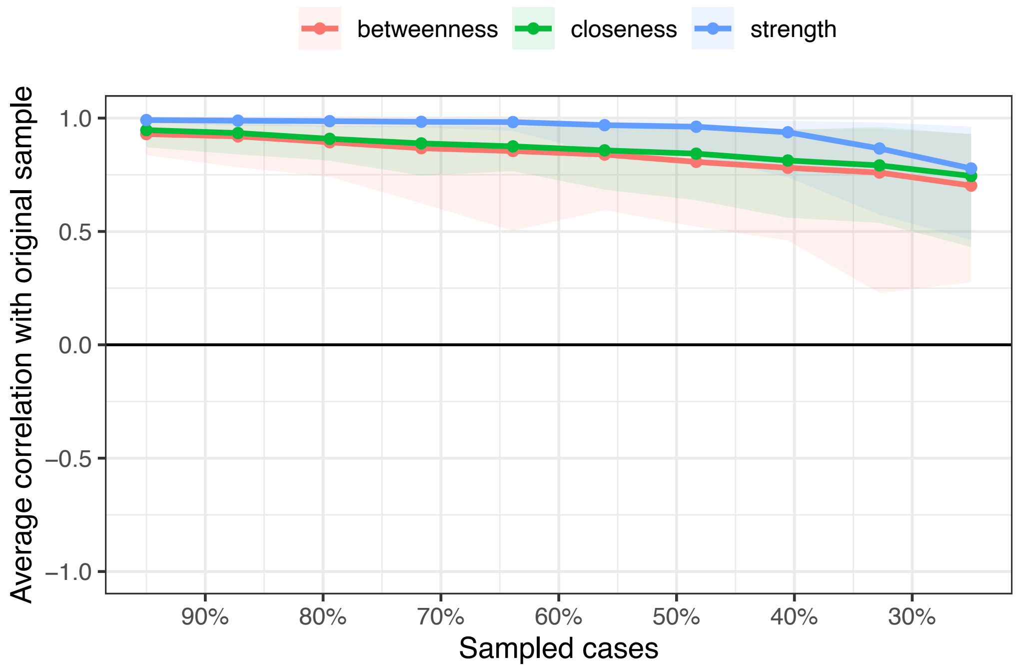
The assessment of the accuracy of estimated network connections demonstrated that many edge-weights significantly differ from one-another. Supplemental Figure 1 shows the bootstrapped confidence intervals of estimated edge-weights for the estimated overall network. The red line indicates the sample values and the gray area represent the bootstrapped confidence intervals. Each horizontal line represents one edge of the network, ordered from the edge with the highest edge-weight to the edge with the lowest edge-weight. The y-axis labels have been removed to avoid cluttering. There were narrow bootstrapped confidence intervals (narrowest 95% CI = -0.013; 0.013 for alcohol—fighting; widest 95% CI = -4.112; -1.211 for ECIG—dual CIG + ECIG) around the estimated edge-weights allowing for valid interpretation of edge-weights in the network.



Supplemental Figure 4.2: Results from the Edge-Weights Significant Difference Test for the Overall Sample Network

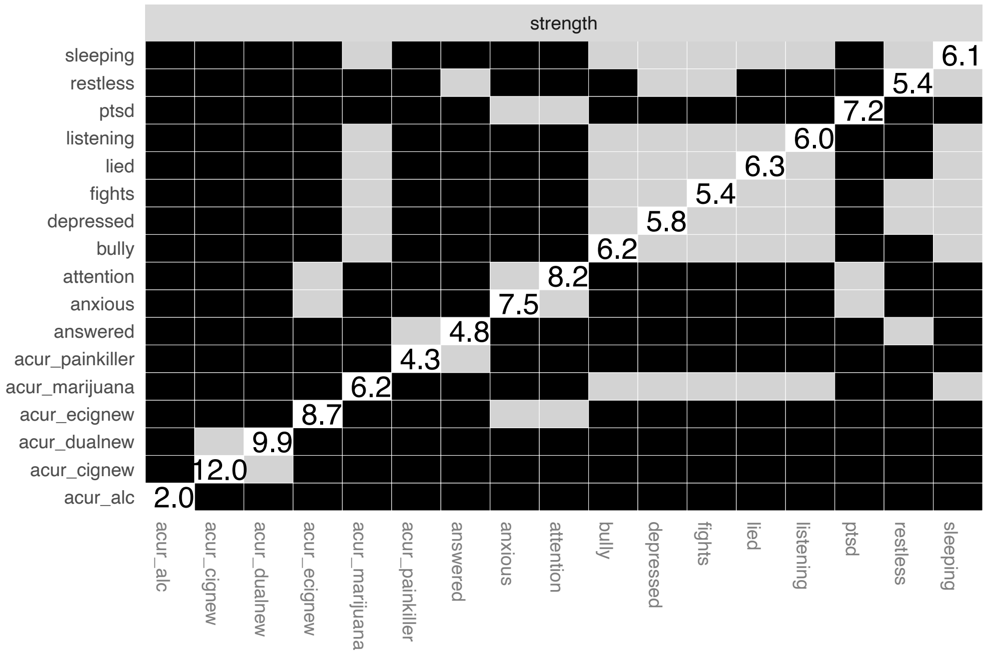
To test for significant difference between edges, a confidence interval was constructed on the difference of two edges and the test was deemed significant if zero was not in this confidence interval (represented as a black square in the grid). Supplemental Figure 2 shows the bootstrapped difference test (alpha = 0.05) between edges weights that were non-zero in the estimated network. Gray boxes indicate edges that do not differ significantly from one-another and black boxes represent edges that do differ significantly from one-another. Colored boxes correspond to the direction of the edge's magnitude (i.e., the negative “Dual CIG + ECIG” and “CIG” edge is red, the positive “Attention” and “Listening” edge is blue). The labels have been removed to avoid cluttering.

Supplemental Table 4.1: Edge Matrix for the Overall Sample																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. CIG	0																
2. ECIG	-4.74	0															
3. Dual CIG + ECIG	-4.6	-2.66	0														
4. Alcohol	0	0	0	0													
5. Marijuana	0.78	0.62	0.83	1	0												
6. PDNP	0.54	0.43	0.78	0	0.62	0											
7. Depressed	0.12	0	0	0	0.28	0.11	0										
8. Sleeping	0.16	0.22	0.15	0.04	0	0.53	1.16	0									
9. Anxious	0.17	0	0.27	0	0	0.31	1.5	1.39	0								
10. Distressed/Past	0.24	0	0.25	0	0.17	0.31	1.25	0.77	1.63	0							
11. Lied	-0.1	0	-0.14	0.2	0.6	0.21	0.35	0.11	0.25	0.71	0						
12. Attention	-0.22	0	0.15	0.18	0.13	0	0.47	0.53	0.49	0.36	0.71	0					
13. Listening	0	0	0	0	0	0	0.13	0.33	0.36	0.2	0.35	3.47	0				
14. Bully	0.23	0	0	0	0.13	0	0.23	0	0.44	0.39	0.91	0.25	0.42	0			
15. Fights	0	0	0	0	0.54	0.36	0	0	0	0.45	0.76	0	0	2.4	0		
16. Restless	-0.09	0	0	0.11	0.37	0	0.19	0.37	0.37	0.29	0.33	0.53	0.4	0.3	0.92	0	
17. Answered	-0.05	0	0.1	0.48	0.11	0.07	0	0.35	0.27	0.17	0.59	0.69	0.34	0.48	0	1.1	0



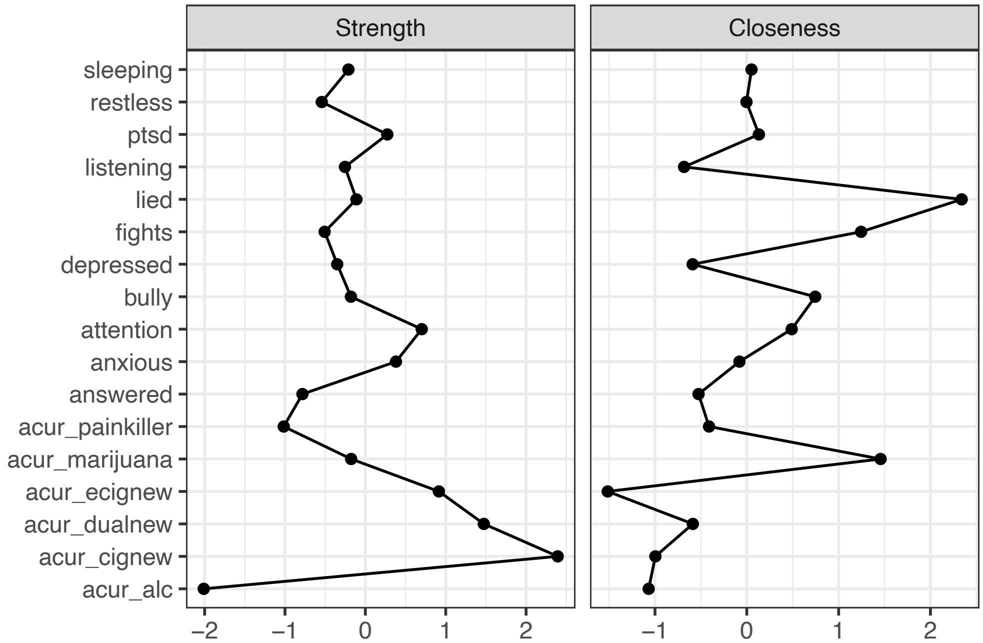
Supplemental Figure 4.3: Results from Centrality Stability Test for the Overall Sample Network

Supplemental Figure 4.3 shows the average correlations between centrality indices of networks samples with persons dropped from the original sample to establish the stability in the centrality indices. Lines represent the means of the centrality indices and shaded areas indicate the range from the 2.5th quantile and the 97.5th quantile.

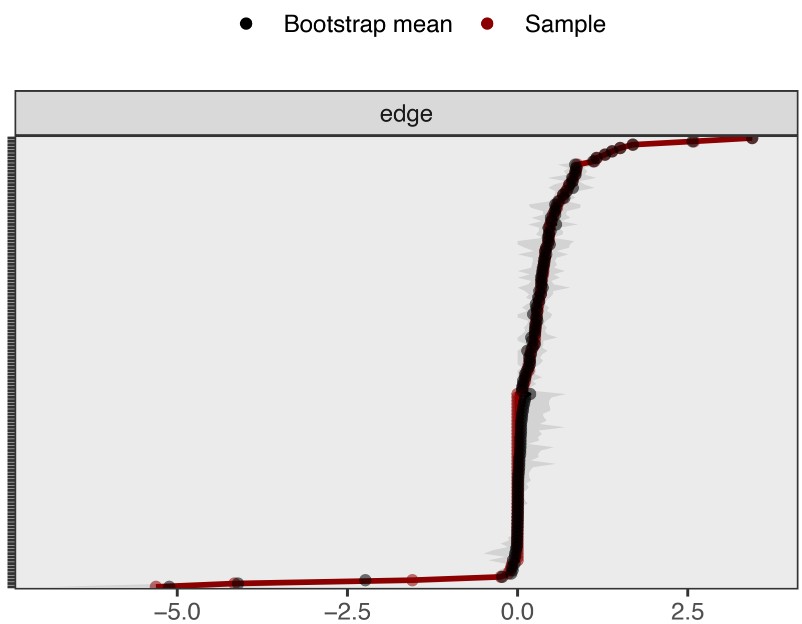


Supplemental Figure 4.4: Results from the Centrality (Node Strength) Significant Difference Test for the Overall Sample Network

Significant differences between node strength were also tested. Supplemental Figure 4.4 shows the bootstrapped difference tests (alpha = 0.05) between node strength of the 17 nodes. Gray boxes indicate nodes that did not differ significantly from one-another and black boxes represent nodes that do differ significantly from one-another (e.g., the node strength of sleeping is significantly different from the node strength of alcohol use). White boxes show the value of node strength.

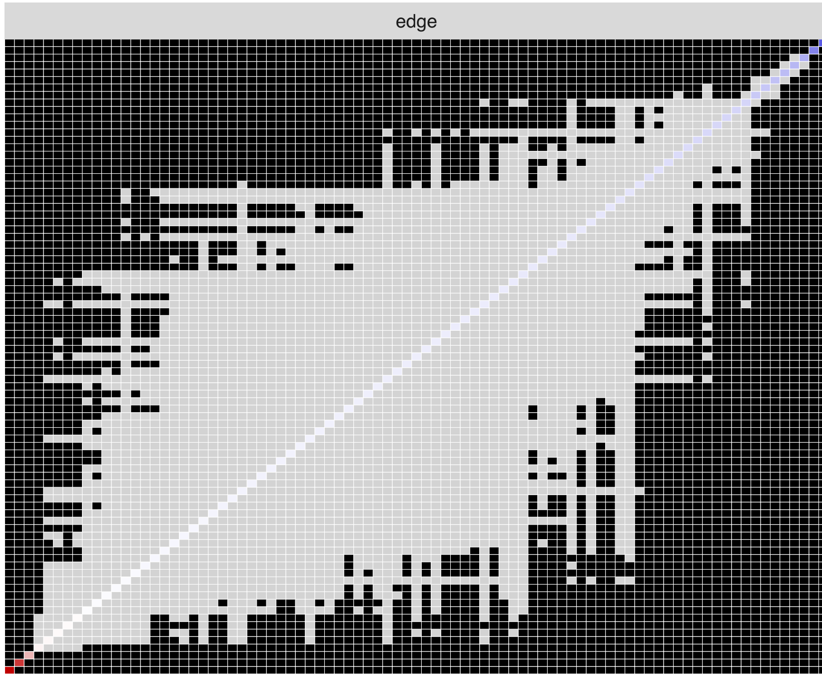


Supplemental Figure 4.5: Strength and Closeness Centrality Indices as Z-scores for the Overall Sample Network



Supplemental Figure 4.6: Results from Edge-Weight Accuracy Test for the Men-Only Network

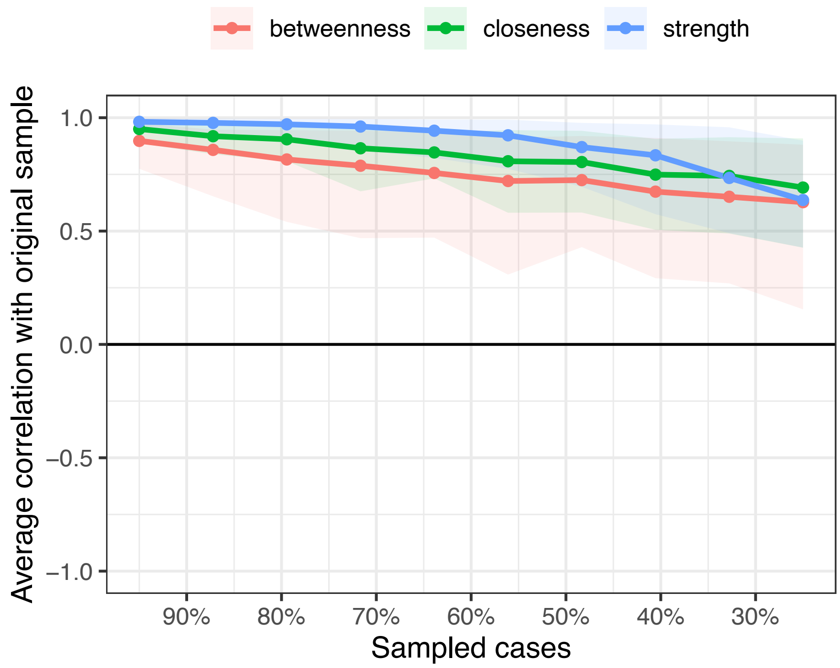
The assessment of the accuracy of estimated network connections demonstrated that many edge-weights significantly differ from one-another. Supplemental Figure 4.6 shows the bootstrapped confidence intervals of estimated edge-weights for the estimated overall network. The red line indicates the sample values and the gray area represent the bootstrapped confidence intervals. Each horizontal line represents one edge of the network, ordered from the edge with the highest edge-weight to the edge with the lowest edge-weight. The y-axis labels have been removed to avoid cluttering.



Supplemental Figure 4.7: Results from the Edge-Weights Significant Difference Test for the Men-Only Network

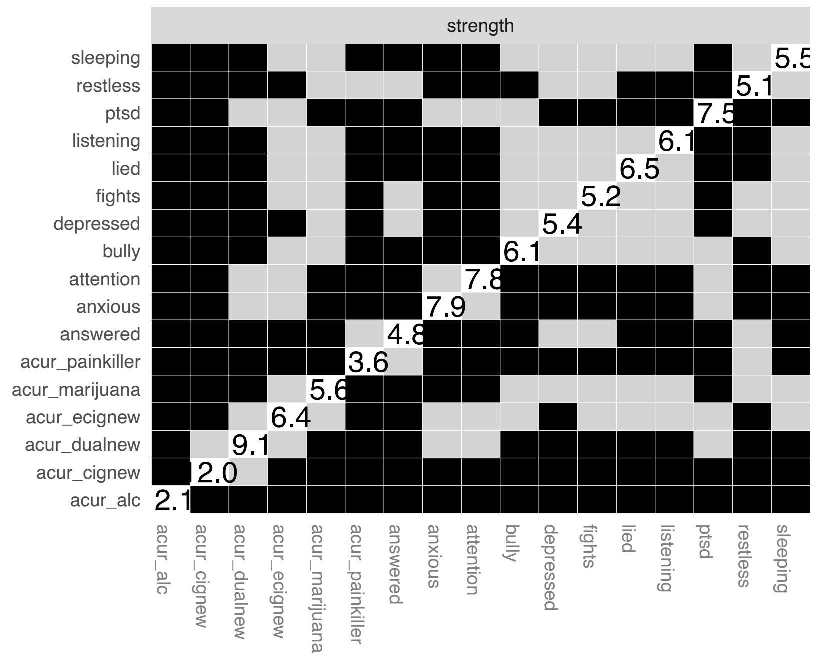
Supplemental Figure 4.7 shows the bootstrapped difference test (alpha = 0.05) between edges weights that were non-zero in the estimated network. Gray boxes indicate edges that do not differ significantly from one-another and black boxes represent edges that do differ significantly from one-another. Colored boxes correspond to the color of the edge (i.e., the negative “Dual CIG + ECIG” and “CIG” edge is red, the positive “Attention” and “Listening” edge is blue). The labels have been removed to avoid cluttering.

Supplemental Table 4.2: Edge Matrix for the Men-Only Sample																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. CIG	0																
2. ECIG	-4.16	0															
3. Dual CIG + ECIG	-5.31	-1.55	0														
4. Alcohol	0	0	0	0													
5. Marijuana	0.81	0.48	0.76	0.87	0												
6. PDNP	0.52	0	0.74	0	0.66	0											
7. Depressed	0.11	0	0	0	0.22	0	0										
8. Sleeping	0	0	0	0.08	0	0.47	1.16	0									
9. Anxious	0.16	0	0.2	-0.08	0	0.34	1.51	1.39	0								
10. Distressed/Past	0.27	0	0.44	0	0.16	0.27	1.28	0.84	1.69	0							
11. Lied	-0.14	0.25	0	0.31	0.47	0.29	0.32	0.07	0.37	0.71	0						
12. Attention	-0.25	0	0	0.25	0.12	0	0.38	0.5	0.4	0.35	0.86	0					
13. Listening	0	0	0	-0.11	0	0	0.26	0.34	0.46	0.18	0.27	3.45	0				
14. Bully	0.28	0	0	0	0.11	0	0	0	0.58	0.42	0.88	0	0.38	0			
15. Fights	0	0	0	0	0.4	0.34	0	0	0	0.46	0.58	0	0	2.59	0		
16. Restless	0	0	0	0	0.35	0	0.18	0.29	0.41	0.19	0.4	0.55	0.39	0.36	0.82	0	
17. Answered	-0.09	0	0.09	0.46	0.16	0	0	0.31	0.27	0.24	0.61	0.67	0.26	0.52	0	1.13	0



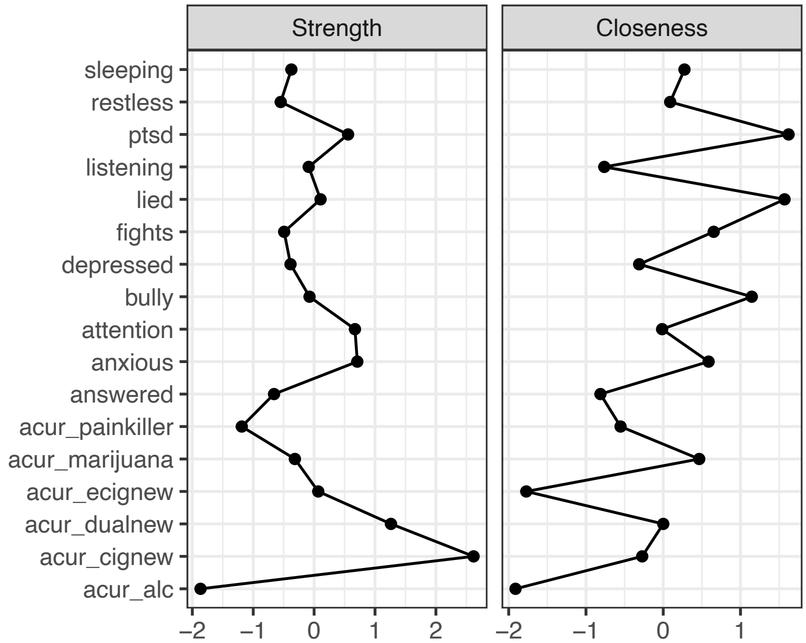
Supplemental Figure 4.8: Results from Centrality Stability Test for the Men-Only Network

Supplemental Figure 4.8 shows the average correlations between centrality indices of networks samples with persons dropped from the original sample to establish the stability in the centrality indices.



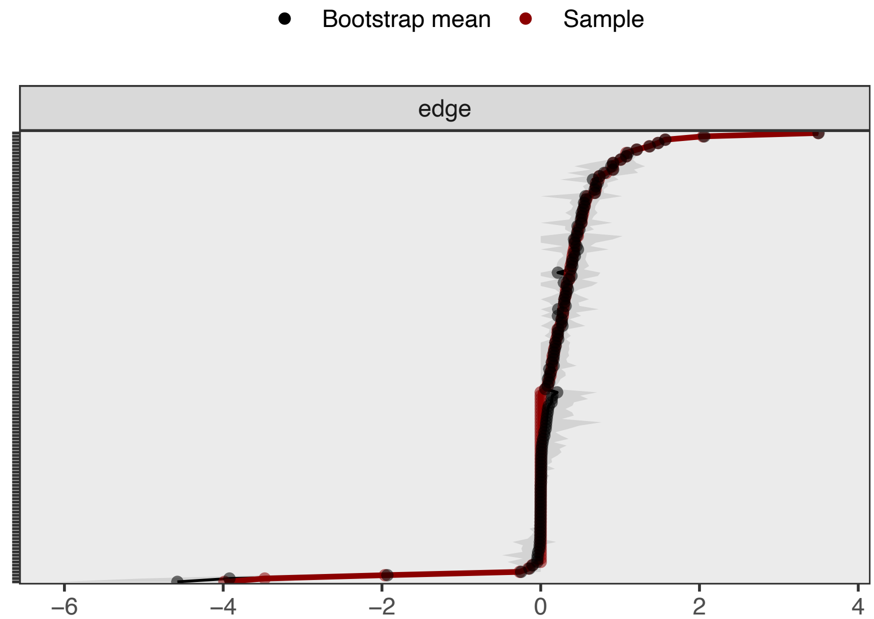
Supplemental Figure 4.9: Results from the Centrality (Node Strength) Significant Difference Test for the Men-Only Network

Significant differences between node strength were also tested. Supplemental Figure 4.9 shows the bootstrapped difference tests (alpha = 0.05) between node strength of the 17 nodes. Gray boxes indicate nodes that do not differ significantly from one-another and black boxes represent nodes that do differ significantly from one-another (e.g., the node strength of restless is significantly different from the node strength of alcohol use). White boxes show the value of node strength.



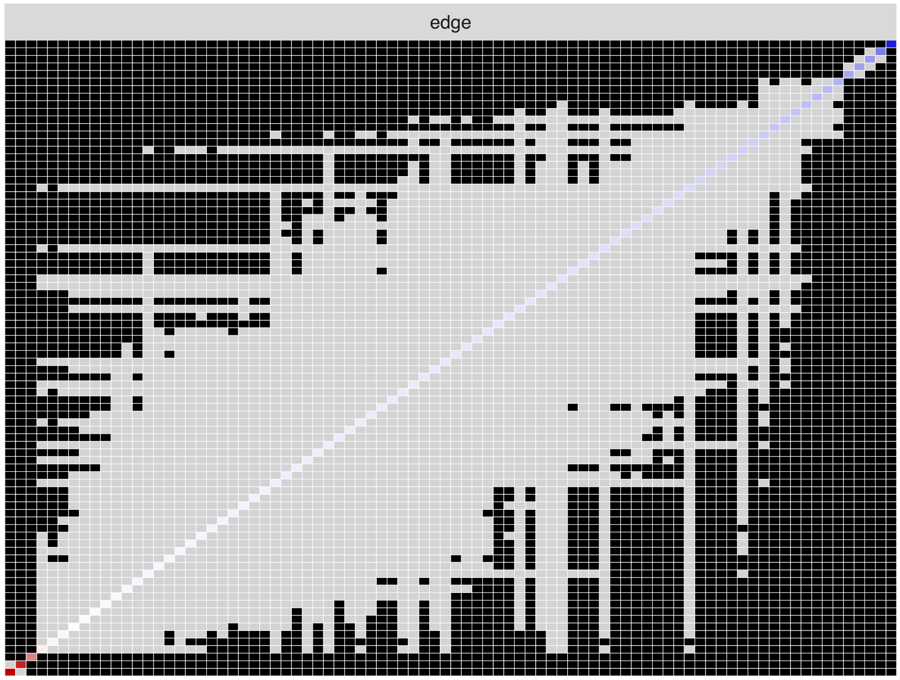
Supplemental Figure 4.10: Strength and Closeness Centrality Indices as Z-scores for the Men-Only Network

Supplemental Figure 4.10 shows the men-only network’s corresponding centrality indices. Centrality indices are shown as z-scores.



Supplemental Figure 4.11: Results from Edge-Weight Accuracy Test for the Women-Only Network

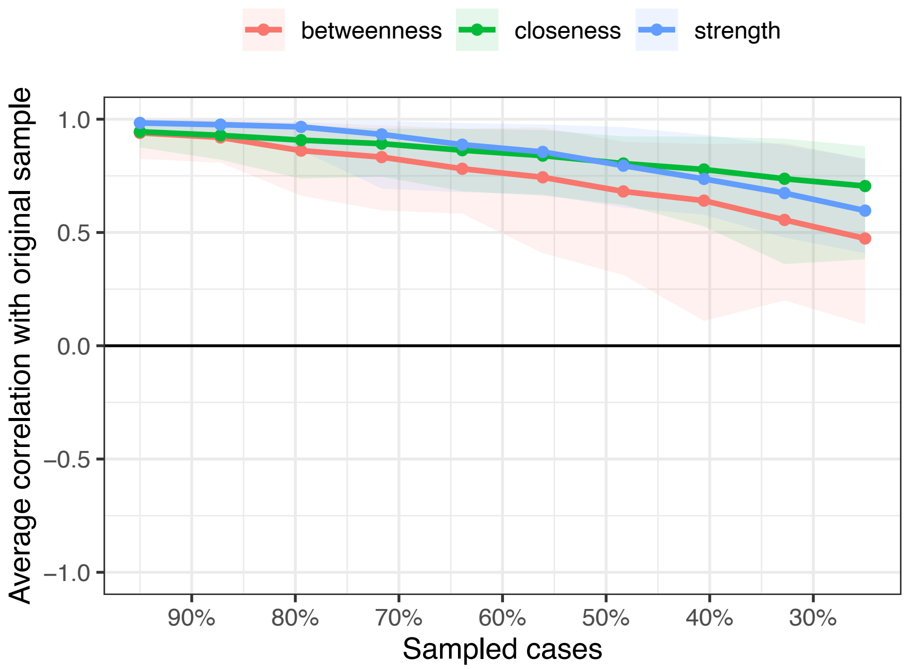
The assessment of the accuracy of estimated network connections demonstrated that many edge-weights significantly differ from one-another. Supplemental Figure 4.11 shows the bootstrapped confidence intervals of estimated edge-weights for the estimated overall network. The red line indicates the sample values and the gray area represent the bootstrapped confidence intervals. Each horizontal line represents one edge of the network, ordered from the edge with the highest edge-weight to the edge with the lowest edge-weight. The y-axis labels have been removed to avoid cluttering.



Supplemental Figure 4.12: Results from the Edge-Weights Significant Difference Test for the Women-Only Network

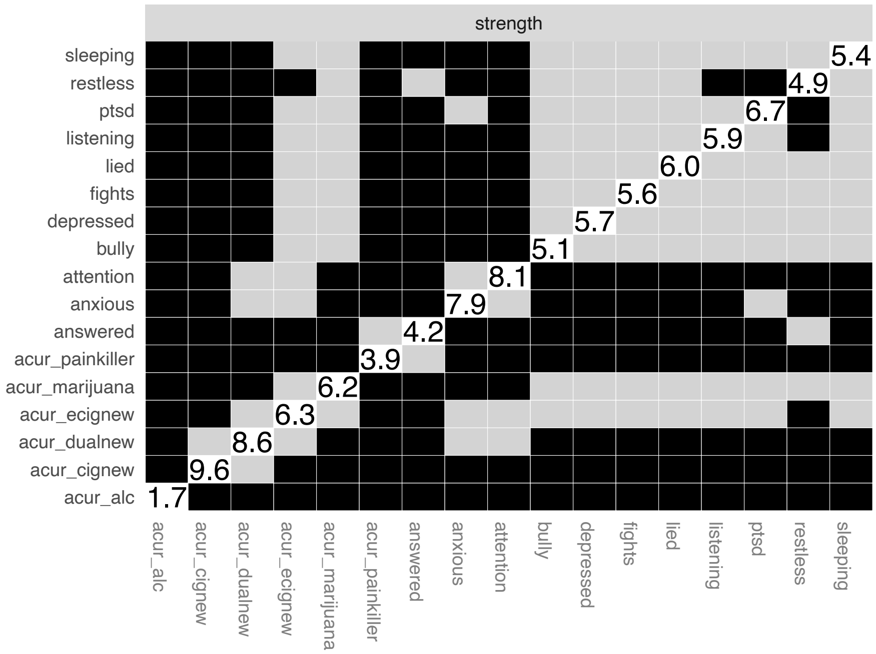
Supplemental Figure 4.12 shows the bootstrapped difference test (alpha = 0.05) between edges weights that were non-zero in the estimated network. Gray boxes indicate edges that do not differ significantly from one-another and black boxes represent edges that do differ significantly from one-another. Colored boxes correspond to the color of the edge (i.e., the negative “Dual CIG + ECIG” and “CIG” edge is red, the positive “Attention” and “Listening” edge is blue). The labels have been removed to avoid cluttering.

Supplemental Table 4.3: Edge Matrix for the Women-Only Sample																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. CIG	0																
2. ECIG	-3.48	0															
3. Dual CIG + ECIG	-3.98	-1.96	0														
4. Alcohol	0	0	0	0													
5. Marijuana	0.7	0.51	0.91	1.08	0												
6. PDNP	0.48	0.36	0.73	0	0.57	0											
7. Depressed	0.09	0	0	0	0.32	0.15	0										
8. Sleeping	0.2	0	0	0	0	0.47	1.08	0									
9. Anxious	0.21	0	0.43	0	0.15	0.3	1.48	1.37	0								
10. Distressed/Past	0.21	0	0	0	0.18	0.31	1.21	0.68	1.57	0							
11. Lied	0	0	-0.26	0	0.69	0.12	0.37	0.12	0.19	0.74	0						
12. Attention	-0.16	0	0.28	0	0.14	0	0.52	0.52	0.54	0.36	0.56	0					
13. Listening	0	0	0	0	0	0	0	0.29	0.28	0.21	0.43	3.49	0				
14. Bully	0	0	0	0	0	0	0.36	0	0.32	0.3	0.91	0.33	0.42	0			
15. Fights	0	0	0	0	0.59	0.29	0	0	0.47	0.43	0.91	0	0	2.04	0		
16. Restless	-0.11	0	0	0.14	0.26	0	0.17	0.41	0.39	0.4	0.17	0.53	0.38	0.14	0.82	0	
17. Answered	0	0	0	0.47	0.05	0.11	0	0.32	0.26	0.09	0.55	0.69	0.4	0.25	0	1.01	0



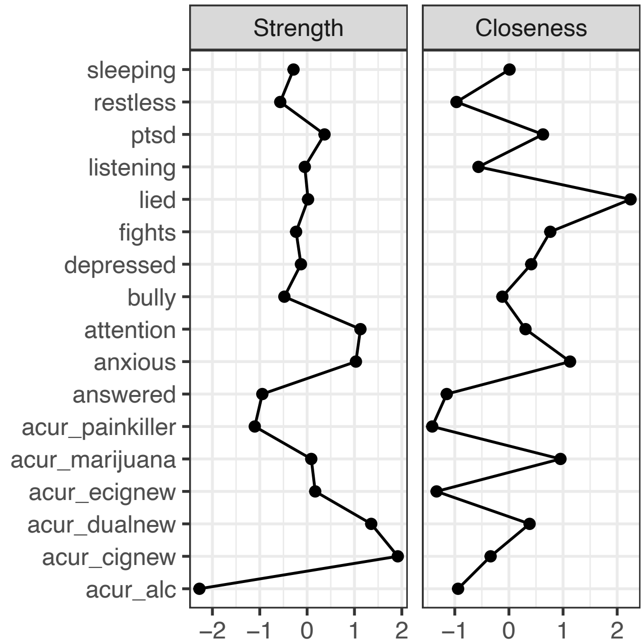
Supplemental Figure 4.13: Results from Centrality Stability Test for the Women-Only Network

Supplemental Figure 4.13 shows the average correlations between centrality indices of networks samples with persons dropped from the original sample to establish the stability in the centrality indices.



Supplemental Figure 4.14. Results from the Centrality (Node Strength) Significant Difference Test for the Women-Only Network

Significant differences between node strength were also tested. Supplemental Figure 4.14 shows the bootstrapped difference tests (alpha = 0.05) between node strength of the 17 nodes. Gray boxes indicate nodes that do not differ significantly from one-another and black boxes represent nodes that do differ significantly from one-another (e.g., the node strength of lied is significantly different from the node strength of alcohol use). White boxes show the value of node strength.



Supplemental Figure 4.15: Strength and Closeness Centrality as Z-scores for the Women-Only Network

Supplemental Figure 4.15 shows the women-only network’s corresponding strength and closeness. Strength and closeness are shown as z-scores.

APPENDIX D: CHAPTER 5

Supplemental Table 5.1: Edge Matrix for the Wave 1 Sample																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. CIG	0																
2. ECIG	-4.74	0															
3. Dual CIG + ECIG	-4.6	-2.66	0														
4. Alcohol	0	0	0	0													
5. Marijuana	0.78	0.62	0.83	1	0												
6. PDNP	0.54	0.43	0.78	0	0.62	0											
7. Depressed	0.12	0	0	0	0.28	0.11	0										
8. Sleeping	0.16	0.22	0.15	0.04	0	0.53	1.16	0									
9. Anxious	0.17	0	0.27	0	0	0.31	1.5	1.39	0								
10. Distressed/Past	0.24	0	0.25	0	0.17	0.31	1.25	0.77	1.63	0							
11. Lied	-0.1	0	-0.14	0.2	0.6	0.21	0.35	0.11	0.25	0.71	0						
12. Attention	-0.22	0	0.15	0.18	0.13	0	0.47	0.53	0.49	0.36	0.71	0					
13. Listening	0	0	0	0	0	0	0.13	0.33	0.36	0.2	0.35	3.47	0				
14. Bully	0.23	0	0	0	0.13	0	0.23	0	0.44	0.39	0.91	0.25	0.42	0			
15. Fights	0	0	0	0	0.54	0.36	0	0	0	0.45	0.76	0	0	2.4	0		
16. Restless	-0.09	0	0	0.11	0.37	0	0.19	0.37	0.37	0.29	0.33	0.53	0.4	0.3	0.92	0	
17. Answered	-0.05	0	0.1	0.48	0.11	0.07	0	0.35	0.27	0.17	0.59	0.69	0.34	0.48	0	1.1	0

Supplemental Table 5.2: Edge Matrix for the Wave 2 Sample																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. CIG	0																
2. ECIG	-5.41	0															
3. Dual CIG + ECIG	-6.17	-3.76	0														
4. Alcohol	0	0	0	0													
5. Marijuana	0.91	0.68	1.01	0.91	0												
6. PDNP	0.6	0.55	0.78	0	0.37	0											
7. Depressed	0.06	0	0	0	0.27	0.12	0										
8. Sleeping	0	0	0	0	0	0.44	1.19	0									
9. Anxious	0.25	0	0.34	0	0.07	0.24	1.51	1.4	0								
10. Distressed/Past	0.33	0.42	0.42	0	0.12	0.17	1.25	0.8	1.72	0							
11. Lied	0	0	0	0	0.54	0.27	0.27	0.08	0.46	0.83	0						
12. Attention	-0.16	0	0	0.15	0.09	0	0.51	0.67	0.55	0.25	0.66	0					
13. Listening	0	0	0	0	0	0	0.02	0.31	0.26	0.46	0.35	3.72	0				
14. Bully	0.31	0	0.3	0	0	0	0.23	0	0.29	0.47	0.97	0	0.63	0			
15. Fights	0	0	0	0	0	0	0	0	0	0.58	0.68	0	0	2.8	0		
16. Restless	0	0	0	0	0.48	0	0.14	0.09	0.33	0.43	0.21	0.67	0.42	0.44	0.89	0	
17. Answered	-0.13	0.08	0	0.44	0.05	0.22	0.04	0.45	0.28	0.12	0.52	0.69	0.37	0.38	0	1.24	0

Supplemental Table 5.3: Edge Matrix for the Wave 3 Sample																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1. CIG	0																
2. ECIG	-4.73	0															
3. Dual CIG + ECIG	-5.42	-2.44	0														
4. Alcohol	-0.07	0	0	0													
5. Marijuana	0.93	0.88	0.99	0.93	0												
6. PDNP	0.58	0.32	0.81	0	0.32	0											
7. Depressed	0.07	0	0	-0.09	0.28	0	0										
8. Sleeping	0	0	0	0.07	-0.09	0.52	1.19	0									
9. Anxious	0.26	0	0.13	0	0.16	0.36	1.59	1.41	0								
10. Distressed/Past	0.32	0	0.46	-0.1	0.12	0.08	1.37	0.76	1.77	0							
11. Lied	0	0	0	0	0.39	0.11	0.19	0.3	0.34	0.88	0						
12. Attention	-0.17	0	0	0.13	0.17	0	0.54	0.69	0.53	0.34	0.68	0					
13. Listening	0	0.2	0.22	0	-0.2	0	0.19	0.27	0.31	0.4	0.42	3.66	0				
14. Bully	0.18	0	0	0	0.16	0	0.35	0	0.32	0.56	0.74	0.2	0.4	0			
15. Fights	0.47	0	0	0	0.36	0	0.27	0	0	0.56	1.02	0	0.4	2.88	0		
16. Restless	-0.29	0	0	0.07	0.45	0	0.44	0.14	0.41	0.23	0.22	0.63	0.35	0.36	1.12	0	
17. Answered	-0.13	0	0	0.55	0.14	0.09	0	0.39	0.32	0.18	0.57	0.69	0.48	0.65	0	1.24	0

Supplemental Table 5.4: Edge Matrix for Wave 1, Wave 2, Wave 3 Network by Wave 1

	W1 CIG	W1 ECIG	W1 Dual CIG + ECIG	W1 Alcohol	W1 Marijuana	W1 PDNP	W1 Depressed	W1 Sleeping	W1 Anxious	W1 Distressed/ Past	W1 Lied	W1 Attention	W1 Listening	W1 Bully	W1 Fights	W1 Restless	W1 Answered
W1 CIG	0.00	-2.28	-2.82	0.00	0.68	0.43	0.00	0.08	0.19	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 ECIG	-2.28	0.00	-0.86	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Dual CIG + ECIG	-2.82	-0.86	0.00	0.00	0.68	0.58	0.00	0.00	0.21	0.30	0.00	0.16	0.00	0.00	0.00	0.00	0.00
W1 Alcohol	0.00	0.00	0.00	0.00	0.95	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.42
W1 Marijuana	0.68	0.50	0.68	0.95	0.00	0.59	0.22	0.00	0.00	0.23	0.55	0.11	0.00	0.00	0.47	0.27	0.07
W1 PDNP	0.43	0.00	0.58	0.00	0.59	0.00	0.00	0.47	0.28	0.24	0.18	0.00	0.00	0.00	0.34	0.00	0.00
W1 Depressed	0.00	0.00	0.00	0.00	0.22	0.00	0.00	1.06	1.52	1.26	0.31	0.46	0.17	0.20	0.00	0.14	0.00
W1 Sleeping	0.08	0.00	0.00	0.00	0.00	0.47	1.06	0.00	1.42	0.79	0.10	0.51	0.34	0.00	0.00	0.31	0.32
W1 Anxious	0.19	0.00	0.21	0.00	0.00	0.28	1.52	1.42	0.00	1.59	0.30	0.52	0.30	0.39	0.24	0.39	0.26
W1 Distressed/ Past	0.17	0.00	0.30	0.00	0.23	0.24	1.26	0.79	1.59	0.00	0.65	0.36	0.18	0.39	0.29	0.28	0.16
W1 Lied	0.00	0.00	0.00	0.12	0.55	0.18	0.31	0.10	0.30	0.65	0.00	0.69	0.37	0.95	0.81	0.38	0.55
W1 Attention	0.00	0.00	0.16	0.00	0.11	0.00	0.46	0.51	0.52	0.36	0.69	0.00	3.36	0.22	0.00	0.45	0.67
W1 Listening	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.34	0.30	0.18	0.37	3.36	0.00	0.48	0.00	0.41	0.34
W1 Bully	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.00	0.39	0.39	0.95	0.22	0.48	0.00	2.30	0.24	0.27

W1 Fights	0.00	0.00	0.00	0.00	0.47	0.34	0.00	0.00	0.24	0.29	0.81	0.00	0.00	2.30	0.00	0.79	0.00
W1 Restless	0.00	0.00	0.00	0.00	0.27	0.00	0.14	0.31	0.39	0.28	0.38	0.45	0.41	0.24	0.79	0.00	0.98
W1 Answered	0.00	0.00	0.00	0.42	0.07	0.00	0.00	0.32	0.26	0.16	0.55	0.67	0.34	0.27	0.00	0.98	0.00
W2 CIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Dual CIG + ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Alcohol	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Marijuana	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 PDNP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.14	0.00	0.00	0.00	0.00	0.00	0.00
W2 Depressed	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Sleeping	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Anxious	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Distressed/ Past	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Lied	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Attention	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

W2 Listening	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Bully	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Fights	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Restless	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Answered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 CIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Dual CIG + ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Alcohol	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Marijuana	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 PDNP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Depressed	0.00	0.00	0.00	0.00	-0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Sleeping	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Anxious	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Distressed/Past	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Lied	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

W3 Attention	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Listening	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Bully	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Fights	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Restless	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Answered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Supplemental Table 5.5: Edge Matrix for Wave 1, Wave 2, Wave 3 Network by Wave 2																	
	W2 CIG	W2 ECIG	W2 Dual CIG + ECIG	W2 Alcohol	W2 Marijuana	W2 PDNP	W2 Depressed	W2 Sleeping	W2 Anxious	W2 Distressed/ Past	W2 Lied	W2 Attention	W2 Listening	W2 Bully	W2 Fights	W2 Restless	W2 Answered
W1 CIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Dual CIG + ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Alcohol	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Marijuana	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 PDNP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Depressed	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Sleeping	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Anxious	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Distressed/ Past	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Lied	0.00	0.00	0.00	0.00	0.00	-0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Attention	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Listening	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Bully	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

W1 Fights	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Restless	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Answered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 CIG	0.00	-2.58	-3.18	0.00	0.82	0.48	0.00	0.00	0.16	0.31	0.00	-0.09	0.00	0.00	0.00	-0.03	0.00
W2 ECIG	-2.58	0.00	-1.69	0.00	0.51	0.39	0.00	0.00	0.00	0.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Dual CIG + ECIG	-3.18	-1.69	0.00	0.00	0.90	0.61	0.00	0.00	0.00	0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Alcohol	0.00	0.00	0.00	0.00	0.86	0.00	0.00	0.00	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.08	0.39
W2 Marijuana	0.82	0.51	0.90	0.86	0.00	0.31	0.28	0.00	0.00	0.12	0.53	0.10	0.00	0.00	0.00	0.46	0.00
W2 PDNP	0.48	0.39	0.61	0.00	0.31	0.00	0.00	0.41	0.15	0.00	0.22	0.10	0.00	0.00	0.00	0.00	0.09
W2 Depressed	0.00	0.00	0.00	0.00	0.28	0.00	0.00	1.17	1.52	1.23	0.23	0.48	0.00	0.00	0.00	0.15	0.00
W2 Sleeping	0.00	0.00	0.00	0.00	0.00	0.41	1.17	0.00	1.36	0.77	0.00	0.71	0.28	0.00	0.00	0.00	0.43
W2 Anxious	0.16	0.00	0.00	0.00	0.00	0.15	1.52	1.36	0.00	1.74	0.47	0.60	0.16	0.00	0.00	0.28	0.17
W2 Distressed/ Past	0.31	0.37	0.40	0.00	0.12	0.00	1.23	0.77	1.74	0.00	0.81	0.18	0.55	0.56	0.50	0.40	0.10
W2 Lied	0.00	0.00	0.00	0.00	0.53	0.22	0.23	0.00	0.47	0.81	0.00	0.66	0.31	0.93	0.77	0.20	0.51
W2 Attention	-0.09	0.00	0.00	0.13	0.10	0.10	0.48	0.71	0.60	0.18	0.66	0.00	3.69	0.00	0.14	0.70	0.65

W2 Listening	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.28	0.16	0.55	0.31	3.69	0.00	0.59	0.00	0.40	0.38
W2 Bully	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.56	0.93	0.00	0.59	0.00	2.65	0.49	0.20
W2 Fights	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.77	0.14	0.00	2.65	0.00	0.93	0.00
W2 Restless	-0.03	0.00	0.00	0.08	0.46	0.00	0.15	0.00	0.28	0.40	0.20	0.70	0.40	0.49	0.93	0.00	1.21
W2 Answered	0.00	0.00	0.00	0.39	0.00	0.09	0.00	0.43	0.17	0.10	0.51	0.65	0.38	0.20	0.00	1.21	0.00
W3 CIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Dual CIG + ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Alcohol	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Marijuana	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 PDNP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Depressed	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Sleeping	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Anxious	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Distressed/ Past	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Lied	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

W3 Attention	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Listening	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Bully	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Fights	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Restless	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Answered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Supplemental Table 5.6: Edge Matrix for Wave 1, Wave 2, Wave 3 Network by Wave 3																	
	W3 CIG	W3 ECIG	W3 Dual CIG + ECIG	W3 Alcohol	W3 Marijuana	W3 PDNP	W3 Depressed	W3 Sleeping	W3 Anxious	W3 Distressed/ Past	W3 Lied	W3 Attention	W3 Listening	W3 Bully	W3 Fights	W3 Restless	W3 Answered
W1 CIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Dual CIG + ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Alcohol	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Marijuana	0.00	0.00	0.00	0.00	0.00	0.00	-0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 PDNP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Depressed	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Sleeping	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Anxious	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Distressed/ Past	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Lied	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Attention	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Listening	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Bully	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

W1 Fights	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Restless	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W1 Answered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 CIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Dual CIG + ECIG	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Alcohol	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Marijuana	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 PDNP	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Depressed	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Sleeping	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Anxious	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Distressed/ Past	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Lied	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Attention	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

W2 Listening	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Bully	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Fights	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Restless	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W2 Answered	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 CIG	0.00	-2.79	-3.00	-0.04	0.84	0.51	0.00	0.00	0.19	0.27	0.00	-0.06	0.00	0.00	0.29	-0.14	0.00
W3 ECIG	-2.79	0.00	-1.33	0.00	0.75	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Dual CIG + ECIG	-3.00	-1.33	0.00	0.00	0.79	0.75	0.00	0.00	0.00	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00
W3 Alcohol	-0.04	0.00	0.00	0.00	0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.00	0.00	0.00	0.00	0.51
W3 Marijuana	0.84	0.75	0.79	0.85	0.00	0.28	0.20	0.00	0.11	0.09	0.38	0.07	0.00	0.11	0.21	0.38	0.09
W3 PDNP	0.51	0.22	0.75	0.00	0.28	0.00	0.00	0.50	0.31	0.09	0.00	0.00	0.00	0.00	0.00	0.00	0.05
W3 Depressed	0.00	0.00	0.00	0.00	0.20	0.00	0.00	1.19	1.59	1.37	0.17	0.47	0.16	0.25	0.13	0.43	0.00
W3 Sleeping	0.00	0.00	0.00	0.00	0.00	0.50	1.19	0.00	1.38	0.73	0.24	0.69	0.21	0.00	0.00	0.00	0.41
W3 Anxious	0.19	0.00	0.00	0.00	0.11	0.31	1.59	1.38	0.00	1.72	0.34	0.52	0.35	0.29	0.00	0.36	0.32
W3 Distressed/ Past	0.27	0.00	0.30	0.00	0.09	0.09	1.37	0.73	1.72	0.00	0.89	0.32	0.38	0.54	0.48	0.27	0.15
W3 Lied	0.00	0.00	0.00	0.00	0.38	0.00	0.17	0.24	0.34	0.89	0.00	0.65	0.41	0.80	0.96	0.19	0.51

W3 Attention	-0.06	0.00	0.00	0.08	0.07	0.00	0.47	0.69	0.52	0.32	0.65	0.00	3.63	0.00	0.00	0.67	0.66
W3 Listening	0.00	0.00	0.00	0.00	0.00	0.00	0.16	0.21	0.35	0.38	0.41	3.63	0.00	0.23	0.22	0.27	0.42
W3 Bully	0.00	0.00	0.00	0.00	0.11	0.00	0.25	0.00	0.29	0.54	0.80	0.00	0.23	0.00	2.98	0.37	0.56
W3 Fights	0.29	0.00	0.00	0.00	0.21	0.00	0.13	0.00	0.00	0.48	0.96	0.00	0.22	2.98	0.00	1.06	0.00
W3 Restless	-0.14	0.00	0.00	0.00	0.38	0.00	0.43	0.00	0.36	0.27	0.19	0.67	0.27	0.37	1.06	0.00	1.17
W3 Answered	0.00	0.00	0.00	0.51	0.09	0.05	0.00	0.41	0.32	0.15	0.51	0.66	0.42	0.56	0.00	1.17	0.00

REFERENCES

1. Substance Abuse and Mental Health Services Administration. Mental Health and Substance Use Disorders.
2. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed.; 2013. doi:https://doi.org/10.1176/appi.books.9780890425596
3. Maddux. *Psychopathology: Foundations for a Contemporary Understanding*. 4th ed. Routledge; 2016.
4. Rehm J, Dawson D, Frick U, et al. Burden of disease associated with alcohol use disorders in the United States. *Alcoholism, clinical and experimental research*. 2014;38(4):1068-1077. doi:10.1111/acer.12331
5. SAMHSA. *Key Substance Use and Mental Health Indicators in the United States: Results from the 2019 National Survey on Drug Use and Health*.; 2020. https://www.samhsa.gov/data/sites/default/files/reports/rpt29393/2019NSDUHFFRPDFWHTML/2019NSDUHFFR1PDFW090120.pdf
6. Volkow ND. Fiscal Year 2013 Budget Request. Presented at the: 2012. https://web.archive.org/web/20180124111243/http:/www.drugabuse.gov/about-nida/legislative-activities/testimony-to-congress/2012/03/fiscal-year-2013-budget-request/
7. Center for Behavioral Health Statistics and Quality. *Behavioral Health Trends in the United States: Results from the 2014 National Survey on Drug Use and Health*.; 2015. https://www.samhsa.gov/data/sites/default/files/NSDUH-FRR1-2014/NSDUH-FRR1-2014.pdf
8. McLellan AT. Substance Misuse and Substance use Disorders: Why do they Matter in Healthcare? *Transactions of the American Clinical and Climatological Association*. 2017;128:112-130.
9. National Institute on Drug Abuse. Common Comorbidities with Substance Use Disorders. 2018. https://www.drugabuse.gov/publications/research-reports/common-comorbidities-substance-use-disorders/what-are-some-approaches-to-diagnosis
10. Connor JP, Gullo MJ, White A, Kelly AB. Polysubstance use: diagnostic challenges, patterns of use and health. *Current opinion in psychiatry*. 2014;27(4):269-275. doi:10.1097/YCO.0000000000000069
11. Morley KI, Lynskey MT, Moran P, Borschmann R, Winstock AR. Polysubstance use, mental health and high-risk behaviours: Results from the 2012 Global Drug Survey. *Drug and alcohol review*. 2015;34(4):427-437. doi:10.1111/dar.12263
12. Salom CL, Betts KS, Williams GM, Najman JM, Alati R. Predictors of comorbid polysubstance use and mental health disorders in young adults-a latent class analysis. *Addiction (Abingdon, England)*. 2016;111(1):156-164. doi:10.1111/add.13058
13. Hartz SM, Pato CN, Medeiros H, et al. Comorbidity of severe psychotic disorders with measures of substance use. *JAMA psychiatry*. 2014;71(3):248-254. doi:10.1001/jamapsychiatry.2013.3726
14. de Alwis D, Lynskey MT, Reiersen AM, Agrawal A. Attention-deficit/hyperactivity disorder subtypes and substance use and use disorders in NESARC. *Addictive behaviors*. 2014;39(8):1278-1285. doi:10.1016/j.addbeh.2014.04.003
15. Lynskey MT, Hall W. Attention deficit hyperactivity disorder and substance use disorders: Is there a causal link? *Addiction (Abingdon, England)*. 2001;96(6):815-822. doi:10.1080/09652140020050988

16. Hasin DS, Grant BF. The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) Waves 1 and 2: review and summary of findings. *Social psychiatry and psychiatric epidemiology*. 2015;50(11):1609-1640. doi:10.1007/s00127-015-1088-0
17. Maslowsky J, Schulenberg JE, O'Malley PM, Kloska DD. Depressive symptoms, conduct problems, and risk for polysubstance use among adolescents: Results from US national surveys. *Mental health and substance use : dual diagnosis*. 2013;7(2):157-169. doi:10.1080/17523281.2013.786750
18. Morisano D, Babor TF, Robaina KA. Co-Occurrence of Substance use Disorders with other Psychiatric Disorders: Implications for Treatment Services. *Nordic Studies on Alcohol and Drugs*. 2014;31(1):5-25. doi:10.2478/nsad-2014-0002
19. Cox, Cynthia & Sawyer B. What do we know about the burden of disease in the U.S.? Kaiser Family Foundation. Published 2017. https://www.healthsystemtracker.org/chart-collection/know-burden-disease-u-s/#item-start
20. Substance Abuse and Mental Health Services Administration [SAMHSA]. *Projections of National Expenditures for Treatment of Mental and Substance Use Disorders, 2010–2020.* .; 2014.
21. Moller H-J, Bandelow B, Volz H-P, Barnikol UB, Seifritz E, Kasper S. The relevance of "mixed anxiety and depression" as a diagnostic category in clinical practice. *European archives of psychiatry and clinical neuroscience*. 2016;266(8):725-736. doi:10.1007/s00406-016-0684-7
22. Cosgrove VE, Rhee SH, Gelhorn HL, et al. Structure and etiology of co-occurring internalizing and externalizing disorders in adolescents. *Journal of abnormal child psychology*. 2011;39(1):109-123. doi:10.1007/s10802-010-9444-8
23. Turgay A. Treatment of comorbidity in conduct disorder with attention-deficit hyperactivity disorder (ADHD). *Essential psychopharmacology*. 2005;6(5):277-290.
24. American Psychological Association. ICD vs. DSM. Published 2009. Accessed April 4, 2021. https://www.apa.org/monitor/2009/10/icd-dsm#:~:text=The%20ICD%20is%20produced%20by,disease%20burden%20of%20mental%20disorders.
25. American Academy of Pediatrics. ICD-10-CM used to document diagnoses but also affects payment. AAP News. Published February 6, 2019. Accessed April 4, 2021. https://www.aappublications.org/news/2019/02/06/coding020619#:~:text=ICD%2D10%2DCM%20is%20the,outcomes%2C%20mortality%20statistics%20and%20billing
26. National Institute on Drug Abuse. Screening and Assessment Tools Chart. Published March 23, 2021. Accessed April 4, 2021. https://www.drugabuse.gov/nidamed-medical-health-professionals/screening-tools-resources/chart-screening-tools
27. Strong DR, Pearson J, Ehlke S, et al. Indicators of dependence for different types of tobacco product users: Descriptive findings from Wave 1 (2013-2014) of the Population Assessment of Tobacco and Health (PATH) study. *Drug and alcohol dependence*. 2017;178:257-266. doi:10.1016/j.drugalcdep.2017.05.010
28. Dennis ML, Chan YF, Funk RR. Development and validation of the GAIN Short Screener (GSS) for internalizing, externalizing and substance use disorders and crime/violence problems among adolescents and adults. *The American journal on addictions*. 2006;15 Suppl 1:80-91. doi:G007886P85267424 [pii]

29. National Institute on Alcohol Abuse and Alcoholism. Alcohol Use in the United States. Alcohol Facts and Statistics. Published March 2021. Accessed April 4, 2021. https://www.niaaa.nih.gov/publications/brochures-and-fact-sheets/alcohol-facts-and-statistics
30. National Academies of Sciences E and MH and MDB on PH and PHP. *Committee on the Health Effects of Marijuana: An Evidence Review and Research Agenda*. National Academies Press (US); 2017.
31. Abramson A. Substance use during the pandemic. American Psychological Association. Published March 1, 2021. Accessed April 4, 2021. https://www.apa.org/monitor/2021/03/substance-use-pandemic
32. Hossain MM, Tasnim S, Sultana A, et al. Epidemiology of mental health problems in COVID-19: a review. *F1000Research*. 2020;9:636. doi:10.12688/f1000research.24457.1
33. Boschloo L, Borkulo CD van, Rhemtulla M, Keyes KM. The Network Structure of Symptoms of the Diagnostic and Statistical Manual of Mental Disorders. 2015;(Icd):1-12. doi:10.1371/journal.pone.0137621
34. Rhemtulla M, Fried EI, Aggen SH, Tuerlinckx F, Kendler KS, Borsboom D. Network analysis of substance abuse and dependence symptoms. *Drug and alcohol dependence*. 2016;161:230-237. doi:10.1016/j.drugalcdep.2016.02.005
35. Timko C, Ilgen M, Haverfield M, Shelley A, Breland JY. Polysubstance use by psychiatry inpatients with co-occurring mental health and substance use disorders. *Drug and Alcohol Dependence*. 2017;180:319-322. doi:10.1016/j.drugalcdep.2017.08.018
36. Wu L-T, Ling W, Burchett B, et al. Use of item response theory and latent class analysis to link poly-substance use disorders with addiction severity, HIV risk, and quality of life among opioid-dependent patients in the Clinical Trials Network. *Drug and alcohol dependence*. 2011;118(2):186-193. doi:10.1016/j.drugalcdep.2011.03.018
37. Liu Y, Elliott AL, Serdarevic M, Leeman RF, Cottler LB. A latent class analysis of the past-30-day substance use patterns among lifetime cocaine users: Findings from a community sample in North Central Florida. *Addictive Behaviors Reports*. 2019;9:100170. doi:https://doi.org/10.1016/j.abrep.2019.100170
38. Krueger RF. The structure of common mental disorders. *Archives of general psychiatry*. 1999;56(10):921-926. doi:10.1001/archpsyc.56.10.921
39. Krueger RF, McGue M, Iacono WG. The higher-order structure of common DSM mental disorders: internalization, externalization, and their connections to personality. *Personality and Individual Differences*. 2001;30(7):1245-1259. doi:https://doi.org/10.1016/S0191-8869(00)00106-9
40. Krueger RF, Caspi A, Moffitt TE, Silva PA. The structure and stability of common mental disorders (DSM-III-R): A longitudinal-epidemiological study. *Journal of Abnormal Psychology*. 1998;107(2):216-227. doi:10.1037/0021-843X.107.2.216
41. El-Gabalawy R, Tsai J, Harpaz-Rotem I, Hoff R, Sareen J, Pietrzak RH. Predominant typologies of psychopathology in the United States: a latent class analysis. *Journal of psychiatric research*. 2013;47(11):1649-1657. doi:10.1016/j.jpsychires.2013.07.028
42. Liu Y, Elliott AL, Serdarevic M, Leeman RF, Cottler LB. A latent class analysis of the past-30-day substance use patterns among lifetime cocaine users: Findings from a community sample in North Central Florida. *Addictive behaviors reports*. 2019;9:100170. doi:10.1016/j.abrep.2019.100170

43. Maccoun RJ. The Puzzling Unidimensionality of DSM-5 Substance Use Disorder Diagnoses. *Frontiers in psychiatry*. 2013;4:153. doi:10.3389/fpsyt.2013.00153
44. Engel GL. The need for a new medical model: a challenge for biomedicine. *Science (New York, NY)*. 1977;196(4286):129-136. doi:10.1126/science.847460
45. Borsboom D. A network theory of mental disorders. *World psychiatry : official journal of the World Psychiatric Association (WPA)*. 2017;16(1):5-13. doi:10.1002/wps.20375
46. Borsboom D, Cramer OJ. Network Analysis : An Integrative Approach to the Structure of Psychopathology. Published online 2013. doi:10.1146/annurev-clinpsy-050212-185608
47. Kendler KS, Zachar P, Craver C. What kinds of things are psychiatric disorders? *Psychological medicine*. 2011;41(6):1143-1150. doi:10.1017/S0033291710001844
48. KENDLER KS. The dappled nature of causes of psychiatric illness: replacing the organic-functional/hardware-software dichotomy with empirically based pluralism. *Molecular psychiatry*. 2012;17(4):377-388. doi:10.1038/mp.2011.182
49. Fried EI, van Borkulo CD, Cramer AOJ, Boschloo L, Schoevers RA, Borsboom D. Mental disorders as networks of problems: a review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*. 2017;52(1):1-10. doi:10.1007/s00127-016-1319-z
50. Fried EI, Cramer AOJ. Moving Forward: Challenges and Directions for Psychopathological Network Theory and Methodology. *Perspectives on Psychological Science*. 2017;12(6):999-1020. doi:10.1177/1745691617705892
51. Gordis Leon 1934-2015. *Epidemiology*. 4th ed.. Elsevier/Saunders; 2009.
52. Tomczyk S, Pedersen A, Hanewinkel R, Isensee B, Morgenstern M. Polysubstance use patterns and trajectories in vocational students--a latent transition analysis. *Addictive behaviors*. 2016;58:136-141. doi:10.1016/j.addbeh.2016.02.027
53. Kandel DB, Yamaguchi K, Chen K. Stages of progression in drug involvement from adolescence to adulthood: further evidence for the gateway theory. *Journal of studies on alcohol*. 1992;53(5):447—457. doi:10.15288/jsa.1992.53.447
54. Maldonado-Molina MM, Lanza ST. A Framework to Examine Gateway Relations in Drug Use: An Application of Latent Transition Analysis. *Journal of drug issues*. 2010;40(4):901-924. doi:10.1177/002204261004000407
55. CORYELL W, ENDICOTT J, KELLER M. Outcome of patients with chronic affective disorder: a five-year follow- up. *The American journal of psychiatry*. 1990;147(12):1627-1633. doi:10.1176/ajp.147.12.1627
56. Klein DN, Taylor EB, Dickstein S, Harding K. Primary Early-Onset Dysthymia: Comparison With Primary Nonbipolar Nonchronic Major Depression on Demographic, Clinical, Familial, Personality, and Socioenvironmental Characteristics and Short-Term Outcome. *Journal of abnormal psychology (1965)*. 1988;97(4):387-398. doi:10.1037/0021-843X.97.4.387
57. Groenman PhD AP, Janssen PhD TWP, Oosterlaan PhD J. Childhood Psychiatric Disorders as Risk Factor for Subsequent Substance Abuse: A Meta-Analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2017;56(7):556-569. doi:10.1016/j.jaac.2017.05.004
58. Miettunen J, Murray GK, Jones PB, et al. Longitudinal associations between childhood and adulthood externalizing and internalizing psychopathology and adolescent substance use. *Psychological medicine*. 2014;44(8):1727-1738. doi:10.1017/S0033291713002328
59. MONSHOUWER K, van DORSSELAER S, VERDURMEN J, BOGT TOMTER, de GRAAF RON, VOLLEBERGH W. Cannabis use and mental health in secondary school

children: Findings from a Dutch survey. *British journal of psychiatry*. 2006;188(2):148-153. doi:10.1192/bjp.188.2.148
60. Maggs JL, Patrick ME, Feinstein L. Childhood and adolescent predictors of alcohol use and problems in adolescence and adulthood in the National Child Development Study. *Addiction (Abingdon, England)*. 2008;103(s1):7-22. doi:10.1111/j.1360-0443.2008.02173.x
61. Colder CR, Scalco M, Trucco EM, et al. Prospective associations of internalizing and externalizing problems and their co-occurrence with early adolescent substance use. *Journal of abnormal child psychology*. 2013;41(4):667-677. doi:10.1007/s10802-012-9701-0 [doi]
62. Kendler KS, Edwards A, Myers J, Cho S bin, Adkins A, Dick D. The predictive power of family history measures of alcohol and drug problems and internalizing disorders in a college population. *American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics*. 2015;168B(5):337-346. doi:10.1002/ajmg.b.32320
63. Kendler KS, Prescott CA, Myers J, Neale MC. The structure of genetic and environmental risk factors for common psychiatric and substance use disorders in men and women. *Archives of general psychiatry*. 2003;60(9):929-937. doi:10.1001/archpsyc.60.9.929
64. Edwards AC, Maes HH, Pedersen NL, Kendler KS. A population-based twin study of the genetic and environmental relationship of major depression, regular tobacco use and nicotine dependence. *Psychological medicine*. 2011;41(2):395-405. doi:10.1017/S0033291710000589
65. Shoal GD, Giancola PR. The relation between social problems and substance use in adolescent boys: an investigation of potential moderators. *Experimental and clinical psychopharmacology*. 2005;13(4):357-366. doi:10.1037/1064-1297.13.4.357
66. Trenz RC, Scherer M, Harrell P, Zur J, Sinha A, Latimer W. Early onset of drug and polysubstance use as predictors of injection drug use among adult drug users. *Addictive behaviors*. 2012;37(4):367-372. doi:10.1016/j.addbeh.2011.11.011
67. Kim H, Eaton NR. A Hierarchical Integration of Person-Centered Comorbidity Models: Structure, Stability, and Transition Over Time. *Clinical psychological science*. 2017;5(4):595-612. doi:10.1177/2167702617704018
68. Hawke JM, Kaminer Y, Burke R, Burleson JA. Stability of Comorbid Psychiatric Diagnosis Among Youths in Treatment and Aftercare for Alcohol Use Disorders. *Substance abuse*. 2008;29(2):33-41. doi:10.1080/08897070802093015
69. Wittchen H-U, Lieb R, Pfister H, Schuster P. The waxing and waning of mental disorders: Evaluating the stability of syndromes of mental disorders in the population. *Comprehensive psychiatry*. 2000;41(2):122-132. doi:10.1016/S0010-440X(00)80018-8
70. Nylund KL. Latent transition analysis: Modeling extensions and an application to peer victimization. Published online 2007.
71. Abuse USD of H and HServicesNI of HealthNI on D, Products USD of H and HServicesF and DAdministrationC for T. Population Assessment of Tobacco and Health (PATH) Study [United States] Public-Use Files. Published online 2019. doi:10.3886/ICPSR36498.v10
72. Hyland A, Ambrose BK, Conway KP, et al. Design and methods of the Population Assessment of Tobacco and Health (PATH) Study. *Tobacco control*. 2017;26(4):371-378. doi:10.1136/tobaccocontrol-2016-052934 [doi]

73. Conway KP, Green VR, Kasza KA, et al. Co-occurrence of tobacco product use, substance use, and mental health problems among adults: Findings from Wave 1 (2013-2014) of the Population Assessment of Tobacco and Health (PATH) Study. *Drug and alcohol dependence*. 2017;177:104-111. doi:S0376-8716(17)30222-3 [pii]
74. Miller S, Pike J, Shono Y, Beleva Y, Xie B, Stacy AW. The role of negative affect in the persistence of nicotine dependence among alternative high school students: A latent growth curve analysis. *Drug and alcohol dependence*. 2020;209:107883. doi:10.1016/j.drugalcdep.2020.107883
75. Bobo JK, Husten C. Sociocultural influences on smoking and drinking. *Alcohol research & health : the journal of the National Institute on Alcohol Abuse and Alcoholism*. 2000;24(4):225-232. https://pubmed.ncbi.nlm.nih.gov/15986717
76. National Institute on Alcohol Abuse and Alcoholism. *Alcohol and Tobacco*.; 2007. https://pubs.niaaa.nih.gov/publications/aa71/aa71.htm
77. Centers for Disease Control and Prevention. Current Cigarette Smoking Among Adults in the United States. Published 2019. https://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/index.htm#:~:text=In 2018%2C nearly 14 of,with a smoking-related disease.
78. Creamer MR, Wang TW, Babb S et al. *Tobacco Product Use and Cessation Indicators Among Adults — United States, 2018*.; 2019. doi:http://dx.doi.org/10.15585/mmwr.mm6845a2
79. Substance Abuse and Mental Health Services Administration (SAMHSA). 2019 National Survey on Drug Use and Health (NSDUH). *Table 2.1B—Tobacco Product and Alcohol Use in Lifetime, Past Year, and Past Month among Persons Aged 12 or Older, by Age Group: Percentages, 2018 and 2019*. https://www.samhsa.gov/data/sites/default/files/cbhsq-reports/NSDUHDetailedTabs2018R2/NSDUHDetTabsSect2pe2018.htm#tab2-1b
80. Adams S. Psychopharmacology of Tobacco and Alcohol Comorbidity: a Review of Current Evidence. *Current addiction reports*. 2017;4(1):25-34. doi:10.1007/s40429-017-0129-z
81. Cross SJ, Lotfipour S, Leslie FM. Mechanisms and genetic factors underlying co-use of nicotine and alcohol or other drugs of abuse. *The American Journal of Drug and Alcohol Abuse*. 2017;43(2):171-185. doi:10.1080/00952990.2016.1209512 [doi]
82. Verplaetse TL, McKee SA. An overview of alcohol and tobacco/nicotine interactions in the human laboratory. *The American journal of drug and alcohol abuse*. 2017;43(2):186-196. doi:10.1080/00952990.2016.1189927
83. Maglia M, Caponnetto P, di Piazza J, la Torre D, Polosa R. Dual use of electronic cigarettes and classic cigarettes: a systematic review. *Addiction Research & Theory*. 2018;26(4):330-338. doi:10.1080/16066359.2017.1388372
84. Villarroel MA, Ph D, Cha AE, Ph D, Vahratian A, Ph D. Electronic Cigarette Use Among U . S . Adults , 2018. 2020;(365):1-8.
85. Schoenborn CA, Gindi RM. Electronic Cigarette Use Among Adults: United States, 2014. *NCHS data brief*. 2015;(217):1-8.
86. Hershberger AR, VanderVeen JD, Karyadi KA, Cyders MA. Transitioning From Cigarettes to Electronic Cigarettes Increases Alcohol Consumption. *Substance use & misuse*. 2016;51(14):1838-1845. doi:10.1080/10826084.2016.1197940

87. Roberts W, Moore KE, Peltier MR, et al. Electronic Cigarette Use and Risk of Harmful Alcohol Consumption in the U.S. Population. *Alcoholism, clinical and experimental research*. 2018;42(12):2385-2393. doi:10.1111/acer.13889
88. Roberts W, Verplaetse T, Peltier MKR, Moore KE, Gueorguieva R, McKee SA. Prospective association of e-cigarette and cigarette use with alcohol use in two waves of the Population Assessment of Tobacco and Health. *Addiction (Abingdon, England)*. Published online January 2020. doi:10.1111/add.14980
89. McClernon FJ, Kollins SH. ADHD and smoking: from genes to brain to behavior. *Annals of the New York Academy of Sciences*. 2008;1141:131-147. doi:10.1196/annals.1441.016
90. Smith PH, Mazure CM, McKee SA. Smoking and mental illness in the U.S. population. *Tobacco control*. 2014;23(e2):e147-53. doi:10.1136/tobaccocontrol-2013-051466 [doi]
91. Ziedonis D, Hitsman B, Beckham JC, et al. Tobacco use and cessation in psychiatric disorders: National Institute of Mental Health report. *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco*. 2008;10(12):1691-1715. doi:10.1080/14622200802443569
92. Luger TM, Suls J, Weg MW vander. How robust is the association between smoking and depression in adults? A meta-analysis using linear mixed-effects models. *Addictive Behaviors*. 2014;39(10):1418-1429. doi:10.1016/j.addbeh.2014.05.011 [doi]
93. Mathew AR, Hogarth L, Leventhal AM, Cook JW, Hitsman B. Cigarette smoking and depression comorbidity: systematic review and proposed theoretical model. *Addiction (Abingdon, England)*. 2017;112(3):401-412. doi:10.1111/add.13604
94. Grant JE, Lust K, Fridberg DJ, King AC, Chamberlain SR. E-cigarette use (vaping) is associated with illicit drug use, mental health problems, and impulsivity in university students. *Annals of clinical psychiatry : official journal of the American Academy of Clinical Psychiatrists*. 2019;31(1):27-35.
95. Hefner KR, Sollazzo A, Mullaney S, Coker KL, Sofuoglu M. E-cigarettes, alcohol use, and mental health: Use and perceptions of e-cigarettes among college students, by alcohol use and mental health status. *Addictive behaviors*. 2019;91:12-20. doi:10.1016/j.addbeh.2018.10.040
96. Vallone DM, Cuccia AF, Briggs J, Xiao H, Schillo BA, Hair EC. Electronic Cigarette and JUUL Use Among Adolescents and Young Adults. *JAMA pediatrics*. Published online January 2020. doi:10.1001/jamapediatrics.2019.5436
97. Wong S-W, Lin H-C, Piper ME, Siddiqui A, Buu A. Measuring characteristics of e-cigarette consumption among college students. *Journal of American college health : J of ACH*. 2019;67(4):338-347. doi:10.1080/07448481.2018.1481075
98. Prevention NC for HSC for DC and. National Health Interview Survey [Data set].
99. Garner BR, Belur VK, Dennis ML. The GAIN Short Screener (GSS) as a Predictor of Future Arrest or Incarceration Among Youth Presenting to Substance Use Disorder (SUD) Treatment. *Substance abuse : research and treatment*. 2013;7:199-208. doi:10.4137/SART.S13152
100. Goodwin RD, Zvolensky MJ, Keyes KM, Hasin DS. Mental disorders and cigarette use among adults in the United States. *The American journal on addictions*. 2012;21(5):416-423. doi:10.1111/j.1521-0391.2012.00263.x
101. Grant BF, Hasin DS, Chou SP, Stinson FS, Dawson DA. Nicotine dependence and psychiatric disorders in the United States: results from the national epidemiologic survey

on alcohol and related conditions. *Archives of general psychiatry*. 2004;61(11):1107-1115. doi:10.1001/archpsyc.61.11.1107
102. Drobes DJ. *Concurrent Alcohol and Tobacco Dependence, Mechanisms and Treatment*.; 2002. https://pubs.niaaa.nih.gov/publications/arh26-2/136-142.htm
103. Jankowski M, Krzystanek M, Zejda JE, et al. E-Cigarettes are More Addictive than Traditional Cigarettes-A Study in Highly Educated Young People. *International journal of environmental research and public health*. 2019;16(13). doi:10.3390/ijerph16132279
104. Rostron BL, Schroeder MJ, Ambrose BK. Dependence symptoms and cessation intentions among US adult daily cigarette, cigar, and e-cigarette users, 2012-2013. *BMC public health*. 2016;16(1):814. doi:10.1186/s12889-016-3510-2
105. Ames SC, Stevens SR, Chudley E, et al. Students : Alcohol and Tobacco Use The Association of Alcohol Consumption with Tobacco Use in Black and White College Students. 2010;6084:1230-1244. doi:10.3109/10826080903554192
106. Bizzarri J v, Casetti V, Panzani P, et al. Risky use and misuse of alcohol and cigarettes in psychiatric inpatients: a screening questionnaire study. *Comprehensive psychiatry*. 2016;70:9-16. doi:10.1016/j.comppsych.2016.05.011 [doi]
107. Cance JD, Talley AE, Morgan-Lopez A, Fromme K. Longitudinal Conjoint Patterns of Alcohol and Tobacco Use throughout Emerging Adulthood. *Substance use & misuse*. 2017;52(3):373-382. doi:10.1080/10826084.2016.1228677
108. Caton CLM, Xie H, Drake RE, McHugo G. Gender differences in psychotic disorders with concurrent substance use. *Journal of Dual Diagnosis*. 2014;10(4):177-186. doi:10.1080/15504263.2014.961882
109. Choi NG, DiNitto DM, Marti CN. Alcohol and other substance use, mental health treatment use, and perceived unmet treatment need: Comparison between baby boomers and older adults. *The American journal on addictions*. 2015;24(4):299-307. doi:10.1111/ajad.12225 [doi]
110. Galea S, Ahern J, Tracy M, Vlahov D. Neighborhood Income and Income Distribution and the Use of Cigarettes, Alcohol, and Marijuana. *American journal of preventive medicine*. 2007;32(6 Suppl):S195-202. doi:10.1016/j.amepre.2007.04.003
111. Hrywna M, Bover Manderski MT, Delnevo CD. Sex differences in the association of psychological distress and tobacco use. *American journal of health behavior*. 2014;38(4):570-576. doi:10.5993/AJHB.38.4.10
112. Karriker-jaffe KJ. Neighborhood socioeconomic status and substance use by U . S . adults. *Drug and Alcohol Dependence*. 2013;133(1):212-221. doi:10.1016/j.drugalcdep.2013.04.033
113. Keyes KM, Vo T, Wall MM, et al. Racial/ethnic differences in use of alcohol, tobacco, and marijuana: is there a cross-over from adolescence to adulthood? *Social science & medicine (1982)*. 2015;124:132-141. doi:10.1016/j.socscimed.2014.11.035
114. Peiper N, Rodu B. Evidence of sex differences in the relationship between current tobacco use and past-year serious psychological distress: 2005-2008 National Survey on Drug Use and Health. *Social Psychiatry and Psychiatric Epidemiology*. 2013;48(8):1261-1271. doi:10.1007/s00127-012-0644-0
115. Andreas JB, Lauritzen G, Nordfjaern T. Co-occurrence between mental distress and poly-drug use: a ten year prospective study of patients from substance abuse treatment. *Addictive Behaviors*. 2015;48:71-78. doi:10.1016/j.addbeh.2015.05.001 [doi]

116. Bandiera FC, Anteneh B, Le T, Delucchi K, Guydish J. Tobacco-related mortality among persons with mental health and substance abuse problems. *PloS one*. 2015;10(3):e0120581. doi:10.1371/journal.pone.0120581 [doi]
117. Carragher N, Krueger RF, Eaton NR, Slade T. Disorders without borders: current and future directions in the meta-structure of mental disorders. *Social psychiatry and psychiatric epidemiology*. 2015;50(3):339-350. doi:10.1007/s00127-014-1004-z
118. Eaton NR, Rodriguez-Seijas C, Carragher N, Krueger RF. Transdiagnostic factors of psychopathology and substance use disorders: a review. *Social psychiatry and psychiatric epidemiology*. 2015;50(2):171-182. doi:10.1007/s00127-014-1001-2
119. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software; Vol 1, Issue 2 (2012)* . Published online May 24, 2012. https://www.jstatsoft.org/v048/i02
120. Babson KA, Feldner MT, Sachs-Ericsson N, Schmidt NB, Zvolensky MJ. Nicotine dependence mediates the relations between insomnia and both panic and posttraumatic stress disorder in the NCS-R sample. *Depression and anxiety*. 2008;25(8):670-679. doi:10.1002/da.20374
121. Maes HH, Sullivan PF, Bulik CM, et al. A twin study of genetic and environmental influences on tobacco initiation, regular tobacco use and nicotine dependence. *Psychological medicine*. 2004;34(7):1251-1261. doi:10.1017/s0033291704002405
122. Rodriguez-Seijas C, Stohl M, Hasin DS, Eaton NR. Transdiagnostic Factors and Mediation of the Relationship Between Perceived Racial Discrimination and Mental Disorders. *JAMA psychiatry*. 2015;72(7):706-713. doi:10.1001/jamapsychiatry.2015.0148 [doi]
123. Maddux. *Psychopathology: Foundations for a Contemporary Understanding*. 4th ed. Routledge; 2016.
124. Olino TM, Klein DN, Farmer RF, Seeley JR, Lewinsohn PM. Examination of the structure of psychopathology using latent class analysis. *Comprehensive psychiatry*. 2012;53(4):323-332. doi:10.1016/j.comppsych.2011.05.008
125. Truth Initiative. E-cigarettes: Facts, stats and regulations. Published April 2, 2021. Accessed April 4, 2021. https://truthinitiative.org/research-resources/emerging-tobacco-products/e-cigarettes-facts-stats-and-regulations
126. National Institute on Drug Abuse. *Comorbidity: Addiction and Other Mental Illnesses* .; 2010. Accessed April 4, 2021. https://www.drugabuse.gov/sites/default/files/rrcomorbidity.pdf
127. Dobkin PL, De CM, Paraherakis A, Gill K. The role of functional social support in treatment retention and outcomes among outpatient adult substance abusers. *Addiction (Abingdon, England)*. 2002;97(3):347-356. doi:10.1046/j.1360-0443.2002.00083.x
128. Stevens E, Jason LA, Ram D, Light J. Investigating Social Support and Network Relationships in Substance Use Disorder Recovery. *Substance Abuse*. 2015;36(4):396-399. doi:10.1080/08897077.2014.965870
129. Stucky BD, Edelen MO, Ramchand R. A psychometric assessment of the GAIN individual severity scale (GAIN-GISS) and short screeners (GAIN-SS) among adolescents in outpatient treatment programs. *Journal of substance abuse treatment*. 2014;46(2):165-173. doi:10.1016/j.jsat.2013.07.007
130. McCutcheon Allan L. 1950-. *Latent Class Analysis*. Sage Publications; 1987.

131. Clark SL, Muthén B, Kaprio J, D'Onofrio BM, Viken R, Rose RJ. Models and Strategies for Factor Mixture Analysis: An Example Concerning the Structure Underlying Psychological Disorders. *Structural equation modeling*. 2013;20(4):681-703. doi:10.1080/10705511.2013.824786
132. Forbes D, Elhai JD, Miller MW, Creamer M. Internalizing and externalizing classes in posttraumatic stress disorder: a latent class analysis. *Journal of traumatic stress*. 2010;23(3):340-349. doi:10.1002/jts.20526
133. Tein J-Y, Coxe S, Cham H. Statistical Power to Detect the Correct Number of Classes in Latent Profile Analysis. *Structural equation modeling*. 2013;20(4):640-657. doi:10.1080/10705511.2013.824781
134. Lo Y, Mendell NR, Rubin DB. Testing the Number of Components in a Normal Mixture. *Biometrika*. 2001;88(3):767-778. doi:10.1093/biomet/88.3.767
135. Holt JA, Macready GB. A Simulation Study of the Difference Chi-Square Statistic for Comparing Latent Class Models Under Violation of Regularity Conditions. *Applied psychological measurement*. 1989;13(3):221-231. doi:10.1177/014662168901300301
136. AKAIKE H. Factor analysis and AIC. *Psychometrika*. 1987;52(3):317-332. doi:10.1007/BF02294359
137. Schwarz G. Estimating the Dimension of a Model. *The Annals of statistics*. 1978;6(2):461-464. doi:10.1214/aos/1176344136
138. Schreiber JB. Latent Class Analysis: An example for reporting results. *Research in social and administrative pharmacy*. 2017;13(6):1196-1201. doi:10.1016/j.sapharm.2016.11.011
139. SCLOVE SL. Application of model-selection criteria to some problems in multivariate analysis. *Psychometrika*. 1987;52(3):333-343. doi:10.1007/BF02294360
140. CELEUX G, SOROMENHO G. An entropy criterion for assessing the number of clusters in a mixture model. *Journal of classification*. 1996;13(2):195-212. doi:10.1007/BF01246098
141. Muthén B. Latent Variable Analysis: Growth Mixture Modeling and Related Techniques for Longitudinal Data. Published online 2004. doi:10.4135/9781412986311.n19
142. Asparouhov T, Muthén B. Auxiliary Variables in Mixture Modeling: Three-Step Approaches Using Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*. 2014;21(3):329-341. doi:10.1080/10705511.2014.915181
143. deCastro BR. Cumulative ROC curves for discriminating three or more ordinal outcomes with cutpoints on a shared continuous measurement scale. *PloS one*. 2019;14(8):e0221433. doi:10.1371/journal.pone.0221433
144. Hoo ZH, Candlish J, Teare D. What is an ROC curve? *Emergency medicine journal : EMJ*. 2017;34(6):357-359. doi:10.1136/emermed-2017-206735
145. Hoo ZH, Candlish J, Teare D. What is an ROC curve? *Emergency medicine journal : EMJ*. 2017;34(6):357-359. doi:10.1136/emermed-2017-206735
146. Muller CJ, MacLehose RF. Estimating predicted probabilities from logistic regression: different methods correspond to different target populations. *International journal of epidemiology*. 2014;43(3):962-970. doi:10.1093/ije/dyu029
147. Muthén LK, & MBO. Mplus User's Guide.
148. Blondino CT, Clifford JS, Lu J, Prom-Wormley EC. The Association between Internalizing and Externalizing Severity with Current Use of Cigarettes, E-cigarettes, and Alcohol in Adults: Wave 1 of the Population Assessment of Tobacco and Health (PATH)

Study. *Addictive Behaviors*. Published online 2021:106890. doi:https://doi.org/10.1016/j.addbeh.2021.106890
149. Buckner JD, Schmidt NB, Lang AR, Small JW, Schlauch RC, Lewinsohn PM. Specificity of social anxiety disorder as a risk factor for alcohol and cannabis dependence. *Journal of psychiatric research*. 2008;42(3):230-239. doi:10.1016/j.jpsychires.2007.01.002
150. Hasin DS, Kerridge BT, Saha TD, et al. Prevalence and Correlates of DSM-5 Cannabis Use Disorder, 2012-2013: Findings from the National Epidemiologic Survey on Alcohol and Related Conditions-III. *The American journal of psychiatry*. 2016;173(6):588-599. doi:10.1176/appi.ajp.2015.15070907
151. Kedzior KK, Laeber LT. A positive association between anxiety disorders and cannabis use or cannabis use disorders in the general population--a meta-analysis of 31 studies. *BMC psychiatry*. 2014;14:136. doi:10.1186/1471-244X-14-136
152. Rasic D, Weerasinghe S, Asbridge M, Langille DB. Longitudinal associations of cannabis and illicit drug use with depression, suicidal ideation and suicidal attempts among Nova Scotia high school students. *Drug and alcohol dependence*. 2013;129(1-2):49-53. doi:10.1016/j.drugalcdep.2012.09.009
153. Bassir Nia A, Mann C, Kaur H, Ranganathan M. Cannabis Use: Neurobiological, Behavioral, and Sex/Gender Considerations. *Current behavioral neuroscience reports*. 2018;5(4):271-280. doi:10.1007/s40473-018-0167-4
154. Price RK, Risk NK, Haden AH, Lewis CE, Spitznagel EL. Post-traumatic stress disorder, drug dependence, and suicidality among male Vietnam veterans with a history of heavy drug use. *Drug and alcohol dependence*. 2004;76 Suppl:S31-43. doi:10.1016/j.drugalcdep.2004.08.005
155. Goldner EM, Lusted A, Roerecke M, Rehm J, Fischer B. Prevalence of Axis-1 psychiatric (with focus on depression and anxiety) disorder and symptomatology among non-medical prescription opioid users in substance use treatment: systematic review and meta-analyses. *Addictive behaviors*. 2014;39(3):520-531. doi:10.1016/j.addbeh.2013.11.022
156. Trujillo CA, Obando D, Trujillo A. An examination of the association between early initiation of substance use and interrelated multilevel risk and protective factors among adolescents. *PloS one*. 2019;14(12):e0225384-e0225384. doi:10.1371/journal.pone.0225384
157. SUITE DH, la BRIL R, PRIMM A, HARRISON-ROSS P. Beyond misdiagnosis, misunderstanding and mistrust : Relevance of the historical perspective in the medical and mental health treatment of people of color. *Journal of the National Medical Association*. 2007;99(8):879-885.
158. Erving CL, Thomas CS, Frazier C. Is the Black-White Mental Health Paradox Consistent Across Gender and Psychiatric Disorders? *American journal of epidemiology*. 2019;188(2):314-322. doi:10.1093/aje/kwy224
159. Sareen J, Afifi TO, McMillan KA, Asmundson GJG. Relationship Between Household Income and Mental Disorders: Findings From a Population-Based Longitudinal Study. *Archives of General Psychiatry*. 2011;68(4):419-427. doi:10.1001/archgenpsychiatry.2011.15
160. Cohen S, Wills TA. Stress, Social Support, and the Buffering Hypothesis. *Psychological Bulletin*. Published online 1985. doi:10.1037/0033-2909.98.2.310

161. Bennette C, Vickers A. Against quantiles: categorization of continuous variables in epidemiologic research, and its discontents. *BMC Medical Research Methodology*. 2012;12(1):21. doi:10.1186/1471-2288-12-21
162. Lanza ST, Rhoades BL. Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. *Prevention science : the official journal of the Society for Prevention Research*. 2013;14(2):157-168. doi:10.1007/s11121-011-0201-1
163. Reboussin BA, Ip EH, Wolfson M. Locally dependent latent class models with covariates: an application to under-age drinking in the USA. *Journal of the Royal Statistical Society Series A, (Statistics in Society)*. 2008;171(4):877-897. doi:10.1111/j.1467-985X.2008.00544.x
164. Clark SL, Muthén B, Kaprio J, D'Onofrio BM, Viken R, Rose RJ. Models and Strategies for Factor Mixture Analysis: An Example Concerning the Structure Underlying Psychological Disorders. *Structural equation modeling : a multidisciplinary journal*. 2013;20(4):10.1080/10705511.2013.824786. doi:10.1080/10705511.2013.824786
165. Nylund KL, Asparouhov T, Muthén BO. Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal*. 2007;14(4):535-569. doi:10.1080/10705510701575396
166. NAMI. *Health Reform & Mental Illness*. Accessed May 5, 2021. https://www.nami.org/getattachment/Get-Involved/NAMI-National-Convention/Convention-Program-Schedule/Hill-Day-2017/FINAL-Hill-Day-17-Leave-Behind-all-(1).pdf
167. Merrell KW. Understanding Internalizing Problems: Depression and Anxiety in Children and Adolescents. In: *Helping Students Overcome Depression and Anxiety, Second Edition: A Practical Guide*. ; 2008.
168. Liu J. Childhood externalizing behavior: theory and implications. *Journal of child and adolescent psychiatric nursing : official publication of the Association of Child and Adolescent Psychiatric Nurses, Inc*. 2004;17(3):93-103. doi:10.1111/j.1744-6171.2004.tb00003.x
169. Duncan B, Rees DI. Effect of smoking on depressive symptomatology: a reexamination of data from the National Longitudinal Study of Adolescent Health. *American journal of epidemiology*. 2005;162(5):461-470. doi:10.1093/aje/kwi219
170. Fluharty M, Taylor AE, Grabski M, Munafo MR. The Association of Cigarette Smoking With Depression and Anxiety: A Systematic Review. *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco*. 2017;19(1):3-13. doi:10.1093/ntr/ntw140
171. Moylan S, Jacka FN, Pasco JA, Berk M. Cigarette smoking, nicotine dependence and anxiety disorders: a systematic review of population-based, epidemiological studies. *BMC medicine*. 2012;10:123. doi:10.1186/1741-7015-10-123
172. Rhodes JD, Pelham WE, Gnagy EM, Shiffman S, Derefinko KJ, Molina BSG. Cigarette smoking and ADHD: An examination of prognostically relevant smoking behaviors among adolescents and young adults. *Psychology of addictive behaviors : journal of the Society of Psychologists in Addictive Behaviors*. 2016;30(5):588-600. doi:10.1037/adb0000188

173. Anker JJ, Kushner MG. Co-Occurring Alcohol Use Disorder and Anxiety: Bridging Psychiatric, Psychological, and Neurobiological Perspectives. *Alcohol research : current reviews*. 2019;40(1). doi:10.35946/arcr.v40.1.03
174. Boden JM, Fergusson DM. Alcohol and depression. *Addiction (Abingdon, England)*. 2011;106(5):906-914. doi:10.1111/j.1360-0443.2010.03351.x
175. Schuckit MA. Alcohol-use disorders. *Lancet (London, England)*. 2009;373(9662):492-501. doi:10.1016/S0140-6736(09)60009-X
176. Charach A, Yeung E, Climans T, Lillie E. Childhood attention-deficit/hyperactivity disorder and future substance use disorders: comparative meta-analyses. *Journal of the American Academy of Child and Adolescent Psychiatry*. 2011;50(1):9-21. doi:10.1016/j.jaac.2010.09.019
177. Ghosh A, Malhotra S, Basu D. Are childhood externalizing disorders the harbinger of early-onset alcohol dependence? *The Indian journal of medical research*. 2016;144(3):385-392. doi:10.4103/0971-5916.198673
178. Kroon E, Kuhns L, Hoch E, Cousijn J. Heavy cannabis use, dependence and the brain: a clinical perspective. *Addiction (Abingdon, England)*. 2020;115(3):559-572. doi:10.1111/add.14776
179. Lev-Ran S, le Foll B, McKenzie K, George TP, Rehm J. Cannabis use and cannabis use disorders among individuals with mental illness. *Comprehensive psychiatry*. 2013;54(6):589-598. doi:10.1016/j.comppsych.2012.12.021
180. Stinson FS, Ruan WJ, Pickering R, Grant BF. Cannabis use disorders in the USA: prevalence, correlates and co-morbidity. *Psychological medicine*. 2006;36(10):1447-1460. doi:10.1017/S0033291706008361
181. Becker WC, Fiellin DA, Desai RA. Non-medical use, abuse and dependence on sedatives and tranquilizers among U.S. adults: psychiatric and socio-demographic correlates. *Drug and alcohol dependence*. 2007;90(2-3):280-287. doi:10.1016/j.drugalcdep.2007.04.009
182. McCabe SE, Cranford JA, West BT. Trends in prescription drug abuse and dependence, co-occurrence with other substance use disorders, and treatment utilization: results from two national surveys. *Addictive behaviors*. 2008;33(10):1297-1305. doi:10.1016/j.addbeh.2008.06.005
183. Conway KP, Compton W, Stinson FS, Grant BF. Lifetime comorbidity of DSM-IV mood and anxiety disorders and specific drug use disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. *The Journal of clinical psychiatry*. 2006;67(2):247-257. doi:10.4088/jcp.v67n0211
184. Blanco C, Alderson D, Ogburn E, et al. Changes in the prevalence of non-medical prescription drug use and drug use disorders in the United States: 1991-1992 and 2001-2002. *Drug and alcohol dependence*. 2007;90(2-3):252-260. doi:10.1016/j.drugalcdep.2007.04.005
185. Katz C, El-Gabalawy R, Keyes KM, Martins SS, Sareen J. Risk factors for incident nonmedical prescription opioid use and abuse and dependence: results from a longitudinal nationally representative sample. *Drug and alcohol dependence*. 2013;132(1-2):107-113. doi:10.1016/j.drugalcdep.2013.01.010
186. Gudin JA, Mogali S, Jones JD, Comer SD. Risks, management, and monitoring of combination opioid, benzodiazepines, and/or alcohol use. *Postgraduate medicine*. 2013;125(4):115-130. doi:10.3810/pgm.2013.07.2684

187. Jones JD, Mogali S, Comer SD. Polydrug abuse: a review of opioid and benzodiazepine combination use. *Drug and alcohol dependence*. 2012;125(1-2):8-18. doi:10.1016/j.drugalcdep.2012.07.004
188. Mcelroy E, Shevlin M, Murphy J, Mcbride O. Co - occurring internalizing and externalizing psychopathology in childhood and adolescence : a network approach. *European Child & Adolescent Psychiatry*. 2018;27(11):1449-1457. doi:10.1007/s00787-018-1128-x
189. Khan S, Okuda M, Hasin DS, et al. Gender differences in lifetime alcohol dependence: results from the national epidemiologic survey on alcohol and related conditions. *Alcoholism, clinical and experimental research*. 2013;37(10):1696-1705. doi:10.1111/acer.12158
190. McHugh RK, Votaw VR, Sugarman DE, Greenfield SF. Sex and gender differences in substance use disorders. *Clinical psychology review*. 2018;66:12-23.
191. Greenfield SF, Brooks AJ, Gordon SM, et al. Substance abuse treatment entry, retention, and outcome in women: a review of the literature. *Drug and alcohol dependence*. 2007;86(1):1-21. doi:10.1016/j.drugalcdep.2006.05.012
192. Lydon-Staley DM, Schnoll RA, Hitsman B, Bassett DS. The Network Structure of Tobacco Withdrawal in a Community Sample of Smokers Treated With Nicotine Patch and Behavioral Counseling. *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco*. 2020;22(3):408-414. doi:10.1093/ntr/nty250
193. van Borkulo C, Epskamp S. Package ‘ IsingFit .’ Published online 2016:1-6. doi:10.1038/srep05918.
194. Barber RF, Drton M. High-dimensional Ising model selection with Bayesian information criteria. *Electron J Statist*. 2015;9(1):567-607. doi:10.1214/15-EJS1012
195. Chen J, Chen Z. Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*. 2008;95(3):759-771. doi:10.1093/biomet/asn034
196. van Borkulo CD, Borsboom D, Epskamp S, et al. A new method for constructing networks from binary data. *Scientific reports*. 2014;4:5918. doi:10.1038/srep05918
197. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*. 2018;50(1):195-212. doi:10.3758/s13428-017-0862-1
198. Epskamp, S., Costantini, G., Haslbeck, J., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., & Borsboom D. Package ‘qgraph’. Graph Plotting Methods, Psychometric Data Visualization and Graphical Model Estimation. Published online 2020:62. doi:10.18637/jss.v048.i04>.URL
199. Ravikumar P, Wainwright MJ, Lafferty JD. High-dimensional Ising model selection using 1 -regularized logistic regression. *Ann Statist*. 2010;38(3):1287-1319. doi:10.1214/09-AOS691
200. Borkulo C v, Boschloo L, Kossakowski JJ, et al. Comparing network structures on three aspects: A permutation test. In: ; 2017.
201. van Borkulo C, Epskamp S, Jones P, Haslbeck J, Millner A. Package ‘ NetworkComparisonTest .’ Published online 2019. doi:10.13140/RG.2.2.29455.38569
202. Haslbeck JMB, Waldorp LJ. How well do network models predict observations? On the importance of predictability in network models. *Behavior research methods*. 2018;50(2):853-861. doi:10.3758/s13428-017-0910-x

203. Haslbeck J. Package ‘mgm.’ Published online February 2, 2021. Accessed April 5, 2021. https://cran.r-project.org/web/packages/mgm/mgm.pdf
204. Epskamp S, Fried EI. Package ‘ bootnet .’ Published online 2020. doi:10.3758/s13428-017-0862-1>.
205. Becker JB, McClellan ML, Reed BG. Sex differences, gender and addiction. *Journal of neuroscience research*. 2017;95(1-2):136-147. doi:10.1002/jnr.23963
206. NIDA. Sex and Gender Differences in Substance Use. Published 2020. https://www.drugabuse.gov/publications/research-reports/substance-use-in-women/sex-gender-differences-in-substance-use
207. Stein MD, Friedmann PD. Disturbed sleep and its relationship to alcohol use. *Substance abuse*. 2005;26(1):1-13. doi:10.1300/j465v26n01_01
208. Britton A, Fat LN, Neligan A. The association between alcohol consumption and sleep disorders among older people in the general population. *Scientific Reports*. 2020;10(1):5275. doi:10.1038/s41598-020-62227-0
209. NIDA. Are there gender differences in tobacco smoking? Published January 2020. Accessed April 5, 2021. https://www.drugabuse.gov/publications/research-reports/tobacco-nicotine-e-cigarettes/are-there-gender-differences-in-tobacco-smoking
210. Warner LA, Kessler RC, Hughes M, Anthony JC, Nelson CB. Prevalence and Correlates of Drug Use and Dependence in the United States: Results From the National Comorbidity Survey. *Archives of General Psychiatry*. 1995;52(3):219-229. doi:10.1001/archpsyc.1995.03950150051010
211. Keyes KM, Grant BF, Hasin DS. Evidence for a closing gender gap in alcohol use, abuse, and dependence in the United States population. *Drug and alcohol dependence*. 2008;93(1-2):21-29. doi:10.1016/j.drugalcdep.2007.08.017
212. Helzer JE, Pryzbeck TR. The co-occurrence of alcoholism with other psychiatric disorders in the general population and its impact on treatment. *Journal of studies on alcohol*. 1988;49(3):219—224. doi:10.15288/jsa.1988.49.219
213. Shaw P, Stringaris A, Nigg J, Leibenluft E. Emotional dysregulation and Attention-Deficit/Hyperactivity Disorder. *The American journal of psychiatry*. 2014;171(3):276-293. doi:10.1176/appi.ajp.2013.13070966
214. Leibenluft E, Stoddard J. The developmental psychopathology of irritability. *Development and psychopathology*. 2013;25(4 Pt 2):1473-1487. doi:10.1017/S0954579413000722
215. Smith BH, Molina BSG, Pelham Jr WE. The Clinically Meaningful Link Between Alcohol Use and Attention Deficit Hyperactivity Disorder. *Alcohol Research & Health*. 2002;26(2):122-129. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6683828/
216. GLASS K, FLORY K. Are Symptoms of ADHD Related to Substance Use Among College Students? *Psychology of addictive behaviors*. 2012;26(1):124-132. doi:10.1037/a0024215
217. Kolla NJ, van der Maas M, Toplak ME, et al. Adult attention deficit hyperactivity disorder symptom profiles and concurrent problems with alcohol and cannabis: sex differences in a representative, population survey. *BMC psychiatry*. 2016;16(49):50. doi:10.1186/s12888-016-0746-4
218. Roberts W, Peters JR, Adams ZW, Lynam DR, Milich R. Identifying the facets of impulsivity that explain the relation between ADHD symptoms and substance use in a nonclinical sample. *Addictive behaviors*. 2014;39(8):1272-1277. doi:10.1016/j.addbeh.2014.04.005

219. Bidwell LC, Henry EA, Willcutt EG, Kinnear MK, Ito TA. Childhood and current ADHD symptom dimensions are associated with more severe cannabis outcomes in college students. *Drug and alcohol dependence*. 2013;135:88-94. doi:10.1016/j.drugalcdep.2013.11.013
220. Ofrat S, Krueger RF, Eaton NR, et al. Nonmedical Prescription Drug Use Comorbidity: Developing a Cohesive Risk Model. *Journal of Psychopathology and Behavioral Assessment*. 2014;36(3):371-379. doi:10.1007/s10862-014-9409-2
221. Herman AM, Duka T. Facets of impulsivity and alcohol use: What role do emotions play? *Neuroscience and biobehavioral reviews*. 2019;106:202-216. doi:10.1016/j.neubiorev.2018.08.011
222. Lord S, Marsch L. Emerging trends and innovations in the identification and management of drug use among adolescents and young adults. *Adolescent medicine: state of the art reviews*. 2011;22(3):649-xiv. https://pubmed.ncbi.nlm.nih.gov/22423469
223. Hulvershorn LA, Quinn PD, Scott EL. Treatment of Adolescent Substance Use Disorders and Co-Occurring Internalizing Disorders: A Critical Review and Proposed Model. *Current drug abuse reviews*. 2015;8(1):41-49. doi:10.2174/1874473708666150514102745
224. Silveira ML, Conway KP, Green VR, et al. Longitudinal associations between youth tobacco and substance use in waves 1 and 2 of the Population Assessment of Tobacco and Health (PATH) Study. *Drug and alcohol dependence*. 2018;191:25-36. doi:10.1016/j.drugalcdep.2018.06.018
225. Merrin GJ, Leadbeater B. Do Classes of Polysubstance Use in Adolescence Differentiate Growth in Substances Used in the Transition to Young Adulthood? *Substance use & misuse*. 2018;53(13):2112-2124. doi:10.1080/10826084.2018.1455702
226. Lanza ST, Patrick ME, Maggs JL. Latent Transition Analysis: Benefits of a Latent Variable Approach to Modeling Transitions in Substance Use. *Journal of drug issues*. 2010;40(1):93-120. doi:10.1177/002204261004000106
227. Choi HJ, Lu Y, Schulte M, Temple JR. Adolescent substance use: Latent class and transition analysis. *Addictive behaviors*. 2018;77:160-165. doi:10.1016/j.addbeh.2017.09.022
228. Dyer ML, Easey KE, Heron J, Hickman M, Munafò MR. Associations of child and adolescent anxiety with later alcohol use and disorders: a systematic review and meta-analysis of prospective cohort studies. *Addiction (Abingdon, England)*. 2019;114(6):968-982. doi:10.1111/add.14575
229. Armenta BE, Sittner Hartshorn KJ, Whitbeck LB, Crawford DM, Hoyt DR. A longitudinal examination of the measurement properties and predictive utility of the Center for Epidemiologic Studies Depression Scale among North American Indigenous adolescents. *Psychological assessment*. 2014;26(4):1347-1355. doi:10.1037/a0037608
230. McMillen RC, Gottlieb MA, Shaefer RMW, Winickoff JP, Klein JD. Trends in Electronic Cigarette Use Among U.S. Adults: Use is Increasing in Both Smokers and Nonsmokers. *Nicotine & tobacco research*. 2015;17(10):1195-1202. doi:10.1093/ntr/ntu213
231. Onukwugha E, Onukwugha E, Bergtold J, Bergtold J, Jain R, Jain R. A Primer on Marginal Effects—Part I: Theory and Formulae. *PharmacoEconomics*. 2015;33(1):25-30. doi:10.1007/s40273-014-0210-6
232. Rapier R, McKernan S, Stauffer CS. An inverse relationship between perceived social support and substance use frequency in socially stigmatized populations. *Addictive behaviors reports*. 2019;10:100188. doi:10.1016/j.abrep.2019.100188

233. Stone AL, Becker LG, Huber AM, Catalano RF. Review of risk and protective factors of substance use and problem use in emerging adulthood. *Addictive behaviors*. 2012;37(7):747-775. doi:10.1016/j.addbeh.2012.02.014
234. Krieger J, Allen C, Cheadle A, et al. Using Community-Based Participatory Research to Address Social Determinants of Health: Lessons Learned from Seattle Partners for Healthy Communities. *Health Education & Behavior*. 2002;29(3):361-382. doi:10.1177/109019810202900307
235. Michael YL, Farquhar SA, Wiggins N, Green MK. Findings from a Community-based Participatory Prevention Research Intervention Designed to Increase Social Capital in Latino and African American Communities. *Journal of Immigrant and Minority Health*. 2008;10(3):281-289. doi:10.1007/s10903-007-9078-2
236. Cacioppo JT, Cacioppo S. Social Relationships and Health: The Toxic Effects of Perceived Social Isolation. *Social and Personality Psychology Compass*. 2014;8(2):58-72. doi:https://doi.org/10.1111/spc3.12087
237. Hawkley LC, Cacioppo JT. Loneliness matters: a theoretical and empirical review of consequences and mechanisms. *Annals of behavioral medicine : a publication of the Society of Behavioral Medicine*. 2010;40(2):218-227. doi:10.1007/s12160-010-9210-8
238. Reblin M, Uchino BN. Social and emotional support and its implication for health. *Current opinion in psychiatry*. 2008;21(2):201-205. doi:10.1097/YCO.0b013e3282f3ad89
239. Ozbay F, Johnson DC, Dimoulas E, Morgan CA, Charney D, Southwick S. Social support and resilience to stress: from neurobiology to clinical practice. *Psychiatry (Edgmont (Pa : Township))*. 2007;4(5):35-40. https://pubmed.ncbi.nlm.nih.gov/20806028
240. Husky MM, Mazure CM, Paliwal P, McKee SA. Gender differences in the comorbidity of smoking behavior and major depression. *Drug and alcohol dependence*. 2008;93(1-2):176-179. doi:10.1016/j.drugalcdep.2007.07.015
241. Needham BL. Gender differences in trajectories of depressive symptomatology and substance use during the transition from adolescence to young adulthood. *Social science & medicine (1982)*. 2007;65(6):1166-1179. doi:10.1016/j.socscimed.2007.04.037
242. He H, Tang J, Liu T, Hao W, Liao Y. Gender Differences in Sleep Problems Among Drug Users. *Frontiers in psychiatry*. 2020;11:808. doi:10.3389/fpsyt.2020.00808
243. Suh S, Cho N, Zhang J. Sex Differences in Insomnia: from Epidemiology and Etiology to Intervention. *Current psychiatry reports*. 2018;20(9):1-12. doi:10.1007/s11920-018-0940-9
244. Epskamp S, Rhemtulla M, Borsboom D. Generalized Network Psychometrics: Combining Network and Latent Variable Models. *Psychometrika*. 2017;82(4):904-927. doi:10.1007/s11336-017-9557-x
245. Epskamp S. Psychometric network models from time-series and panel data. *Psychometrika*. 2020;85(1):206-231. doi:10.1007/s11336-020-09697-3
246. Acion L, Kelmansky D, van der Laan M, Sahker E, Jones D, Arndt S. Use of a machine learning framework to predict substance use disorder treatment success. *PloS one*. 2017;12(4):e0175383-e0175383. doi:10.1371/journal.pone.0175383
247. Jing Y, Hu Z, Fan P, et al. Analysis of substance use and its outcomes by machine learning I. Childhood evaluation of liability to substance use disorder. *Drug and alcohol dependence*. 2020;206:107605. doi:10.1016/j.drugalcdep.2019.107605
248. Lo-Ciganic W-H, Huang JL, Zhang HH, et al. Evaluation of Machine-Learning Algorithms for Predicting Opioid Overdose Risk Among Medicare Beneficiaries With

Opioid Prescriptions. *JAMA Network Open*. 2019;2(3):e190968-e190968. doi:10.1001/jamanetworkopen.2019.0968

249. O'Connor M, Hanlon A, Mauer E, et al. Identifying distinct risk profiles to predict adverse events among community-dwelling older adults. *Geriatric nursing (New York, NY)*. 2017;38(6):510-519. doi:10.1016/j.gerinurse.2017.03.013
250. Chapman BP, Lin F, Roy S, Benedict RHB, Lyness JM. Health risk prediction models incorporating personality data: Motivation, challenges, and illustration. *Personality disorders*. 2019;10(1):46-58. doi:10.1037/per0000300
251. Hogan BE, Linden W, Najarian B. Social support interventions: Do they work? *Clinical Psychology Review*. 2002;22(3):381-440. doi:https://doi.org/10.1016/S0272-7358(01)00102-7
252. Humphreys K, Mankowski ES, Moos RH, Finney JW. Do enhanced friendship networks and active coping mediate the effect of self-help groups on substance abuse? *Annals of behavioral medicine*. 1999;21(1):54-60. doi:10.1007/BF02895034
253. KELLER DS, GALANTER M. Technology transfer of network therapy to community-based addictions counselors. *Journal of substance abuse treatment*. 1999;16(2):183-189. doi:10.1016/S0740-5472(98)00044-0

SAS File name: LCA W1 4132021

```
*In the ICPSR_36498 folder, select DS1001 and open the data file (36498-1001-Data) which is
a SAS Cport Transport file. Once this is open, formats are in, and can begin data management;

libname LCA "U:\CourtneyBlondino\PhD Epidemiology\LCA\Data Management";

*Recoding Missings;
data LCA.W1;
set da36498p1001;


        *Current User Cigarette;
        *R01_AC1002: Ever smoked a cigarette, even one or two puffs;
        *R01_AC1005: Number of cigarettes smoked in your entire life;
        *R01_AC1003: Now smoke cigarettes;

        if R01_AC1002 = 1 AND R01_AC1005=6 AND R01_AC1003 in (1 2) then acur_cig = 1;
        else if R01_AC1002 = 2 OR R01_AC1003=3 OR (R01_AC1003 in (1,2,.) AND R01_AC1005 in
(1,2,3,4,5)) then acur_cig=0;
        else if R01_AC1002 = . OR R01_AC1003=. OR R01_AC1005=. then acur_cig = .;

        *Current E-cigarette user;
        *R01_AE1002: Ever used an e-cigarette, even one or two times;
        *R01_AE1100: Ever smoked e-cigarettes fairly regularly;
        *R01_AE1003: Now use e-cigarettes;

        if R01_AE1002 = 1 AND R01_AE1100=1 AND R01_AE1003 in (1,2) then acur_ecig = 1;
        else if R01_AE1001=2 OR R01_AE1002 = 2 OR R01_AE1003 = 3 OR (R01_AE1003 in (1,2,.) AND
R01_AE1100 = 2) then acur_ecig=0;
        else if R01_AE1002 = . OR R01_AE1001=. OR R01_AE1003 = . OR R01_AE1100 = . then acur_ecig
= .;

        ***NOT USING FOR LCA************
        ********************************
        *Current Traditional cigar user;

        *if R01_AG9003 = 1 AND R01_AG1100TC=1 AND R01_AG1003TC in (1,2) then acur_cigr = 1;
        *else if R01_AG1001=2 OR R01_AG9002_01 = 2 OR R01_AG9003= 2 OR R01_AG1003TC= 3 OR
(R01_AG1003TC in
        *(1,2,.) AND R01_AG1100TC = 2) THEN acur_cigr = 0;
        *ELSE IF R01_AG1001 = . OR R01_AG9003 = . OR R01_AG1100TC = . OR R01_AG1003TC = . OR
R01_AG9002_01 = . THEN
        *acur_cigr = .;

        *Current Cigarillo user;

        *IF R01_AG9004=1 AND (R01_AG9009_01=1 OR R01_AG9009_03=1) AND R01_AG1100CG = 1 AND
R01_AG1003CG in
        (1, 2) THEN acur_cigrlo= 1;
        *ELSE IF R01_AG9004= 2 OR R01_AG1001=2 OR R01_AG9002_02 = 2 OR R01_AG1003CG=3 OR
R01_AG1100CG=2 OR (R01_AG9009_01=2 AND R01_AG9009_03=2) OR ((R01_AG9009_01=1
        *OR R01_AG9009_03=1) AND R01_AG1100CG= 2 AND R01_AG1003CG=.) OR ((R01_AG9009_01=1 OR
        *R01_AG9009_03=1) AND R01_AG1100CG=. AND R01_AG1003CG= 3) THEN acur_cigrlo= 0;
        *ELSE IF R01_AG1001 = . OR R01_AG9004 = . OR R01_AG9009_03 = . OR R01_AG9009_01 = . OR
        R01_AG1100CG = . OR R01_AG1003CG = . OR R01_AG9002_02 = . THEN acur_cigrlo = .;

        *Current Filtered Cigar user;

        *IF R01_AG9004=1 AND R01_AG9009_02=1 AND R01_AG1100FC = 1 AND R01_AG1003FC in (1, 2) THEN
        acur_filcigr= 1;
        *ELSE IF R01_AG9004= 2 OR R01_AG1001=2 OR R01_AG9002_02 = 2 OR R01_AG1003FC=3 OR
R01_AG1100FC=2 OR R01_AG9009_02=2
        OR (R01_AG9009_02=1 AND R01_AG1100FC= 2 AND R01_AG1003FC=.) OR (R01_AG9009_02=1 AND
R01_AG1100FC=. AND
        R01_AG1003FC= 3) THEN acur_filcigr=0;
        *ELSE IF R01_AG9004 = . OR R01_AG9009_02 =. OR R01_AG1100FC = . OR R01_AG1003FC = . OR
R01_AG1001 = . OR R01_AG9002_02 = . THEN
        *acur_filcigr = .;

        *Current Use Any Cigar/Cigarillo;
```

```
*IF (acur_cigr = 1 OR acur_cigrlo = 1 OR acur_filcigr = 1) THEN acur_fullcigr = 1;
*ELSE IF (acur_cigr = 0 AND acur_cigrlo = 0 AND acur_filcigr= 0) THEN acur_fullcigr = 0;
*ELSE IF acur_cigr = . OR acur_cigrlo = . OR acur_filcigr = . THEN acur_fullcigr = .;

*Current Pipe user;

*IF R01_AP1002 = 1 AND R01_AP1100=1 AND R01_AP1003 in (1,2) THEN acur_pipe= 1;
*ELSE IF R01_AP1001=2 OR R01_AP1002= 2 OR R01_AP1003= 3 OR (R01_AP1003 in (1,2,.) AND
R01_AP1100 = 2)
THEN acur_pipe=0;
*ELSE IF R01_AP1001 = . OR R01_AP1002 = . OR R01_AP1003 = . OR R01_AP1100 = . THEN
acur_pipe= .;

*Current Hookah User;

*IF R01_AH1002 = 1 AND R01_AH1100=1 AND R01_AH1003 in (1, 2) THEN acur_hook= 1;
*ELSE IF R01_AH1001=2 OR R01_AH1002= 2 OR R01_AH1003= 3 OR (R01_AH1003 in (1,2,.) AND
R01_AH1100 = 2)
*THEN acur_hook=0;
*ELSE IF R01_AH1002=. OR R01_AH1001=. OR R01_AH1003=. OR R01_AH1100=.
*THEN acur_hook=.;

*Current User Smokeless;

*IF (R01_AS1002_02=1 OR R01_AU1003 in (1,2)) AND R01_AS1100SM = 1 AND R01_AS1003SM in (1,
2) THEN acur_smls= 1;
*ELSE IF R01_AS1001=2 OR R01_AS1002_03=1 OR (R01_AS1002_02=2 AND R01_AU1003 in
(2,3,.)) OR R01_AS1003SM= 3 OR (R01_AS1003SM in (1,2,.) AND R01_AS1100SM = 2) THEN
acur_smls=0;
*ELSE IF R01_AS1002_02 = . OR R01_AU1003 = . OR R01_AS1100SM = . OR
R01_AS1003SM = . OR R01_AS1001 = . THEN acur_smls = .;

*Current User Snus;

*IF R01_AS1002_01=1 AND R01_AU1003 in (2, 3) AND R01_AS1100SU= 1 AND R01_AS1003SU in
(1,2) THEN acur_snus= 1;
*ELSE IF R01_AS1001=2 OR R01_AS1002_03=1 OR (R01_AS1002_01=2 AND R01_AS1002_02=1)
OR (R01_AS1002_01=1 AND R01_AU1003=1) OR (R01_AU1003 in (2,3) AND R01_AS1003SU= 3) OR
(R01_AU1003 in (2,3) AND
R01_AS1003SU in (1,2,.) AND R01_AS1100SU = 2) THEN acur_snus= 0;
*ELSE IF R01_AS1002_01 = . OR R01_AS1002_02 = . OR R01_AS1002_03 = . OR R01_AU1003 = . OR
R01_AS1100SU = . OR R01_AS1003SU =.
*OR R01_AS1001 = . THEN acur_snus=.;

*Current Use Any Smokeless/Snus;

*IF (acur_smls = 1 OR acur_snus = 1) THEN acur_fullsmkls = 1;
*ELSE IF (acur_smls = 0 AND acur_snus = 0) THEN acur_fullsmkls = 0;
*ELSE IF acur_smls = . OR acur_snus = . THEN acur_fullsmkls = .;

*Current User Dissolvable;

*IF R01_AD1002 = 1 AND R01_AD1100=1 AND R01_AD1003 in (1,2) THEN acur_diss= 1;
*ELSE IF R01_AD1001=2 OR R01_AD1002= 2 OR R01_AD1003= 3 OR (R01_AD1003 in (1,2,.) AND
R01_AD1100 = 2) THEN acur_diss=0;
*ELSE IF R01_AD1001 = . OR R01_AD1002 = . OR R01_AD1003 = . OR R01_AD1100 = . THEN
acur_diss = .;

*******************************
*******************************
NEW SUBSTANCES ADDED;

*Current Use Alcohol;
*R01_AX0084 is ever used alcohol
*R01_AX0073 is how long since last used alcohol and 1 is within the past 30 days;
if R01_AX0084 = 1 AND R01_AX0073 = 1 then acur_alc=1;
else if R01_AX0084 = 2 OR R01_AX0073 in (2,3) then acur_alc=0;
else if R01_AX0084= . OR R01_AX0073= . then acur_alc=.;

*Current User Marijuana;
```

```
        *R01_AX0085 is ever used marijuana - look at measures spreadsheet in oral dis proposal
folder;
        *R01_AX0078 is how long since last used marijuana and 1 is within the past 30 days;
        if R01_AX0078 = 1 then acur_marijuana=1;
        else if R01_AX0078 in (-1,2,3) then acur_marijuana=0;
        else if R01_AX0078= . then acur_marijuana=.;

        *Current User Ritalin or Adderall (prescription drugs not prescribed to you);
        *R01_AX0089_01 is ever used ritalin or adderal
        *R01_AX0081_01 is how long since last used ritalin or adderall and 1 is within the past
30 days;
        *if R01_AX0089_01 = 1 AND R01_AX0081_01 = 1 then acur_ritadder=1;
        *else if R01_AX0089_01 = 2 OR R01_AX0081_01 in (2,3) then acur_ritadder=0;
        *else if R01_AX0089_01 = . OR R01_AX0081_01= . then acur_ritadder=.;

        *Current User Painkillers, Sedatives, or Tranquilizers (prescription drugs not prescribed
to you);
        *R01_AX0089_02 is ever used painkillers
        *R01_AX0081_02 is how long since last used painkillers and 1 is within the past 30 days;
        if R01_AX0089_02 = 1 AND R01_AX0081_02 = 1 then acur_painkiller=1;
        else if R01_AX0089_02 = 2 OR R01_AX0081_02 in (2,3) then acur_painkiller=0;
        else if R01_AX0089_02 = . OR R01_AX0081_02= . then acur_painkiller=.;

        *Current User Cocaine or Crack
        *R01_AX0220_01 is ever used cocaine or crack
        *R01_AX0081_03 is how long since last used cocaine or crack and 1 is within the past 30
days;
        *if R01_AX0220_01 = 1 AND R01_AX0081_03 = 1 then acur_cocaine=1;
        *else if R01_AX0220_01 = 2 OR R01_AX0081_03 in (2,3) then acur_cocaine=0;
        *else if R01_AX0220_01 = . OR R01_AX0081_03= . then acur_cocaine=.;

        *Curent User Meth or Speed
        *R01_AX0220_02 is ever used meth or speed
        *R01_AX0081_04 is how long since last used meth or speed and 1 is within the past 30
days;
        *if R01_AX0220_02 = 1 AND R01_AX0081_04 = 1 then acur_meth=1;
        *else if R01_AX0220_02 = 2 OR R01_AX0081_04 in (2,3) then acur_meth=0;
        *else if R01_AX0220_02 = . OR R01_AX0081_04= . then acur_meth=.;


        *Current User Heroin, Inhalents, Solvents, Hallucinogens
        *R01_AX0220_03 is ever used heroin, inhalents, solvents, hallucinogens
        *R01_AX0081_05 is how long since last used heroin.... and 1 is within the past 30 days;
        *if R01_AX0220_03 = 1 AND R01_AX0081_05 = 1 then acur_heroinplus=1;
        *else if R01_AX0220_03 = 2 OR R01_AX0081_05 in (2,3) then acur_heroinplus=0;
        *else if R01_AX0220_03 = . OR R01_AX0081_05 = . then acur_heroinplus=.;


*RACE;
*R01R_A_RACECAT3: DERIVED - Race from the interview (3 levels): 1 = white alone, 2 = black alone,
3 = other;
*R01R_A_HISP: DERIVED - Hispanic origin from the interview (2 levels): 1 = hispanic, 2 = not
hispanic;
NUMRACES = 0 ;
if R01R_A_RACECAT3 = 1 then NUMRACES = NUMRACES + 1 ;
if R01R_A_RACECAT3 = 2 then NUMRACES= NUMRACES + 1 ;
if R01R_A_RACECAT3 = 3 then NUMRACES = NUMRACES + 1 ;
if R01R_A_HISP = 1 then NUMRACES = NUMRACES + 1;
if (NUMRACES = 1 and R01R_A_RACECAT3 = 1 AND R01R_A_HISP=2) then R01R_A_ETHRACECAT7= 1 ; *NH
White;
if (NUMRACES = 1 and R01R_A_RACECAT3 = 2 AND R01R_A_HISP=2) then R01R_A_ETHRACECAT7= 2 ; *NH AA;
if (NUMRACES = 1 and R01R_A_RACECAT3 = 3 AND R01R_A_HISP=2) then R01R_A_ETHRACECAT7= 3 ; *NH
Other;
if (NUMRACES = 1 and R01R_A_HISP=1) then R01R_A_ETHRACECAT7= 4; *Hispanic Only;
if (NUMRACES > 1 and R01R_A_HISP=2) then R01R_A_ETHRACECAT7= 5; *NH Multiracial;
if (NUMRACES > 1 and R01R_A_HISP=1) then R01R_A_ETHRACECAT7= 6; *Hispanic Multiracial;
ELSE IF R01R_A_HISP=. OR R01R_A_RACECAT3 = . THEN R01R_A_ETHRACECAT7=.;

*AGE;
if R01R_A_AGECAT7=1 then age=1; *18-24;
else if R01R_A_AGECAT7=2 then age=2; *25-34;
```

```
else if R01R_A_AGECAT7=3 then age=3; *35-44;
else if R01R_A_AGECAT7=4 then age=4; *45-54;
else if R01R_A_AGECAT7=5 then age=5; *55-64;
else if R01R_A_AGECAT7 in (6 7) then age=6; *65 and older;
else age=.;

*EDUCATION;
if R01R_A_AM0018=1 then education=1; *less than high school;
else if R01R_A_AM0018 in (2 3) then education=2; *GED/high school graduate;
else if R01R_A_AM0018=4 then education=3; *Some college (no degree) or associates degree;
else if R01R_A_AM0018=5 then education=4; *Bachelor's degree;
else if R01R_A_AM0018=6 then education=5; *Advanced degree;
else education=.;

*LIMIT ALL MH VARIABLES TO PAST 30 DAYS;

******INTERNALIZING*********;

*R01_AX0161: Last time you had significant problems with: feeling very trapped, lonely, sad,
blue,
depressed or hopeless about the future;
if R01_AX0161 in (2, 3, 4) then depressed=0;
else if R01_AX0161 in (1) then depressed=1;
else if R01_AX0161 = . then depressed= .;

*R01_AX0162: Last time you had significant problems with: Sleep trouble - such as bad
dreams, sleeping restlessly or falling asleep during the day;
if R01_AX0162 in (2, 3, 4) then sleeping=0;
else if R01_AX0162 in (1) then sleeping=1;
else if R01_AX0162 = . then sleeping=.;

*R01_AX0163: Last time you had significant problems with: feeling very anxious, nervous, tense,
panicked or like something bad was going to happen;
if R01_AX0163 in (2, 3, 4) then anxious=0;
else if R01_AX0163 in (1) then anxious=1;
else if R01_AX0163 = . then anxious=.;

*R01_AX0164: Last time you had significant problems with: Becoming very distressed and
upset when something reminded you of the past;
if R01_AX0164 in (2, 3, 4) then ptsd=0;
else if R01_AX0164 in (1) then ptsd=1;
else if R01_AX0164 = . then ptsd=.;

******EXTERNALIZING*********;

*R01_AX0165: Last time you lied or conned to get something;
if R01_AX0165 in (2, 3, 4) then lied=0;
else if R01_AX0165 in (1) then lied=1;
else if R01_AX0165 = . then lied=.;

*R01_AX0166: Last time you did the following two or more times:
had a hard time paying attention at school, work or home;
if R01_AX0166 in (2, 3, 4) then attention=0;
else if R01_AX0166 in (1) then attention=1;
else if R01_AX0166 = . then attention=.;

*R01_AX0167: Last time you did the following two or more times: had a hard
time listening to instructions at school, work or home;
if R01_AX0167 in (2, 3, 4) then listening=0;
else if R01_AX0167 in (1) then listening=1;
else if R01_AX0167 = . then listening= .;

*R01_AX0168: Last time you did the following two or more times:
were a bully or threatened other people;
if R01_AX0168 in (2, 3, 4) then bully=0;
else if R01_AX0168 in (1) then bully=1;
else if R01_AX0168 = . then bully= .;

*R01_AX0169: Last time you did the following two or more times:
started physical fights with other people;
if R01_AX0169 in (2, 3, 4) then fights=0;
```

```
else if R01_AX0169 in (1) then fights=1;
else if R01_AX0169 = . then fights= .;

*R01_AX0250: Last time felt restless/need to climb on things;
if R01_AX0250 in (2, 3, 4) then restless=0;
if R01_AX0250 in (1) then restless=1;
else if R01_AX0250 = . then restless=.;

*R01_AX0251: Last time gave answers before question was finished;
if R01_AX0251 in (2, 3, 4) then answered=0;
if R01_AX0251 in (1) then answered=1;
else if R01_AX0251 = . then answered=.;

******SUBSTANCE USE DISORDERS**********;

*R01_AX0170: Last time used alcohol/drugs weekly or more often;
if R01_AX0170 in (2, 3, 4) then weeklyuse=0;
if R01_AX0170 in (1) then weeklyuse=1;
else if R01_AX0170 = . then weeklyuse=.;

*R01_AX0171: Last time spent a lot of time getting alcohol/drugs;
if R01_AX0171 in (2, 3, 4) then timegetting=0;
if R01_AX0171 in (1) then timegetting=1;
else if R01_AX0171 = . then timegetting=.;

*R01_AX0193: Last time you spent a lot of time using or recovering from alcohol or other drugs;
if R01_AX0193 in (2, 3, 4) then timeusing=0;
if R01_AX0193 in (1) then timeusing=1;
else if R01_AX0193 = . then timeusing=.;

*R01_AX0172: Last time that you kept using alcohol or other drugs even though it was causing
social problems, leading to fights, or getting you into trouble with other people;
if R01_AX0172 in (2, 3, 4) then socialprob=0;
if R01_AX0172 in (1) then socialprob=1;
else if R01_AX0172 = . then socialprob=.;

*R01_AX0173: Last time that your use of alcohol or other drugs reduced your involvement in
activities at work, school, home or social events;
if R01_AX0173 in (2, 3, 4) then reducedact=0;
if R01_AX0173 in (1) then reducedact=1;
else if R01_AX0173 = . then reducedact=.;

*R01_AX0174: Last time that you had withdrawal problems such as shaky hands, throwing up,
having trouble sitting still or sleeping;
if R01_AX0174 in (2, 3, 4) then withdraw=0;
if R01_AX0174 in (1) then withdraw=1;
else if R01_AX0174 = . then withdraw=.;

*R01_AX0194: Use of alcohol/drugs to avoid withdrawal;
if R01_AX0194 in (2, 3, 4) then usetoavoid=0;
if R01_AX0194 in (1) then usetoavoid=1;
else if R01_AX0194 = . then usetoavoid=.;

*ALL PAST 30 DAY;
sud_score = sum(weeklyuse, timegetting, timeusing, socialprob, reducedact, withdraw, usetoavoid);

*OLD;
*SUD is 3 levels- no/low, moderate, and high;
*if sud_score in (0,1) then sud=0;
*if sud_score in (2,3) then sud=1;
*if sud_score in (4,5,6,7) then sud=2;
*if sud_score = . then sud=.;

*NEW = 1/16/20;
*SUD is 3 levels- no/low, moderate, and high;
if sud_score in (0) then sud=0;
if sud_score in (1,2) then sud=1;
if sud_score in (3,4,5,6,7) then sud=2;
if sud_score = . then sud=.;

*Dichotomize by 0 = no/low, 1 = moderate/high;
```

```
*if sud in (0) then sudbin1=0;
*if sud in (1, 2) then sudbin1=1;
*if sud = . then sudbin1 = .;

*Dichotomize by 0 = no/low/moderate, 1 = high;
*if sud in (0,1) then sudbin2=0;
*if sud in (2) then sudbin2=1;
*if sud = . then sudbin2 = .;


*R01_AX0170 + R01_AX0171 + R01_AX0172 + R01_AX0173 + R01_AX0174 + R01_AX0193 + R01_AX0194;
*Last time used alcohol/drugs weekly or more often;
*if R01_AX0170 = 1 then sudcon1=3;
*if R01_AX0170 = 2 then sudcon1=2;
*if R01_AX0170 = 3 then sudcon1=1;
*if R01_AX0170 = 4 then sudcon1=0;
*else if R01_AX0170 = . then sudcon1=.;
*Last time spent a lot of time getting alcohol/drugs;
*if R01_AX0171 = 1 then sudcon2=3;
*if R01_AX0171 = 2 then sudcon2=2;
*if R01_AX0171 = 3 then sudcon2=1;
*if R01_AX0171 = 4 then sudcon2=0;
*else if R01_AX0171 = . then sudcon2=.;
*Last time spent a lot of time using or recovering;
*if R01_AX0172 = 1 then sudcon3=3;
*if R01_AX0172 = 2 then sudcon3=2;
*if R01_AX0172 = 3 then sudcon3=1;
*if R01_AX0172 = 4 then sudcon3=0;
*else if R01_AX0172 = . then sudcon3=.;
*Last time alcohol/drugs causing social problems;
*if R01_AX0173 = 1 then sudcon4=3;
*if R01_AX0173 = 2 then sudcon4=2;
*if R01_AX0173 = 3 then sudcon4=1;
*if R01_AX0173 = 4 then sudcon4=0;
*else if R01_AX0173 = . then sudcon4=.;
*Reduced involvement with activities;
*if R01_AX0174 = 1 then sudcon5=3;
*if R01_AX0174 = 2 then sudcon5=2;
*if R01_AX0174 = 3 then sudcon5=1;
*if R01_AX0174 = 4 then sudcon5=0;
*else if R01_AX0174 = . then sudcon5=.;
*Withdrawal problems;
*if R01_AX0193 = 1 then sudcon6=3;
*if R01_AX0193 = 2 then sudcon6=2;
*if R01_AX0193 = 3 then sudcon6=1;
*if R01_AX0193 = 4 then sudcon6=0;
*else if R01_AX0193 = . then sudcon6=.;
*Use of alcohol/drugs to avoid withdrawal;
*if R01_AX0194 = 1 then sudcon7=3;
*if R01_AX0194 = 2 then sudcon7=2;
*if R01_AX0194 = 3 then sudcon7=1;
*if R01_AX0194 = 4 then sudcon7=0;
*else if R01_AX0194 = . then sudcon7=.;

*sudconscore = sum(sudcon1, sudcon2, sudcon3, sudcon4, sudcon5, sudcon6, sudcon7);


*****DUMMY CODING FOR THE COVARIATES******;

IF R01R_A_SEX=1 THEN SEXMALE_1=1;
ELSE SEXMALE_1=0;

IF R01R_A_SEX=2 THEN SEXFEMALE_2=1;
ELSE SEXFEMALE_2=0;

IF age=1 THEN AGE1824_1=1;
ELSE AGE1824_1=0;

IF age=2 THEN AGE2534_2=1;
ELSE AGE2534_2=0;
```

```
IF age=3 THEN AGE3544_3=1;
ELSE AGE3544_3=0;

IF age=4 THEN AGE4554_4=1;
ELSE AGE4554_4=0;

IF age=5 THEN AGE5564_5=1;
ELSE AGE5564_5=0;

IF age=6 THEN AGE65_6=1;
ELSE AGE65_6=0;

IF R01R_A_ETHRACECAT7=1 THEN RACEWH_1=1;
ELSE RACEWH_1=0;

IF R01R_A_ETHRACECAT7=2 THEN RACEBL_2=1;
ELSE RACEBL_2=0;

IF R01R_A_ETHRACECAT7=3 THEN RACEOT_3=1;
ELSE RACEOT_3=0;

IF R01R_A_ETHRACECAT7=6 THEN RACEHI_6=1;
ELSE RACEHI_6=0;

IF education=1 THEN EDU_1=1;
ELSE EDU_1=0;

IF education=2 THEN EDU_2=1;
ELSE EDU_2=0;

IF education=3 THEN EDU_3=1;
ELSE EDU_3=0;

IF education=4 THEN EDU_4=1;
ELSE EDU_4=0;

IF education=5 THEN EDU_5=1;
ELSE EDU_5=0;

IF R01R_A_AM0030=1 THEN INC_1=1;
ELSE INC_1=0;

IF R01R_A_AM0030=2 THEN INC_2=1;
ELSE INC_2=0;

IF R01R_A_AM0030=3 THEN INC_3=1;
ELSE INC_3=0;

IF R01R_A_AM0030=4 THEN INC_4=1;
ELSE INC_4=0;

IF R01R_A_AM0030=5 THEN INC_5=1;
ELSE INC_5=0;

*extremely satisfied =1;
IF R01_AX0092=1 THEN SOC_1=1;
ELSE SOC_1=0;

IF R01_AX0092=2 THEN SOC_2=1;
ELSE SOC_2=0;

IF R01_AX0092=3 THEN SOC_3=1;
ELSE SOC_3=0;

IF R01_AX0092=4 THEN SOC_4=1;
ELSE SOC_4=0;

*not at all satisfied =5;
IF R01_AX0092=5 THEN SOC_5=1;
ELSE SOC_5=0;
```

```
        array change _numeric_;
        do over change;
        if change=-97777 then change=.;
        else if change=-99999 then change=.;
        else if change=-99988 then change=.;
        else if change=-99977 then change=.;
        else if change=-99955 then change=.;
        else if change=-99911 then change=.;
        else if change=-9 then change=.;
        else if change=-8 then change=.;
        else if change=-7 then change=.;
        else if change=-1 then change=.;
        else if change=-5 then change=.;
        end;

run;


*******************NOW CHECK********************;


*check dummies;
proc freq data=lca.w1;
table   R01R_A_SEX*SEXMALE_1
               R01R_A_SEX*SEXFEMALE_2
               age*AGE1824_1
               age*AGE2534_2
               age*AGE3544_3
               age*AGE4554_4
               age*AGE5564_5
               age*AGE65_6
               R01R_A_ETHRACECAT7*RACEWH_1
               R01R_A_ETHRACECAT7*RACEBL_2
               R01R_A_ETHRACECAT7*RACEOT_3
               R01R_A_ETHRACECAT7*RACEHI_6
               education*EDU_1
               education*EDU_2
               education*EDU_3
               education*EDU_4
               education*EDU_5
               R01R_A_AM0030*INC_1
               R01R_A_AM0030*INC_2
               R01R_A_AM0030*INC_3
               R01R_A_AM0030*INC_4
               R01R_A_AM0030*INC_5
               R01_AX0092*SOC_1
               R01_AX0092*SOC_2
               R01_AX0092*SOC_3
               R01_AX0092*SOC_4
               R01_AX0092*SOC_5;
run;

*check cig and ecig;
proc freq data=lca.w1;
table R01R_A_CUR_ESTD_CIGS*acur_cig;
table R01R_A_CUR_ESTD_ECIG*acur_ecig;
run;

proc freq data=lca.w1;
table R01_AX0078;
run;

***********************************************
*JULY 1 2020 - MAKING CC AND EC EXCLUSIVE VARS;
data lca.w1july20;
set lca.w1;

*first do multinomial - 4 levels;
```

```
if acur_cig=0 and acur_ecig=0 then acur_dual=0;
else if acur_cig=1 and acur_ecig=0 then acur_dual=1;
else if acur_cig=0 and acur_ecig=1 then acur_dual=2;
else if acur_cig=1 and acur_ecig=1 then acur_dual=3;
else acur_dual=.;
run;

*check;
proc freq data=lca.w1july20;
table acur_cig*acur_dual;
table acur_ecig*acur_dual;
run;

*then do the dummies;
data lca.w1july20b;
set lca.w1july20;
if acur_dual = 1 then acur_cignew=1;
else acur_cignew=0;
if acur_dual = 2 then acur_ecignew=1;
else acur_ecignew=0;
if acur_dual = 3 then acur_dualnew=1;
else acur_dualnew=0;
run;

*check;
proc freq data=lca.w1july20b;
table acur_cignew*acur_dual;
table acur_ecignew*acur_dual;
table acur_dualnew*acur_dual;
run;

*confirm marijuana is good;
proc freq data=lca.w1july20b;
table acur_marijuana;
run;



*check all substances;
proc freq data=lca.w1;
table acur_cig acur_ecig acur_alc acur_marijuana acur_painkiller;
run;


*check sud;
proc freq data=lca.w1;
table sud_score*sud;
run;

*check int/ext/sud;
proc freq data=lca.w1;
table acur_cig acur_ecig acur_alc acur_marijuana acur_painkiller
         R01R_A_SEX age R01R_A_ETHRACECAT7 education R01R_A_AM0030  R01_AX0092
      depressed sleeping anxious ptsd
         lied attention listening bully fights restless answered
         weeklyuse timegetting timeusing socialprob reducedact withdraw usetoavoid
         sud;
run;



*Identify all variable want to keep;

*Now limit to the main variables that we want to keep;
data LCA.W1mplus;
set LCA.W1 (keep = caseid personid R01_A_PWGT
                              acur_cig acur_ecig acur_alc acur_marijuana acur_painkiller
                               R01R_A_SEX age R01R_A_ETHRACECAT7 education R01R_A_AM0030
R01_AX0092
                               SEXMALE_1 SEXFEMALE_2
                           AGE1824_1 AGE2534_2 AGE3544_3 AGE4554_4 AGE5564_5 AGE65_6
```

```
                              RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
                              EDU_1 EDU_2 EDU_3 EDU_4 EDU_5
                              INC_1 INC_2 INC_3 INC_4 INC_5
                              SOC_1 SOC_2 SOC_3 SOC_4 SOC_5
                              depressed sleeping anxious ptsd
                              lied attention listening bully fights restless answered
                              weeklyuse timegetting timeusing socialprob reducedact withdraw
usetoavoid
                              sud );
*rename missings;
array change _numeric_;
do over change;
if change =. then change = -99999;
end;
run;

*Check frequencies;
proc contents data=LCA.W1mplus;
run;

*Maybe later- add weights back in;
proc surveyfreq data=LCA.W1 varmethod=BRR (fay=0.3);
table
acur_cig acur_ecig acur_alc acur_marijuana acur_painkiller
R01R_A_SEX age R01R_A_ETHRACECAT7 education R01R_A_AM0030
R01_AX0092
depressed sleeping anxious ptsd
lied attention listening bully fights restless answered
sud
/row chisq(secondorder);
weight R01_A_PWGT;
        repweights R01_A_PWGT1 - R01_A_PWGT100;
        run;
        quit;
run;

*So far so good, let's pull this datset into MPlus and try LCA;



proc freq data=LCA.W1;
table acur_painkiller*acur_cig/chisq oddsratio plcorr;
run;

proc freq data=LCA.W1;
table acur_painkiller*sleeping/chisq oddsratio plcorr;
run;


proc freq data=LCA.W1;
table acur_painkiller*attention/chisq oddsratio plcorr;
run;


proc freq data=LCA.W1;
table acur_painkiller*sleeping*attention/chisq oddsratio plcorr;
run;


proc print data=LCA.W1;
var SOC_5 depressed;
run;

proc freq data=LCA.W1;
table SOC_5*depressed/ chisq oddsratio plcorr;
run;


*current use - conventional cigarette prevalence;
proc surveyfreq data=LCA.W1 varmethod=BRR (fay=0.3);
table
```

```
acur_cig ;
weight R01_A_PWGT;
        repweights R01_A_PWGT1 - R01_A_PWGT100;
        run;
        quit;
run;




*July 2 2020 re-run with new exclusive CC, exclusive EC, and dual variables;
data LCA.W1mplusJuly2020;
set LCA.w1july20b (keep = caseid personid R01_A_PWGT
                                acur_cignew acur_ecignew acur_dualnew
                                acur_alc acur_marijuana acur_painkiller
                                R01R_A_SEX age R01R_A_ETHRACECAT7 education R01R_A_AM0030
R01_AX0092
                                SEXMALE_1 SEXFEMALE_2
                            AGE1824_1 AGE2534_2 AGE3544_3 AGE4554_4 AGE5564_5 AGE65_6
                                RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
                                EDU_1 EDU_2 EDU_3 EDU_4 EDU_5
                                INC_1 INC_2 INC_3 INC_4 INC_5
                                SOC_1 SOC_2 SOC_3 SOC_4 SOC_5
                                depressed sleeping anxious ptsd
                                lied attention listening bully fights restless answered
                                weeklyuse timegetting timeusing socialprob reducedact withdraw
usetoavoid
                                sud );
*rename missings;
array change _numeric_;
do over change;
if change =. then change = -99999;
end;
run;


proc freq data=LCA.W1mplusJuly2020;
table acur_marijuana;
run;


*New summary stats;


data LCA.W1mplusJuly2020weights;
set LCA.w1july20b (keep = caseid personid R01_A_PWGT
                                acur_cignew acur_ecignew acur_dualnew
                                acur_alc acur_marijuana acur_painkiller
                                R01R_A_SEX age R01R_A_ETHRACECAT7 education R01R_A_AM0030
R01_AX0092
                                SEXMALE_1 SEXFEMALE_2
                            AGE1824_1 AGE2534_2 AGE3544_3 AGE4554_4 AGE5564_5 AGE65_6
                                RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
                                EDU_1 EDU_2 EDU_3 EDU_4 EDU_5
                                INC_1 INC_2 INC_3 INC_4 INC_5
                                SOC_1 SOC_2 SOC_3 SOC_4 SOC_5
                                depressed sleeping anxious ptsd
                                lied attention listening bully fights restless answered
                                weeklyuse timegetting timeusing socialprob reducedact withdraw
usetoavoid
                                sud
                            R01_A_PWGT1 - R01_A_PWGT100);
                                run;
*rename missings;
*array change _numeric_;
*do over change;
*if change =. then change = -99999;
*end;
*run;
```

```
proc freq data=LCA.W1mplusJuly2020weights;
table acur_marijuana;
run;

proc surveyfreq data= LCA.W1mplusJuly2020weights varmethod=BRR (fay=0.3);
table acur_marijuana /row chisq(secondorder);
weight R01_A_PWGT;
       repweights R01_A_PWGT1 - R01_A_PWGT100;
       run;
       quit;
run;

proc surveyfreq data= LCA.W1mplusJuly2020weights varmethod=BRR (fay=0.3);
table acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana acur_painkiller
R01R_A_SEX age R01R_A_ETHRACECAT7 education R01R_A_AM0030
R01_AX0092
depressed sleeping anxious ptsd
lied attention listening bully fights restless answered
sud
/row chisq(secondorder);
weight R01_A_PWGT;
       repweights R01_A_PWGT1 - R01_A_PWGT100;
       run;
       quit;
run;



proc surveyfreq data= LCA.W1 varmethod=BRR (fay=0.3);
table acur_marijuana
/row chisq(secondorder);
weight R01_A_PWGT;
       repweights R01_A_PWGT1 - R01_A_PWGT100;
       run;
       quit;
run;

*confirm marijuana is good;
proc surveyfreq data=lca.w1july20b varmethod=BRR (fay=0.3);
table acur_marijuana
/row chisq(secondorder);
weight R01_A_PWGT;
       repweights R01_A_PWGT1 - R01_A_PWGT100;
       run;
       quit;
run;



proc mi data=lca.w1july20b seed=14832 nimpute=0 simple;
var acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana acur_painkiller
R01R_A_SEX age R01R_A_ETHRACECAT7 education R01R_A_AM0030
R01_AX0092
depressed sleeping anxious ptsd
lied attention listening bully fights restless answered
sud ;
run;



****4/13/2021;

data LCA.W1mplusJuly2020weights4132021;
set LCA.w1july20b (keep = caseid personid R01_A_PWGT
                               acur_cignew acur_ecignew acur_dualnew
                               acur_alc acur_marijuana acur_painkiller
                               R01R_A_SEX age R01R_A_ETHRACECAT7 education R01R_A_AM0030
R01_AX0092
                               SEXMALE_1 SEXFEMALE_2
                            AGE1824_1 AGE2534_2 AGE3544_3 AGE4554_4 AGE5564_5 AGE65_6
```

```
                              RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
                              EDU_1 EDU_2 EDU_3 EDU_4 EDU_5
                              INC_1 INC_2 INC_3 INC_4 INC_5
                              SOC_1 SOC_2 SOC_3 SOC_4 SOC_5
                              depressed sleeping anxious ptsd
                              lied attention listening bully fights restless answered
                              weeklyuse timegetting timeusing socialprob reducedact withdraw
usetoavoid
                              sud
                           R01_A_PWGT1 - R01_A_PWGT100);
*rename missings;
array change _numeric_;
do over change;
if change =. then change = -99999;
end;
run;

proc surveyfreq data=lca.W1mplusJuly2020weights4132021 varmethod=BRR (fay=0.3);
table  acur_cignew acur_ecignew acur_dualnew
                              acur_alc acur_marijuana acur_painkiller
                              depressed sleeping anxious ptsd
                              lied attention listening bully fights restless answered
/row chisq(secondorder);
weight R01_A_PWGT;
       repweights R01_A_PWGT1 - R01_A_PWGT100;
       run;
       quit;
run;

*export lca.W1mplusJuly2020weights4132021
1 - run lca in mplus
*need to rerun the summary stats for wave 1 from this data set =
LCA.W1mplusJuly2020weights4132021;!!!
2 - take results import into sas for prediction
3 - take that into network??
;



***Missing vs nonmissing for W1;
proc contents data=lca.w1july20b;
run;

data lca.w1missingtest;
set lca.w1july20b;
if (acur_cignew=.) or (acur_ecignew=.) or (acur_dualnew=.) or (acur_alc=.) or
(acur_marijuana=.) or (acur_painkiller=.) or
(R01R_A_SEX=.)  or (age=.) or (R01R_A_ETHRACECAT7=.) or (education=.) or (R01R_A_AM0030=.) or
(R01_AX0092=.) or
(depressed=.) or (sleeping=.) or (anxious=.) or (ptsd=.) or
(lied=.) or (attention=.) or (listening=.) or (bully=.) or (fights=.) or
(restless=.) or (answered=.) or
(sud=.) then compare=0;
else compare=1;
run;


ods pdf;
proc freq data=lca.w1missingtest;
table compare;
run;
*complete data/analytic sample (compare = 1) = 24039;
*missing (compare = 0) = 8281;

********************************
*compare missing and nonmissing;
*look at column percent;
proc freq data=lca.w1missingtest;
table acur_cignew*compare/chisq;
run;
```

```
*sig different: analytic sample engages in more cig use , chi sq = <.0001;

proc freq data=lca.w1missingtest;
table acur_ecignew*compare/chisq;
run;
*sig different: analytic sample engages in more ecig use , chi sq = 0.002;

proc freq data=lca.w1missingtest;
table acur_dualnew*compare/chisq;
run;
*sig different: analytic sample engages in more dual use , chi sq = <.0001;

proc freq data=lca.w1missingtest;
table acur_alc*compare/chisq;
run;
*sig different: analytic sample engages in more alcohol use , chi sq = <.0001;

proc freq data=lca.w1missingtest;
table acur_marijuana*compare/chisq;
run;
*sig different: analytic sample engages in more, chi sq =<.0001;

proc freq data=lca.w1missingtest;
table acur_painkiller*compare/chisq;
run;
*sig different: analytic sample engages in more, chi sq=<.0001;

*demos;
proc freq data=lca.w1missingtest;
table R01R_A_SEX*compare
      age*compare
      R01R_A_ETHRACECAT7*compare
      education*compare
      R01R_A_AM0030*compare
      R01_AX0092*compare/chisq;
run;
*sig difference sex: more males, less women in analytic sample;
*sig difference by age: more in categories 2, 3, 4 (25-54) in analytic sample;
*sig difference by race: more white, less other cats in analytic sample;
*sig difference by edu: higher edu levels in analytic sample;
*sig difference by income: higher income levels in analytic sample;
*sig difference by social: missing had more extremely and very satisfied;

*internalizing;
proc freq data=lca.w1missingtest;
table depressed*compare
      sleeping*compare
      anxious*compare
      ptsd*compare/chisq;
run;
*sig diff for all: analytic sample has higher endorsement of all 4 symptoms;

*externalizing;
proc freq data=lca.w1missingtest;
table lied*compare
      attention*compare
      listening*compare
      bully*compare
      fights*compare
      restless*compare
      answered*compare/chisq;
run;
*sig diff for all except fights: all others - analytic sample has higher endorsement of the other
6 symptoms;

*sud;
proc freq data=lca.w1missingtest;
table sud*compare/chisq;
run;
*sig diff: analytic sample has higher endorsement of moderate and high sud severity;
ods pdf close;
```

MPLUS File name: WAVE 1 RUN 4132021 4 CLASS

```
TITLE: WAVE 1 MODEL 4-13-2021 with weights added and new tobacco variables : fixing the missing;
   DATA: FILE IS newwave4132021editnoheader.csv;
   VARIABLE: NAMES ARE CASEID PERSONID weight
         acur_cignew acur_ecignew acur_dualnew
       acur_alc acur_marijuana acur_painkiller
         R01R_A_ETHRACECAT7 age education
         depressed sleeping anxious ptsd
         lied attention listening bully fights        restless answered
         weeklyuse timegetting timeusing socialprob reducedactwithdraw usetoavoid
         sud
         SEXMALE_1 SEXFEMALE_2
         AGE1824_1 AGE2534_2 AGE3544_3 AGE4554_4 AGE5564_5 AGE65_6
         RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
         EDU_1          EDU_2 EDU_3 EDU_4 EDU_5
         INC_1 INC_2 INC_3 INC_4 INC_5
         SOC_1          SOC_2 SOC_3 SOC_4 SOC_5;
     USEVARIABLES = acur_cignew acur_ecignew acur_dualnew
          acur_alc acur_marijuana acur_painkiller
          depressed sleeping anxious ptsd
          lied attention listening bully fights restless answered;
     IDVARIABLE = CASEID;
     MISSING ARE ALL (-99999);
     CLASSES = c(4);
     CATEGORICAL = acur_cignew acur_ecignew acur_dualnew
          acur_alc acur_marijuana acur_painkiller
          depressed sleeping anxious ptsd
          lied attention listening bully fights restless answered;
     AUXILIARY = SEXMALE_1 (R3STEP)
           AGE1824_1 (R3STEP) AGE2534_2 (R3STEP) AGE3544_3 (R3STEP)
           AGE4554_4 (R3STEP) AGE5564_5 (R3STEP)
           RACEBL_2 (R3STEP) RACEOT_3 (R3STEP) RACEHI_6 (R3STEP)
           EDU_1 (R3STEP) EDU_2 (R3STEP) EDU_3(R3STEP)
          EDU_4 (R3STEP) INC_1 (R3STEP) INC_2 (R3STEP)
           INC_3 (R3STEP) INC_4 (R3STEP)
            SOC_2 (R3STEP) SOC_3 (R3STEP) SOC_4 (R3STEP) SOC_5 (R3STEP);
    WEIGHT is weight;
    ANALYSIS: TYPE = MIXTURE;
          STARTS = 100 10;
          OPTSEED = 991329;
          LRTSTARTS = 0 0 150 40;
    SAVEDATA: file is w14class4132021.csv;
         save = Cprob;
    OUTPUT: TECH1 TECH8 TECH10 TECH11 TECH14;
```

SAS File name: Wave 1 4 class prediction 4142021

```
*Prediction Model - Wave 1 4 Class Solution;
*Data into Mplus is from LCA W1 4132021 (newwave4132021editnoheader.csv);
*MPLUS Output = wave 1 run 4132021 4 class;
*CSV = = w14class4132021;

libname pred "U:\CourtneyBlondino\PhD Epidemiology\April Re Run\Wave 1 Prediction 4142021";
```

```
data pred.w14classprob4142021;
input
ACUR_CIG       ACUR_ECI       ACUR_DUA      ACUR_ALC       ACUR_MAR       ACUR_PAI
DEPRESS        SLEEPING       ANXIOUS       PTSD
LIED           ATTENTIO       LISTENING     BULLY          FIGHTS         RESTLESS       ANSWERED
SEXMALE_
AGE1824_
    AGE2534_
    AGE3544_
    AGE4554_
    AGE5564_
    RACEBL_2
    RACEOT_3
    RACEHI_6
    EDU_1
    EDU_2
    EDU_3
    EDU_4
    INC_1
    INC_2
    INC_3
    INC_4
    SOC_2
    SOC_3
    SOC_4
    SOC_5
    CPROB1
    CPROB2
    CPROB3
    CPROB4
    C
    WEIGHT
    CASEID;
datalines;
******COPY PASTE OUTPUT DATA FROM MPLUS*****
run;
```

SAS File name: Wave 1 4 class prediction analysis 4142021

```
*Run analyses;

libname pred "U:\CourtneyBlondino\PhD Epidemiology\April Re Run\Wave 1 Prediction 4142021";

proc contents data=pred.w14classprob4142021;
run;

*check freqs;
proc surveyfreq data=pred.w14classprob4142021;
table ACUR_MAR/row chisq(secondorder);
weight weight;
run;
*the weighted freqs match with the mplus output;

*now need to merge sud into the dataset using idvariable to get sud outcome in same dataset;
proc sort data=pred.w14classprob4142021;
by caseid;
run;

libname LCA "U:\CourtneyBlondino\PhD Epidemiology\LCA\Data Management";

proc sort data=LCA.W1mplusJuly2020weights4132021;
by caseid;
run;

data pred.w14classprobmerge;
merge pred.w14classprob4142021 LCA.W1mplusJuly2020weights4132021;
by caseid;
array change _numeric_;
```

```
do over change;
if change =-99999 then change = .;
end;
run;

proc print data= pred.w14classprobmerge (obs=20);
run;

****confirming data;
proc contents data=pred.w14classprobmerge;
run;

proc surveyfreq data=pred.w14classprobmerge;
table R01R_A_SEX/row chisq(secondorder);
weight weight;
run;

proc freq data=pred.w14classprobmerge;
table acur_cig*acur_cignew;
*acur_alc;
table acur_dua*acur_dualnew;
table acur_eci*acur_ecignew;
table acur_mar*acur_marijuana;
table acur_pai*acur_painkiller;
*table answered*;
*table anxious;
table attentio*attention;
*table bully;
table depress*depressed;
run;


*Just regression;
proc surveylogistic data=pred.w14classprobmerge;
class c (ref='3')/param=ref;
model sud (descending) = c/ link=glogit;
*output predprobs=(I) out=pred.probs72220;
weight weight;
run;

ods pdf;
*Trying ordinal regression;
proc surveylogistic data=pred.w14classprobmerge;
class c (ref='3')/param=ref order=internal;
model sud (descending) = c/lackfit;
*output predprobs=(I) out=pred.probs72220;
weight weight;
run;

*use lackfit to test get pvalue for prop odds assumption;
proc logistic data=pred.w14classprobmerge;
class c (ref='3')/param=ref order=internal;
model sud (descending) = c/lackfit;
*output predprobs=(I) out=pred.probs72220;
weight weight;
run;
ods pdf close;

libname cb "U:\CourtneyBlondino\PhD Epidemiology\LCA\LizHelp";

/* Mode-selective macro wrapper */
%MACRO
cumRoc3(_yOut,_xPred,_vsLbl,_cutFmt,_dsn,_dir00,_dirOut,_dirPng,_dateOut,_libNm=cb,_propOdds=PO,_
yOrd=A,_macMode=1,_macComp=YES,_outCntnts=YES,_outRtf=NO,_debug0=NO) ;
    /* Compile supporting macros */
    %IF       %UPCASE(&_macComp)= YES %THEN %DO ;
        %INCLUDE "U:\CourtneyBlondino\PhD Epidemiology\LCA\LizHelp\words_MAC.sas"          ;
        %INCLUDE "U:\CourtneyBlondino\PhD Epidemiology\LCA\LizHelp\00_preCheck_MAC.sas"    ;
        %INCLUDE "U:\CourtneyBlondino\PhD Epidemiology\LCA\LizHelp\01_dataPre_MAC.sas"     ;
        %INCLUDE "U:\CourtneyBlondino\PhD Epidemiology\LCA\LizHelp\02_cr3_1Logit_MAC.sas"  ;
        %INCLUDE "U:\CourtneyBlondino\PhD Epidemiology\LCA\LizHelp\03_cr3_2ROC_MAC.sas"    ;
```

```
        %INCLUDE "U:\CourtneyBlondino\PhD Epidemiology\LCA\LizHelp\04_cut3Base_MAC.sas"    ;
        %INCLUDE "U:\CourtneyBlondino\PhD Epidemiology\LCA\LizHelp\05_cut3Parmx_MAC.sas"   ;
        %INCLUDE "U:\CourtneyBlondino\PhD Epidemiology\LCA\LizHelp\06_parmx95_MAC.sas"     ;
        %INCLUDE "U:\CourtneyBlondino\PhD Epidemiology\LCA\LizHelp\07_cr3Results_MAC.sas"  ;
    %END ;

    %GLOBAL _poTitle _fileSfx ;

    /* For portrait with 10pt font */
    OPTIONS LINESIZE= 95
            PAGESIZE= 54
    ;
    OPTIONS FORMCHAR='|----|+|---+=|-/\<>*';
    ODS ESCAPECHAR= "^" ;

    /* Check ternary ordinal outcome encoding is compatible with macro */
    %preCheck ;

    %IF &_yOK EQ PASS %THEN %DO ;
        %IF       %UPCASE(&_propOdds)= PO %THEN %DO ;
            %LET _poTitle= %STR(Proportional Odds) ;
            %LET _fileSfx= &_YOUT._&_XPRED._PO  ;
        %END ;
        %ELSE %IF %UPCASE(&_propOdds)= NPO %THEN %DO ;
            %LET _poTitle= %STR(Non-Proportional Odds) ;
            %LET _fileSfx= &_YOUT._&_XPRED._NPO ;
        %END ;

        /* Discard previous temporary datasets */
        PROC DATASETS LIBRARY= WORK NOLIST NOPRINT ;
            DELETE  _inDsn _cutParmx _parmx95 ;
        RUN ; QUIT ;

        /* Discard previous permanent output datasets */
        PROC DATASETS LIBRARY= &_LIBNM NOLIST NOPRINT ;
            DELETE  PARMS4VAR_&_fileSfx
                    COVB_&_fileSfx
                    CUMLOGPARM_&_fileSfx
                    CUMLOGTABLE_&_fileSfx
                    CUMLOGPRED_&_fileSfx
                    ASSOC_&_fileSfx
                    ROC_&_fileSfx
                    AUC_&_fileSfx
                    CUTBASE_&_fileSfx
                    CUTPARMX_&_fileSfx
                    CUMROC3_&_fileSfx
            ;
        RUN ; QUIT ;

        /*  MACRO MODE
            1:  Complete procedure: analysis, criteria and parametric cutpoint calculation,
reporting
            2:  Analysis and criteria and parametric cutpoint calculation only
            3:  Reporting only: requires 1 or 2 to have been run previously */
        %IF &_macMode= 1 OR &_macMode= 2 %THEN %DO ;
            %dataPre ;
            %cr3_1Logit ;
            %cr3_2ROC ;
            %cut3Base ;
            %cut3Parmx ; %parmx95 ;
        %END ;

        %IF &_macMode= 1 OR &_macMode= 3
            %THEN %cr3Results(CUTPARMX CUTBASE) ;

        /* Clean up */
        %IF %upCase(&_debug0)= NO
            AND
            (&_macMode= 1 OR &_macMode= 2)
            %THEN %DO ;
                PROC DATASETS library= WORK NOLIST NOPRINT ;
```

```
                    DELETE  _inDsn _cutParmx _parmx95 ;
                RUN ; QUIT ;
            %END ;
    %END ;
%MEND cumRoc3 ;


*Create dataset for class and SUD to run the cum ROC;
data cb.macrotestclass4142021;
set pred.w14classprobmerge (keep = sud c);
run;

*ODS HTML Close ; *ODS HTML ;
ODS PDF;
    /* Macro debugging: ENABLED */
    OPTIONS MLOGIC MPRINT SYMBOLGEN ;
    %cumRoc3(sud,c,SUD,%STR(BESTD8.3),macrotestclass4142021,
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),

        2019_DEMO,_macMode=1,_macComp=YES,
        _outCntnts=YES,_outRtf=NO) ;
    /* Macro debugging: DISABLED */
    OPTIONS noMLOGIC noMPRINT noSYMBOLGEN ;
*ODS HTML Close ; *ODS HTML ;
ODS PDF CLOSE;


*Create internal score;
data pred.internalscore;
set pred.w14classprobmerge;
*create our continuous internalizing "predictive variable";
int_score = sum (depressed, sleeping, anxious, ptsd);
run;

proc freq data=pred.internalscore;
table int_score;
table sud;
run;

*Create dataset for internalizing and SUD to run the cum ROC;
data cb.macrotestintscore4142021;
set pred.internalscore (keep = sud int_score);
run;

proc contents data=cb.macrotestintscore4142021;
run;

*ODS HTML Close ; *ODS HTML ;
ODS PDF;
    /* Macro debugging: ENABLED */
    OPTIONS MLOGIC MPRINT SYMBOLGEN ;
    %cumRoc3(sud,int_score,SUD,%STR(BESTD8.3),macrotestintscore4142021,
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),

        2019_DEMO,_macMode=1,_macComp=YES,
        _outCntnts=YES,_outRtf=NO) ;
    /* Macro debugging: DISABLED */
    OPTIONS noMLOGIC noMPRINT noSYMBOLGEN ;
*ODS HTML Close ; *ODS HTML ;
ODS PDF CLOSE;



*Create external score;
data pred.externalscore;
set pred.w14classprobmerge;
*create our continuous  "predictive variable";
```

```
ext_score = sum (lied, attention, listening, bully, fights, restless, answered);
run;

proc freq data=pred.externalscore;
table ext_score;
table sud;
run;

*Create dataset for externalizing and SUD to run the cum ROC;
data cb.macrotestextscore4142021;
set pred.externalscore (keep = sud ext_score);
run;

proc contents data=cb.macrotestextscore4142021;
run;

*ODS HTML Close ; *ODS HTML ;
ODS PDF;
    /* Macro debugging: ENABLED */
    OPTIONS MLOGIC MPRINT SYMBOLGEN ;
    %cumRoc3(sud,ext_score,SUD,%STR(BESTD8.3),macrotestextscore4142021,
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),

        2019_DEMO,_macMode=1,_macComp=YES,
        _outCntnts=YES,_outRtf=NO) ;
    /* Macro debugging: DISABLED */
    OPTIONS noMLOGIC noMPRINT noSYMBOLGEN ;
*ODS HTML Close ; *ODS HTML ;
ODS PDF CLOSE;


proc freq data=pred.w14classprobmerge;
table acur_marijuana;
run;


*Create substance use score;
data pred.subsscore;
set pred.w14classprobmerge;
*create our continuous  "predictive variable";
subs_score = sum (acur_cignew, acur_ecignew, acur_dualnew, acur_alc, acur_marijuana,
acur_painkiller);
run;

proc freq data=pred.subsscore;
table subs_score;
table sud;
run;

*Create dataset for substance use and SUD to run the cum ROC;
data cb.macrotestsubsscore4142021;
set pred.subsscore (keep = sud subs_score);
run;

proc contents data=cb.macrotestsubsscore4142021;
run;

*ODS HTML Close ; *ODS HTML ;
ODS PDF;
    /* Macro debugging: ENABLED */
    OPTIONS MLOGIC MPRINT SYMBOLGEN ;
    %cumRoc3(sud,subs_score,SUD,%STR(BESTD8.3),macrotestsubsscore4142021,
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),

        2019_DEMO,_macMode=1,_macComp=YES,
        _outCntnts=YES,_outRtf=NO) ;
    /* Macro debugging: DISABLED */
```

```
    OPTIONS noMLOGIC noMPRINT noSYMBOLGEN ;
*ODS HTML Close ; *ODS HTML ;
ODS PDF CLOSE;







*Now create datasets for IP_2 IP_1 IP_0 to run in macro from pred.probs72220;

*********************************************IP_2;
data pred.macroip2;
set pred.probs72220 (keep = sud IP_2);
run;

*move the dataset into cb;
*ODS HTML Close ; *ODS HTML ;
ODS PDF;
    /* Macro debugging: ENABLED */
    OPTIONS MLOGIC MPRINT SYMBOLGEN ;
    %cumRoc3(sud,IP_2,SUD,%STR(BESTD8.3),macroip2,
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),

        2019_DEMO,_macMode=1,_macComp=YES,
        _outCntnts=YES,_outRtf=NO) ;
    /* Macro debugging: DISABLED */
    OPTIONS noMLOGIC noMPRINT noSYMBOLGEN ;
*ODS HTML Close ; *ODS HTML ;
ODS PDF CLOSE;


*********************************************IP_1;
data pred.macroip1;
set pred.probs72220 (keep = sud IP_1);
run;

*move the dataset into cb;
*ODS HTML Close ; *ODS HTML ;
ODS PDF;
    /* Macro debugging: ENABLED */
    OPTIONS MLOGIC MPRINT SYMBOLGEN ;
    %cumRoc3(sud,IP_1,SUD,%STR(BESTD8.3),macroip1,
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),

        2019_DEMO,_macMode=1,_macComp=YES,
        _outCntnts=YES,_outRtf=NO) ;
    /* Macro debugging: DISABLED */
    OPTIONS noMLOGIC noMPRINT noSYMBOLGEN ;
*ODS HTML Close ; *ODS HTML ;
ODS PDF CLOSE;


*********************************************IP_0;
data pred.macroip0;
set pred.probs72220 (keep = sud IP_0);
run;

*move the dataset into cb;
*ODS HTML Close ; *ODS HTML ;
ODS PDF;
    /* Macro debugging: ENABLED */
    OPTIONS MLOGIC MPRINT SYMBOLGEN ;
    %cumRoc3(sud,IP_0,SUD,%STR(BESTD8.3),macroip0,
        %STR(C:\Users\blondinoct\Documents\LCA),
        %STR(C:\Users\blondinoct\Documents\LCA),
```

```
        %STR(C:\Users\blondinoct\Documents\LCA),

        2019_DEMO,_macMode=1,_macComp=YES,
        _outCntnts=YES,_outRtf=NO) ;
    /* Macro debugging: DISABLED */
    OPTIONS noMLOGIC noMPRINT noSYMBOLGEN ;
*ODS HTML Close ; *ODS HTML ;
ODS PDF CLOSE;
```

SAS File name: Network W1 4142021

```
****SA 2 - Network Analysis
****Making datasets for overall Wave 1 then by sex
****NEW!!!! APRIL 14 2021;


libname LCA "U:\CourtneyBlondino\PhD Epidemiology\LCA\Data Management";


proc contents data=LCA.W1mplusJuly2020weights4132021;
run;

data lca.newwave1fornet442021;
set LCA.W1mplusJuly2020weights4132021 (keep = R01R_A_SEX CASEID PERSONID
                               acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana
acur_painkiller
                               depressed sleeping anxious ptsd
                               lied attention listening bully fights restless answered);
 array change _numeric_;
do over change;
if change =-99999 then change = .;
end;
run;

proc contents data=lca.newwave1fornet442021;
run;

proc freq data=lca.newwave1fornet442021;
table R01R_A_SEX ;
run;

proc freq data=lca.newwave1fornet442021;
table R01R_A_SEX
acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana acur_painkiller
                               depressed sleeping anxious ptsd
                               lied attention listening bully fights restless answered;
run;


libname net "U:\CourtneyBlondino\PhD Epidemiology\April Re Run\Network Wave 1 Data Management";

*OVERALL WAVE 1;
data net.overallwave14142021;
set lca.newwave1fornet442021 (keep =
                               acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana
acur_painkiller
                               depressed sleeping anxious ptsd
                               lied attention listening bully fights restless answered);
                               run;


*************
*male wave 1;
data net.malesubset4142021;
set lca.newwave1fornet442021;
if R01R_A_SEX=1 then output;
run;
```

```
proc contents data=net.malesubset4142021;
run;

*use this one;
data net.malewave14142021;
set net.malesubset4142021(keep = acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana
acur_painkiller
                               depressed sleeping anxious ptsd
                               lied attention listening bully fights restless answered);
run;




**************
*female wave 1;
data net.femalesubset4142021;
set lca.newwave1fornet442021;
if R01R_A_SEX="2" then output;
run;

proc contents data=net.femalesubset4142021;
run;

*use this one;
data net.femalewave14142021;
set net.femalesubset4142021(keep = acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana
acur_painkiller
                               depressed sleeping anxious ptsd
                               lied attention listening bully fights restless answered);
run;





**correlations;;

proc contents data=net.overallwave14142021;
run;

proc corr data=net.overallwave14142021;
var ACUR_CIGNEW       ACUR_ECIGNEW   ACUR_DUALNEW   ACUR_ALC       ACUR_MARIJUANA
       ACUR_PAINKILLER
DEPRESSED      SLEEPING      ANXIOUS       PTSD
LIED    ATTENTION      LISTENING      BULLY   FIGHTS  RESTLESS       ANSWERED;
run;


ods pdf;
*overall cc*substances;
proc freq data=net.overallwave14142021;
table
ACUR_CIGNEW*ACUR_ECIGNEW
ACUR_CIGNEW*ACUR_DUALNEW
ACUR_CIGNEW*ACUR_ALC
ACUR_CIGNEW*ACUR_MARIJUANA
ACUR_CIGNEW*ACUR_PAINKILLER
/plcorr chisq;
run;

*overall ec*substances;
proc freq data=net.overallwave14142021;
table
ACUR_ECIGNEW*ACUR_CIGNEW
ACUR_ECIGNEW*ACUR_DUALNEW
ACUR_ECIGNEW*ACUR_ALC
ACUR_ECIGNEW*ACUR_MARIJUANA
ACUR_ECIGNEW*ACUR_PAINKILLER
/plcorr chisq;
run;
```

```
*overall dual*substances;
proc freq data=net.overallwave14142021;
table
ACUR_DUALNEW*ACUR_CIGNEW
ACUR_DUALNEW*ACUR_ECIGNEW
ACUR_DUALNEW*ACUR_ALC
ACUR_DUALNEW*ACUR_MARIJUANA
ACUR_DUALNEW*ACUR_PAINKILLER
/plcorr chisq;
run;

*overall alcohol*substances;
proc freq data=net.overallwave14142021;
table
ACUR_ALC*ACUR_CIGNEW
ACUR_ALC*ACUR_ECIGNEW
ACUR_ALC*ACUR_DUALNEW
ACUR_ALC*ACUR_MARIJUANA
ACUR_ALC*ACUR_PAINKILLER
/plcorr chisq;
run;

*overall marijuana*pain;
proc freq data=net.overallwave14142021;
table
ACUR_MARIJUANA*ACUR_PAINKILLER
/plcorr chisq;
run;

proc freq data=net.overallwave14142021;
table
ACUR_CIGNEW*DEPRESSED
ACUR_ECIGNEW*DEPRESSED
ACUR_DUALNEW*DEPRESSED
ACUR_ALC*DEPRESSED
ACUR_MARIJUANA*DEPRESSED
ACUR_PAINKILLER*DEPRESSED
/plcorr chisq;
run;


proc freq data=net.overallwave14142021;
table
ACUR_CIGNEW*SLEEPING
ACUR_ECIGNEW*SLEEPING
ACUR_DUALNEW*SLEEPING
ACUR_ALC*SLEEPING
ACUR_MARIJUANA*SLEEPING
ACUR_PAINKILLER*SLEEPING

ACUR_CIGNEW*ANXIOUS
ACUR_ECIGNEW*ANXIOUS
ACUR_DUALNEW*ANXIOUS
ACUR_ALC*ANXIOUS
ACUR_MARIJUANA*ANXIOUS
ACUR_PAINKILLER*ANXIOUS

ACUR_CIGNEW*PTSD
ACUR_ECIGNEW*PTSD
ACUR_DUALNEW*PTSD
ACUR_ALC*PTSD
ACUR_MARIJUANA*PTSD
ACUR_PAINKILLER*PTSD
/plcorr chisq;
run;

proc freq data=net.overallwave14142021;
table SLEEPING*DEPRESSED
ANXIOUS*DEPRESSED
ANXIOUS*SLEEPING
```

```
PTSD*DEPRESSED
PTSD*SLEEPING
PTSD*ANXIOUS/plcorr chisq;
run;

proc freq data=net.overallwave14142021;
table
ACUR_CIGNEW*LIED
ACUR_ECIGNEW*LIED
ACUR_DUALNEW*LIED
ACUR_ALC*LIED
ACUR_MARIJUANA*LIED
ACUR_PAINKILLER*LIED

ACUR_CIGNEW*ATTENTION
ACUR_ECIGNEW*ATTENTION
ACUR_DUALNEW*ATTENTION
ACUR_ALC*ATTENTION
ACUR_MARIJUANA*ATTENTION
ACUR_PAINKILLER*ATTENTION

ACUR_CIGNEW*LISTENING
ACUR_ECIGNEW*LISTENING
ACUR_DUALNEW*LISTENING
ACUR_ALC*LISTENING
ACUR_MARIJUANA*LISTENING
ACUR_PAINKILLER*LISTENING

ACUR_CIGNEW*BULLY
ACUR_ECIGNEW*BULLY
ACUR_DUALNEW*BULLY
ACUR_ALC*BULLY
ACUR_MARIJUANA*BULLY
ACUR_PAINKILLER*BULLY

ACUR_CIGNEW*FIGHTS
ACUR_ECIGNEW*FIGHTS
ACUR_DUALNEW*FIGHTS
ACUR_ALC*FIGHTS
ACUR_MARIJUANA*FIGHTS
ACUR_PAINKILLER*FIGHTS

ACUR_CIGMEW*RESTLESS
ACUR_ECIGNEW*RESTLESS
ACUR_DUALNEW*RESTLESS
ACUR_ALC*RESTLESS
ACUR_MARIJUANA*RESTLESS
ACUR_PAINKILLER*RESTLESS

ACUR_CIGNEW*ANSWERED
ACUR_ECIGNEW*ANSWERED
ACUR_DUALNEW*ANSWERED
ACUR_ALC*ANSWERED
ACUR_MARIJUANA*ANSWERED
ACUR_PAINKILLER*ANSWERED
/plcorr chisq;
run;


proc freq data=net.overallwave14142021;
table
LIED*DEPRESSED
LIED*SLEEPING
LIED*ANXIOUS
LIED*PTSD

ATTENTION*DEPRESSED
ATTENTION*SLEEPING
ATTENTION*ANXIOUS
ATTENTION*PTSD
ATTENTION*LIED
```

```
LISTENING*DEPRESSED
LISTENING*SLEEPING
LISTENING*ANXIOUS
LISTENING*PTSD
LISTENING*LIED
LISTENING*ATTENTION

BULLY*DEPRESSED
BULLY*SLEEPING
BULLY*ANXIOUS
BULLY*PTSD
BULLY*LIED
BULLY*ATTENTION
BULLY*LISTENING

FIGHTS*DEPRESSED
FIGHTS*SLEEPING
FIGHTS*ANXIOUS
FIGHTS*PTSD
FIGHTS*LIED
FIGHTS*ATTENTION
FIGHTS*LISTENING
FIGHTS*BULLY

RESTLESS*DEPRESSED
RESTLESS*SLEEPING
RESTLESS*ANXIOUS
RESTLESS*PTSD
RESTLESS*LIED
RESTLESS*ATTENTION
RESTLESS*LISTENING
RESTLESS*BULLY
RESTLESS*FIGHTS

ANSWERED*DEPRESSED
ANSWERED*SLEEPING
ANSWERED*ANXIOUS
ANSWERED*PTSD
ANSWERED*LIED
ANSWERED*ATTENTION
ANSWERED*LISTENING
ANSWERED*BULLY
ANSWERED*FIGHTS
ANSWERED*RESTLESS

/plcorr chisq;
run;
ods pdf close;



**correlations by sex;

ods pdf;
**male - su;
proc freq data=net.malewave14142021;
table
ACUR_CIGNEW*ACUR_ECIGNEW
ACUR_CIGNEW*ACUR_DUALNEW
ACUR_CIGNEW*ACUR_ALC
ACUR_CIGNEW*ACUR_MARIJUANA
ACUR_CIGNEW*ACUR_PAINKILLER

ACUR_ECIGNEW*ACUR_CIGNEW
ACUR_ECIGNEW*ACUR_DUALNEW
ACUR_ECIGNEW*ACUR_ALC
ACUR_ECIGNEW*ACUR_MARIJUANA
ACUR_ECIGNEW*ACUR_PAINKILLER

ACUR_DUALNEW*ACUR_CIGNEW
```

```
ACUR_DUALNEW*ACUR_ECIGNEW
ACUR_DUALNEW*ACUR_ALC
ACUR_DUALNEW*ACUR_MARIJUANA
ACUR_DUALNEW*ACUR_PAINKILLER

ACUR_ALC*ACUR_CIGNEW
ACUR_ALC*ACUR_ECIGNEW
ACUR_ALC*ACUR_DUALNEW
ACUR_ALC*ACUR_MARIJUANA
ACUR_ALC*ACUR_PAINKILLER

ACUR_MARIJUANA*ACUR_PAINKILLER
/plcorr chisq;
run;

**male - su and int;
proc freq data=net.malewave14142021;
table
ACUR_CIGNEW*DEPRESSED
ACUR_ECIGNEW*DEPRESSED
ACUR_DUALNEW*DEPRESSED
ACUR_ALC*DEPRESSED
ACUR_MARIJUANA*DEPRESSED
ACUR_PAINKILLER*DEPRESSED

ACUR_CIGNEW*SLEEPING
ACUR_ECIGNEW*SLEEPING
ACUR_DUALNEW*SLEEPING
ACUR_ALC*SLEEPING
ACUR_MARIJUANA*SLEEPING
ACUR_PAINKILLER*SLEEPING

ACUR_CIGNEW*ANXIOUS
ACUR_ECIGNEW*ANXIOUS
ACUR_DUALNEW*ANXIOUS
ACUR_ALC*ANXIOUS
ACUR_MARIJUANA*ANXIOUS
ACUR_PAINKILLER*ANXIOUS

ACUR_CIGNEW*PTSD
ACUR_ECIGNEW*PTSD
ACUR_DUALNEW*PTSD
ACUR_ALC*PTSD
ACUR_MARIJUANA*PTSD
ACUR_PAINKILLER*PTSD
/plcorr chisq;
run;

*int;
proc freq data=net.malewave14142021;
table SLEEPING*DEPRESSED
ANXIOUS*DEPRESSED
ANXIOUS*SLEEPING
PTSD*DEPRESSED
PTSD*SLEEPING
PTSD*ANXIOUS/plcorr chisq;
run;

**male - su and ext;
proc freq data=net.malewave14142021;
table
ACUR_CIGNEW*LIED
ACUR_ECIGNEW*LIED
ACUR_DUALNEW*LIED
ACUR_ALC*LIED
ACUR_MARIJUANA*LIED
ACUR_PAINKILLER*LIED

ACUR_CIGNEW*ATTENTION
ACUR_ECIGNEW*ATTENTION
ACUR_DUALNEW*ATTENTION
```

```
ACUR_ALC*ATTENTION
ACUR_MARIJUANA*ATTENTION
ACUR_PAINKILLER*ATTENTION

ACUR_CIGNEW*LISTENING
ACUR_ECIGNEW*LISTENING
ACUR_DUALNEW*LISTENING
ACUR_ALC*LISTENING
ACUR_MARIJUANA*LISTENING
ACUR_PAINKILLER*LISTENING

ACUR_CIGNEW*BULLY
ACUR_ECIGNEW*BULLY
ACUR_DUALNEW*BULLY
ACUR_ALC*BULLY
ACUR_MARIJUANA*BULLY
ACUR_PAINKILLER*BULLY

ACUR_CIGNEW*FIGHTS
ACUR_ECIGNEW*FIGHTS
ACUR_DUALNEW*FIGHTS
ACUR_ALC*FIGHTS
ACUR_MARIJUANA*FIGHTS
ACUR_PAINKILLER*FIGHTS

ACUR_CIGNEW*RESTLESS
ACUR_ECIGNEW*RESTLESS
ACUR_DUALNEW*RESTLESS
ACUR_ALC*RESTLESS
ACUR_MARIJUANA*RESTLESS
ACUR_PAINKILLER*RESTLESS

ACUR_CIGNEW*ANSWERED
ACUR_ECIGNEW*ANSWERED
ACUR_DUALNEW*ANSWERED
ACUR_ALC*ANSWERED
ACUR_MARIJUANA*ANSWERED
ACUR_PAINKILLER*ANSWERED
/plcorr chisq;
run;


*male int and ext;
proc freq data=net.malewave14142021;
table
LIED*DEPRESSED
LIED*SLEEPING
LIED*ANXIOUS
LIED*PTSD

ATTENTION*DEPRESSED
ATTENTION*SLEEPING
ATTENTION*ANXIOUS
ATTENTION*PTSD
ATTENTION*LIED

LISTENING*DEPRESSED
LISTENING*SLEEPING
LISTENING*ANXIOUS
LISTENING*PTSD
LISTENING*LIED
LISTENING*ATTENTION

BULLY*DEPRESSED
BULLY*SLEEPING
BULLY*ANXIOUS
BULLY*PTSD
BULLY*LIED
BULLY*ATTENTION
BULLY*LISTENING
```

```
FIGHTS*DEPRESSED
FIGHTS*SLEEPING
FIGHTS*ANXIOUS
FIGHTS*PTSD
FIGHTS*LIED
FIGHTS*ATTENTION
FIGHTS*LISTENING
FIGHTS*BULLY

RESTLESS*DEPRESSED
RESTLESS*SLEEPING
RESTLESS*ANXIOUS
RESTLESS*PTSD
RESTLESS*LIED
RESTLESS*ATTENTION
RESTLESS*LISTENING
RESTLESS*BULLY
RESTLESS*FIGHTS

ANSWERED*DEPRESSED
ANSWERED*SLEEPING
ANSWERED*ANXIOUS
ANSWERED*PTSD
ANSWERED*LIED
ANSWERED*ATTENTION
ANSWERED*LISTENING
ANSWERED*BULLY
ANSWERED*FIGHTS
ANSWERED*RESTLESS

/plcorr chisq;
run;
ods pdf close;

****************Female;

ods pdf;
**female - su;
proc freq data=net.femalewave14142021;
table
ACUR_CIGNEW*ACUR_ECIGNEW
ACUR_CIGNEW*ACUR_DUALNEW
ACUR_CIGNEW*ACUR_ALC
ACUR_CIGNEW*ACUR_MARIJUANA
ACUR_CIGNEW*ACUR_PAINKILLER

ACUR_ECIGNEW*ACUR_CIGNEW
ACUR_ECIGNEW*ACUR_DUALNEW
ACUR_ECIGNEW*ACUR_ALC
ACUR_ECIGNEW*ACUR_MARIJUANA
ACUR_ECIGNEW*ACUR_PAINKILLER

ACUR_DUALNEW*ACUR_CIGNEW
ACUR_DUALNEW*ACUR_ECIGNEW
ACUR_DUALNEW*ACUR_ALC
ACUR_DUALNEW*ACUR_MARIJUANA
ACUR_DUALNEW*ACUR_PAINKILLER

ACUR_ALC*ACUR_CIGNEW
ACUR_ALC*ACUR_ECIGNEW
ACUR_ALC*ACUR_DUALNEW
ACUR_ALC*ACUR_MARIJUANA
ACUR_ALC*ACUR_PAINKILLER

ACUR_MARIJUANA*ACUR_PAINKILLER
/plcorr chisq;
run;

**female - su and int;
proc freq data=net.femalewave14142021;
table
```

```
ACUR_CIGNEW*DEPRESSED
ACUR_ECIGNEW*DEPRESSED
ACUR_DUALNEW*DEPRESSED
ACUR_ALC*DEPRESSED
ACUR_MARIJUANA*DEPRESSED
ACUR_PAINKILLER*DEPRESSED

ACUR_CIGNEW*SLEEPING
ACUR_ECIGNEW*SLEEPING
ACUR_DUALNEW*SLEEPING
ACUR_ALC*SLEEPING
ACUR_MARIJUANA*SLEEPING
ACUR_PAINKILLER*SLEEPING

ACUR_CIGNEW*ANXIOUS
ACUR_ECIGNEW*ANXIOUS
ACUR_DUALNEW*ANXIOUS
ACUR_ALC*ANXIOUS
ACUR_MARIJUANA*ANXIOUS
ACUR_PAINKILLER*ANXIOUS

ACUR_CIGNEW*PTSD
ACUR_ECIGNEW*PTSD
ACUR_DUALNEW*PTSD
ACUR_ALC*PTSD
ACUR_MARIJUANA*PTSD
ACUR_PAINKILLER*PTSD
/plcorr chisq;
run;

*int;
proc freq data=net.femalewave14142021;
table SLEEPING*DEPRESSED
ANXIOUS*DEPRESSED
ANXIOUS*SLEEPING
PTSD*DEPRESSED
PTSD*SLEEPING
PTSD*ANXIOUS/plcorr chisq;
run;

**female - su and ext;
proc freq data=net.femalewave14142021;
table
ACUR_CIGNEW*LIED
ACUR_ECIGNEW*LIED
ACUR_DUALNEW*LIED
ACUR_ALC*LIED
ACUR_MARIJUANA*LIED
ACUR_PAINKILLER*LIED

ACUR_CIGNEW*ATTENTION
ACUR_ECIGNEW*ATTENTION
ACUR_DUALNEW*ATTENTION
ACUR_ALC*ATTENTION
ACUR_MARIJUANA*ATTENTION
ACUR_PAINKILLER*ATTENTION

ACUR_CIGNEW*LISTENING
ACUR_ECIGNEW*LISTENING
ACUR_DUALNEW*LISTENING
ACUR_ALC*LISTENING
ACUR_MARIJUANA*LISTENING
ACUR_PAINKILLER*LISTENING

ACUR_CIGNEW*BULLY
ACUR_ECIGNEW*BULLY
ACUR_DUALNEW*BULLY
ACUR_ALC*BULLY
ACUR_MARIJUANA*BULLY
ACUR_PAINKILLER*BULLY
```

```
ACUR_CIGNEW*FIGHTS
ACUR_ECIGNEW*FIGHTS
ACUR_DUALNEW*FIGHTS
ACUR_ALC*FIGHTS
ACUR_MARIJUANA*FIGHTS
ACUR_PAINKILLER*FIGHTS

ACUR_CIGNEW*RESTLESS
ACUR_ECIGNEW*RESTLESS
ACUR_DUALNEW*RESTLESS
ACUR_ALC*RESTLESS
ACUR_MARIJUANA*RESTLESS
ACUR_PAINKILLER*RESTLESS

ACUR_CIGNEW*ANSWERED
ACUR_ECIGNEW*ANSWERED
ACUR_DUALNEW*ANSWERED
ACUR_ALC*ANSWERED
ACUR_MARIJUANA*ANSWERED
ACUR_PAINKILLER*ANSWERED
/plcorr chisq;
run;


*female int and ext;
proc freq data=net.femalewave14142021;
table
LIED*DEPRESSED
LIED*SLEEPING
LIED*ANXIOUS
LIED*PTSD

ATTENTION*DEPRESSED
ATTENTION*SLEEPING
ATTENTION*ANXIOUS
ATTENTION*PTSD
ATTENTION*LIED

LISTENING*DEPRESSED
LISTENING*SLEEPING
LISTENING*ANXIOUS
LISTENING*PTSD
LISTENING*LIED
LISTENING*ATTENTION

BULLY*DEPRESSED
BULLY*SLEEPING
BULLY*ANXIOUS
BULLY*PTSD
BULLY*LIED
BULLY*ATTENTION
BULLY*LISTENING

FIGHTS*DEPRESSED
FIGHTS*SLEEPING
FIGHTS*ANXIOUS
FIGHTS*PTSD
FIGHTS*LIED
FIGHTS*ATTENTION
FIGHTS*LISTENING
FIGHTS*BULLY

RESTLESS*DEPRESSED
RESTLESS*SLEEPING
RESTLESS*ANXIOUS
RESTLESS*PTSD
RESTLESS*LIED
RESTLESS*ATTENTION
RESTLESS*LISTENING
RESTLESS*BULLY
RESTLESS*FIGHTS
```

```
ANSWERED*DEPRESSED
ANSWERED*SLEEPING
ANSWERED*ANXIOUS
ANSWERED*PTSD
ANSWERED*LIED
ANSWERED*ATTENTION
ANSWERED*LISTENING
ANSWERED*BULLY
ANSWERED*FIGHTS
ANSWERED*RESTLESS

/plcorr chisq;
run;
ods pdf close;

*tobacco;
proc freq data=net.overallwave14142021;
table
ACUR_CIG*ACUR_ECI
ACUR_CIG*ACUR_DUA
ACUR_ECI*ACUR_DUA
/plcorr chisq;
run;

ods pdf close;
```

R File name: New W1, M, W Network Analysis 4142021

```
#PATH WAVE 1 - Network Analysis (Specific Aim 2)


##########################################
# Starting with Overall Wave 1 Sample  #
##########################################

#OVERALL WAVE 1#
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files/Network Chapter")
getwd()
overall<-read.csv("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files/Network Chapter/OverallWave1-4142021.csv", header=T, sep=',')
names(overall)

#checking distributions
table(overall$acur_cignew)
table(overall$acur_ecignew)
table(overall$ACUR_DUA)
table(overall$ACUR_ALC)
table(overall$acur_marijuana)
table(overall$ACUR_PAI)
table(overall$DEPRESS)
table(overall$SLEEPING)
table(overall$ANXIOUS)
table(overall$PTSD)
table(overall$LIED)
table(overall$ATTENTIO)
table(overall$LISTENING)
table(overall$BULLY)
```

```
table(overall$FIGHTS)
table(overall$RESTLESS)
table(overall$ANSWERED)

#rename variables so they look nice on the network
names(overall)[names(overall) == "ACUR_CIG"] <- "CIG"
names(overall)[names(overall) == "ACUR_ECI"] <- "ECIG"
names(overall)[names(overall) == "ACUR_DUA"] <- "Dual CIG + ECIG"
names(overall)[names(overall) == "ACUR_ALC"] <- "Alcohol"
names(overall)[names(overall) == "ACUR_MAR"] <- "Marijuana"
names(overall)[names(overall) == "ACUR_PAI"] <- "PDNP"
names(overall)[names(overall) == "DEPRESS"] <- "Depressed"
names(overall)[names(overall) == "SLEEPING"] <- "Sleeping"
names(overall)[names(overall) == "ANXIOUS"] <- "Anxious"
names(overall)[names(overall) == "PTSD"] <- "Distressed/Past"
names(overall)[names(overall) == "LIED"] <- "Lied"
names(overall)[names(overall) == "ATTENTIO"] <- "Attention"
names(overall)[names(overall) == "LISTENING"] <- "Listening"
names(overall)[names(overall) == "BULLY"] <- "Bully"
names(overall)[names(overall) == "FIGHTS"] <- "Fights"
names(overall)[names(overall) == "RESTLESS"] <- "Restless"
names(overall)[names(overall) == "ANSWERED"] <- "Answered"


require(ggplot2)
require(bootnet)
require(IsingFit)
require(IsingSampler)
require(qgraph)


#########
#IsingFit
OverallNetworkIF <-estimateNetwork(overall, default="IsingFit", missing="listwise")
#try a network with spring layout
plot(OverallNetworkIF, layout = "spring", vsize = 10, cex=8)

OverallNetworkIF

#########
OverallNetworkIF$labels

Names<- c("CIG", "ECIG" , "Dual CIG + ECIG", "Alcohol", "Marijuana",  "PDNP",
        "Depressed", "Sleeping" , "Anxious", "Distressed/Past" , "Lied",
        "Attention" , "Listening" , "Bully" ,"Fights"  , "Restless",
        "Answered")

Traits <- rep(c(
  'Substance Use',
  'Negative Affect',
  'Externalizing'
), times=c(6,4,7))


#BLACK EDGES
#plot(OverallNetworkIF,
```

```
#    layout="spring",
#    cut=0,
#    theme="colorblind",
#    groups=Traits,
#    labels=Names,
     #nodeNames=Names,
#    edge.color="black",
#    label.scale.equal=TRUE,
#    label.cex= 1.2,
#    legend.cex = 0.4)

layout(1)
#COLORED EDGES
plot(OverallNetworkIF,
     layout="spring",
     cut=0,
     theme="colorblind",
     groups=Traits,
     labels=Names,
     #nodeNames=Names,
     #edge.color="black",
     label.scale.equal=TRUE,
     label.cex= 3.0,
     legend.cex = 0.4)

     #title= "Overall Wave 1 Sample")

#edges
OverallEdges <- OverallNetworkIF$graph
print(OverallEdges)
write.csv(OverallEdges, file="Overall_W1_Edges.csv")

View(OverallNetworkIF$graph)
#write(OverallNetworkIF$graph, file="OverallEdges.csv", sep=" ")

##########################################################################
#Accuracy, Stability, and Replicabiity from PNASS PRACTICALS          #
#TRY THIS WITH OVERALL SAMPLE using IsingFit Model (OverallNetworkIF) #
##########################################################################

#First, test accuracy of connections by obtaining confidence interval around
#estimated edge weight using non-parametric bootstrapping (on original sample and in smaller sample)


library(bootnet)


#Perform a non-parametric bootstrap on the estimated network, and
#plot the confidence intervals of the edge-weights
OverallBoot <- bootnet(OverallNetworkIF, nCores=8)
#note that the default is not listed here but in the notes, they are
plot(OverallBoot, order="sample")
plot(OverallBoot, order="sample", plot="interval", split0=TRUE, labels=FALSE)
plot(OverallBoot, order="sample", labels=FALSE)
plot(OverallBoot, order="sample", plot="interval", split0=TRUE)
```

```
OverallBootSummary <- summary(OverallBoot)
write.csv(OverallBootSummary, file="Overall_Boot_Summary.csv")

OverallInclusion<-bootInclude(OverallBoot, verbose=TRUE)
plot(OverallInclusion)
#plot bootstrapped edge CIs
plot(OverallBoot, labels=FALSE, order="sample")
#plot significant differences (alpha=0.05) of edges
plot(OverallBoot, "edge", plot="difference", onlyNonZero = TRUE,
    order="sample")


#removing edges (setting them to 0) based on significance alpha=0.05
# Threshold network:
OverallNetwork_thresholded <- bootThreshold(OverallBoot)
# Plot:
plot(OverallNetwork_thresholded)
OverallNetwork_thresholded$results


#print on PDF so you can read which edge and how many times
#make sure dimensions are quite long
#was it included in the bootstrap
Overall_Ising_threshold <- bootThreshold(OverallBoot, alpha=0.01)
Overall_Ising_threshold$results

L<- averageLayout(OverallNetworkIF, Overall_Ising_threshold)
layout(t(1:2))
plot(OverallNetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    edge.color="black",
    label.scale.equal=TRUE,
    label.cex= 1.2,
    legend.cex = 0.4,
    title="Ising Fit Overall Sample")
plot(Overall_Ising_threshold,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    edge.color="black",
    label.scale.equal=TRUE,
    label.cex= 1.2,
    legend.cex = 0.4,
    title= "Ising Threshold")

OverallNetworkIF$results
Overall_Ising_threshold$results

#edges
Edges <- OverallNetworkIF$graph
```

```
print(Edges)
write(Edges, file="OverallEdges.csv", sep=" ")

OverallSigDifEdge <- summary(OverallBoot)
write(OverallSigDifEdge, file="OverallSigDifEdge.csv", sep=" ")

#Second, investigate stability of centrality indices by case-dropping subset bootstrap
#and get the CS-coefficient


#Perform a case-drop bootstrap on the network, and plot the stability
#of centrality indices. Remember that the default values have now changes
#and do not automatically include stability estimates of 'closeness'
#and 'betweenness'. If you do wish to inspect these, you must include
#statistics = c("strength", "closeness", "betweenness")
OverallBoot2 <- bootnet(OverallNetworkIF, nCores=8, type="case",
                   statistics=c("strength", "closeness", "betweenness"))
plot(OverallBoot2, statistics = c("strength", "closeness", "betweenness"))


plot(OverallBoot2, statistics = c("strength", "closeness"))


plot(OverallBoot2, statistics = c("strength", "closeness", "betweenness"),
   CIstyle="quantiles")

differenceTest(OverallBoot2, "ACUR_CIG", "ACUR_ECIG", "strength")


#Give the CS-coefficient of the three centrality indices, and explain how
#this measure can be interpreted
corStability(OverallBoot2)
#CS-coefficient for
#betweenness= 0.206 (below 0.25- not good) - should not interpret betweenness values because CS
coefficient is not stable
#closeness= 0.517 this is ok
#strength= 0.594 this is good, above 0.5

centralityPlot(OverallNetworkIF, include=c("Strength", "Closeness","Betweenness"))
centralityPlot(OverallNetworkIF, include=c("Strength", "Closeness"))

centralityTable(OverallNetworkIF)

summary(OverallNetworkIF)
summary(OverallBoot)
OverallBoot
summary(OverallBoot2)
OverallBoot2$type

#Third, test whether network connections (step1) and centrality metrics (step2)
#for different variables significantly differ from each other using bootstrapped difference test
#can do the edge weight difference test and the centrality difference test

plot(OverallBoot, "edge", plot="difference", onlyNonZero = TRUE, order="sample", labels=FALSE)

differenceTest(OverallBoot, 3, 17, "strength")
```

```
differenceTest(OverallBoot, 1, 1, "strength")
plot(OverallBoot, "strength")
#plot(OverallBoot, statistics = c("betweenness", "closeness", "strength"), plot =
#      "difference")
# ^ only gave me strength
OverallBoot$bootTable
```

```
######################
##WAVE 1 - MALE ONLY##
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation
Files/Network Chapter")
getwd()
male<-read.csv("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation
Files/Network Chapter/MaleWave1-4142021.csv", header=T, sep=',')
dim(male)
names(male)

#rename variables so they look nice on the network
names(male)[names(male) == "ACUR_CIG"] <- "CIG"
names(male)[names(male) == "ACUR_ECI"] <- "ECIG"
names(male)[names(male) == "ACUR_DUA"] <- "Dual CIG + ECIG"
names(male)[names(male) == "ACUR_ALC"] <- "Alcohol"
names(male)[names(male) == "ACUR_MAR"] <- "Marijuana"
names(male)[names(male) == "ACUR_PAI"] <- "PDNP"
names(male)[names(male) == "DEPRESS"] <- "Depressed"
names(male)[names(male) == "SLEEPING"] <- "Sleeping"
names(male)[names(male) == "ANXIOUS"] <- "Anxious"
names(male)[names(male) == "PTSD"] <- "PTSD"
names(male)[names(male) == "LIED"] <- "Lied"
names(male)[names(male) == "ATTENTIO"] <- "Attention"
names(male)[names(male) == "LISTENING"] <- "Listening"
names(male)[names(male) == "BULLY"] <- "Bully"
names(male)[names(male) == "FIGHTS"] <- "Fights"
names(male)[names(male) == "RESTLESS"] <- "Restless"
names(male)[names(male) == "ANSWERED"] <- "Answered"


#########
#IsingFit
MaleNetworkIF <-estimateNetwork(male, default="IsingFit", missing="listwise")
#try a network with spring layout
plot(MaleNetworkIF, layout = "spring", vsize = 10, cex=8)


#########
MaleNetworkIF$labels

Names<- c("CIG", "ECIG" , "Dual CIG + ECIG", "Alcohol", "Marijuana",  "PDNP",
        "Depressed", "Sleeping" , "Anxious", "Distressed/Past" , "Lied",
        "Attention" , "Listening" , "Bully" ,"Fights"  , "Restless",
        "Answered")
```

```
Traits <- rep(c(
  'Substance Use',
  'Internalizing',
  'Externalizing'
), times=c(6,4,7))


plot(MaleNetworkIF,
     layout="spring",
     cut=0,
     theme="colorblind",
     #title="Wave 1 - Men Only",
     groups=Traits,
     labels=Names,
     #nodeNames=Names,
     label.scale.equal=TRUE,
     label.cex= 4,
     legend.cex = 0.4)

View(MaleNetworkIF$graph)
MaleNetworkIF$results
MaleNetworkIF

#edges
MenEdges <- MaleNetworkIF$graph
print(MenEdges)


########################################################################
#Accuracy, Stability, and Replicabiity from PNASS PRACTICALS          #
#TRY THIS WITH OVERALL SAMPLE using IsingFit Model (MaleNetworkIF)    #
########################################################################
library(bootnet)
#Network <- estimateNetwork(bfiData, default="ggmModSelect",
#                           stepwise=FALSE, corMethod="cor")
#plot(Network)

#Perform a non-parametric bootstrap on the estimated network, and
#plot the confidence intervals of the edge-weights
MaleBoot1 <- bootnet(MaleNetworkIF, nCores=8)
#note that the default is not listed here but in the notes, they are
plot(MaleBoot1, order="sample")
plot(MaleBoot1, order="sample", plot="interval", split0=TRUE)
#print on PDF so you can read which edge and how many times
#make sure dimensions are quite long
#was it included in the bootstrap
plot(MaleBoot1, order="sample", labels=FALSE)

#Perform a case-drop bootstrap on the network, and plot the stability
#of centrality indices. Remember that the default values have now changes
#and do not automatically include stability estimates of 'closeness'
#and 'betweenness'. If you do wish to inspect these, you must include
#statistics = c("strength", "closeness", "betweenness")
MaleBoot2 <- bootnet(MaleNetworkIF, nCores=8, type="case",
              statistics=c("strength", "closeness", "betweenness"))
```

```
plot(MaleBoot2, statistics = c("strength", "closeness", "betweenness"))

#Give the CS-coefficient of the three centrality indices, and explain how
#this measure can be interpreted
corStability(MaleBoot2)
#CS-coefficient for
#betweenness= 0.128 (below 0.25- not good)
#closeness= 0.361  (don't think this is good but check)
#strength= 0.517 this is good, before 0.5

centralityPlot(MaleNetworkIF)
centralityPlot(MaleNetworkIF, include=c("Strength", "Closeness","Betweenness"))

centralityPlot(MaleNetworkIF, include=c("Strength", "Closeness"))
centralityPlot(MaleNetworkIF, include=c("Strength"))

centralityTable(MaleNetworkIF)
summary(MaleNetworkIF)

#plot significant differences (alpha=0.05) of edges
plot(MaleBoot1, "edge", plot="difference", onlyNonZero = TRUE,
    order="sample", labels=FALSE)
#plot node strength difference
plot(MaleBoot1, "strength")




#######################
##WAVE 1 - FEMALE ONLY##
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation
Files/Network Chapter")
getwd()
female<-read.csv("/Users/courtneyblondino/Library/Mobile
Documents/com~apple~CloudDocs/Dissertation Files/Network Chapter/FemaleWave1-4142021.csv",
header=T, sep=',')
dim(female)
names(female)

#rename variables so they look nice on the network
names(female)[names(female) == "ACUR_CIG"] <- "CIG"
names(female)[names(female) == "ACUR_ECI"] <- "ECIG"
names(female)[names(female) == "ACUR_DUA"] <- "Dual CIG + ECIG"
names(female)[names(female) == "ACUR_ALC"] <- "Alcohol"
names(female)[names(female) == "ACUR_MAR"] <- "Marijuana"
names(female)[names(female) == "ACUR_PAI"] <- "PDNP"
names(female)[names(female) == "DEPRESS"] <- "Depressed"
names(female)[names(female) == "SLEEPING"] <- "Sleeping"
names(female)[names(female) == "ANXIOUS"] <- "Anxious"
names(female)[names(female) == "PTSD"] <- "PTSD"
names(female)[names(female) == "LIED"] <- "Lied"
names(female)[names(female) == "ATTENTIO"] <- "Attention"
names(female)[names(female) == "LISTENING"] <- "Listening"
names(female)[names(female) == "BULLY"] <- "Bully"
names(female)[names(female) == "FIGHTS"] <- "Fights"
```

```
names(female)[names(female) == "RESTLESS"] <- "Restless"
names(female)[names(female) == "ANSWERED"] <- "Answered"


#########
#IsingFit
FemaleNetworkIF <-estimateNetwork(female, default="IsingFit", missing="listwise")
#try a network with spring layout
plot(FemaleNetworkIF, layout = "spring", vsize = 10, cex=8)

FemaleNetworkIF

#########
FemaleNetworkIF$labels

Names<- c("CIG", "ECIG" , "Dual CIG + ECIG", "Alcohol", "Marijuana",  "PDNP",
        "Depressed", "Sleeping" , "Anxious", "Distressed/Past" , "Lied",
        "Attention" , "Listening" , "Bully" ,"Fights"  , "Restless",
        "Answered")

Traits <- rep(c(
  'Substance Use',
  'Internalizing',
  'Externalizing'
), times=c(6,4,7))


plot(FemaleNetworkIF,
    layout="spring",
    cut=0,
    theme="colorblind",
    #title="Wave 1 - Women Only",
    groups=Traits,
    labels=Names,
    #nodeNames=Names,
    label.scale.equal=TRUE,
    label.cex= 4,
    legend.cex = 0.4)


View(FemaleNetworkIF$graph)
FemaleNetworkIF$results

#edges
WomenEdges <- FemaleNetworkIF$graph
print(WomenEdges)

##########################################################################
#Accuracy, Stability, and Replicabiity from PNASS PRACTICALS          #
#TRY THIS WITH OVERALL SAMPLE using IsingFit Model (FemaleNetworkIF)    #
##########################################################################
library(bootnet)
#Network <- estimateNetwork(bfiData, default="ggmModSelect",
#                        stepwise=FALSE, corMethod="cor")
#plot(Network)
```

```
#Perform a non-parametric bootstrap on the estimated network, and
#plot the confidence intervals of the edge-weights
FemaleBoot1 <- bootnet(FemaleNetworkIF, nCores=8)
#note that the default is not listed here but in the notes, they are
plot(FemaleBoot1, order="sample", labels=FALSE)
plot(FemaleBoot1, order="sample", plot="interval", split0=TRUE)
#print on PDF so you can read which edge and how many times
#make sure dimensions are quite long
#was it included in the bootstrap

#Perform a case-drop bootstrap on the network, and plot the stability
#of centrality indices. Remember that the default values have now changes
#and do not automatically include stability estimates of 'closeness'
#and 'betweenness'. If you do wish to inspect these, you must include
#statistics = c("strength", "closeness", "betweenness")
FemaleBoot2 <- bootnet(FemaleNetworkIF, nCores=8, type="case",
                   statistics=c("strength", "closeness", "betweenness"))
plot(FemaleBoot2, statistics = c("strength", "closeness", "betweenness"))

#Give the CS-coefficient of the three centrality indices, and explain how
#this measure can be interpreted
corStability(FemaleBoot2)
#CS-coefficient for
#betweenness= 0.128 (below 0.25- not good)
#closeness= 0.439 (don't think this is good but check)
#strength= 0.361 not good - all below 0.5 so not good

centralityPlot(FemaleNetworkIF)
centralityPlot(FemaleNetworkIF, include=c("Strength", "Closeness","Betweenness"))

centralityPlot(FemaleNetworkIF, include=c("Strength"))

FemaleNetworkIF$graph
centralityTable(FemaleNetworkIF)

summary(FemaleNetworkIF)

#plot significant differences (alpha=0.05) of edges
plot(FemaleBoot1, "edge", plot="difference", onlyNonZero = TRUE,
    order="sample", labels=FALSE)
#plot node strength difference
plot(FemaleBoot1, "strength")




#Network Comparisons


library("qgraph")
L<-averageLayout(MaleNetworkIF, FemaleNetworkIF)
Max<- max(abs(c(getWmat(MaleNetworkIF), getWmat(FemaleNetworkIF))))
layout(t(1:2))
plot(MaleNetworkIF, layout=L, title="Men", maximum=Max)
```

```
plot(FemaleNetworkIF, layout=L, title="Women", maximum=Max)


library("qgraph")
L<-averageLayout(MaleNetworkIF, FemaleNetworkIF)
Max<- max(abs(c(getWmat(MaleNetworkIF), getWmat(FemaleNetworkIF))))
layout(t(1:2))
plot(MaleNetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    labels=Names,
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    label.scale.equal=TRUE,
    label.cex=4.0,
    legend.cex = 0.4,
    legend=FALSE,
    title= "Men",
    maximum=Max)
plot(FemaleNetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    labels=Names,
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    label.scale.equal=TRUE,
    label.cex= 4.0,
    legend.cex = 0.4,
    legend = FALSE,
    title= "Women",
    maximum=Max)

MaleW1Edges <- MaleNetworkIF$graph
write.csv(MaleW1Edges, file="Male_W1_Edges.csv")

FemaleW1Edges <- FemaleNetworkIF$graph
write.csv(FemaleW1Edges, file="Female_W1_Edges.csv")



library("devtools")
install_github("cvborkulo/NetworkComparisonTest")
library("NetworkComparisonTest")
#perform NCT and interpret results
NCTres<- NCT(MaleNetworkIF, FemaleNetworkIF, test.edges=TRUE,
          it=100)

#difference in global strength between the networks of the observed data sets
NCTres$glstrinv.real
#2.478884

#global strength values of individual networks
```

```
NCTres$glstrinv.sep
#53.41989 vs 50.94101

#Difference in global strength p-value
NCTres$glstrinv.pval
#0.4554455 - so not significantly different from one another in regard to global strength

#Value of the max difference in edge weights of observed networks
NCTres$nwinv.real
#1.327559

#Maximum difference in edge weights
NCTres$nwinv.pval
#0.3168317 - so not significantly different from one another in regard to number of edge weights

#Which edges significantly differ?
NCTres$einv.pvals[which(NCTres$einv.pvals[,3]<0.05),]
#          Var1          Var2    p-value
#72        acur_alc acur_marijuana 0.01980198
#123       acur_alc       sleeping 0.00990099
#141 acur_marijuana        anxious 0.02970297
#172   acur_ecignew           lied 0.03960396
#174       acur_alc           lied 0.00990099
#191       acur_alc      attention 0.01980198
#198           lied      attention 0.03960396
#208       acur_alc      listening 0.04950495



#NCTresCen<- NCT(MaleNetworkIF, FemaleNetworkIF, test.centrality=TRUE, centrality=c("strength"),
#              nodes="all",it=20)
#NCTresCen$diffcen.pvals[which(NCTresCen$diffcen.pvals[,3]<0.05),]


centralityTable(OverallNetworkIF)
centralityTable(MaleNetworkIF)
centralityTable(FemaleNetworkIF)





#########################################################################
#Accuracy, Stability, and Replicabiity from PNASS PRACTICALS          #
#TRY THIS WITH OVERALL SAMPLE using IsingFit Model (OverallNetworkIF) #
#########################################################################

#First, test accuracy of connections by obtaining confidence interval around
#estimated edge weight using non-parametric bootstrapping (on original sample and in smaller sample)


library(bootnet)
#Network <- estimateNetwork(bfiData, default="ggmModSelect",
#                         stepwise=FALSE, corMethod="cor")
#plot(Network)
```

```
#Perform a non-parametric bootstrap on the estimated network, and
#plot the confidence intervals of the edge-weights
OverallBoot <- bootnet(OverallNetworkIF, nCores=8)
#note that the default is not listed here but in the notes, they are
plot(OverallBoot, order="sample")
plot(OverallBoot, order="sample", plot="interval", split0=TRUE, labels=FALSE)
plot(OverallBoot, order="sample", labels=FALSE)
plot(OverallBoot, order="sample", plot="interval", split0=TRUE)

summary(OverallBoot)
OverallInclusion<-bootInclude(OverallBoot, verbose=TRUE)
plot(OverallInclusion)
#plot bootstrapped edge CIs
plot(OverallBoot, labels=FALSE, order="sample")
#plot significant differences (alpha=0.05) of edges
plot(OverallBoot, "edge", plot="difference", onlyNonZero = TRUE,
    order="sample")


#removing edges (setting them to 0) based on significance alpha=0.05
# Threshold network:
OverallNetwork_thresholded <- bootThreshold(OverallBoot)
# Plot:
plot(OverallNetwork_thresholded)
OverallNetwork_thresholded$results


#print on PDF so you can read which edge and how many times
#make sure dimensions are quite long
#was it included in the bootstrap
Overall_Ising_threshold <- bootThreshold(OverallBoot, alpha=0.01)
Overall_Ising_threshold$results

L<- averageLayout(OverallNetworkIF, Overall_Ising_threshold)
layout(t(1:2))
plot(OverallNetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    edge.color="black",
    label.scale.equal=TRUE,
    label.cex= 1.2,
    legend.cex = 0.4,
    title="Ising Fit Overall Sample")
plot(Overall_Ising_threshold,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    edge.color="black",
    label.scale.equal=TRUE,
    label.cex= 1.2,
```

```
    legend.cex = 0.4,
    title= "Ising Threshold")

OverallNetworkIF$results
Overall_Ising_threshold$results




#Second, investigate stability of centrality indices by case-dropping subset bootstrap
#and get the CS-coefficient


#Perform a case-drop bootstrap on the network, and plot the stability
#of centrality indices. Remember that the default values have now changes
#and do not automatically include stability estimates of 'closeness'
#and 'betweenness'. If you do wish to inspect these, you must include
#statistics = c("strength", "closeness", "betweenness")
OverallBoot2 <- bootnet(OverallNetworkIF, nCores=8, type="case",
                    statistics=c("strength", "closeness", "betweenness"))
plot(OverallBoot2, statistics = c("strength", "closeness", "betweenness"))


plot(OverallBoot2, statistics = c("strength", "closeness"))


plot(OverallBoot2, statistics = c("strength", "closeness", "betweenness"),
    CIstyle="quantiles")

differenceTest(OverallBoot2, "ACUR_CIG", "ACUR_ECIG", "strength")


#Give the CS-coefficient of the three centrality indices, and explain how
#this measure can be interpreted
corStability(OverallBoot2)
#CS-coefficient for
#betweenness= 0.128 (below 0.25- not good) - should not interpret betweenness values because CS
coefficient is not stable
#closeness= 0.594 this is ok
#strength= 0.594 this is good, above 0.5

centralityPlot(OverallNetworkIF, include=c("Strength", "Closeness","Betweenness"))
centralityPlot(OverallNetworkIF, include=c("Strength", "Closeness"))

centralityTable(OverallNetworkIF)

summary(OverallNetworkIF)
summary(OverallBoot)
OverallBoot
summary(OverallBoot2)
OverallBoot2$type

#Third, test whether network connections (step1) and centrality metrics (step2)
#for different variables significantly differ from each other using bootstrapped difference test
#can do the edge weight difference test and the centrality difference test
```

```
plot(OverallBoot, "edge", plot="difference", onlyNonZero = TRUE, order="sample", labels=FALSE)

differenceTest(OverallBoot, 3, 17, "strength")
differenceTest(OverallBoot, 1, 1, "strength")
plot(OverallBoot, "strength")
#plot(OverallBoot, statistics = c("betweenness", "closeness", "strength"), plot =
#       "difference")
# ^ only gave me strength




#Fifth, compare networks visually AND using network comparison test (NCT)
```

```
#######################################################################
#Accuracy, Stability, and Replicabiity from PNASS PRACTICALS          #
#TRY THIS WITH OVERALL SAMPLE using IsingFit Model (MaleNetworkIF)    #
#######################################################################
library(bootnet)
#Network <- estimateNetwork(bfiData, default="ggmModSelect",
#                           stepwise=FALSE, corMethod="cor")
#plot(Network)

#Perform a non-parametric bootstrap on the estimated network, and
#plot the confidence intervals of the edge-weights
Boot1 <- bootnet(MaleNetworkIF, nCores=8)
#note that the default is not listed here but in the notes, they are
plot(Boot1, order="sample")
plot(Boot1, order="sample", plot="interval", split0=TRUE)
#print on PDF so you can read which edge and how many times
#make sure dimensions are quite long
#was it included in the bootstrap
plot(Boot1, order="sample", labels=FALSE)

#Perform a case-drop bootstrap on the network, and plot the stability
```

```
#of centrality indices. Remember that the default values have now changes
#and do not automatically include stability estimates of 'closeness'
#and 'betweenness'. If you do wish to inspect these, you must include
#statistics = c("strength", "closeness", "betweenness")
Boot2 <- bootnet(MaleNetworkIF, nCores=8, type="case",
            statistics=c("strength", "closeness", "betweenness"))
plot(Boot2, statistics = c("strength", "closeness", "betweenness"))

#Give the CS-coefficient of the three centrality indices, and explain how
#this measure can be interpreted
corStability(Boot2)
#CS-coefficient for
#betweenness= 0.128 (below 0.25- not good)
#closeness= 0.283 (don't think this is good but check)
#strength= 0.517 this is good, before 0.5

centralityPlot(MaleNetworkIF)
centralityPlot(MaleNetworkIF, include=c("Strength", "Closeness","Betweenness"))

centralityPlot(MaleNetworkIF, include=c("Strength", "Closeness"))

centralityTable(MaleNetworkIF)
summary(MaleNetworkIF)

#plot significant differences (alpha=0.05) of edges
plot(Boot1, "edge", plot="difference", onlyNonZero = TRUE,
  order="sample", labels=FALSE)
#plot node strength difference
plot(Boot1, "strength")




##########################################################################
#Accuracy, Stability, and Replicabiity from PNASS PRACTICALS          #
#TRY THIS WITH OVERALL SAMPLE using IsingFit Model (FemaleNetworkIF)    #
##########################################################################
library(bootnet)
#Network <- estimateNetwork(bfiData, default="ggmModSelect",
#                         stepwise=FALSE, corMethod="cor")
#plot(Network)

#Perform a non-parametric bootstrap on the estimated network, and
#plot the confidence intervals of the edge-weights
FemaleBoot1 <- bootnet(FemaleNetworkIF, nCores=8)
#note that the default is not listed here but in the notes, they are
plot(FemaleBoot1, order="sample", labels=FALSE)
plot(FemaleBoot1, order="sample", plot="interval", split0=TRUE)
#print on PDF so you can read which edge and how many times
#make sure dimensions are quite long
#was it included in the bootstrap

#Perform a case-drop bootstrap on the network, and plot the stability
#of centrality indices. Remember that the default values have now changes
#and do not automatically include stability estimates of 'closeness'
```

```
#and 'betweenness'. If you do wish to inspect these, you must include
#statistics = c("strength", "closeness", "betweenness")
FemaleBoot2 <- bootnet(FemaleNetworkIF, nCores=8, type="case",
                  statistics=c("strength", "closeness", "betweenness"))
plot(FemaleBoot2, statistics = c("strength", "closeness", "betweenness"))

#Give the CS-coefficient of the three centrality indices, and explain how
#this measure can be interpreted
corStability(FemaleBoot2)
#CS-coefficient for
#betweenness= 0.05 (below 0.25- not good)
#closeness= 0.05 (don't think this is good but check)
#strength= 0.439 not good

centralityPlot(FemaleNetworkIF)
centralityPlot(FemaleNetworkIF, include=c("Strength", "Closeness","Betweenness"))

centralityPlot(FemaleNetworkIF, include=c("Strength"))

FemaleNetworkIF$graph
centralityTable(FemaleNetworkIF)

summary(FemaleNetworkIF)

#plot significant differences (alpha=0.05) of edges
plot(FemaleBoot1, "edge", plot="difference", onlyNonZero = TRUE,
   order="sample", labels=FALSE)
#plot node strength difference
plot(FemaleBoot1, "strength")
```

R File name: Nodewise Predictability 4162021

```
#Nodewise predictability 4162021

#OVERALL WAVE 1#
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation
Files/Network Chapter")
getwd()
overall<-read.csv("/Users/courtneyblondino/Library/Mobile
Documents/com~apple~CloudDocs/Dissertation Files/Network Chapter/OverallWave1-4142021.csv",
header=T, sep=',')
names(overall)

#rename variables so they look nice on the network
names(overall)[names(overall) == "acur_cignew"] <- "CIG"
names(overall)[names(overall) == "acur_ecignew"] <- "ECIG"
names(overall)[names(overall) == "acur_dualnew"] <- "Dual CIG + ECIG"
names(overall)[names(overall) == "acur_alc"] <- "Alcohol"
names(overall)[names(overall) == "acur_marijuana"] <- "Marijuana"
names(overall)[names(overall) == "acur_painkiller"] <- "PDNP"
names(overall)[names(overall) == "depressed"] <- "Depressed"
names(overall)[names(overall) == "sleeping"] <- "Sleeping"
names(overall)[names(overall) == "anxious"] <- "Anxious"
names(overall)[names(overall) == "ptsd"] <- "PTSD"
names(overall)[names(overall) == "lied"] <- "Lied"
```

```
names(overall)[names(overall) == "attention"] <- "Attention"
names(overall)[names(overall) == "listening"] <- "Listening"
names(overall)[names(overall) == "bully"] <- "Bully"
names(overall)[names(overall) == "fights"] <- "Fights"
names(overall)[names(overall) == "restless"] <- "Restless"
names(overall)[names(overall) == "answered"] <- "Answered"


require(ggplot2)
#require(bootnet)
require(IsingFit)
require(IsingSampler)
require(qgraph)
require(mgm)


#delete obs with missing data
overall_complete_cases <- overall[complete.cases(overall),]

#make into matrix
overall_matrix <- data.matrix(overall_complete_cases)


OverallNetworkMGM <- mgm (data = overall_matrix,
                    type =  c("c", "c","c","c","c","c","c","c","c","c",
                              "c","c","c","c","c","c","c"),
                    level = c(2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2,
                              2, 2, 2),
                    ruleReg = "OR",
                    k = 2,
                    binarySign = TRUE)

OverallNetworkMGM$pairwise


#########
Names<- c("CIG", "ECIG" , "Dual CIG + ECIG", "Alcohol", "Marijuana",  "PDNP",
        "Depressed", "Sleeping" , "Anxious", "Distressed/Past" , "Lied",
        "Attention" , "Listening" , "Bully" ,"Fights"  , "Restless",
        "Answered")

Traits <- rep(c(
  'Substance Use',
  'Negative Affect',
  'Externalizing'
), times=c(6,4,7))

#this won't run...
plot(OverallNetworkMGM,
     layout="spring",
     cut=0,
     theme="colorblind",
     groups=Traits,
     #nodeNames=Names,
     #edge.color="black",
     minimum=0,
```

```
    label.scale.equal=TRUE,
    label.cex= 2.0,
    legend.cex = 0.6,
    title= "Overall Wave 1 Sample")


################################################################################
# Step 2: Predict the given node A by its neighbors - "Making Predictions #
################################################################################


predModel <- predict(OverallNetworkMGM, overall_matrix,
                errorCat = c("CC", "nCC", "CCmarg"))

predModel$errors

#created a columb list with CCmarg values
error_list_me <- list ()
for(i in 1:17) error_list_me[[i]] <- predModel$errors[i,4]
error_list_me

#created the beyond marg values
#beyondmarg_me <- predModel$errors[1:17,2]-predModel$errors[1:17,4]
#beyondmarg_me

#need to combine ccmarg values with beyond marg values in 2 columns, 1 list

beyondmarg_list_me <- list ()
for(i in 1:17) beyondmarg_list_me[[i]] <- (predModel$errors[i,2]-predModel$errors[i,4])
beyondmarg_list_me


#for (i in 1:17) error_list_me[[i]] <- c(predModel$errors[4], beyondmarg_me)
#new_error_list_me  <- c(error_list_me, beyondmarg_list_me)
#new_error_list_me

new_error_list_me <- Map(c, error_list_me, beyondmarg_list_me)
new_error_list_me

color_list_me <- list ()
for(i in 1:17) color_list_me[[i]] <- c("#ffa500", "#ff4300")
color_list_me


#error_list_CC <- list()
#for (i in 1:17) error_list_CC[[i]] <- predModel$errors[i,2]

#error_list_NCC <- list()
#for (i in 1:17) error_list_NCC[[i]] <- predModel$errors[i,3]

#error_list_CCmarg <- list()
#for (i in 1:17) error_list_CCmarg[[i]] <- predModel$errors[i,4]

#color_list <- list ()
#for (i in 1:17) color_list[[i]] <- "#90B4D4"
```

```
#pieColor <- c(rep("#90B4D4", 17), rep("#EB9446", 1))
#pieColor

########################################################################
# Step 3: Quantify how close predictions are to actual values of A #
########################################################################
OverallNetworkMGM$pairwise

layout(t(1))

library(qgraph)
set.seed(1)

OGpred <- qgraph(OverallNetworkMGM$pairwise$wadj, pie = new_error_list_me,
             layout="spring", labels = Names,
             theme="colorblind",
             groups=Traits,
             pieColor = color_list_me,
             label.scale.equal=TRUE,
             label.cex= 4.0,
             legend.cex = 0.4,
             edge.color = OverallNetworkMGM$pairwise$edgecolor,
             curveAll = TRUE, curveDefault = .6,
             cut = 0)
```

SAS File name: READ in W1 4 CLASS 4142021

```
*Read in W1;
*MPLUS Output = wave 1 run 4132021 4 class;
*CSV = = w14class4132021;

libname aim3 "U:\CourtneyBlondino\PhD Epidemiology\April Re Run\LCA Wave 2 and 3 - 4 class -
4142021";

data aim3.w14classprob4142021;
input
W1_ACUR_CIG       W1_ACUR_ECI       W1_ACUR_DUA      W1_ACUR_ALC       W1_ACUR_MAR
W1_ACUR_PAI
W1_DEPRESS        W1_SLEEPING       W1_ANXIOUS       W1_PTSD
W1_LIED           W1_ATTENTIO       W1_LISTENING     W1_BULLY          W1_FIGHTS
W1_RESTLESS       W1_ANSWERED
W1_SEXMALE_
W1_AGE1824_
    W1_AGE2534_
    W1_AGE3544_
    W1_AGE4554_
    W1_AGE5564_
    W1_RACEBL_2
    W1_RACEOT_3
    W1_RACEHI_6
    W1_EDU_1
    W1_EDU_2
    W1_EDU_3
    W1_EDU_4
    W1_INC_1
    W1_INC_2
    W1_INC_3
    W1_INC_4
    W1_SOC_2
```

```
    W1_SOC_3
    W1_SOC_4
    W1_SOC_5
    W1_CPROB1
    W1_CPROB2
    W1_CPROB3
    W1_CPROB4
    W1_C
    W1_WEIGHT
    CASEID;
datalines;
```

*************COPY AND PASTE OUTPUT FROM MPLUS************

```
RUN;
```

SAS File name: LCA W2

```
*In the ICPSR_36498 folder, select DS2001 and open the data file (36498-2001-Data) which is
a SAS Cport Transport file. Once this is open, formats are in, and can begin data management;

libname LCA "U:\CourtneyBlondino\PhD Epidemiology\LCA\Wave 2\Data Management";

*Recoding Missings;
data LCA.W2;
set da36498p2001;


        *Current User Cigarette;
        *R02_AC1002_12M: In past 12 months, smoked a cigarette, even one or two puffs;
        *R02_AC1005: Number of cigarettes smoked in your entire life;
        *R02_AC1003: Now smoke cigarettes;


        *if R02_AC1002_12M = 1 AND R02_AC1005=6 AND R02_AC1003 in (1 2) then acur_cig = 1;
        *else if R02_AC1002_12M = 2 OR R02_AC1003=3 OR (R02_AC1003 in (1,2,.) AND R02_AC1005 in
(1,2,3,4,5)) then acur_cig=0;
        *else if R02_AC1002_12M = . OR R02_AC1003=. OR R02_AC1005=. then acur_cig = .;

        if R02R_A_CUR_ESTD_CIGS=1 then acur_cig=1;
        else if R02R_A_CUR_ESTD_CIGS=2 then acur_cig=0;
        else if R02R_A_CUR_ESTD_CIGS=. then acur_cig=.;

        *Current E-cigarette user;
        *R02_AO9035_01: Ever used the following electronic nicotine product: E-cigarette;
        *R02_AE1100: Ever used e-cigarettes fairly regularly;
        *R02_AO1003C: Now use e-cigarettes;

        *if R02_AO9035_01 = 1 AND R02_AE1100=1 AND R02_AO1003C in (1,2) then acur_ecig = 1;
        *else if R02_AO9035_01=2 OR R02_AE1100 = 2 OR R02_AO1003C = 3 OR (R02_AE1100 in (1,2,.)
AND R02_AO1003C = 2) then acur_ecig=0;
        *else if R02_AO9035_01 = . OR R02_AE1100=. OR R02_AO1003C = . OR R02_AE1100 = . then
acur_ecig = .;

        if R02R_A_CUR_ESTD_ECIG=1 then acur_ecig=1;
        else if R02R_A_CUR_ESTD_ECIG=2 then acur_ecig=0;
        else if R02R_A_CUR_ESTD_ECIG=. then acur_ecig=.;

        ***NOT USING FOR LCA************
        ********************************
        *Current Traditional cigar user

        if R02_AG9003 = 1 AND R02_AG1100TC=1 AND R02_AG1003TC in (1,2) then acur_cigr = 1
        else if R02_AG1001=2 OR R02_AG9002_01 = 2 OR R02_AG9003= 2 OR R02_AG1003TC= 3 OR
(R02_AG1003TC in
        (1,2,.) AND R02_AG1100TC = 2) THEN acur_cigr = 0
        ELSE IF R02_AG1001 = . OR R02_AG9003 = . OR R02_AG1100TC = . OR R02_AG1003TC = . OR
R02_AG9002_01 = . THEN
        acur_cigr = .;

        *Current Cigarillo user
```

```
        IF R02_AG9004=1 AND (R02_AG9009_01=1 OR R02_AG9009_03=1) AND R02_AG1100CG = 1 AND
R02_AG1003CG in
        (1, 2) THEN acur_cigrlo= 1
        ELSE IF R02_AG9004= 2 OR R02_AG1001=2 OR R02_AG9002_02 = 2 OR R02_AG1003CG=3 OR
R02_AG1100CG=2 OR (R02_AG9009_01=2 AND R02_AG9009_03=2) OR ((R02_AG9009_01=1
        OR R02_AG9009_03=1) AND R02_AG1100CG= 2 AND R02_AG1003CG=.) OR ((R02_AG9009_01=1 OR
        R02_AG9009_03=1) AND R02_AG1100CG=. AND R02_AG1003CG= 3) THEN acur_cigrlo= 0
        ELSE IF R02_AG1001 = . OR R02_AG9004 = . OR R02_AG9009_03 = . OR R02_AG9009_01 = . OR
        R02_AG1100CG = . OR R02_AG1003CG = . OR R02_AG9002_02 = . THEN acur_cigrlo = .;

        *Current Filtered Cigar user

        IF R02_AG9004=1 AND R02_AG9009_02=1 AND R02_AG1100FC = 1 AND R02_AG1003FC in (1, 2) THEN
        acur_filcigr= 1
        ELSE IF R02_AG9004= 2 OR R02_AG1001=2 OR R02_AG9002_02 = 2 OR R02_AG1003FC=3 OR
R02_AG1100FC=2 OR R02_AG9009_02=2
        OR (R02_AG9009_02=1 AND R02_AG1100FC= 2 AND R02_AG1003FC=.) OR (R02_AG9009_02=1 AND
R02_AG1100FC=. AND
        R02_AG1003FC= 3) THEN acur_filcigr=0
        ELSE IF R02_AG9004 = . OR R02_AG9009_02 =. OR R02_AG1100FC = . OR R02_AG1003FC = . OR
R02_AG1001 = . OR R02_AG9002_02 = . THEN
        acur_filcigr = .;

        *Current Use Any Cigar/Cigarillo

        IF (acur_cigr = 1 OR acur_cigrlo = 1 OR acur_filcigr = 1) THEN acur_fullcigr = 1
        ELSE IF (acur_cigr = 0 AND acur_cigrlo = 0 AND acur_filcigr= 0) THEN acur_fullcigr = 0
        ELSE IF acur_cigr = . OR acur_cigrlo = . OR acur_filcigr = . THEN acur_fullcigr = .;

        *Current Pipe user

        IF R01_AP1002 = 1 AND R01_AP1100=1 AND R01_AP1003 in (1,2) THEN acur_pipe= 1
        ELSE IF R01_AP1001=2 OR R01_AP1002= 2 OR R01_AP1003= 3 OR (R01_AP1003 in (1,2,.) AND
R01_AP1100 = 2)
        THEN acur_pipe=0
        ELSE IF R01_AP1001 = . OR R01_AP1002 = . OR R01_AP1003 = . OR R01_AP1100 = . THEN
acur_pipe= .;

        *Current Hookah User

        IF R02_AH1002 = 1 AND R02_AH1100=1 AND R02_AH1003 in (1, 2) THEN acur_hook= 1
        ELSE IF R02_AH1001=2 OR R02_AH1002= 2 OR R02_AH1003= 3 OR (R02_AH1003 in (1,2,.) AND
R02_AH1100 = 2)
        THEN acur_hook=0
        ELSE IF R02_AH1002=. OR R02_AH1001=. OR R02_AH1003=. OR R02_AH1100=.
        THEN acur_hook=.;

        *Current User Smokeless

        IF (R02_AS1002_02=1 OR R02_AU1003 in (1,2)) AND R02_AS1100SM = 1 AND R02_AS1003SM in (1,
2) THEN acur_smls= 1
        ELSE IF R02_AS1001=2 OR R02_AS1002_03=1 OR (R02_AS1002_02=2 AND R02_AU1003 in
        (2,3,.)) OR R02_AS1003SM= 3 OR (R02_AS1003SM in (1,2,.) AND R02_AS1100SM = 2) THEN
acur_smls=0
        ELSE IF R02_AS1002_02 = . OR R02_AU1003 = . OR R02_AS1100SM = . OR
        R02_AS1003SM = . OR R02_AS1001 = . THEN acur_smls = .;

        *Current User Snus

        IF R02_AS1002_01=1 AND R02_AU1003 in (2, 3) AND R02_AS1100SU= 1 AND R02_AS1003SU in (1,2)
THEN acur_snus= 1
        ELSE IF R02_AS1001=2 OR R02_AS1002_03=1 OR (R02_AS1002_01=2 AND R02_AS1002_02=1)
        OR (R02_AS1002_01=1 AND R02_AU1003=1) OR (R02_AU1003 in (2,3) AND R02_AS1003SU= 3) OR
(R02_AU1003 in (2,3) AND
        R02_AS1003SU in (1,2,.) AND R02_AS1100SU = 2) THEN acur_snus= 0
        ELSE IF R02_AS1002_01 = . OR R02_AS1002_02 = . OR R02_AS1002_03 = . OR R02_AU1003 = . OR
R02_AS1100SU = . OR R02_AS1003SU =.
        OR R02_AS1001 = . THEN acur_snus=.;

        *Current Use Any Smokeless/Snus
```

```
	IF (acur_smls = 1 OR acur_snus = 1) THEN acur_fullsmkls = 1
	ELSE IF (acur_smls = 0 AND acur_snus = 0) THEN acur_fullsmkls = 0
	ELSE IF acur_smls = . OR acur_snus = . THEN acur_fullsmkls = .;

	*Current User Dissolvable

	IF R02_AD1002 = 1 AND R02_AD1100=1 AND R02_AD1003 in (1,2) THEN acur_diss= 1
	ELSE IF R02_AD1001=2 OR R02_AD1002= 2 OR R02_AD1003= 3 OR (R02_AD1003 in (1,2,.) AND
R02_AD1100 = 2) THEN acur_diss=0
	ELSE IF R02_AD1001 = . OR R02_AD1002 = . OR R02_AD1003 = . OR R02_AD1100 = . THEN
acur_diss = .;



	*******************************
	*******************************
	NEW SUBSTANCES ADDED;

	*Current Use Alcohol;
	*R02_AX0084_12M: In past 12 months, used alcohol, including small tastes or sips
	*R02_AX0673: In past 30 days, used alcohol;
	if R02_AX0084_12M = 1 AND R02_AX0673 = 1 then acur_alc=1;
	else if R02_AX0084_12M = 2 OR R02_AX0673 = 2 then acur_alc=0;
	else if R02_AX0084_12M= . OR R02_AX0673= . then acur_alc=.;

	*Current User Marijuana;
	*R02_AX0085_12M: In past 12 months, used marijuana, hash, THC, grass, pot or weed;
	*R02_AX0675: In past 30 days, used marijuana, hash, THC, grass, pot or weed;
	if R02_AX0675 = 1 then acur_marijuana=1;
	else if  R02_AX0675 in (2  -1) then acur_marijuana=0;
	else if R02_AX0675= . then acur_marijuana=.;

	*Current User Ritalin or Adderall (prescription drugs not prescribed to you);
	*R02_AX0089_12M_01: In past 12 months, used prescription drugs not prescribed to you:
Ritalin or Adderall;
	*R02_AX0676_01: In past 30 days, used: Ritalin or Adderall;
	*if R02_AX0089_12M_01 = 1 AND R02_AX0676_01 = 1 then acur_ritadder=1;
	*else if R02_AX0089_12M_01 = 2 OR R02_AX0676_01 = 2 then acur_ritadder=0;
	*else if R02_AX0089_12M_01 = . OR R02_AX0676_01= . then acur_ritadder=.;

	*Current User Painkillers, Sedatives, or Tranquilizers (prescription drugs not prescribed
to you);
	*R02_AX0089_12M_02: In past 12 months, used prescription drugs not prescribed to you:
Painkillers, sedatives or tranquilizers;
	*R02_AX0676_02: In past 30 days, used: Painkillers, sedatives or tranquilizers;
	if R02_AX0089_12M_02= 1 AND R02_AX0676_02 = 1 then acur_painkiller=1;
	else if R02_AX0089_12M_02 = 2 OR R02_AX0676_02 = 2 then acur_painkiller=0;
	else if R02_AX0089_12M_02 = . OR R02_AX0676_02= . then acur_painkiller=.;

	*Current User Cocaine or Crack
	*R02_AX0220_12M_01: In past 12 months, used substance: Cocaine or crack;
	*R02_AX0676_03: In past 30 days, used: Cocaine or crack;
	*if R02_AX0220_12M_01 = 1 AND R02_AX0676_03 = 1 then acur_cocaine=1;
	*else if R02_AX0220_12M_01 = 2 OR R02_AX0676_03 = 2 then acur_cocaine=0;
	*else if R02_AX0220_12M_01 = . OR R02_AX0676_03= . then acur_cocaine=.;

	*Curent User Meth or Speed
	*R02_AX0220_12M_02: In past 12 months, used substance: Stimulants like methamphetamine or
speed;
	*R02_AX0676_04: In past 30 days, used Stimulants like methamphetamine or speed;
	*if R02_AX0220_12M_02 = 1 AND R02_AX0676_04 = 1 then acur_meth=1;
	*else if R02_AX0220_12M_02 = 2 OR R02_AX0676_04= 2 then acur_meth=0;
	*else if R02_AX0220_12M_02 = . OR R02_AX0676_04= . then acur_meth=.;


	*Current User Heroin, Inhalents, Solvents, Hallucinogens
	*R02_AX0220_12M_03: In past 12 months, used substance: Any other drugs like heroin,
inhalants, solvents or hallucinogens;
	*R02_AX0676_05: In past 30 days, used: Any other drugs like heroin, inhalants, solvents
or hallucinogens;
```

```
        *if R02_AX0220_12M_03 = 1 AND R02_AX0676_05 = 1 then acur_heroinplus=1;
        *else if R02_AX0220_12M_03 = 2 OR R02_AX0676_05 = 2 then acur_heroinplus=0;
        *else if R02_AX0220_12M_03 = . OR R02_AX0676_05 = . then acur_heroinplus=.;


*Create new variables;
*********************
*********************
NEED TO COME BACK AND
ADD SUBSTANCE USE VARS
AND SUICIDE QUESTIONS
TO NEXT DATASET;
********************;

*RACE;
*R02R_A_RACECAT3: DERIVED - Race from the interview (3 levels): 1 = white alone, 2 = black alone,
3 = other;
*R02R_A_HISP: DERIVED - Wave 2 Adult Hispanic Origin (2 levels): 1 = hispanic, 2 = not hispanic;
NUMRACES = 0 ;
if R02R_A_RACECAT3 = 1 then NUMRACES = NUMRACES + 1 ;
if R02R_A_RACECAT3 = 2 then NUMRACES= NUMRACES + 1 ;
if R02R_A_RACECAT3 = 3 then NUMRACES = NUMRACES + 1 ;
if R02R_A_HISP = 1 then NUMRACES = NUMRACES + 1;
if (NUMRACES = 1 and R02R_A_RACECAT3 = 1 AND R02R_A_HISP=2) then R02R_A_ETHRACECAT7= 1 ; *NH
White;
if (NUMRACES = 1 and R02R_A_RACECAT3 = 2 AND R02R_A_HISP=2) then R02R_A_ETHRACECAT7= 2 ; *NH AA;
if (NUMRACES = 1 and R02R_A_RACECAT3 = 3 AND R02R_A_HISP=2) then R02R_A_ETHRACECAT7= 3 ; *NH
Other;
if (NUMRACES = 1 and R02R_A_HISP=1) then R02R_A_ETHRACECAT7= 4; *Hispanic Only;
if (NUMRACES > 1 and R02R_A_HISP=2) then R02R_A_ETHRACECAT7= 5; *NH Multiracial;
if (NUMRACES > 1 and R02R_A_HISP=1) then R02R_A_ETHRACECAT7= 6; *Hispanic Multiracial;
ELSE IF R02R_A_HISP=. OR R02R_A_RACECAT3 = . THEN R02R_A_ETHRACECAT7=.;

*AGE;
*R02R_A_AGECAT7: DERIVED - Age range when interviewed (7 levels);
if R02R_A_AGECAT7=1 then age=1; *18-24;
else if R02R_A_AGECAT7=2 then age=2; *25-34;
else if R02R_A_AGECAT7=3 then age=3; *35-44;
else if R02R_A_AGECAT7=4 then age=4; *45-54;
else if R02R_A_AGECAT7=5 then age=5; *55-64;
else if R02R_A_AGECAT7 in (6 7) then age=6; *65 and older;
else age=.;

*EDUCATION;
*R02R_A_AM0018: DERIVED - Highest grade or level of school completed (6 levels);
if R02R_A_AM0018=1 then education=1; *less than high school;
else if R02R_A_AM0018 in (2 3) then education=2; *GED/high school graduate;
else if R02R_A_AM0018=4 then education=3; *Some college (no degree) or associates degree;
else if R02R_A_AM0018=5 then education=4; *Bachelor's degree;
else if R02R_A_AM0018=6 then education=5; *Advanced degree;
else education=.;

*LIMIT ALL MH VARIABLES TO PAST 30 DAYS;

******INTERNALIZING**********;

*R02_AX0161: Last time you had significant problems with: Feeling very trapped, lonely, sad,
blue, depressed or hopeless about the future;
if R02_AX0161 in (2, 3, 4) then depressed=0;
else if R02_AX0161 in (1) then depressed=1;
else if R02_AX0161 = . then depressed= .;

*R02_AX0162: Last time you had significant problems with: Sleep trouble - such as bad
dreams, sleeping restlessly or falling asleep during the day;
if R02_AX0162 in (2, 3, 4) then sleeping=0;
else if R02_AX0162 in (1) then sleeping=1;
else if R02_AX0162 = . then sleeping=.;

*R02_AX0163: Last time you had significant problems with: Feeling very anxious, nervous,
tense, scared, panicked or like something bad was going to happen;
if R02_AX0163 in (2, 3, 4) then anxious=0;
```

```
else if R02_AX0163 in (1) then anxious=1;
else if R02_AX0163 = . then anxious=.;

*R02_AX0164: Last time you had significant problems with: Becoming very distressed and
upset when something reminded you of the past;
if R02_AX0164 in (2, 3, 4) then ptsd=0;
else if R02_AX0164 in (1) then ptsd=1;
else if R02_AX0164 = . then ptsd=.;

******EXTERNALIZING**********;

*R02_AX0165: Last time you did the following two or more times: Lied or conned to get things
you wanted or to avoid having to do something;
if R02_AX0165 in (2, 3, 4) then lied=0;
else if R02_AX0165 in (1) then lied=1;
else if R02_AX0165 = . then lied=.;

*R02_AX0166: Last time you did the following two or more times: Had a hard time paying
attention at school, work or home;
if R02_AX0166 in (2, 3, 4) then attention=0;
else if R02_AX0166 in (1) then attention=1;
else if R02_AX0166 = . then attention=.;

*R02_AX0167: Last time you did the following two or more times: Had a hard time listening to
instructions at school, work or home;
if R02_AX0167 in (2, 3, 4) then listening=0;
else if R02_AX0167 in (1) then listening=1;
else if R02_AX0167 = . then listening= .;

*R02_AX0168: Last time you did the following two or more times: Were a bully or threatened
other people;
if R02_AX0168 in (2, 3, 4) then bully=0;
else if R02_AX0168 in (1) then bully=1;
else if R02_AX0168 = . then bully= .;

*R02_AX0169: Last time you did the following two or more times: Started physical fights with
other people;
if R02_AX0169 in (2, 3, 4) then fights=0;
else if R02_AX0169 in (1) then fights=1;
else if R02_AX0169 = . then fights= .;

*R02_AX0250: Last time you did the following two or more times: Felt restless or the need to
run around or climb on things;
if R02_AX0250 in (2, 3, 4) then restless=0;
if R02_AX0250 in (1) then restless=1;
else if R02_AX0250 = . then restless=.;

*R02_AX0251: Last time you did the following two or more times: Gave answers before the
other person finished asking the question;
if R02_AX0251 in (2, 3, 4) then answered=0;
if R02_AX0251 in (1) then answered=1;
else if R02_AX0251 = . then answered=.;

******SUBSTANCE USE DISORDERS**********;

*R02_AX0170: Last time that you used alcohol or other drugs weekly or more often;
if R02_AX0170 in (2, 3, 4) then weeklyuse=0;
if R02_AX0170 in (1) then weeklyuse=1;
else if R02_AX0170 = . then weeklyuse=.;

*R02_AX0171: Last time that you spent a lot of time getting alcohol or other drugs;
if R02_AX0171 in (2, 3, 4) then timegetting=0;
if R02_AX0171 in (1) then timegetting=1;
else if R02_AX0171 = . then timegetting=.;

*R02_AX0193: Last time you spent a lot of time using or recovering from alcohol or other drugs;
if R02_AX0193 in (2, 3, 4) then timeusing=0;
if R02_AX0193 in (1) then timeusing=1;
else if R02_AX0193 = . then timeusing=.;

*R02_AX0172: Last time that you kept using alcohol or other drugs even though it was causing
```

```
social problems, leading to fights, or getting you into trouble with other people;
if R02_AX0172 in (2, 3, 4) then socialprob=0;
if R02_AX0172 in (1) then socialprob=1;
else if R02_AX0172 = . then socialprob=.;

*R02_AX0173: Last time that your use of alcohol or other drugs reduced your involvement in
activities at work, school, home or social events;
if R02_AX0173 in (2, 3, 4) then reducedact=0;
if R02_AX0173 in (1) then reducedact=1;
else if R02_AX0173 = . then reducedact=.;

*R02_AX0174: Last time that you had withdrawal problems such as shaky hands, throwing up,
having trouble sitting still or sleeping;
if R02_AX0174 in (2, 3, 4) then withdraw=0;
if R02_AX0174 in (1) then withdraw=1;
else if R02_AX0174 = . then withdraw=.;

*R02_AX0194: Last time you used any alcohol or other drugs to stop being sick or avoid
withdrawal problems;
if R02_AX0194 in (2, 3, 4) then usetoavoid=0;
if R02_AX0194 in (1) then usetoavoid=1;
else if R02_AX0194 = . then usetoavoid=.;

*ALL PAST 30 DAY;
sud_score = sum(weeklyuse, timegetting, timeusing, socialprob, reducedact, withdraw, usetoavoid);

*OLD
*SUD is 3 levels- no/low, moderate, and high;
*if sud_score in (0,1) then sud=0;
*if sud_score in (2,3) then sud=1;
*if sud_score in (4,5,6,7) then sud=2;
*if sud_score = . then sud=.;

*NEW = 1/16/20;
*SUD is 3 levels- no/low, moderate, and high;
if sud_score in (0) then sud=0;
if sud_score in (1,2) then sud=1;
if sud_score in (3,4,5,6,7) then sud=2;
if sud_score = . then sud=.;

*Dichotomize by 0 = no/low, 1 = moderate/high;
*if sud in (0) then sudbin1=0;
*if sud in (1, 2) then sudbin1=1;
*if sud = . then sudbin1 = .;

*Dichotomize by 0 = no/low/moderate, 1 = high;
*if sud in (0,1) then sudbin2=0;
*if sud in (2) then sudbin2=1;
*if sud = . then sudbin2 = .;


*****DUMMY CODING FOR THE COVARIATES******;

IF R02R_A_SEX=1 THEN SEXMALE_1=1;
ELSE SEXMALE_1=0;

IF R02R_A_SEX=2 THEN SEXFEMALE_2=1;
ELSE SEXFEMALE_2=0;

IF age=1 THEN AGE1824_1=1;
ELSE AGE1824_1=0;

IF age=2 THEN AGE2534_2=1;
ELSE AGE2534_2=0;

IF age=3 THEN AGE3544_3=1;
ELSE AGE3544_3=0;

IF age=4 THEN AGE4554_4=1;
ELSE AGE4554_4=0;
```

```
IF age=5 THEN AGE5564_5=1;
ELSE AGE5564_5=0;

IF age=6 THEN AGE65_6=1;
ELSE AGE65_6=0;

IF R02R_A_ETHRACECAT7=1 THEN RACEWH_1=1;
ELSE RACEWH_1=0;

IF R02R_A_ETHRACECAT7=2 THEN RACEBL_2=1;
ELSE RACEBL_2=0;

IF R02R_A_ETHRACECAT7=3 THEN RACEOT_3=1;
ELSE RACEOT_3=0;

IF R02R_A_ETHRACECAT7=6 THEN RACEHI_6=1;
ELSE RACEHI_6=0;

IF education=1 THEN EDU_1=1;
ELSE EDU_1=0;

IF education=2 THEN EDU_2=1;
ELSE EDU_2=0;

IF education=3 THEN EDU_3=1;
ELSE EDU_3=0;

IF education=4 THEN EDU_4=1;
ELSE EDU_4=0;

IF education=5 THEN EDU_5=1;
ELSE EDU_5=0;

IF R02R_A_AM0030=1 THEN INC_1=1;
ELSE INC_1=0;

IF R02R_A_AM0030=2 THEN INC_2=1;
ELSE INC_2=0;

IF R02R_A_AM0030=3 THEN INC_3=1;
ELSE INC_3=0;

IF R02R_A_AM0030=4 THEN INC_4=1;
ELSE INC_4=0;

IF R02R_A_AM0030=5 THEN INC_5=1;
ELSE INC_5=0;

*extremely satisfied =1;
IF R02_AX0092=1 THEN SOC_1=1;
ELSE SOC_1=0;

IF R02_AX0092=2 THEN SOC_2=1;
ELSE SOC_2=0;

IF R02_AX0092=3 THEN SOC_3=1;
ELSE SOC_3=0;

IF R02_AX0092=4 THEN SOC_4=1;
ELSE SOC_4=0;

*not at all satisfied =5;
IF R02_AX0092=5 THEN SOC_5=1;
ELSE SOC_5=0;


        array change _numeric_;
        do over change;
        if change=-97777 then change=.;
        else if change=-99999 then change=.;
```

```
        else if change=-99988 then change=.;
        else if change=-99977 then change=.;
        else if change=-99955 then change=.;
        else if change=-99911 then change=.;
        else if change=-9 then change=.;
        else if change=-8 then change=.;
        else if change=-7 then change=.;
        else if change=-1 then change=.;
        else if change=-5 then change=.;
        end;


run;


*check dummies;
proc freq data=lca.w2;
table acur_cig acur_ecig acur_alc acur_marijuana acur_painkiller;
run;


proc freq data=lca.w2;
table   R02R_A_SEX*SEXMALE_1
               R02R_A_SEX*SEXFEMALE_2
               age*AGE1824_1
               age*AGE2534_2
               age*AGE3544_3
               age*AGE4554_4
               age*AGE5564_5
               age*AGE65_6
               R02R_A_ETHRACECAT7*RACEWH_1
               R02R_A_ETHRACECAT7*RACEBL_2
               R02R_A_ETHRACECAT7*RACEOT_3
               R02R_A_ETHRACECAT7*RACEHI_6
               education*EDU_1
               education*EDU_2
               education*EDU_3
               education*EDU_4
               education*EDU_5
               R02R_A_AM0030*INC_1
               R02R_A_AM0030*INC_2
               R02R_A_AM0030*INC_3
               R02R_A_AM0030*INC_4
               R02R_A_AM0030*INC_5
               R02_AX0092*SOC_1
               R02_AX0092*SOC_2
               R02_AX0092*SOC_3
               R02_AX0092*SOC_4
               R02_AX0092*SOC_5;
run;

*check cig/ecig;
proc freq data=lca.w2;
table R02R_A_CUR_ESTD_CIGS*acur_cig
      R02R_A_CUR_ESTD_ECIG*acur_ecig;
run;
*derived variables have info from wave 1 so use these not the ones I created;

*check other subs;
proc freq data=lca.w2;
table acur_alc acur_marijuana acur_painkiller;;
run;
*marijuana is still weird;

*check sud;
proc freq data=lca.w2;
table sud_score*sud;
run;

*check int/ext/sud;
proc freq data=lca.w2;
```

```
table   depressed sleeping anxious ptsd
         lied attention listening bully fights restless answered
        weeklyuse timegetting timeusing socialprob reducedact withdraw usetoavoid
               sud;
run;

*only select people from wave 1;
proc freq data=lca.w2;
table R02_CONTINUING_ADULT_LD;
run;

data lca.w2contadult;
set lca.w2;
if R02_CONTINUING_ADULT_LD=1;
run;

proc freq data=lca.w2contadult;
table R02_CONTINUING_ADULT_LD;
run;

proc freq data=lca.w2contadult;
table acur_marijuana;
run;

*Identify all variable want to keep;
proc freq data=lca.w2contadult;
table
acur_cig acur_ecig acur_alc acur_marijuana acur_painkiller
                               R02R_A_SEX age R02R_A_ETHRACECAT7 education R02R_A_AM0030
R02_AX0092
                               depressed sleeping anxious ptsd
                               lied attention listening bully fights restless answered
                               weeklyuse timegetting timeusing socialprob reducedact withdraw
usetoavoid
                               sud;
                               run;


*Now limit to the main variables that we want to keep;
data LCA.w2mpluscontadult;
set LCA.w2contadult (keep = caseid personid acur_cig acur_ecig acur_alc acur_marijuana
acur_painkiller
                               R02R_A_SEX age R02R_A_ETHRACECAT7 education R02R_A_AM0030
R02_AX0092
                               SEXMALE_1 SEXFEMALE_2
                               AGE1824_1 AGE2534_2 AGE3544_3 AGE4554_4 AGE5564_5 AGE65_6
                               RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
                               EDU_1 EDU_2 EDU_3 EDU_4 EDU_5
                               INC_1 INC_2 INC_3 INC_4 INC_5
                               SOC_1 SOC_2 SOC_3 SOC_4 SOC_5
                               depressed sleeping anxious ptsd
                               lied attention listening bully fights restless answered
                               weeklyuse timegetting timeusing socialprob reducedact withdraw
usetoavoid
                               sud);
*rename missings;
array change _numeric_;
do over change;
if change =. then change = -99999;
end;
run;


*Check frequencies;
proc contents data=LCA.W2mpluscontadult;
run;



*So far so good, let's pull this datset into MPlus and try LCA;
```

```
***Missing vs nonmissing for W2;
***using contadultc data because it has new tobacco vars;
proc contents data=lca.w2contadultc;
run;

proc freq data=lca.w2contadultc;
table acur_cignew acur_ecignew acur_dualnew
acur_alc acur_marijuana acur_painkiller;
run;


data lca.w2missingtest;
set lca.w2contadultc;
if (acur_cignew=.) or (acur_ecignew=.) or (acur_dualnew=.) or (acur_alc=.) or
(acur_marijuana=.) or (acur_painkiller=.) or
(R02R_A_SEX=.)  or (age=.) or (R02R_A_ETHRACECAT7=.) or (education=.) or (R02R_A_AM0030=.) or
(R02_AX0092=.) or
(depressed=.) or (sleeping=.) or (anxious=.) or (ptsd=.) or
(lied=.) or (attention=.) or (listening=.) or (bully=.) or (fights=.) or
(restless=.) or (answered=.) or
(sud=.) then compare=0;
else compare=1;
run;


ods pdf;
proc freq data=lca.w2missingtest;
table compare;
run;
*complete data/analytic sample (compare = 1) = 21508;
*missing (compare = 0) = 4936;

********************************
*compare missing and nonmissing;
*look at column percent;

*subs;
proc freq data=lca.w2missingtest;
table acur_cignew*compare
      acur_ecignew*compare
      acur_dualnew*compare
      acur_alc*compare
      acur_marijuana*compare
      acur_painkiller*compare/chisq;
run;
*sig diff for all: analytic sample has higher endorsement of all subs;

*demos;
proc freq data=lca.w2missingtest;
table R02R_A_SEX*compare
      age*compare
      R02R_A_ETHRACECAT7*compare
      education*compare
      R02R_A_AM0030*compare
      R02_AX0092*compare/chisq;
run;
*sig difference sex: more males, less women in analytic sample;
*sig difference by age: more in categories 2, 3, 4 (25-54) in analytic sample;
*sig difference by race: more white, less other cats in analytic sample;
*sig difference by edu: higher edu levels in analytic sample;
*sig difference by income: higher income levels in analytic sample;
*sig difference by social: missing had slightly more extremely and very satisfied and also not at
all satisfied;

*internalizing;
proc freq data=lca.w2missingtest;
table depressed*compare
      sleeping*compare
      anxious*compare
      ptsd*compare/chisq;
```

```
run;
*sig diff for all: analytic sample has higher endorsement of all 4 symptoms;

*externalizing;
proc freq data=lca.w2missingtest;
table lied*compare
      attention*compare
      listening*compare
      bully*compare
      fights*compare
      restless*compare
      answered*compare/chisq;
run;
*sig diff for all except bully and fights: all others - analytic sample has higher endorsement of
the other 5 symptoms;
*no sig diff for bully or fights;

*sud;
proc freq data=lca.w2missingtest;
table sud*compare/chisq;
run;
*sig diff: analytic sample has higher endorsement of moderate and high sud severity;
ods pdf close;
```

MPLUS File name: w2 4 class 4142021

```
TITLE: WAVE 2 MODEL 4 CLASS -- APRIL 14 2021;
    DATA: FILE IS w2dataformplus232021_noheader.csv;
    VARIABLE: NAMES ARE CASEID PERSONID weight
        acur_cignew acur_ecignew acur_dualnew
          acur_alc acur_marijuana acur_painkiller
        R02R_A_ETHRACECAT7 age education
        depressed sleeping anxious ptsd
        lied attention listening bully fights      restless answered
        weeklyuse timegetting  timeusing socialprob reducedactwithdraw usetoavoid
                sud
        SEXMALE_1 SEXFEMALE_2
        AGE1824_1 AGE2534_2 AGE3544_3  AGE4554_4 AGE5564_5 AGE65_6
        RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
        EDU_1          EDU_2 EDU_3 EDU_4 EDU_5
        INC_1 INC_2 INC_3 INC_4 INC_5
        SOC_1          SOC_2 SOC_3 SOC_4 SOC_5;
      USEVARIABLES = acur_cignew acur_ecignew acur_dualnew
            acur_alc acur_marijuana acur_painkiller
          depressed sleeping anxious ptsd
          lied attention listening bully fights restless answered;
      IDVARIABLE = CASEID;
      MISSING ARE ALL (-99999);
      CLASSES = c(4);
      CATEGORICAL = acur_cignew acur_ecignew acur_dualnew
            acur_alc acur_marijuana acur_painkiller
          depressed sleeping anxious ptsd
          lied attention listening bully fights restless answered;
      AUXILIARY = SEXMALE_1 (R3STEP)
              AGE1824_1 (R3STEP) AGE2534_2 (R3STEP) AGE3544_3 (R3STEP)
              AGE4554_4 (R3STEP) AGE5564_5 (R3STEP)
              RACEBL_2 (R3STEP) RACEOT_3 (R3STEP) RACEHI_6 (R3STEP)
              EDU_1 (R3STEP) EDU_2 (R3STEP) EDU_3(R3STEP)
                  EDU_4 (R3STEP) INC_1 (R3STEP) INC_2 (R3STEP)
              INC_3 (R3STEP) INC_4 (R3STEP)
```

```
        SOC_2 (R3STEP) SOC_3 (R3STEP) SOC_4 (R3STEP) SOC_5 (R3STEP);
WEIGHT is weight;
ANALYSIS: TYPE = MIXTURE;
         STARTS = 100 10;
         OPTSEED = 991329;
         LRTSTARTS = 0 0 150 40;
SAVEDATA: file is w24classweight414.csv;
        save = Cprob;
OUTPUT: TECH1 TECH8 TECH10 TECH11 TECH14;
```

SAS File name: Read in W2 4 CLASS 4142021

```
*WAVE 2 4 CLASS SOLUTION - import to compare with W1 and W3;
*MPLUS OUTPUT = w2 4 class 4142021;
*CSV = w24classweight414;

libname aim3 "U:\CourtneyBlondino\PhD Epidemiology\April Re Run\LCA Wave 2 and 3 - 4 class -
4142021";

data aim3.w24classprob4142021;
input
 W2_ACUR_CIG
    W2_ACUR_ECI
    W2_ACUR_DUA
    W2_ACUR_ALC
    W2_ACUR_MAR
    W2_ACUR_PAI
    W2_DEPRESS
    W2_SLEEPING
    W2_ANXIOUS
    W2_PTSD
    W2_LIED
    W2_ATTENTIO
    W2_LISTENIN
    W2_BULLY
    W2_FIGHTS
    W2_RESTLESS
    W2_ANSWERED
    W2_SEXMALE_
    W2_AGE1824_
    W2_AGE2534_
    W2_AGE3544_
    W2_AGE4554_
    W2_AGE5564_
    W2_RACEBL_2
    W2_RACEOT_3
    W2_RACEHI_6
    W2_EDU_1
    W2_EDU_2
    W2_EDU_3
    W2_EDU_4
    W2_INC_1
    W2_INC_2
    W2_INC_3
    W2_INC_4
    W2_SOC_2
    W2_SOC_3
    W2_SOC_4
    W2_SOC_5
    W2_CPROB1
    W2_CPROB2
    W2_CPROB3
    W2_CPROB4
    W2_C
    W2_WEIGHT
    CASEID;
datalines;
```

*************COPY AND PASTE OUTPUT FROM MPLUS*********

```
run;
```

SAS File name: LCA W3

```
*In the ICPSR_36498 folder, select DS3001 and open the data file (36498-3001-Data) which is
a SAS Cport Transport file. Once this is open, formats are in, and can begin data management;

libname LCA "U:\CourtneyBlondino\PhD Epidemiology\LCA\Wave 3\Data Management";

*Recoding Missings;
data LCA.W3;
set da36498p3001;


        *Current User Cigarette;
        *R03_AC1002_12M: In past 12 months, smoked a cigarette, even one or two puffs;
        *R03_AC1005: Number of cigarettes smoked in your entire life;
        *R03_AC1003: Now smoke cigarettes;

        *if R03_AC1002_12M = 1 AND R03_AC1005=6 AND R03_AC1003 in (1 2) then acur_cig = 1;
        *else if R03_AC1002_12M = 2 OR R03_AC1003=3 OR (R03_AC1003 in (1,2,.) AND R03_AC1005 in
(1,2,3,4,5)) then acur_cig=0;
        *else if R03_AC1002_12M = . OR R03_AC1003=. OR R03_AC1005=. then acur_cig = .;

        if R03R_A_CUR_ESTD_CIGS=1 then acur_cig=1;
        else if R03R_A_CUR_ESTD_CIGS=2 then acur_cig=0;
        else if R03R_A_CUR_ESTD_CIGS=. then acur_cig=.;

        *Current E-cigarette user;
        *R03_AV1002_12M: Ever used the following electronic nicotine product: E-cigarette;
        *R03_AV1100: Ever used e-cigarettes fairly regularly;
        *R03_AV1003EC: Now use e-cigarettes;

        *if R03_AV1002_12M = 1 AND R03_AV1100=1 AND R03_AV1003EC in (1,2) then acur_ecig = 1;
        *else if R03_AV1002_12M=2 OR R03_AV1100 = 2 OR R03_AV1003EC = 3 OR (R03_AV1100 in (1,2,.)
AND R03_AV1003EC = 2) then acur_ecig=0;
        *else if R03_AV1002_12M= . OR R03_AV1100=. OR R03_AV1003EC = . OR R03_AV1100 = . then
acur_ecig = .;

        if R03R_A_CUR_ESTD_EPRODS=1 then acur_ecig=1;
        else if R03R_A_CUR_ESTD_EPRODS=2 then acur_ecig=0;
        else if R03R_A_CUR_ESTD_EPRODS=. then acur_ecig=.;

        *Current Use Alcohol;
        *R03_AX0084_12M: In past 12 months, used alcohol, including small tastes or sips
        *R03_AX0673: In past 30 days, used alcohol;

        if R03_AX0084_12M = 1 AND R03_AX0673 = 1 then acur_alc=1;
        else if R03_AX0084_12M = 2 OR R03_AX0673 = 2 then acur_alc=0;
        else if R03_AX0084_12M = . OR R03_AX0673 = . then acur_alc=.;

        *Current User Marijuana;
        *R03_AX0085_12M: In past 12 months, used marijuana, hash, THC, grass, pot or weed;
        *R03_AX0675: In past 30 days, used marijuana, hash, THC, grass, pot or weed;

        if R03_AX0675 = 1 then acur_marijuana=1;
        else if R03_AX0675 in (-1, 2) then acur_marijuana=0;
        else if R03_AX0675 = . then acur_marijuana=.;

        *Current User Painkillers, Sedatives, or Tranquilizers (prescription drugs not prescribed
to you);
        *R03_AX0089_12M_02: In past 12 months, used prescription drugs not prescribed to you:
Painkillers, sedatives or tranquilizers;
        *R03_AX0676_02: In past 30 days, used: Painkillers, sedatives or tranquilizers;

        if R03_AX0089_12M_02 = 1 AND R03_AX0676_02 = 1 then acur_painkiller=1;
        else if R03_AX0089_12M_02 = 2 OR R03_AX0676_02 = 2 then acur_painkiller=0;
        else if R03_AX0089_12M_02 = . OR R03_AX0676_02 = . then acur_painkiller=.;
```

```
*RACE;
*R03R_A_RACECAT3: DERIVED - Race from the interview (3 levels): 1 = white alone, 2 = black alone,
3 = other;
*R03R_A_HISP: DERIVED - Wave 2 Adult Hispanic Origin (2 levels): 1 = hispanic, 2 = not hispanic;
NUMRACES = 0 ;
if R03R_A_RACECAT3 = 1 then NUMRACES = NUMRACES + 1 ;
if R03R_A_RACECAT3 = 2 then NUMRACES= NUMRACES + 1 ;
if R03R_A_RACECAT3 = 3 then NUMRACES = NUMRACES + 1 ;
if R03R_A_HISP = 1 then NUMRACES = NUMRACES + 1;
if (NUMRACES = 1 and R03R_A_RACECAT3 = 1 AND R03R_A_HISP=2) then R03R_A_ETHRACECAT7= 1 ; *NH
White;
if (NUMRACES = 1 and R03R_A_RACECAT3 = 2 AND R03R_A_HISP=2) then R03R_A_ETHRACECAT7= 2 ; *NH AA;
if (NUMRACES = 1 and R03R_A_RACECAT3 = 3 AND R03R_A_HISP=2) then R03R_A_ETHRACECAT7= 3 ; *NH
Other;
if (NUMRACES = 1 and R03R_A_HISP=1) then R03R_A_ETHRACECAT7= 4; *Hispanic Only;
if (NUMRACES > 1 and R03R_A_HISP=2) then R03R_A_ETHRACECAT7= 5; *NH Multiracial;
if (NUMRACES > 1 and R03R_A_HISP=1) then R03R_A_ETHRACECAT7= 6; *Hispanic Multiracial;
ELSE IF R03R_A_HISP=. OR R03R_A_RACECAT3 = . THEN R03R_A_ETHRACECAT7=.;

*AGE;
*R03R_A_AGECAT7: DERIVED - Age range when interviewed (7 levels);
if R03R_A_AGECAT7=1 then age=1; *18-24;
else if R03R_A_AGECAT7=2 then age=2; *25-34;
else if R03R_A_AGECAT7=3 then age=3; *35-44;
else if R03R_A_AGECAT7=4 then age=4; *45-54;
else if R03R_A_AGECAT7=5 then age=5; *55-64;
else if R03R_A_AGECAT7 in (6 7) then age=6; *65 and older;
else age=.;

*EDUCATION;
*R03R_A_AM0018: DERIVED - Highest grade or level of school completed (6 levels);
if R03R_A_AM0018=1 then education=1; *less than high school;
else if R03R_A_AM0018 in (2 3) then education=2; *GED/high school graduate;
else if R03R_A_AM0018=4 then education=3; *Some college (no degree) or associates degree;
else if R03R_A_AM0018=5 then education=4; *Bachelor's degree;
else if R03R_A_AM0018=6 then education=5; *Advanced degree;
else education=.;

******INTERNALIZING*********;

*R03_AX0161: Last time you had significant problems with: Feeling very trapped, lonely, sad,
blue, depressed or hopeless about the future;
if R03_AX0161 in (2, 3, 4) then depressed=0;
else if R03_AX0161 in (1) then depressed=1;
else if R03_AX0161 = . then depressed= .;

*R03_AX0162: Last time you had significant problems with: Sleep trouble - such as bad
dreams, sleeping restlessly or falling asleep during the day;
if R03_AX0162 in (2, 3, 4) then sleeping=0;
else if R03_AX0162 in (1) then sleeping=1;
else if R03_AX0162 = . then sleeping=.;

*R03_AX0163: Last time you had significant problems with: Feeling very anxious, nervous,
tense, scared, panicked or like something bad was going to happen;
if R03_AX0163 in (2, 3, 4) then anxious=0;
else if R03_AX0163 in (1) then anxious=1;
else if R03_AX0163 = . then anxious=.;

*R03_AX0164: Last time you had significant problems with: Becoming very distressed and
upset when something reminded you of the past;
if R03_AX0164 in (2, 3, 4) then ptsd=0;
else if R03_AX0164 in (1) then ptsd=1;
else if R03_AX0164 = . then ptsd=.;


******EXTERNALIZING*********;

*R03_AX0165: Last time you did the following two or more times: Lied or conned to get things
you wanted or to avoid having to do something;
```

```
if R03_AX0165 in (2, 3, 4) then lied=0;
else if R03_AX0165 in (1) then lied=1;
else if R03_AX0165 = . then lied=.;

*R03_AX0166: Last time you did the following two or more times: Had a hard time paying
attention at school, work or home;
if R03_AX0166 in (2, 3, 4) then attention=0;
else if R03_AX0166 in (1) then attention=1;
else if R03_AX0166 = . then attention=.;

*R03_AX0167: Last time you did the following two or more times: Had a hard time listening to
instructions at school, work or home;
if R03_AX0167 in (2, 3, 4) then listening=0;
else if R03_AX0167 in (1) then listening=1;
else if R03_AX0167 = . then listening= .;

*R03_AX0168: Last time you did the following two or more times: Were a bully or threatened
other people;
if R03_AX0168 in (2, 3, 4) then bully=0;
else if R03_AX0168 in (1) then bully=1;
else if R03_AX0168 = . then bully= .;

*R03_AX0169: Last time you did the following two or more times: Started physical fights with
other people;
if R03_AX0169 in (2, 3, 4) then fights=0;
else if R03_AX0169 in (1) then fights=1;
else if R03_AX0169 = . then fights= .;

*R03_AX0250: Last time you did the following two or more times: Felt restless or the need to
run around or climb on things;
if R03_AX0250 in (2, 3, 4) then restless=0;
if R03_AX0250 in (1) then restless=1;
else if R03_AX0250 = . then restless=.;

*R03_AX0251: Last time you did the following two or more times: Gave answers before the
other person finished asking the question;
if R03_AX0251 in (2, 3, 4) then answered=0;
if R03_AX0251 in (1) then answered=1;
else if R03_AX0251 = . then answered=.;


******SUBSTANCE USE DISORDERS**********;

*R03_AX0170: Last time that you used alcohol or other drugs weekly or more often;
if R03_AX0170 in (2, 3, 4) then weeklyuse=0;
if R03_AX0170 in (1) then weeklyuse=1;
else if R03_AX0170 = . then weeklyuse=.;

*R03_AX0171: Last time that you spent a lot of time getting alcohol or other drugs;
if R03_AX0171 in (2, 3, 4) then timegetting=0;
if R03_AX0171 in (1) then timegetting=1;
else if R03_AX0171 = . then timegetting=.;

*R03_AX0193: Last time you spent a lot of time using or recovering from alcohol or other drugs;
if R03_AX0193 in (2, 3, 4) then timeusing=0;
if R03_AX0193 in (1) then timeusing=1;
else if R03_AX0193 = . then timeusing=.;

*R03_AX0172: Last time that you kept using alcohol or other drugs even though it was causing
social problems, leading to fights, or getting you into trouble with other people;
if R03_AX0172 in (2, 3, 4) then socialprob=0;
if R03_AX0172 in (1) then socialprob=1;
else if R03_AX0172 = . then socialprob=.;

*R03_AX0173: Last time that your use of alcohol or other drugs reduced your involvement in
activities at work, school, home or social events;
if R03_AX0173 in (2, 3, 4) then reducedact=0;
if R03_AX0173 in (1) then reducedact=1;
else if R03_AX0173 = . then reducedact=.;

*R03_AX0174: Last time that you had withdrawal problems such as shaky hands, throwing up,
```

```
having trouble sitting still or sleeping;
if R03_AX0174 in (2, 3, 4) then withdraw=0;
if R03_AX0174 in (1) then withdraw=1;
else if R03_AX0174 = . then withdraw=.;

*R03_AX0194: Last time you used any alcohol or other drugs to stop being sick or avoid
withdrawal problems;
if R03_AX0194 in (2, 3, 4) then usetoavoid=0;
if R03_AX0194 in (1) then usetoavoid=1;
else if R03_AX0194 = . then usetoavoid=.;

*ALL PAST 30 DAY;
sud_score = sum(weeklyuse, timegetting, timeusing, socialprob, reducedact, withdraw, usetoavoid);

*SUD is 3 levels- no/low, moderate, and high;
if sud_score in (0) then sud=0;
if sud_score in (1,2) then sud=1;
if sud_score in (3,4,5,6,7) then sud=2;
if sud_score = . then sud=.;

*****DUMMY CODING FOR THE COVARIATES******;

IF R03R_A_SEX=1 THEN SEXMALE_1=1;
ELSE SEXMALE_1=0;

IF R03R_A_SEX=2 THEN SEXFEMALE_2=1;
ELSE SEXFEMALE_2=0;

IF age=1 THEN AGE1824_1=1;
ELSE AGE1824_1=0;

IF age=2 THEN AGE2534_2=1;
ELSE AGE2534_2=0;

IF age=3 THEN AGE3544_3=1;
ELSE AGE3544_3=0;

IF age=4 THEN AGE4554_4=1;
ELSE AGE4554_4=0;

IF age=5 THEN AGE5564_5=1;
ELSE AGE5564_5=0;

IF age=6 THEN AGE65_6=1;
ELSE AGE65_6=0;

IF R03R_A_ETHRACECAT7=1 THEN RACEWH_1=1;
ELSE RACEWH_1=0;

IF R03R_A_ETHRACECAT7=2 THEN RACEBL_2=1;
ELSE RACEBL_2=0;

IF R03R_A_ETHRACECAT7=3 THEN RACEOT_3=1;
ELSE RACEOT_3=0;

IF R03R_A_ETHRACECAT7=6 THEN RACEHI_6=1;
ELSE RACEHI_6=0;

IF education=1 THEN EDU_1=1;
ELSE EDU_1=0;

IF education=2 THEN EDU_2=1;
ELSE EDU_2=0;

IF education=3 THEN EDU_3=1;
ELSE EDU_3=0;

IF education=4 THEN EDU_4=1;
ELSE EDU_4=0;

IF education=5 THEN EDU_5=1;
```

```
ELSE EDU_5=0;

IF R03R_A_AM0030=1 THEN INC_1=1;
ELSE INC_1=0;

IF R03R_A_AM0030=2 THEN INC_2=1;
ELSE INC_2=0;

IF R03R_A_AM0030=3 THEN INC_3=1;
ELSE INC_3=0;

IF R03R_A_AM0030=4 THEN INC_4=1;
ELSE INC_4=0;

IF R03R_A_AM0030=5 THEN INC_5=1;
ELSE INC_5=0;

*extremely satisfied =1;
IF R03_AX0092=1 THEN SOC_1=1;
ELSE SOC_1=0;

IF R03_AX0092=2 THEN SOC_2=1;
ELSE SOC_2=0;

IF R03_AX0092=3 THEN SOC_3=1;
ELSE SOC_3=0;

IF R03_AX0092=4 THEN SOC_4=1;
ELSE SOC_4=0;

*not at all satisfied =5;
IF R03_AX0092=5 THEN SOC_5=1;
ELSE SOC_5=0;


        array change _numeric_;
        do over change;
        if change=-97777 then change=.;
        else if change=-99999 then change=.;
        else if change=-99988 then change=.;
        else if change=-99977 then change=.;
        else if change=-99955 then change=.;
        else if change=-99911 then change=.;
        else if change=-9 then change=.;
        else if change=-8 then change=.;
        else if change=-7 then change=.;
        else if change=-1 then change=.;
        else if change=-5 then change=.;
        end;

run;



*********confirming recodes worked;

*check dummies;
proc freq data=lca.w3;
table   R03R_A_SEX*SEXMALE_1
               R03R_A_SEX*SEXFEMALE_2
               age*AGE1824_1
               age*AGE2534_2
               age*AGE3544_3
               age*AGE4554_4
               age*AGE5564_5
               age*AGE65_6
               R03R_A_ETHRACECAT7*RACEWH_1
               R03R_A_ETHRACECAT7*RACEBL_2
               R03R_A_ETHRACECAT7*RACEOT_3
               R03R_A_ETHRACECAT7*RACEHI_6
               education*EDU_1
```

```
               education*EDU_2
               education*EDU_3
               education*EDU_4
               education*EDU_5
               R03R_A_AM0030*INC_1
               R03R_A_AM0030*INC_2
               R03R_A_AM0030*INC_3
               R03R_A_AM0030*INC_4
               R03R_A_AM0030*INC_5
               R03_AX0092*SOC_1
               R03_AX0092*SOC_2
               R03_AX0092*SOC_3
               R03_AX0092*SOC_4
               R03_AX0092*SOC_5;
run;

*check sud;
proc freq data=lca.w3;
table sud_score*sud;
run;

*int/ext/sud;
proc freq data=lca.w3;
table  acur_alc acur_marijuana acur_painkiller
               depressed sleeping anxious ptsd
               lied attention listening bully fights restless answered
               weeklyuse timegetting timeusing socialprob reducedact withdraw usetoavoid
               sud;
run;

proc freq data=lca.w3;
table R03R_A_CUR_ESTD_CIGS*acur_cig
      R03R_A_CUR_ESTD_EPRODS*acur_ecig;
run;

proc freq data=lca.w3;
table acur_alc acur_marijuana acur_painkiller;
run;
*marijuana is weird again for past 30 day--
because R03_AX0675 is either smoked traditional cigar, cigarillo, or filtered cigar with
marijuana in the past 12 months
OR
have you used marijuana in the past 12 months;


*only select people from wave 1 and wave 2;
proc freq data=lca.w3;
table R03_ADULTTYPE;
run;

data lca.w3contadult;
set lca.w3;
if R03_ADULTTYPE=1;
run;

proc freq data=lca.w3contadult;
table R03_ADULTTYPE;
run;

proc freq data=lca.w3contadult;
table acur_cig acur_ecig acur_alc acur_marijuana acur_painkiller
         R03R_A_SEX age R03R_A_ETHRACECAT7 education R03R_A_AM0030 R03_AX0092
         depressed sleeping anxious ptsd
                              lied attention listening bully fights restless answered
                              sud;
                              run;


*Identify all variable want to keep;
```

```
*Now limit to the main variables that we want to keep;
data LCA.W3mpluscontadult;
set LCA.W3contadult (keep = caseid personid acur_cig acur_ecig acur_alc acur_marijuana
acur_painkiller
                               R03R_A_SEX age R03R_A_ETHRACECAT7 education R03R_A_AM0030
R03_AX0092
                               SEXMALE_1 SEXFEMALE_2
                 AGE1824_1 AGE2534_2 AGE3544_3 AGE4554_4 AGE5564_5 AGE65_6
                               RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
                               EDU_1 EDU_2 EDU_3 EDU_4 EDU_5
                               INC_1 INC_2 INC_3 INC_4 INC_5
                               SOC_1 SOC_2 SOC_3 SOC_4 SOC_5
                               depressed sleeping anxious ptsd
                               lied attention listening bully fights restless answered
                               weeklyuse timegetting timeusing socialprob reducedact withdraw
usetoavoid
                               sud);
*rename missings;
array change _numeric_;
do over change;
if change =. then change = -99999;
end;
run;

*Check frequencies;
proc contents data=LCA.W3mpluscontadult;
run;




***Missing vs nonmissing for W3;
***using contadultc data because it has new tobacco vars;
proc contents data=lca.w3contadultc;
run;

proc freq data=lca.w3contadultc;
table acur_cignew acur_ecignew acur_dualnew
acur_alc acur_marijuana acur_painkiller;
run;


data lca.w3missingtest;
set lca.w3contadultc;
if (acur_cignew=.) or (acur_ecignew=.) or (acur_dualnew=.) or (acur_alc=.) or
(acur_marijuana=.) or (acur_painkiller=.) or
(R03R_A_SEX=.)  or (age=.) or (R03R_A_ETHRACECAT7=.) or (education=.) or (R03R_A_AM0030=.) or
(R03_AX0092=.) or
(depressed=.) or (sleeping=.) or (anxious=.) or (ptsd=.) or
(lied=.) or (attention=.) or (listening=.) or (bully=.) or (fights=.) or
(restless=.) or (answered=.) or
(sud=.) then compare=0;
else compare=1;
run;


ods pdf;
proc freq data=lca.w3missingtest;
table compare;
run;
*complete data/analytic sample (compare = 1) = 21628;
*missing (compare = 0) = 4611;

********************************
*compare missing and nonmissing;
*look at column percent;

*subs;
proc freq data=lca.w3missingtest;
```

```
table acur_cignew*compare
      acur_ecignew*compare
      acur_dualnew*compare
      acur_alc*compare
      acur_marijuana*compare
      acur_painkiller*compare/chisq;
run;
*sig diff for all: analytic sample has higher endorsement of all subs;

*demos;
proc freq data=lca.w3missingtest;
table R03R_A_SEX*compare
      age*compare
      R03R_A_ETHRACECAT7*compare
      education*compare
      R03R_A_AM0030*compare
      R03_AX0092*compare/chisq;
run;
*sig difference sex: more males, less women in analytic sample;
*sig difference by age: more in categories 2, 3, 4 (25-54) in analytic sample;
*sig difference by race: more white, less other cats in analytic sample;
*sig difference by edu: higher edu levels in analytic sample;
*sig difference by income: higher income levels in analytic sample;
*sig difference by social: missing had more extremely and also not at all satisfied;

*internalizing;
proc freq data=lca.w3missingtest;
table depressed*compare
      sleeping*compare
      anxious*compare
      ptsd*compare/chisq;
run;
*sig diff for all: analytic sample has higher endorsement of all 4 symptoms;

*externalizing;
proc freq data=lca.w3missingtest;
table lied*compare
      attention*compare
      listening*compare
      bully*compare
      fights*compare
      restless*compare
      answered*compare/chisq;
run;
*sig diff for all except bully and fights: all others - analytic sample has higher endorsement of
the other 5 symptoms;
*no sig diff for bully or fights;

*sud;
proc freq data=lca.w3missingtest;
table sud*compare/chisq;
run;
*sig diff: analytic sample has higher endorsement of moderate and high sud severity;
ods pdf close;
```

MPLUS File name: w3 4 class 4142021

```
TITLE: WAVE 3 MODEL 4 CLASS -- APRIL 14 2021;
      DATA: FILE IS w3dataformplus332021_noheader.csv;
      VARIABLE: NAMES ARE CASEID PERSONID weight
          acur_cignew acur_ecignew acur_dualnew
            acur_alc acur_marijuana acur_painkiller
          R03R_A_ETHRACECAT7 age education
          depressed sleeping anxious ptsd
          lied attention listening bully fights      restless answered
```

```
    weeklyuse timegetting  timeusing socialprob reducedactwithdraw usetoavoid
            sud
    SEXMALE_1 SEXFEMALE_2
    AGE1824_1 AGE2534_2 AGE3544_3  AGE4554_4 AGE5564_5 AGE65_6
    RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
    EDU_1           EDU_2 EDU_3 EDU_4 EDU_5
    INC_1 INC_2 INC_3 INC_4 INC_5
    SOC_1           SOC_2 SOC_3 SOC_4 SOC_5;
  USEVARIABLES = acur_cignew acur_ecignew acur_dualnew
         acur_alc acur_marijuana acur_painkiller
      depressed sleeping anxious ptsd
      lied attention listening bully fights restless answered;
  IDVARIABLE = CASEID;
  MISSING ARE ALL (-99999);
  CLASSES = c(5);
  CATEGORICAL = acur_cignew acur_ecignew acur_dualnew
        acur_alc acur_marijuana acur_painkiller
      depressed sleeping anxious ptsd
      lied attention listening bully fights restless answered;
  AUXILIARY = SEXMALE_1 (R3STEP)
         AGE1824_1 (R3STEP) AGE2534_2 (R3STEP) AGE3544_3 (R3STEP)
         AGE4554_4 (R3STEP) AGE5564_5 (R3STEP)
         RACEBL_2 (R3STEP) RACEOT_3 (R3STEP) RACEHI_6 (R3STEP)
         EDU_1 (R3STEP) EDU_2 (R3STEP) EDU_3(R3STEP)
             EDU_4 (R3STEP) INC_1 (R3STEP) INC_2 (R3STEP)
         INC_3 (R3STEP) INC_4 (R3STEP)
        SOC_2 (R3STEP) SOC_3 (R3STEP) SOC_4 (R3STEP) SOC_5 (R3STEP);
WEIGHT is weight;
ANALYSIS: TYPE = MIXTURE;
        STARTS = 100 10;
        OPTSEED = 991329;
        LRTSTARTS = 0 0 150 40;
SAVEDATA: file is w34classweight414.csv;
      save = Cprob;
OUTPUT: TECH1 TECH8 TECH10 TECH11 TECH14;
```

SAS File name: Read in W3 4 CLASS 4142021

```
*WAVE 3 4 CLASS SOLUTION - import to compare with W1 and W2;
*MPLUS OUTPUT = w3 4 class 4152021;
*CSV = w34classweight415;

libname aim3 "U:\CourtneyBlondino\PhD Epidemiology\April Re Run\LCA Wave 2 and 3 - 4 class -
4142021";

data aim3.w34classprob4152021;
input
   W3_ACUR_CIG
    W3_ACUR_ECI
    W3_ACUR_DUA
    W3_ACUR_ALC
    W3_ACUR_MAR
    W3_ACUR_PAI
    W3_DEPRESS
    W3_SLEEPING
    W3_ANXIOUS
```

```
    W3_PTSD
    W3_LIED
    W3_ATTENTIO
    W3_LISTENIN
    W3_BULLY
    W3_FIGHTS
    W3_RESTLESS
    W3_ANSWERED
    W3_SEXMALE_
    W3_AGE1824_
    W3_AGE2534_
    W3_AGE3544_
    W3_AGE4554_
    W3_AGE5564_
    W3_RACEBL_2
    W3_RACEOT_3
    W3_RACEHI_6
    W3_EDU_1
    W3_EDU_2
    W3_EDU_3
    W3_EDU_4
    W3_INC_1
    W3_INC_2
    W3_INC_3
    W3_INC_4
    W3_SOC_2
    W3_SOC_3
    W3_SOC_4
    W3_SOC_5
    W3_CPROB1
    W3_CPROB2
    W3_CPROB3
    W3_CPROB4
    W3_C
    W3_WEIGHT
    CASEID;
datalines;
```

*****************COPY AND PASTE OUTPUT FROM MPLUS***********

```
run;
```

SAS File name: LCA Comparisons, Transition Tables 4142021

```
*Merge the datasets to look at transition tables;

libname aim3 "U:\CourtneyBlondino\PhD Epidemiology\April Re Run\LCA Wave 2 and 3 - 4 class -
4142021";

*start by checking freqs;

*Wave 1;
proc surveyfreq data=aim3.w14classprob4142021;
table
W1_ACUR_CIG      W1_ACUR_ECI       W1_ACUR_DUA       W1_ACUR_ALC       W1_ACUR_MAR
W1_ACUR_PAI
W1_DEPRESS       W1_SLEEPING       W1_ANXIOUS        W1_PTSD
W1_LIED          W1_ATTENTIO       W1_LISTENING      W1_BULLY          W1_FIGHTS
W1_RESTLESS        W1_ANSWERED   / row chisq (secondorder);
weight W1_weight;
run;

*Wave 2;
proc surveyfreq data=aim3.w24classprob4142021;
table
W2_ACUR_CIG      W2_ACUR_ECI      W2_ACUR_DUA       W2_ACUR_ALC       W2_ACUR_MAR
W2_ACUR_PAI
W2_DEPRESS       W2_SLEEPING      W2_ANXIOUS        W2_PTSD
W2_LIED          W2_ATTENTIO      W2_LISTENIN      W2_BULLY         W2_FIGHTS      W2_RESTLESS
W2_ANSWERED / row chisq (secondorder);
weight W2_weight;
```

```
run;

*Wave 3;
proc surveyfreq data=aim3.W34classprob4152021;
table
W3_ACUR_CIG      W3_ACUR_ECI      W3_ACUR_DUA       W3_ACUR_ALC     W3_ACUR_MAR    W3_ACUR_PAI
W3_DEPRESS       W3_SLEEPING      W3_ANXIOUS        W3_PTSD
W3_LIED          W3_ATTENTIO      W3_LISTENIN      W3_BULLY        W3_FIGHTS      W3_RESTLESS
W3_ANSWERED
/ row chisq (secondorder);
weight W3_weight;
run;

proc surveyfreq data=aim3.W34classprob4152021;
table
w3_C
/ row chisq (secondorder);
weight W3_weight;
run;

*then sort by caseid;
proc sort data=aim3.w14classprob4142021;
by caseid;
run;

proc sort data=aim3.w24classprob4142021;
by caseid;
run;

proc sort data=aim3.W34classprob4152021;
by caseid;
run;

********************need to add in R03_A_AWGT to use when running weighted transition tables;
proc contents data=work.Da36498p3101;
*table R03_A_AWGT;
run;
data aim3.w3allweights;
set WORK.DA36498P3101
(keep = caseid R03_A_AWGT);
run;
proc sort data=aim3.w3allweights;
by caseid;
run;



*then merge;
data aim3.master;
merge aim3.w14classprob4142021 aim3.w24classprob4142021 aim3.W34classprob4152021
aim3.w3allweights;
by caseid;
*array change _numeric_;
*do over change;
*if change =-99999 then change = .;
*end;
run;

proc contents data=aim3.master;
run;

proc print data=aim3.master (obs=20);
run;

*make transition tables;

ods pdf;
*W1 vs W2;
proc freq data=aim3.master;
table w1_c*w2_c;
run;
```

```
*W2 vs W3;
proc freq data=aim3.master;
table w2_c*w3_c;
run;

*W1 vs W3;
proc freq data=aim3.master;
table w1_c*w3_c;
run;
ods pdf close;


*make transition tables - with wave 3 all weights;

ods pdf;
*W1 vs W2;
proc surveyfreq data=aim3.master;
table w1_c*w2_c
/ row chisq (secondorder);
weight R03_A_AWGT;
run;

*W2 vs W3;
proc surveyfreq data=aim3.master;
table w2_c*w3_c
/ row chisq (secondorder);
weight R03_A_AWGT;
run;

*W1 vs W3;
proc surveyfreq data=aim3.master;
table w1_c*w3_c
/ row chisq (secondorder);
weight R03_A_AWGT;
run;
ods pdf close;



*Try to get item response pattern for each comorbidity class for each wave;
*Table 3.7 in Nylund dissertation;






*tetrachoric correlations by wave;
*4/28/2021;
ods pdf;
proc contents data=aim3.master;
run;

******W1 VS W2;
proc freq data=aim3.master;
table
W1_ACUR_CIG*W2_ACUR_CIG
W1_ACUR_CIG*W2_ACUR_ECI
W1_ACUR_CIG*W2_ACUR_DUA
W1_ACUR_CIG*W2_ACUR_ALC
W1_ACUR_CIG*W2_ACUR_MAR
W1_ACUR_CIG*W2_ACUR_PAI
W1_ACUR_CIG*W2_DEPRESS
W1_ACUR_CIG*W2_SLEEPING
W1_ACUR_CIG*W2_ANXIOUS
W1_ACUR_CIG*W2_PTSD
W1_ACUR_CIG*W2_LIED
W1_ACUR_CIG*W2_ATTENTIO
W1_ACUR_CIG*W2_LISTENIN
W1_ACUR_CIG*W2_BULLY
```

W1_ACUR_CIG*W2_FIGHTS
W1_ACUR_CIG*W2_RESTLESS
W1_ACUR_CIG*W2_ANSWERED

W1_ACUR_ECI*W2_ACUR_CIG
W1_ACUR_ECI*W2_ACUR_ECI
W1_ACUR_ECI*W2_ACUR_DUA
W1_ACUR_ECI*W2_ACUR_ALC
W1_ACUR_ECI*W2_ACUR_MAR
W1_ACUR_ECI*W2_ACUR_PAI
W1_ACUR_ECI*W2_DEPRESS
W1_ACUR_ECI*W2_SLEEPING
W1_ACUR_ECI*W2_ANXIOUS
W1_ACUR_ECI*W2_PTSD
W1_ACUR_ECI*W2_LIED
W1_ACUR_ECI*W2_ATTENTIO
W1_ACUR_ECI*W2_LISTENIN
W1_ACUR_ECI*W2_BULLY
W1_ACUR_ECI*W2_FIGHTS
W1_ACUR_ECI*W2_RESTLESS
W1_ACUR_ECI*W2_ANSWERED

W1_ACUR_DUA*W2_ACUR_CIG
W1_ACUR_DUA*W2_ACUR_ECI
W1_ACUR_DUA*W2_ACUR_DUA
W1_ACUR_DUA*W2_ACUR_ALC
W1_ACUR_DUA*W2_ACUR_MAR
W1_ACUR_DUA*W2_ACUR_PAI
W1_ACUR_DUA*W2_DEPRESS
W1_ACUR_DUA*W2_SLEEPING
W1_ACUR_DUA*W2_ANXIOUS
W1_ACUR_DUA*W2_PTSD
W1_ACUR_DUA*W2_LIED
W1_ACUR_DUA*W2_ATTENTIO
W1_ACUR_DUA*W2_LISTENIN
W1_ACUR_DUA*W2_BULLY
W1_ACUR_DUA*W2_FIGHTS
W1_ACUR_DUA*W2_RESTLESS
W1_ACUR_DUA*W2_ANSWERED

W1_ACUR_ALC*W2_ACUR_CIG
W1_ACUR_ALC*W2_ACUR_ECI
W1_ACUR_ALC*W2_ACUR_DUA
W1_ACUR_ALC*W2_ACUR_ALC
W1_ACUR_ALC*W2_ACUR_MAR
W1_ACUR_ALC*W2_ACUR_PAI
W1_ACUR_ALC*W2_DEPRESS
W1_ACUR_ALC*W2_SLEEPING
W1_ACUR_ALC*W2_ANXIOUS
W1_ACUR_ALC*W2_PTSD
W1_ACUR_ALC*W2_LIED
W1_ACUR_ALC*W2_ATTENTIO
W1_ACUR_ALC*W2_LISTENIN
W1_ACUR_ALC*W2_BULLY
W1_ACUR_ALC*W2_FIGHTS
W1_ACUR_ALC*W2_RESTLESS
W1_ACUR_ALC*W2_ANSWERED

W1_ACUR_MAR*W2_ACUR_CIG
W1_ACUR_MAR*W2_ACUR_ECI
W1_ACUR_MAR*W2_ACUR_DUA
W1_ACUR_MAR*W2_ACUR_ALC
W1_ACUR_MAR*W2_ACUR_MAR
W1_ACUR_MAR*W2_ACUR_PAI
W1_ACUR_MAR*W2_DEPRESS
W1_ACUR_MAR*W2_SLEEPING
W1_ACUR_MAR*W2_ANXIOUS

```
W1_ACUR_MAR*W2_PTSD
W1_ACUR_MAR*W2_LIED
W1_ACUR_MAR*W2_ATTENTIO
W1_ACUR_MAR*W2_LISTENIN
W1_ACUR_MAR*W2_BULLY
W1_ACUR_MAR*W2_FIGHTS
W1_ACUR_MAR*W2_RESTLESS
W1_ACUR_MAR*W2_ANSWERED

W1_ACUR_PAI*W2_ACUR_CIG
W1_ACUR_PAI*W2_ACUR_ECI
W1_ACUR_PAI*W2_ACUR_DUA
W1_ACUR_PAI*W2_ACUR_ALC
W1_ACUR_PAI*W2_ACUR_MAR
W1_ACUR_PAI*W2_ACUR_PAI
W1_ACUR_PAI*W2_DEPRESS
W1_ACUR_PAI*W2_SLEEPING
W1_ACUR_PAI*W2_ANXIOUS
W1_ACUR_PAI*W2_PTSD
W1_ACUR_PAI*W2_LIED
W1_ACUR_PAI*W2_ATTENTIO
W1_ACUR_PAI*W2_LISTENIN
W1_ACUR_PAI*W2_BULLY
W1_ACUR_PAI*W2_FIGHTS
W1_ACUR_PAI*W2_RESTLESS
W1_ACUR_PAI*W2_ANSWERED

W1_DEPRESS*W2_ACUR_CIG
W1_DEPRESS*W2_ACUR_ECI
W1_DEPRESS*W2_ACUR_DUA
W1_DEPRESS*W2_ACUR_ALC
W1_DEPRESS*W2_ACUR_MAR
W1_DEPRESS*W2_ACUR_PAI
W1_DEPRESS*W2_DEPRESS
W1_DEPRESS*W2_SLEEPING
W1_DEPRESS*W2_ANXIOUS
W1_DEPRESS*W2_PTSD
W1_DEPRESS*W2_LIED
W1_DEPRESS*W2_ATTENTIO
W1_DEPRESS*W2_LISTENIN
W1_DEPRESS*W2_BULLY
W1_DEPRESS*W2_FIGHTS
W1_DEPRESS*W2_RESTLESS
W1_DEPRESS*W2_ANSWERED

W1_SLEEPING*W2_ACUR_CIG
W1_SLEEPING*W2_ACUR_ECI
W1_SLEEPING*W2_ACUR_DUA
W1_SLEEPING*W2_ACUR_ALC
W1_SLEEPING*W2_ACUR_MAR
W1_SLEEPING*W2_ACUR_PAI
W1_SLEEPING*W2_DEPRESS
W1_SLEEPING*W2_SLEEPING
W1_SLEEPING*W2_ANXIOUS
W1_SLEEPING*W2_PTSD
W1_SLEEPING*W2_LIED
W1_SLEEPING*W2_ATTENTIO
W1_SLEEPING*W2_LISTENIN
W1_SLEEPING*W2_BULLY
W1_SLEEPING*W2_FIGHTS
W1_SLEEPING*W2_RESTLESS
W1_SLEEPING*W2_ANSWERED

W1_ANXIOUS*W2_ACUR_CIG
W1_ANXIOUS*W2_ACUR_ECI
W1_ANXIOUS*W2_ACUR_DUA
W1_ANXIOUS*W2_ACUR_ALC
```

W1_ANXIOUS*W2_ACUR_MAR
W1_ANXIOUS*W2_ACUR_PAI
W1_ANXIOUS*W2_DEPRESS
W1_ANXIOUS*W2_SLEEPING
W1_ANXIOUS*W2_ANXIOUS
W1_ANXIOUS*W2_PTSD
W1_ANXIOUS*W2_LIED
W1_ANXIOUS*W2_ATTENTIO
W1_ANXIOUS*W2_LISTENIN
W1_ANXIOUS*W2_BULLY
W1_ANXIOUS*W2_FIGHTS
W1_ANXIOUS*W2_RESTLESS
W1_ANXIOUS*W2_ANSWERED

W1_PTSD*W2_ACUR_CIG
W1_PTSD*W2_ACUR_ECI
W1_PTSD*W2_ACUR_DUA
W1_PTSD*W2_ACUR_ALC
W1_PTSD*W2_ACUR_MAR
W1_PTSD*W2_ACUR_PAI
W1_PTSD*W2_DEPRESS
W1_PTSD*W2_SLEEPING
W1_PTSD*W2_ANXIOUS
W1_PTSD*W2_PTSD
W1_PTSD*W2_LIED
W1_PTSD*W2_ATTENTIO
W1_PTSD*W2_LISTENIN
W1_PTSD*W2_BULLY
W1_PTSD*W2_FIGHTS
W1_PTSD*W2_RESTLESS
W1_PTSD*W2_ANSWERED

W1_LIED*W2_ACUR_CIG
W1_LIED*W2_ACUR_ECI
W1_LIED*W2_ACUR_DUA
W1_LIED*W2_ACUR_ALC
W1_LIED*W2_ACUR_MAR
W1_LIED*W2_ACUR_PAI
W1_LIED*W2_DEPRESS
W1_LIED*W2_SLEEPING
W1_LIED*W2_ANXIOUS
W1_LIED*W2_PTSD
W1_LIED*W2_LIED
W1_LIED*W2_ATTENTIO
W1_LIED*W2_LISTENIN
W1_LIED*W2_BULLY
W1_LIED*W2_FIGHTS
W1_LIED*W2_RESTLESS
W1_LIED*W2_ANSWERED

W1_ATTENTIO*W2_ACUR_CIG
W1_ATTENTIO*W2_ACUR_ECI
W1_ATTENTIO*W2_ACUR_DUA
W1_ATTENTIO*W2_ACUR_ALC
W1_ATTENTIO*W2_ACUR_MAR
W1_ATTENTIO*W2_ACUR_PAI
W1_ATTENTIO*W2_DEPRESS
W1_ATTENTIO*W2_SLEEPING
W1_ATTENTIO*W2_ANXIOUS
W1_ATTENTIO*W2_PTSD
W1_ATTENTIO*W2_LIED
W1_ATTENTIO*W2_ATTENTIO
W1_ATTENTIO*W2_LISTENIN
W1_ATTENTIO*W2_BULLY
W1_ATTENTIO*W2_FIGHTS
W1_ATTENTIO*W2_RESTLESS
W1_ATTENTIO*W2_ANSWERED

```
W1_LISTENING*W2_ACUR_CIG
W1_LISTENING*W2_ACUR_ECI
W1_LISTENING*W2_ACUR_DUA
W1_LISTENING*W2_ACUR_ALC
W1_LISTENING*W2_ACUR_MAR
W1_LISTENING*W2_ACUR_PAI
W1_LISTENING*W2_DEPRESS
W1_LISTENING*W2_SLEEPING
W1_LISTENING*W2_ANXIOUS
W1_LISTENING*W2_PTSD
W1_LISTENING*W2_LIED
W1_LISTENING*W2_ATTENTIO
W1_LISTENING*W2_LISTENIN
W1_LISTENING*W2_BULLY
W1_LISTENING*W2_FIGHTS
W1_LISTENING*W2_RESTLESS
W1_LISTENING*W2_ANSWERED

W1_BULLY*W2_ACUR_CIG
W1_BULLY*W2_ACUR_ECI
W1_BULLY*W2_ACUR_DUA
W1_BULLY*W2_ACUR_ALC
W1_BULLY*W2_ACUR_MAR
W1_BULLY*W2_ACUR_PAI
W1_BULLY*W2_DEPRESS
W1_BULLY*W2_SLEEPING
W1_BULLY*W2_ANXIOUS
W1_BULLY*W2_PTSD
W1_BULLY*W2_LIED
W1_BULLY*W2_ATTENTIO
W1_BULLY*W2_LISTENIN
W1_BULLY*W2_BULLY
W1_BULLY*W2_FIGHTS
W1_BULLY*W2_RESTLESS
W1_BULLY*W2_ANSWERED

W1_FIGHTS*W2_ACUR_CIG
W1_FIGHTS*W2_ACUR_ECI
W1_FIGHTS*W2_ACUR_DUA
W1_FIGHTS*W2_ACUR_ALC
W1_FIGHTS*W2_ACUR_MAR
W1_FIGHTS*W2_ACUR_PAI
W1_FIGHTS*W2_DEPRESS
W1_FIGHTS*W2_SLEEPING
W1_FIGHTS*W2_ANXIOUS
W1_FIGHTS*W2_PTSD
W1_FIGHTS*W2_LIED
W1_FIGHTS*W2_ATTENTIO
W1_FIGHTS*W2_LISTENIN
W1_FIGHTS*W2_BULLY
W1_FIGHTS*W2_FIGHTS
W1_FIGHTS*W2_RESTLESS
W1_FIGHTS*W2_ANSWERED

W1_RESTLESS*W2_ACUR_CIG
W1_RESTLESS*W2_ACUR_ECI
W1_RESTLESS*W2_ACUR_DUA
W1_RESTLESS*W2_ACUR_ALC
W1_RESTLESS*W2_ACUR_MAR
W1_RESTLESS*W2_ACUR_PAI
W1_RESTLESS*W2_DEPRESS
W1_RESTLESS*W2_SLEEPING
W1_RESTLESS*W2_ANXIOUS
W1_RESTLESS*W2_PTSD
W1_RESTLESS*W2_LIED
W1_RESTLESS*W2_ATTENTIO
W1_RESTLESS*W2_LISTENIN
```

```
W1_RESTLESS*W2_BULLY
W1_RESTLESS*W2_FIGHTS
W1_RESTLESS*W2_RESTLESS
W1_RESTLESS*W2_ANSWERED


W1_ANSWERED*W2_ACUR_CIG
W1_ANSWERED*W2_ACUR_ECI
W1_ANSWERED*W2_ACUR_DUA
W1_ANSWERED*W2_ACUR_ALC
W1_ANSWERED*W2_ACUR_MAR
W1_ANSWERED*W2_ACUR_PAI
W1_ANSWERED*W2_DEPRESS
W1_ANSWERED*W2_SLEEPING
W1_ANSWERED*W2_ANXIOUS
W1_ANSWERED*W2_PTSD
W1_ANSWERED*W2_LIED
W1_ANSWERED*W2_ATTENTIO
W1_ANSWERED*W2_LISTENIN
W1_ANSWERED*W2_BULLY
W1_ANSWERED*W2_FIGHTS
W1_ANSWERED*W2_RESTLESS
W1_ANSWERED*W2_ANSWERED

/plcorr chisq;
run;
ods pdf close;


ods pdf;
******W1 VS W3;
proc freq data=aim3.master;
table
W1_ACUR_CIG*W3_ACUR_CIG
W1_ACUR_CIG*W3_ACUR_ECI
W1_ACUR_CIG*W3_ACUR_DUA
W1_ACUR_CIG*W3_ACUR_ALC
W1_ACUR_CIG*W3_ACUR_MAR
W1_ACUR_CIG*W3_ACUR_PAI
W1_ACUR_CIG*W3_DEPRESS
W1_ACUR_CIG*W3_SLEEPING
W1_ACUR_CIG*W3_ANXIOUS
W1_ACUR_CIG*W3_PTSD
W1_ACUR_CIG*W3_LIED
W1_ACUR_CIG*W3_ATTENTIO
W1_ACUR_CIG*W3_LISTENIN
W1_ACUR_CIG*W3_BULLY
W1_ACUR_CIG*W3_FIGHTS
W1_ACUR_CIG*W3_RESTLESS
W1_ACUR_CIG*W3_ANSWERED


W1_ACUR_ECI*W3_ACUR_CIG
W1_ACUR_ECI*W3_ACUR_ECI
W1_ACUR_ECI*W3_ACUR_DUA
W1_ACUR_ECI*W3_ACUR_ALC
W1_ACUR_ECI*W3_ACUR_MAR
W1_ACUR_ECI*W3_ACUR_PAI
W1_ACUR_ECI*W3_DEPRESS
W1_ACUR_ECI*W3_SLEEPING
W1_ACUR_ECI*W3_ANXIOUS
W1_ACUR_ECI*W3_PTSD
W1_ACUR_ECI*W3_LIED
W1_ACUR_ECI*W3_ATTENTIO
W1_ACUR_ECI*W3_LISTENIN
W1_ACUR_ECI*W3_BULLY
W1_ACUR_ECI*W3_FIGHTS
W1_ACUR_ECI*W3_RESTLESS
W1_ACUR_ECI*W3_ANSWERED
```

```
W1_ACUR_DUA*W3_ACUR_CIG
W1_ACUR_DUA*W3_ACUR_ECI
W1_ACUR_DUA*W3_ACUR_DUA
W1_ACUR_DUA*W3_ACUR_ALC
W1_ACUR_DUA*W3_ACUR_MAR
W1_ACUR_DUA*W3_ACUR_PAI
W1_ACUR_DUA*W3_DEPRESS
W1_ACUR_DUA*W3_SLEEPING
W1_ACUR_DUA*W3_ANXIOUS
W1_ACUR_DUA*W3_PTSD
W1_ACUR_DUA*W3_LIED
W1_ACUR_DUA*W3_ATTENTIO
W1_ACUR_DUA*W3_LISTENIN
W1_ACUR_DUA*W3_BULLY
W1_ACUR_DUA*W3_FIGHTS
W1_ACUR_DUA*W3_RESTLESS
W1_ACUR_DUA*W3_ANSWERED

W1_ACUR_ALC*W3_ACUR_CIG
W1_ACUR_ALC*W3_ACUR_ECI
W1_ACUR_ALC*W3_ACUR_DUA
W1_ACUR_ALC*W3_ACUR_ALC
W1_ACUR_ALC*W3_ACUR_MAR
W1_ACUR_ALC*W3_ACUR_PAI
W1_ACUR_ALC*W3_DEPRESS
W1_ACUR_ALC*W3_SLEEPING
W1_ACUR_ALC*W3_ANXIOUS
W1_ACUR_ALC*W3_PTSD
W1_ACUR_ALC*W3_LIED
W1_ACUR_ALC*W3_ATTENTIO
W1_ACUR_ALC*W3_LISTENIN
W1_ACUR_ALC*W3_BULLY
W1_ACUR_ALC*W3_FIGHTS
W1_ACUR_ALC*W3_RESTLESS
W1_ACUR_ALC*W3_ANSWERED

W1_ACUR_MAR*W3_ACUR_CIG
W1_ACUR_MAR*W3_ACUR_ECI
W1_ACUR_MAR*W3_ACUR_DUA
W1_ACUR_MAR*W3_ACUR_ALC
W1_ACUR_MAR*W3_ACUR_MAR
W1_ACUR_MAR*W3_ACUR_PAI
W1_ACUR_MAR*W3_DEPRESS
W1_ACUR_MAR*W3_SLEEPING
W1_ACUR_MAR*W3_ANXIOUS
W1_ACUR_MAR*W3_PTSD
W1_ACUR_MAR*W3_LIED
W1_ACUR_MAR*W3_ATTENTIO
W1_ACUR_MAR*W3_LISTENIN
W1_ACUR_MAR*W3_BULLY
W1_ACUR_MAR*W3_FIGHTS
W1_ACUR_MAR*W3_RESTLESS
W1_ACUR_MAR*W3_ANSWERED

W1_ACUR_PAI*W3_ACUR_CIG
W1_ACUR_PAI*W3_ACUR_ECI
W1_ACUR_PAI*W3_ACUR_DUA
W1_ACUR_PAI*W3_ACUR_ALC
W1_ACUR_PAI*W3_ACUR_MAR
W1_ACUR_PAI*W3_ACUR_PAI
W1_ACUR_PAI*W3_DEPRESS
W1_ACUR_PAI*W3_SLEEPING
W1_ACUR_PAI*W3_ANXIOUS
W1_ACUR_PAI*W3_PTSD
W1_ACUR_PAI*W3_LIED
W1_ACUR_PAI*W3_ATTENTIO
W1_ACUR_PAI*W3_LISTENIN
W1_ACUR_PAI*W3_BULLY
```

```
W1_ACUR_PAI*W3_FIGHTS
W1_ACUR_PAI*W3_RESTLESS
W1_ACUR_PAI*W3_ANSWERED


W1_DEPRESS*W3_ACUR_CIG
W1_DEPRESS*W3_ACUR_ECI
W1_DEPRESS*W3_ACUR_DUA
W1_DEPRESS*W3_ACUR_ALC
W1_DEPRESS*W3_ACUR_MAR
W1_DEPRESS*W3_ACUR_PAI
W1_DEPRESS*W3_DEPRESS
W1_DEPRESS*W3_SLEEPING
W1_DEPRESS*W3_ANXIOUS
W1_DEPRESS*W3_PTSD
W1_DEPRESS*W3_LIED
W1_DEPRESS*W3_ATTENTIO
W1_DEPRESS*W3_LISTENIN
W1_DEPRESS*W3_BULLY
W1_DEPRESS*W3_FIGHTS
W1_DEPRESS*W3_RESTLESS
W1_DEPRESS*W3_ANSWERED


W1_SLEEPING*W3_ACUR_CIG
W1_SLEEPING*W3_ACUR_ECI
W1_SLEEPING*W3_ACUR_DUA
W1_SLEEPING*W3_ACUR_ALC
W1_SLEEPING*W3_ACUR_MAR
W1_SLEEPING*W3_ACUR_PAI
W1_SLEEPING*W3_DEPRESS
W1_SLEEPING*W3_SLEEPING
W1_SLEEPING*W3_ANXIOUS
W1_SLEEPING*W3_PTSD
W1_SLEEPING*W3_LIED
W1_SLEEPING*W3_ATTENTIO
W1_SLEEPING*W3_LISTENIN
W1_SLEEPING*W3_BULLY
W1_SLEEPING*W3_FIGHTS
W1_SLEEPING*W3_RESTLESS
W1_SLEEPING*W3_ANSWERED


W1_ANXIOUS*W3_ACUR_CIG
W1_ANXIOUS*W3_ACUR_ECI
W1_ANXIOUS*W3_ACUR_DUA
W1_ANXIOUS*W3_ACUR_ALC
W1_ANXIOUS*W3_ACUR_MAR
W1_ANXIOUS*W3_ACUR_PAI
W1_ANXIOUS*W3_DEPRESS
W1_ANXIOUS*W3_SLEEPING
W1_ANXIOUS*W3_ANXIOUS
W1_ANXIOUS*W3_PTSD
W1_ANXIOUS*W3_LIED
W1_ANXIOUS*W3_ATTENTIO
W1_ANXIOUS*W3_LISTENIN
W1_ANXIOUS*W3_BULLY
W1_ANXIOUS*W3_FIGHTS
W1_ANXIOUS*W3_RESTLESS
W1_ANXIOUS*W3_ANSWERED


W1_PTSD*W3_ACUR_CIG
W1_PTSD*W3_ACUR_ECI
W1_PTSD*W3_ACUR_DUA
W1_PTSD*W3_ACUR_ALC
W1_PTSD*W3_ACUR_MAR
W1_PTSD*W3_ACUR_PAI
W1_PTSD*W3_DEPRESS
W1_PTSD*W3_SLEEPING
W1_PTSD*W3_ANXIOUS
```

W1_PTSD*W3_PTSD
W1_PTSD*W3_LIED
W1_PTSD*W3_ATTENTIO
W1_PTSD*W3_LISTENIN
W1_PTSD*W3_BULLY
W1_PTSD*W3_FIGHTS
W1_PTSD*W3_RESTLESS
W1_PTSD*W3_ANSWERED

W1_LIED*W3_ACUR_CIG
W1_LIED*W3_ACUR_ECI
W1_LIED*W3_ACUR_DUA
W1_LIED*W3_ACUR_ALC
W1_LIED*W3_ACUR_MAR
W1_LIED*W3_ACUR_PAI
W1_LIED*W3_DEPRESS
W1_LIED*W3_SLEEPING
W1_LIED*W3_ANXIOUS
W1_LIED*W3_PTSD
W1_LIED*W3_LIED
W1_LIED*W3_ATTENTIO
W1_LIED*W3_LISTENIN
W1_LIED*W3_BULLY
W1_LIED*W3_FIGHTS
W1_LIED*W3_RESTLESS
W1_LIED*W3_ANSWERED

W1_ATTENTIO*W3_ACUR_CIG
W1_ATTENTIO*W3_ACUR_ECI
W1_ATTENTIO*W3_ACUR_DUA
W1_ATTENTIO*W3_ACUR_ALC
W1_ATTENTIO*W3_ACUR_MAR
W1_ATTENTIO*W3_ACUR_PAI
W1_ATTENTIO*W3_DEPRESS
W1_ATTENTIO*W3_SLEEPING
W1_ATTENTIO*W3_ANXIOUS
W1_ATTENTIO*W3_PTSD
W1_ATTENTIO*W3_LIED
W1_ATTENTIO*W3_ATTENTIO
W1_ATTENTIO*W3_LISTENIN
W1_ATTENTIO*W3_BULLY
W1_ATTENTIO*W3_FIGHTS
W1_ATTENTIO*W3_RESTLESS
W1_ATTENTIO*W3_ANSWERED

W1_LISTENING*W3_ACUR_CIG
W1_LISTENING*W3_ACUR_ECI
W1_LISTENING*W3_ACUR_DUA
W1_LISTENING*W3_ACUR_ALC
W1_LISTENING*W3_ACUR_MAR
W1_LISTENING*W3_ACUR_PAI
W1_LISTENING*W3_DEPRESS
W1_LISTENING*W3_SLEEPING
W1_LISTENING*W3_ANXIOUS
W1_LISTENING*W3_PTSD
W1_LISTENING*W3_LIED
W1_LISTENING*W3_ATTENTIO
W1_LISTENING*W3_LISTENIN
W1_LISTENING*W3_BULLY
W1_LISTENING*W3_FIGHTS
W1_LISTENING*W3_RESTLESS
W1_LISTENING*W3_ANSWERED

W1_BULLY*W3_ACUR_CIG
W1_BULLY*W3_ACUR_ECI
W1_BULLY*W3_ACUR_DUA
W1_BULLY*W3_ACUR_ALC

```
W1_BULLY*W3_ACUR_MAR
W1_BULLY*W3_ACUR_PAI
W1_BULLY*W3_DEPRESS
W1_BULLY*W3_SLEEPING
W1_BULLY*W3_ANXIOUS
W1_BULLY*W3_PTSD
W1_BULLY*W3_LIED
W1_BULLY*W3_ATTENTIO
W1_BULLY*W3_LISTENIN
W1_BULLY*W3_BULLY
W1_BULLY*W3_FIGHTS
W1_BULLY*W3_RESTLESS
W1_BULLY*W3_ANSWERED

W1_FIGHTS*W3_ACUR_CIG
W1_FIGHTS*W3_ACUR_ECI
W1_FIGHTS*W3_ACUR_DUA
W1_FIGHTS*W3_ACUR_ALC
W1_FIGHTS*W3_ACUR_MAR
W1_FIGHTS*W3_ACUR_PAI
W1_FIGHTS*W3_DEPRESS
W1_FIGHTS*W3_SLEEPING
W1_FIGHTS*W3_ANXIOUS
W1_FIGHTS*W3_PTSD
W1_FIGHTS*W3_LIED
W1_FIGHTS*W3_ATTENTIO
W1_FIGHTS*W3_LISTENIN
W1_FIGHTS*W3_BULLY
W1_FIGHTS*W3_FIGHTS
W1_FIGHTS*W3_RESTLESS
W1_FIGHTS*W3_ANSWERED

W1_RESTLESS*W3_ACUR_CIG
W1_RESTLESS*W3_ACUR_ECI
W1_RESTLESS*W3_ACUR_DUA
W1_RESTLESS*W3_ACUR_ALC
W1_RESTLESS*W3_ACUR_MAR
W1_RESTLESS*W3_ACUR_PAI
W1_RESTLESS*W3_DEPRESS
W1_RESTLESS*W3_SLEEPING
W1_RESTLESS*W3_ANXIOUS
W1_RESTLESS*W3_PTSD
W1_RESTLESS*W3_LIED
W1_RESTLESS*W3_ATTENTIO
W1_RESTLESS*W3_LISTENIN
W1_RESTLESS*W3_BULLY
W1_RESTLESS*W3_FIGHTS
W1_RESTLESS*W3_RESTLESS
W1_RESTLESS*W3_ANSWERED

W1_ANSWERED*W3_ACUR_CIG
W1_ANSWERED*W3_ACUR_ECI
W1_ANSWERED*W3_ACUR_DUA
W1_ANSWERED*W3_ACUR_ALC
W1_ANSWERED*W3_ACUR_MAR
W1_ANSWERED*W3_ACUR_PAI
W1_ANSWERED*W3_DEPRESS
W1_ANSWERED*W3_SLEEPING
W1_ANSWERED*W3_ANXIOUS
W1_ANSWERED*W3_PTSD
W1_ANSWERED*W3_LIED
W1_ANSWERED*W3_ATTENTIO
W1_ANSWERED*W3_LISTENIN
W1_ANSWERED*W3_BULLY
W1_ANSWERED*W3_FIGHTS
W1_ANSWERED*W3_RESTLESS
W1_ANSWERED*W3_ANSWERED
```

```
/plcorr chisq;
run;
ods pdf close;


ods pdf;
******W2 VS W3;
proc freq data=aim3.master;
table
W2_ACUR_CIG*W3_ACUR_CIG
W2_ACUR_CIG*W3_ACUR_ECI
W2_ACUR_CIG*W3_ACUR_DUA
W2_ACUR_CIG*W3_ACUR_ALC
W2_ACUR_CIG*W3_ACUR_MAR
W2_ACUR_CIG*W3_ACUR_PAI
W2_ACUR_CIG*W3_DEPRESS
W2_ACUR_CIG*W3_SLEEPING
W2_ACUR_CIG*W3_ANXIOUS
W2_ACUR_CIG*W3_PTSD
W2_ACUR_CIG*W3_LIED
W2_ACUR_CIG*W3_ATTENTIO
W2_ACUR_CIG*W3_LISTENIN
W2_ACUR_CIG*W3_BULLY
W2_ACUR_CIG*W3_FIGHTS
W2_ACUR_CIG*W3_RESTLESS
W2_ACUR_CIG*W3_ANSWERED


W2_ACUR_ECI*W3_ACUR_CIG
W2_ACUR_ECI*W3_ACUR_ECI
W2_ACUR_ECI*W3_ACUR_DUA
W2_ACUR_ECI*W3_ACUR_ALC
W2_ACUR_ECI*W3_ACUR_MAR
W2_ACUR_ECI*W3_ACUR_PAI
W2_ACUR_ECI*W3_DEPRESS
W2_ACUR_ECI*W3_SLEEPING
W2_ACUR_ECI*W3_ANXIOUS
W2_ACUR_ECI*W3_PTSD
W2_ACUR_ECI*W3_LIED
W2_ACUR_ECI*W3_ATTENTIO
W2_ACUR_ECI*W3_LISTENIN
W2_ACUR_ECI*W3_BULLY
W2_ACUR_ECI*W3_FIGHTS
W2_ACUR_ECI*W3_RESTLESS
W2_ACUR_ECI*W3_ANSWERED


W2_ACUR_DUA*W3_ACUR_CIG
W2_ACUR_DUA*W3_ACUR_ECI
W2_ACUR_DUA*W3_ACUR_DUA
W2_ACUR_DUA*W3_ACUR_ALC
W2_ACUR_DUA*W3_ACUR_MAR
W2_ACUR_DUA*W3_ACUR_PAI
W2_ACUR_DUA*W3_DEPRESS
W2_ACUR_DUA*W3_SLEEPING
W2_ACUR_DUA*W3_ANXIOUS
W2_ACUR_DUA*W3_PTSD
W2_ACUR_DUA*W3_LIED
W2_ACUR_DUA*W3_ATTENTIO
W2_ACUR_DUA*W3_LISTENIN
W2_ACUR_DUA*W3_BULLY
W2_ACUR_DUA*W3_FIGHTS
W2_ACUR_DUA*W3_RESTLESS
W2_ACUR_DUA*W3_ANSWERED


W2_ACUR_ALC*W3_ACUR_CIG
W2_ACUR_ALC*W3_ACUR_ECI
W2_ACUR_ALC*W3_ACUR_DUA
W2_ACUR_ALC*W3_ACUR_ALC
W2_ACUR_ALC*W3_ACUR_MAR
```

W2_ACUR_ALC*W3_ACUR_PAI
W2_ACUR_ALC*W3_DEPRESS
W2_ACUR_ALC*W3_SLEEPING
W2_ACUR_ALC*W3_ANXIOUS
W2_ACUR_ALC*W3_PTSD
W2_ACUR_ALC*W3_LIED
W2_ACUR_ALC*W3_ATTENTIO
W2_ACUR_ALC*W3_LISTENIN
W2_ACUR_ALC*W3_BULLY
W2_ACUR_ALC*W3_FIGHTS
W2_ACUR_ALC*W3_RESTLESS
W2_ACUR_ALC*W3_ANSWERED

W2_ACUR_MAR*W3_ACUR_CIG
W2_ACUR_MAR*W3_ACUR_ECI
W2_ACUR_MAR*W3_ACUR_DUA
W2_ACUR_MAR*W3_ACUR_ALC
W2_ACUR_MAR*W3_ACUR_MAR
W2_ACUR_MAR*W3_ACUR_PAI
W2_ACUR_MAR*W3_DEPRESS
W2_ACUR_MAR*W3_SLEEPING
W2_ACUR_MAR*W3_ANXIOUS
W2_ACUR_MAR*W3_PTSD
W2_ACUR_MAR*W3_LIED
W2_ACUR_MAR*W3_ATTENTIO
W2_ACUR_MAR*W3_LISTENIN
W2_ACUR_MAR*W3_BULLY
W2_ACUR_MAR*W3_FIGHTS
W2_ACUR_MAR*W3_RESTLESS
W2_ACUR_MAR*W3_ANSWERED

W2_ACUR_PAI*W3_ACUR_CIG
W2_ACUR_PAI*W3_ACUR_ECI
W2_ACUR_PAI*W3_ACUR_DUA
W2_ACUR_PAI*W3_ACUR_ALC
W2_ACUR_PAI*W3_ACUR_MAR
W2_ACUR_PAI*W3_ACUR_PAI
W2_ACUR_PAI*W3_DEPRESS
W2_ACUR_PAI*W3_SLEEPING
W2_ACUR_PAI*W3_ANXIOUS
W2_ACUR_PAI*W3_PTSD
W2_ACUR_PAI*W3_LIED
W2_ACUR_PAI*W3_ATTENTIO
W2_ACUR_PAI*W3_LISTENIN
W2_ACUR_PAI*W3_BULLY
W2_ACUR_PAI*W3_FIGHTS
W2_ACUR_PAI*W3_RESTLESS
W2_ACUR_PAI*W3_ANSWERED

W2_DEPRESS*W3_ACUR_CIG
W2_DEPRESS*W3_ACUR_ECI
W2_DEPRESS*W3_ACUR_DUA
W2_DEPRESS*W3_ACUR_ALC
W2_DEPRESS*W3_ACUR_MAR
W2_DEPRESS*W3_ACUR_PAI
W2_DEPRESS*W3_DEPRESS
W2_DEPRESS*W3_SLEEPING
W2_DEPRESS*W3_ANXIOUS
W2_DEPRESS*W3_PTSD
W2_DEPRESS*W3_LIED
W2_DEPRESS*W3_ATTENTIO
W2_DEPRESS*W3_LISTENIN
W2_DEPRESS*W3_BULLY
W2_DEPRESS*W3_FIGHTS
W2_DEPRESS*W3_RESTLESS
W2_DEPRESS*W3_ANSWERED

```
W2_SLEEPING*W3_ACUR_CIG
W2_SLEEPING*W3_ACUR_ECI
W2_SLEEPING*W3_ACUR_DUA
W2_SLEEPING*W3_ACUR_ALC
W2_SLEEPING*W3_ACUR_MAR
W2_SLEEPING*W3_ACUR_PAI
W2_SLEEPING*W3_DEPRESS
W2_SLEEPING*W3_SLEEPING
W2_SLEEPING*W3_ANXIOUS
W2_SLEEPING*W3_PTSD
W2_SLEEPING*W3_LIED
W2_SLEEPING*W3_ATTENTIO
W2_SLEEPING*W3_LISTENIN
W2_SLEEPING*W3_BULLY
W2_SLEEPING*W3_FIGHTS
W2_SLEEPING*W3_RESTLESS
W2_SLEEPING*W3_ANSWERED

W2_ANXIOUS*W3_ACUR_CIG
W2_ANXIOUS*W3_ACUR_ECI
W2_ANXIOUS*W3_ACUR_DUA
W2_ANXIOUS*W3_ACUR_ALC
W2_ANXIOUS*W3_ACUR_MAR
W2_ANXIOUS*W3_ACUR_PAI
W2_ANXIOUS*W3_DEPRESS
W2_ANXIOUS*W3_SLEEPING
W2_ANXIOUS*W3_ANXIOUS
W2_ANXIOUS*W3_PTSD
W2_ANXIOUS*W3_LIED
W2_ANXIOUS*W3_ATTENTIO
W2_ANXIOUS*W3_LISTENIN
W2_ANXIOUS*W3_BULLY
W2_ANXIOUS*W3_FIGHTS
W2_ANXIOUS*W3_RESTLESS
W2_ANXIOUS*W3_ANSWERED

W2_PTSD*W3_ACUR_CIG
W2_PTSD*W3_ACUR_ECI
W2_PTSD*W3_ACUR_DUA
W2_PTSD*W3_ACUR_ALC
W2_PTSD*W3_ACUR_MAR
W2_PTSD*W3_ACUR_PAI
W2_PTSD*W3_DEPRESS
W2_PTSD*W3_SLEEPING
W2_PTSD*W3_ANXIOUS
W2_PTSD*W3_PTSD
W2_PTSD*W3_LIED
W2_PTSD*W3_ATTENTIO
W2_PTSD*W3_LISTENIN
W2_PTSD*W3_BULLY
W2_PTSD*W3_FIGHTS
W2_PTSD*W3_RESTLESS
W2_PTSD*W3_ANSWERED

W2_LIED*W3_ACUR_CIG
W2_LIED*W3_ACUR_ECI
W2_LIED*W3_ACUR_DUA
W2_LIED*W3_ACUR_ALC
W2_LIED*W3_ACUR_MAR
W2_LIED*W3_ACUR_PAI
W2_LIED*W3_DEPRESS
W2_LIED*W3_SLEEPING
W2_LIED*W3_ANXIOUS
W2_LIED*W3_PTSD
W2_LIED*W3_LIED
W2_LIED*W3_ATTENTIO
W2_LIED*W3_LISTENIN
W2_LIED*W3_BULLY
```

```
W2_LIED*W3_FIGHTS
W2_LIED*W3_RESTLESS
W2_LIED*W3_ANSWERED


W2_ATTENTIO*W3_ACUR_CIG
W2_ATTENTIO*W3_ACUR_ECI
W2_ATTENTIO*W3_ACUR_DUA
W2_ATTENTIO*W3_ACUR_ALC
W2_ATTENTIO*W3_ACUR_MAR
W2_ATTENTIO*W3_ACUR_PAI
W2_ATTENTIO*W3_DEPRESS
W2_ATTENTIO*W3_SLEEPING
W2_ATTENTIO*W3_ANXIOUS
W2_ATTENTIO*W3_PTSD
W2_ATTENTIO*W3_LIED
W2_ATTENTIO*W3_ATTENTIO
W2_ATTENTIO*W3_LISTENIN
W2_ATTENTIO*W3_BULLY
W2_ATTENTIO*W3_FIGHTS
W2_ATTENTIO*W3_RESTLESS
W2_ATTENTIO*W3_ANSWERED


W2_LISTENIN*W3_ACUR_CIG
W2_LISTENIN*W3_ACUR_ECI
W2_LISTENIN*W3_ACUR_DUA
W2_LISTENIN*W3_ACUR_ALC
W2_LISTENIN*W3_ACUR_MAR
W2_LISTENIN*W3_ACUR_PAI
W2_LISTENIN*W3_DEPRESS
W2_LISTENIN*W3_SLEEPING
W2_LISTENIN*W3_ANXIOUS
W2_LISTENIN*W3_PTSD
W2_LISTENIN*W3_LIED
W2_LISTENIN*W3_ATTENTIO
W2_LISTENIN*W3_LISTENIN
W2_LISTENIN*W3_BULLY
W2_LISTENIN*W3_FIGHTS
W2_LISTENIN*W3_RESTLESS
W2_LISTENIN*W3_ANSWERED


W2_BULLY*W3_ACUR_CIG
W2_BULLY*W3_ACUR_ECI
W2_BULLY*W3_ACUR_DUA
W2_BULLY*W3_ACUR_ALC
W2_BULLY*W3_ACUR_MAR
W2_BULLY*W3_ACUR_PAI
W2_BULLY*W3_DEPRESS
W2_BULLY*W3_SLEEPING
W2_BULLY*W3_ANXIOUS
W2_BULLY*W3_PTSD
W2_BULLY*W3_LIED
W2_BULLY*W3_ATTENTIO
W2_BULLY*W3_LISTENIN
W2_BULLY*W3_BULLY
W2_BULLY*W3_FIGHTS
W2_BULLY*W3_RESTLESS
W2_BULLY*W3_ANSWERED


W2_FIGHTS*W3_ACUR_CIG
W2_FIGHTS*W3_ACUR_ECI
W2_FIGHTS*W3_ACUR_DUA
W2_FIGHTS*W3_ACUR_ALC
W2_FIGHTS*W3_ACUR_MAR
W2_FIGHTS*W3_ACUR_PAI
W2_FIGHTS*W3_DEPRESS
W2_FIGHTS*W3_SLEEPING
W2_FIGHTS*W3_ANXIOUS
```

```
W2_FIGHTS*W3_PTSD
W2_FIGHTS*W3_LIED
W2_FIGHTS*W3_ATTENTIO
W2_FIGHTS*W3_LISTENIN
W2_FIGHTS*W3_BULLY
W2_FIGHTS*W3_FIGHTS
W2_FIGHTS*W3_RESTLESS
W2_FIGHTS*W3_ANSWERED


W2_RESTLESS*W3_ACUR_CIG
W2_RESTLESS*W3_ACUR_ECI
W2_RESTLESS*W3_ACUR_DUA
W2_RESTLESS*W3_ACUR_ALC
W2_RESTLESS*W3_ACUR_MAR
W2_RESTLESS*W3_ACUR_PAI
W2_RESTLESS*W3_DEPRESS
W2_RESTLESS*W3_SLEEPING
W2_RESTLESS*W3_ANXIOUS
W2_RESTLESS*W3_PTSD
W2_RESTLESS*W3_LIED
W2_RESTLESS*W3_ATTENTIO
W2_RESTLESS*W3_LISTENIN
W2_RESTLESS*W3_BULLY
W2_RESTLESS*W3_FIGHTS
W2_RESTLESS*W3_RESTLESS
W2_RESTLESS*W3_ANSWERED


W2_ANSWERED*W3_ACUR_CIG
W2_ANSWERED*W3_ACUR_ECI
W2_ANSWERED*W3_ACUR_DUA
W2_ANSWERED*W3_ACUR_ALC
W2_ANSWERED*W3_ACUR_MAR
W2_ANSWERED*W3_ACUR_PAI
W2_ANSWERED*W3_DEPRESS
W2_ANSWERED*W3_SLEEPING
W2_ANSWERED*W3_ANXIOUS
W2_ANSWERED*W3_PTSD
W2_ANSWERED*W3_LIED
W2_ANSWERED*W3_ATTENTIO
W2_ANSWERED*W3_LISTENIN
W2_ANSWERED*W3_BULLY
W2_ANSWERED*W3_FIGHTS
W2_ANSWERED*W3_RESTLESS
W2_ANSWERED*W3_ANSWERED

/plcorr chisq;
run;
ods pdf close;
```

SAS File name: Making W2 and W3 for Network Comparisons

```
****SA 3 - Network Comparisons for W1 vs W2 vs W3
****W1 already generated for SA2 (can be found in PNASS\\Data Management\\CSVs to use in
R\\overallwave1.csv)
****Making datasets for SA3 including overall W2 and W3
    Need to make dual CC + EC variables for W2 and W3
    Keep only adults who have continued from W1 (contadult datasets)
    Change . to -99999
    Export as CSV to R
    ONLY KEEPING 17 NODES FOR NETWORK COMPARISON SECTION OF SA 3
****February 1, 2021;


*********************
*Starting with Wave 2
*********************
```

```
 Pull most recent version of W2 data created in LCA W2 SAS Program (lca.w2contadult)
 Need to download formats for W2;
libname LCA "C:\Users\blondinoct\Documents\LCA\Wave 2\Data Management";

proc contents data=lca.w2contadult;
run;

*Making new tobacco vars;
data lca.w2contadultb;
set lca.w2contadult;

*first do multinomial - 4 levels;
if acur_cig=0 and acur_ecig=0 then acur_dual=0;
else if acur_cig=1 and acur_ecig=0 then acur_dual=1;
else if acur_cig=0 and acur_ecig=1 then acur_dual=2;
else if acur_cig=1 and acur_ecig=1 then acur_dual=3;
else acur_dual=.;
run;

*check;
proc freq data=lca.w2contadultb;
table acur_cig*acur_dual;
table acur_ecig*acur_dual;
run;

*then do the dummies;
data lca.w2contadultc;
set lca.w2contadultb;
if acur_dual = 1 then acur_cignew=1;
else acur_cignew=0;
if acur_dual = 2 then acur_ecignew=1;
else acur_ecignew=0;
if acur_dual = 3 then acur_dualnew=1;
else acur_dualnew=0;
run;

*check;
proc freq data=lca.w2contadultc;
table acur_cignew*acur_dual;
table acur_ecignew*acur_dual;
table acur_dualnew*acur_dual;
run;

*confirm marijuana is good;
proc freq data=lca.w2contadultc;
table acur_marijuana;
run;

***************************
**SUMMARY STATS FOR WAVE 2;
proc surveyfreq data= LCA.w2contadultc varmethod=BRR (fay=0.3);
table acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana acur_painkiller
 R02R_A_SEX age R02R_A_ETHRACECAT7 education R02R_A_AM0030 R02_AX0092
depressed sleeping anxious ptsd
lied attention listening bully fights restless answered
sud
/row chisq(secondorder);
weight R02_A_PWGT;
        repweights R02_A_PWGT1 - R02_A_PWGT100;
        run;
        quit;
run;

******************************************
**data kept to run LCA again then compare;
data LCA.w2contadultd;
set LCA.w2contadultc (keep = caseid personid R02_A_PWGT
                               acur_cignew acur_ecignew acur_dualnew
                 acur_alc acur_marijuana acur_painkiller
                               R02R_A_SEX age R02R_A_ETHRACECAT7 education R02R_A_AM0030
R02_AX0092
```

```
                                SEXMALE_1 SEXFEMALE_2
                                AGE1824_1 AGE2534_2 AGE3544_3 AGE4554_4 AGE5564_5 AGE65_6
                                RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
                                EDU_1 EDU_2 EDU_3 EDU_4 EDU_5
                                INC_1 INC_2 INC_3 INC_4 INC_5
                                SOC_1 SOC_2 SOC_3 SOC_4 SOC_5
                                depressed sleeping anxious ptsd
                                lied attention listening bully fights restless answered
                                weeklyuse timegetting timeusing socialprob reducedact withdraw
usetoavoid
                                sud);
*rename missings;
array change _numeric_;
do over change;
if change =. then change = -99999;
end;
run;

*Check frequencies;
proc contents data=LCA.w2contadultd;
run;
proc freq data=LCA.w2contadultd;
table
acur_cignew acur_ecignew acur_dualnew
acur_alc acur_marijuana acur_painkiller
depressed sleeping anxious ptsd
lied attention listening bully fights restless answered;
run;

*These look good:
*Exported to Aim 3 - Comparisons -> Data Sets as w2dataformplus232021;


**data kept for network comparisons;
data LCA.w2contadulte;
set LCA.w2contadultc (keep =
                                acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana
acur_painkiller
                                depressed sleeping anxious ptsd
                                lied attention listening bully fights restless answered);
*rename missings;
array change _numeric_;
do over change;
if change =. then change = -99999;
end;
run;

proc contents data=LCA.w2contadulte;
run;

*These look good:
*Exported to Aim 3 - Comparisons -> Data Sets as w2datafornetcomp232021;




***************
*Now to Wave 3
***************
 Pull most recent version of W3 data created in LCA W3 SAS Program (lca.w3contadult)
 Need to download formats for W3;
libname LCA "U:\CourtneyBlondino\PhD Epidemiology\LCA\Wave 3\Data Management";

*single w3 weights;
proc contents data=work.Da36498p3102;
*table R03_A_SWGT;
run;

*all waves weights;
proc contents data=work.Da36498p3101;
```

```
*table R03_A_AWGT;
run;


proc contents data=lca.w3;
run;
*proc sort data=lca.w3;
*by personid;
*run;

*********************************************************************************
NEED TO MERGE THIS DATASET WITH WEIGHTS DATASET DS3101 (all) and DS3102 BY PERSONID *
*********************************************************************************;

*all weights;
data lca.w3allweights;
set WORK.DA36498P3101;
run;
proc sort data=lca.w3allweights;
by personid;
run;

*w3 single weights;
data lca.w3singleweights;
set WORK.DA36498P3102;
run;
proc sort data=lca.w3singleweights;
by personid;
run;

*merge weights;
*data lca.w3newweights;
*merge lca.w3allweights lca.w3singleweights;
*by personid;
*run;
*proc contents data = lca.w3newweights
run;


data lca.w3contadultweights;
merge lca.w3contalladultweights lca.w3singleweights lca.w3allweights;
by personid;
run;
proc contents data = lca.w3contadultweights;
run;

*Then limit the sample to only people from Wave 1;
proc freq data=lca.w3contadultweights;
table R03_ADULTTYPE;
run;

data lca.w3contadult;
set lca.w3contadultweights;
if R03_ADULTTYPE=1;
run;

proc contents data=lca.w3contadult;
run;

proc freq data=lca.w3contadult;
table R03_ADULTTYPE;
run;

*Making new tobacco vars;
data lca.w3contadultb;
set lca.w3contadult;

*first do multinomial - 4 levels;
if acur_cig=0 and acur_ecig=0 then acur_dual=0;
else if acur_cig=1 and acur_ecig=0 then acur_dual=1;
else if acur_cig=0 and acur_ecig=1 then acur_dual=2;
```

```
else if acur_cig=1 and acur_ecig=1 then acur_dual=3;
else acur_dual=.;
run;

*check;
proc freq data=lca.w3contadultb;
table acur_cig*acur_dual;
table acur_ecig*acur_dual;
run;

*then do the dummies;
data lca.w3contadultc;
set lca.w3contadultb;
if acur_dual = 1 then acur_cignew=1;
else acur_cignew=0;
if acur_dual = 2 then acur_ecignew=1;
else acur_ecignew=0;
if acur_dual = 3 then acur_dualnew=1;
else acur_dualnew=0;
run;

*check;
proc freq data=lca.w3contadultc;
table acur_cignew*acur_dual;
table acur_ecignew*acur_dual;
table acur_dualnew*acur_dual;
run;

*confirm marijuana is good;
proc freq data=lca.w3contadultc;
table acur_marijuana;
run;

proc contents data=lca.w3contadultc;
run;

*proc print data=lca.w3contadultc;
*var R03_A_SWGT R03_ADULTTYPE;
*run;

*proc freq data=lca.w3contadultc;
*table R03_A_SWGT;
*run;

***************************
**SUMMARY STATS FOR WAVE 3;
ods pdf;
proc surveyfreq data= LCA.w3contadultc varmethod=BRR (fay=0.3);
table acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana acur_painkiller
R03R_A_SEX age R03R_A_ETHRACECAT7 education R03R_A_AM0030 R03_AX0092
depressed sleeping anxious ptsd
lied attention listening bully fights restless answered
sud
/row chisq(secondorder);
weight R03_A_SWGT;
        repweights R03_A_SWGT1 - R03_A_SWGT100;
        run;
        quit;
run;
ods pdf close;

******************************************
**data kept to run LCA again then compare;
data LCA.w3contadultd;
set LCA.w3contadultc (keep = caseid personid R03_A_SWGT
                               acur_cignew acur_ecignew acur_dualnew
                 acur_alc acur_marijuana acur_painkiller
                               R03R_A_SEX age R03R_A_ETHRACECAT7 education R03R_A_AM0030
R03_AX0092
                               SEXMALE_1 SEXFEMALE_2
                               AGE1824_1 AGE2534_2 AGE3544_3 AGE4554_4 AGE5564_5 AGE65_6
```

```
                                RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
                                EDU_1 EDU_2 EDU_3 EDU_4 EDU_5
                                INC_1 INC_2 INC_3 INC_4 INC_5
                                SOC_1 SOC_2 SOC_3 SOC_4 SOC_5
                                depressed sleeping anxious ptsd
                                lied attention listening bully fights restless answered
                                weeklyuse timegetting timeusing socialprob reducedact withdraw
usetoavoid
                                sud);
*rename missings;
array change _numeric_;
do over change;
if change =. then change = -99999;
end;
run;

*Check frequencies;
proc contents data=LCA.w3contadultd;
run;
proc freq data=LCA.w3contadultd;
table
acur_cignew acur_ecignew acur_dualnew
acur_alc acur_marijuana acur_painkiller
depressed sleeping anxious ptsd
lied attention listening bully fights restless answered;
run;

*These look good:
*Exported to Aim 3 - Comparisons -> Data Sets as w3dataformplus332021;


**data kept for network comparisons;
data LCA.w3contadulte;
set LCA.w3contadultc (keep =
                                acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana
acur_painkiller
                                depressed sleeping anxious ptsd
                                lied attention listening bully fights restless answered);
*rename missings;
array change _numeric_;
do over change;
if change =. then change = -99999;
end;
run;

proc contents data=LCA.w3contadulte;
run;

*These look good:
*Exported to Aim 3 - Comparisons -> Data Sets as w3datafornetcomp332021;





*THIS IS THE WRONG ONE;
*LOOK AT LCA W1 4132021;
***************************
**SUMMARY STATS FOR WAVE 1;
*Overall W1 with new exclusive CC and EC, and dual variables for Table 1 Network Paper;

libname LCA "C:\Users\blondinoct\Documents\LCA\Wave 1\Data Management";

proc surveyfreq data= LCA.W1mplusJuly2020weights varmethod=BRR (fay=0.3);
table
R01R_A_SEX age R01R_A_ETHRACECAT7 education R01R_A_AM0030
R01_AX0092
acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana acur_painkiller
depressed sleeping anxious ptsd
lied attention listening bully fights restless answered
sud
```

```
/row chisq(secondorder);
weight R01_A_PWGT;
        repweights R01_A_PWGT1 - R01_A_PWGT100;
        run;
        quit;
run;



*********************************************************************************
*Wave 3 - merging weights with diff dataset to try to figure out missing on R03_A_AWGT
*********************************************************************************
 Pull most recent version of W3 data created in LCA W3 SAS Program (lca.w3contadult)
 Need to download formats for W3;
libname LCA "C:\Users\blondinoct\Documents\LCA\Wave 3\Data Management";

proc freq data=work.Da36498p3101;
table R03_A_AWGT;
run;


proc contents data=lca.w3contadult;
run;
*N = 26239;
proc sort data=lca.w3contadult;
by personid;
run;

****************************************************************
NEED TO MERGE THIS DATASET WITH WEIGHTS DATASET DS3101 BY PERSONID *
****************************************************************;

data lca.w3weights;
set WORK.DA36498P3101;
run;
proc sort data=lca.w3weights;
by personid;
run;


data lca.w3contadultweights;
merge lca.w3contadult lca.w3weights;
by personid;
run;
proc contents data = lca.w3contadultweights;
run;

proc print data = lca.w3contadultweights (obs=20);
var R03_A_AWGT R03_ADULTTYPE;
run;


*Then limit the sample to only people from Wave 1 that have weight info;
proc freq data=lca.w3contadultweights;
table R03_ADULTTYPE;
run;

proc freq data=lca.w3contadultweights;
table R03_A_AWGT;
run;

data lca.w3contadult217;
set lca.w3contadultweights;
if R03_ADULTTYPE=1 and R03_A_AWGT^=.;
run;

proc freq data=lca.w3contadult217;
table R03_ADULTTYPE;
run;
```

```
proc freq data=lca.w3contadult217;
table R03_A_AWGT;
run;

*Making new tobacco vars;
data lca.w3contadultb;
set lca.w3contadult217;

*first do multinomial - 4 levels;
if acur_cig=0 and acur_ecig=0 then acur_dual=0;
else if acur_cig=1 and acur_ecig=0 then acur_dual=1;
else if acur_cig=0 and acur_ecig=1 then acur_dual=2;
else if acur_cig=1 and acur_ecig=1 then acur_dual=3;
else acur_dual=.;
run;

*check;
proc freq data=lca.w3contadultb;
table acur_cig*acur_dual;
table acur_ecig*acur_dual;
run;

*then do the dummies;
data lca.w3contadultc;
set lca.w3contadultb;
if acur_dual = 1 then acur_cignew=1;
else acur_cignew=0;
if acur_dual = 2 then acur_ecignew=1;
else acur_ecignew=0;
if acur_dual = 3 then acur_dualnew=1;
else acur_dualnew=0;
run;

*check;
proc freq data=lca.w3contadultc;
table acur_cignew*acur_dual;
table acur_ecignew*acur_dual;
table acur_dualnew*acur_dual;
run;

*confirm marijuana is good;
proc freq data=lca.w3contadultc;
table acur_marijuana;
run;

proc print data=lca.w3contadultc;
var R03_A_AWGT R03_ADULTTYPE;
run;

proc freq data=lca.w3contadultc;
table R03_A_AWGT;
run;

***************************
**SUMMARY STATS FOR WAVE 3;
proc surveyfreq data= LCA.w3contadultc varmethod=BRR (fay=0.3);
table acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana acur_painkiller
R03R_A_SEX age R03R_A_ETHRACECAT7 education R03R_A_AM0030 R03_AX0092
depressed sleeping anxious ptsd
lied attention listening bully fights restless answered
sud
/row chisq(secondorder);
weight R03_A_AWGT;
        repweights R03_A_AWGT1 - R03_A_AWGT100;
        run;
        quit;
run;

******************************************
**data kept to run LCA again then compare;
data LCA.w3contadultd;
```

```
set LCA.w3contadultc (keep = caseid personid R03_A_AWGT
                              acur_cignew acur_ecignew acur_dualnew
                 acur_alc acur_marijuana acur_painkiller
                              R03R_A_SEX age R03R_A_ETHRACECAT7 education R03R_A_AM0030
R03_AX0092
                              SEXMALE_1 SEXFEMALE_2
                              AGE1824_1 AGE2534_2 AGE3544_3 AGE4554_4 AGE5564_5 AGE65_6
                              RACEWH_1 RACEBL_2 RACEOT_3 RACEHI_6
                              EDU_1 EDU_2 EDU_3 EDU_4 EDU_5
                              INC_1 INC_2 INC_3 INC_4 INC_5
                              SOC_1 SOC_2 SOC_3 SOC_4 SOC_5
                              depressed sleeping anxious ptsd
                              lied attention listening bully fights restless answered
                              weeklyuse timegetting timeusing socialprob reducedact withdraw
usetoavoid
                              sud);
*rename missings;
array change _numeric_;
do over change;
if change =. then change = -99999;
end;
run;

*Check frequencies;
proc contents data=LCA.w3contadultd;
run;
proc freq data=LCA.w3contadultd;
table
acur_cignew acur_ecignew acur_dualnew
acur_alc acur_marijuana acur_painkiller
depressed sleeping anxious ptsd
lied attention listening bully fights restless answered;
run;

*These look good:
*Exported to Aim 3 - Comparisons -> Data Sets as w3dataformplus2172021;


**data kept for network comparisons;
data LCA.w3contadulte;
set LCA.w3contadultc (keep =
                              acur_cignew acur_ecignew acur_dualnew acur_alc acur_marijuana
acur_painkiller
                              depressed sleeping anxious ptsd
                              lied attention listening bully fights restless answered);
*rename missings;
array change _numeric_;
do over change;
if change =. then change = -99999;
end;
run;

proc contents data=LCA.w3contadulte;
run;

*These look good:
*Exported to Aim 3 - Comparisons -> Data Sets as w3datafornetcomp2172021;
```

R File name: Aim 3 – Network Comparisons Waves 1, 2, 3 – 4152021

#PATH Waves 1, 2, 3 - Network Comparisons (Specific Aim 3)

```
##############################################
#    READ IN WAVE 1 DATA                #
##############################################
```

```
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation
Files/Network Chapter")
getwd()
overall<-read.csv("/Users/courtneyblondino/Library/Mobile
Documents/com~apple~CloudDocs/Dissertation Files/Network Chapter/OverallWave1-4142021.csv",
header=T, sep=',')
names(overall)
dim(overall)

#rename variables so they look nice on the network
names(overall)[names(overall) == "acur_cignew"] <- "CIG"
names(overall)[names(overall) == "acur_ecignew"] <- "ECIG"
names(overall)[names(overall) == "acur_dualnew"] <- "Dual CIG + ECIG"
names(overall)[names(overall) == "acur_alc"] <- "Alcohol"
names(overall)[names(overall) == "acur_marijuana"] <- "Marijuana"
names(overall)[names(overall) == "acur_painkiller"] <- "PDNP"
names(overall)[names(overall) == "depressed"] <- "Depressed"
names(overall)[names(overall) == "sleeping"] <- "Sleeping"
names(overall)[names(overall) == "anxious"] <- "Anxious"
names(overall)[names(overall) == "PTSD"] <- "Distressed/Past"
names(overall)[names(overall) == "lied"] <- "Lied"
names(overall)[names(overall) == "attention"] <- "Attention"
names(overall)[names(overall) == "listening"] <- "Listening"
names(overall)[names(overall) == "bully"] <- "Bully"
names(overall)[names(overall) == "fights"] <- "Fights"
names(overall)[names(overall) == "restless"] <- "Restless"
names(overall)[names(overall) == "answered"] <- "Answered"
names(overall)

require(ggplot2)
require(bootnet)
require(IsingFit)
require(IsingSampler)
require(qgraph)


########################################
# Estimate the Network Model  - W1     #
########################################

#########
#IsingFit
Wave1NetworkIF <-estimateNetwork(overall, default="IsingFit", missing="listwise")


Wave1NetworkIF$labels

Names<- c("CIG", "ECIG" , "Dual CIG + ECIG", "Alcohol", "Marijuana",  "PDNP",
        "Depressed", "Sleeping" , "Anxious", "Distressed/Past" , "Lied",
        "Attention" , "Listening" , "Bully" ,"Fights"  , "Restless",
        "Answered")

Traits <- rep(c(
  'Substance Use',
  'Negative Affect',
  'Externalizing'
```

```
), times=c(6,4,7))

layout(t(1))
plot(Wave1NetworkIF,
    layout="spring",
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    minimum=0,
    label.scale.equal=TRUE,
    label.cex= 4.0,
    legend.cex = 0.4,
    title= "Wave 1")

#Results
Wave1NetworkIF$results

#Edges
Edges <- Wave1NetworkIF$graph
print(Edges)
write(Edges, file="OverallEdges.csv", sep=" ")


#Centrality (need to run bootstrap to do accuracy/stability)
centralityTable(Wave1NetworkIF)



##########################################
#    READ IN WAVE 2 DATA              #
##########################################

setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation
Files/Wave Comparison Chapter")
getwd()
Wave2<-read.csv("/Users/courtneyblondino/Library/Mobile
Documents/com~apple~CloudDocs/Dissertation Files/Wave Comparison
Chapter/w2datafornetcomp232021.csv", header=T, sep=',')
dim(Wave2)
names(Wave2)

#Have to tell R what missing means
Wave2$acur_cignew[Wave2$acur_cignew==-99999] <- NA
Wave2$acur_ecignew[Wave2$acur_ecignew==-99999] <- NA
Wave2$acur_dualnew[Wave2$acur_dualnew==-99999] <- NA
Wave2$acur_alc[Wave2$acur_alc==-99999] <- NA
Wave2$acur_marijuana[Wave2$acur_marijuana==-99999] <- NA
Wave2$acur_painkiller[Wave2$acur_painkiller==-99999] <- NA
Wave2$depressed[Wave2$depressed==-99999] <- NA
Wave2$sleeping[Wave2$sleeping==-99999] <- NA
Wave2$anxious[Wave2$anxious==-99999] <- NA
Wave2$ptsd[Wave2$ptsd==-99999] <- NA
Wave2$lied[Wave2$lied==-99999] <- NA
Wave2$attention[Wave2$attention==-99999] <- NA
```

```
Wave2$listening[Wave2$listening==-99999] <- NA
Wave2$bully[Wave2$bully==-99999] <- NA
Wave2$fights[Wave2$fights==-99999] <- NA
Wave2$restless[Wave2$restless==-99999] <- NA
Wave2$answered[Wave2$answered==-99999] <- NA

#rename variables so they look nice on the network
names(Wave2)[names(Wave2) == "acur_cignew"] <- "CIG"
names(Wave2)[names(Wave2) == "acur_ecignew"] <- "ECIG"
names(Wave2)[names(Wave2) == "acur_dualnew"] <- "Dual CIG + ECIG"
names(Wave2)[names(Wave2) == "acur_alc"] <- "Alcohol"
names(Wave2)[names(Wave2) == "acur_marijuana"] <- "Marijuana"
names(Wave2)[names(Wave2) == "acur_painkiller"] <- "PDNP"
names(Wave2)[names(Wave2) == "depressed"] <- "Depressed"
names(Wave2)[names(Wave2) == "sleeping"] <- "Sleeping"
names(Wave2)[names(Wave2) == "anxious"] <- "Anxious"
names(Wave2)[names(Wave2) == "ptsd"] <- "Distressed/Past"
names(Wave2)[names(Wave2) == "lied"] <- "Lied"
names(Wave2)[names(Wave2) == "attention"] <- "Attention"
names(Wave2)[names(Wave2) == "listening"] <- "Listening"
names(Wave2)[names(Wave2) == "bully"] <- "Bully"
names(Wave2)[names(Wave2) == "fights"] <- "Fights"
names(Wave2)[names(Wave2) == "restless"] <- "Restless"
names(Wave2)[names(Wave2) == "answered"] <- "Answered"

########################################
# Estimate the Network Model  - W2     #
########################################

#########
#IsingFit
Wave2NetworkIF <-estimateNetwork(Wave2, default="IsingFit", missing="listwise")

Wave2NetworkIF

Names<- c("CIG", "ECIG" , "Dual CIG + ECIG", "Alcohol", "Marijuana",  "PDNP",
        "Depressed", "Sleeping" , "Anxious", "Distressed/Past" , "Lied",
        "Attention" , "Listening" , "Bully" ,"Fights"  , "Restless",
        "Answered")

Traits <- rep(c(
  'Substance Use',
  'Negative Affect',
  'Externalizing'
), times=c(6,4,7))

plot(Wave2NetworkIF,
     layout="spring",
     cut=0,
     theme="colorblind",
     groups=Traits,
     #nodeNames=Names,
     #edge.color="black",
     minimum=0,
     label.scale.equal=TRUE,
     label.cex= 4.0,
```

```
    legend.cex = 0.4)

#Results
W2Results <- Wave2NetworkIF$results
#write.csv(W2Results, file="W2Results.csv", sep=" ")
#write(Edges, file="OverallEdges.csv", sep=" ")


#Edges
Edges <- Wave2NetworkIF$graph
print(Edges)

#Centrality (need to run bootstrap to do accuracy/stability)
centralityTable(Wave2NetworkIF)


#########################################
#    READ IN WAVE 3 DATA              #
#########################################

setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files/Wave Comparison Chapter")
getwd()
#Wave3<-read.csv("C:\\Users\\blondinoct\\Documents\\Aim 3 - Comparisons\\Data Sets\\w3datafornetcomp232021.csv", header=T, sep=',')
#dim(Wave3)
#names(Wave3)
############NEW WAVE 3 , N = 25382 ################
#Wave3<-read.csv("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files/Wave Comparison Chapter/w3datafornetcomp2172021.csv", header=T, sep=',')
#dim(Wave3)
#names(Wave3)

Wave3<-read.csv("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files/Wave Comparison Chapter/w3datafornetcomp332021.csv", header=T, sep=',')
dim(Wave3)
names(Wave3)

#Have to tell R what missing means
Wave3$acur_cignew[Wave3$acur_cignew==-99999] <- NA
Wave3$acur_ecignew[Wave3$acur_ecignew==-99999] <- NA
Wave3$acur_dualnew[Wave3$acur_dualnew==-99999] <- NA
Wave3$acur_alc[Wave3$acur_alc==-99999] <- NA
Wave3$acur_marijuana[Wave3$acur_marijuana==-99999] <- NA
Wave3$acur_painkiller[Wave3$acur_painkiller==-99999] <- NA
Wave3$depressed[Wave3$depressed==-99999] <- NA
Wave3$sleeping[Wave3$sleeping==-99999] <- NA
Wave3$anxious[Wave3$anxious==-99999] <- NA
Wave3$ptsd[Wave3$ptsd==-99999] <- NA
Wave3$lied[Wave3$lied==-99999] <- NA
Wave3$attention[Wave3$attention==-99999] <- NA
Wave3$listening[Wave3$listening==-99999] <- NA
Wave3$bully[Wave3$bully==-99999] <- NA
Wave3$fights[Wave3$fights==-99999] <- NA
```

```
Wave3$restless[Wave3$restless==-99999] <- NA
Wave3$answered[Wave3$answered==-99999] <- NA

#rename variables so they look nice on the network
names(Wave3)[names(Wave3) == "acur_cignew"] <- "CIG"
names(Wave3)[names(Wave3) == "acur_ecignew"] <- "ECIG"
names(Wave3)[names(Wave3) == "acur_dualnew"] <- "Dual CIG + ECIG"
names(Wave3)[names(Wave3) == "acur_alc"] <- "Alcohol"
names(Wave3)[names(Wave3) == "acur_marijuana"] <- "Marijuana"
names(Wave3)[names(Wave3) == "acur_painkiller"] <- "PDNP"
names(Wave3)[names(Wave3) == "depressed"] <- "Depressed"
names(Wave3)[names(Wave3) == "sleeping"] <- "Sleeping"
names(Wave3)[names(Wave3) == "anxious"] <- "Anxious"
names(Wave3)[names(Wave3) == "ptsd"] <- "Distressed/Past"
names(Wave3)[names(Wave3) == "lied"] <- "Lied"
names(Wave3)[names(Wave3) == "attention"] <- "Attention"
names(Wave3)[names(Wave3) == "listening"] <- "Listening"
names(Wave3)[names(Wave3) == "bully"] <- "Bully"
names(Wave3)[names(Wave3) == "fights"] <- "Fights"
names(Wave3)[names(Wave3) == "restless"] <- "Restless"
names(Wave3)[names(Wave3) == "answered"] <- "Answered"

###########################################
# Estimate the Network Model  - W3     #
###########################################

#########
#IsingFit
Wave3NetworkIF <-estimateNetwork(Wave3, default="IsingFit", missing="listwise")

Wave3NetworkIF

Names<- c("CIG", "ECIG" , "Dual CIG + ECIG", "Alcohol", "Marijuana",  "PDNP",
        "Depressed", "Sleeping" , "Anxious", "Distressed/Past" , "Lied",
        "Attention" , "Listening" , "Bully" ,"Fights"  , "Restless",
        "Answered")

Traits <- rep(c(
  'Substance Use',
  'Negative Affect',
  'Externalizing'
), times=c(6,4,7))

plot(Wave3NetworkIF,
    layout="spring",
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    minimum=0,
    label.scale.equal=TRUE,
    label.cex= 4.0,
    legend.cex = 0.4,
    title= "Wave 3")
```

```
#Results
Wave3NetworkIF$results

#Edges
Wave3Edges <- Wave3NetworkIF$graph
print(Wave3Edges)
write(Wave3Edges, file="Wave3Edges.csv", sep=" ")

#Centrality (need to run bootstrap to do accuracy/stability)
centralityTable(Wave3NetworkIF)



#############################################
# Visually Compare using Average Layout    #
#############################################

L<-averageLayout(Wave1NetworkIF, Wave2NetworkIF, Wave3NetworkIF)
Max<- max(abs(c(getWmat(Wave1NetworkIF),getWmat(Wave2NetworkIF),
getWmat(Wave3NetworkIF))))
layout(t(1:3))
plot(Wave1NetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    #label.scale.equal=TRUE,
    legend=FALSE,
    label.cex= 2.0,
    legend.cex = 0.4,
    title= "Wave 1",
    maximum=Max)
plot(Wave2NetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    #label.scale.equal=TRUE,
    label.cex= 2.0,
    legend.cex = 0.4,
    legend=FALSE,
    title= "Wave 2",
    maximum=Max)
plot(Wave3NetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    #label.scale.equal=TRUE,
    label.cex= 2.0,
```

```
    legend.cex = 0.4,
    title= "Wave 3",
    maximum=Max)


###########################################
#     Compare W1 AND W2                  #
###########################################

L<-averageLayout(Wave1NetworkIF, Wave2NetworkIF)
Max<- max(abs(c(getWmat(Wave1NetworkIF),getWmat(Wave2NetworkIF))))
layout(t(1:2))
plot(Wave1NetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    label.scale.equal=TRUE,
    labels=Names,
    legend=FALSE,
    label.cex= 4.0,
    legend.cex = 0.4,
    title= "Wave 1",
    maximum=Max)
plot(Wave2NetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    label.scale.equal=TRUE,
    label.cex= 4.0,
    legend.cex = 0.4,
    legend=FALSE,
    title= "Wave 2",
    maximum=Max)

library("devtools")
install_github("cvborkulo/NetworkComparisonTest")
library("NetworkComparisonTest")
#perform NCT and interpret results

NCTW1vsW2<- NCT(Wave1NetworkIF, Wave2NetworkIF, test.edges=TRUE,
            it=100)


#difference in global strength between the networks of the observed data sets
NCTW1vsW2$glstrinv.real
#3.307481

#global strength values of individual networks
NCTW1vsW2$glstrinv.sep
#55.99086 VS 59.29835
```

```
NCTW1vsW2$glstrinv.perm
#there's 100 since we did 100 permutations

#Difference in global strength p-value
NCTW1vsW2$glstrinv.pval
#0.2673267- so not significantly different from one another in regard to global strength

#Value of the max difference in edge weights of observed networks
NCTW1vsW2$nwinv.real
#1.563746

NCTW1vsW2$nwinv.perm
#there's 100 since we did 100 permutations

#Maximum difference in edge weights
NCTW1vsW2$nwinv.pval
#0.2277228 - so not significantly different from one another in regard to number of edge weights

#Which edges significantly differ?
NCTW1vsW2$einv.pvals[which(NCTW1vsW2$einv.pvals[,3]<0.05),]
#        Var1      Var2 p-value
#69       CIG Marijuana 0.00990099
#90  Marijuana      PDNP 0.00990099
#120       CIG  Sleeping 0.03960396
#155      ECIG      PTSD 0.02970297
#179   Anxious      Lied 0.02970297
#214      PTSD Listening 0.00990099
#216 Attention Listening 0.01980198
#252     Bully    Fights 0.03960396
#263  Sleeping  Restless 0.00990099


##############################################
#      Compare W1 AND W3                     #
##############################################
L<-averageLayout(Wave1NetworkIF, Wave3NetworkIF)
Max<- max(abs(c(getWmat(Wave1NetworkIF),getWmat(Wave3NetworkIF))))
layout(t(1:2))
plot(Wave1NetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    label.scale.equal=TRUE,
    labels=Names,
    legend=FALSE,
    label.cex= 4.0,
    legend.cex = 0.4,
    title= "Wave 1",
    maximum=Max)
plot(Wave3NetworkIF,
    layout=L,
    cut=0,
```

```
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    label.scale.equal=TRUE,
    label.cex= 4.0,
    legend.cex = 0.4,
    legend=FALSE,
    title= "Wave 3",
    maximum=Max)


NCTW1vsW3<- NCT(Wave1NetworkIF, Wave3NetworkIF, test.edges=TRUE,
            it=100)


#difference in global strength between the networks of the observed data sets
NCTW1vsW3$glstrinv.real
#4.009107

#global strength values of individual networks
NCTW1vsW3$glstrinv.sep
#55.99086 vs 59.99997

NCTW1vsW3$glstrinv.perm
#there's 100 since we did 100 permutations

#Difference in global strength p-value
NCTW1vsW3$glstrinv.pval
#0.2376238 - so not significantly different from one another in regard to global strength

#Value of the max difference in edge weights of observed networks
NCTW1vsW3$nwinv.real
#0.8183753

NCTW1vsW3$nwinv.perm
#there's 100 since we did 100 permutations

#Maximum difference in edge weights
NCTW1vsW3$nwinv.pval
#0.6039604 - so not significantly different from one another in regard to number of edge weights

#Which edges significantly differ?
NCTW1vsW3$einv.pvals[which(NCTW1vsW3$einv.pvals[,3]<0.05),]
#            Var1      Var2 p-value
#52           CIG   Alcohol 0.00990099
#69           CIG Marijuana 0.00990099
#90     Marijuana      PDNP 0.00990099
#106      Alcohol Depressed 0.02970297
#120          CIG  Sleeping 0.03960396
#121         ECIG  Sleeping 0.03960396
#159         PDNP      PTSD 0.03960396
#162      Anxious      PTSD 0.04950495
#174      Alcohol      Lied 0.00990099
#175    Marijuana      Lied 0.01980198
#178     Sleeping      Lied 0.02970297
```

```
#195       Sleeping Attention 0.03960396
#207 Dual CIG + ECIG Listening 0.03960396
#209      Marijuana Listening 0.01980198
#214           PTSD Listening 0.04950495
#216      Attention Listening 0.02970297
#251      Listening    Fights 0.02970297
#252          Bully    Fights 0.02970297
#256            CIG  Restless 0.03960396
#262      Depressed  Restless 0.00990099
#263       Sleeping  Restless 0.01980198
#285      Listening  Answered 0.04950495




###########################################
#     Compare W2 AND W3                   #
###########################################
L<-averageLayout(Wave2NetworkIF, Wave3NetworkIF)
Max<- max(abs(c(getWmat(Wave2NetworkIF),getWmat(Wave3NetworkIF))))
layout(t(1:2))
plot(Wave2NetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    label.scale.equal=TRUE,
    labels=Names,
    legend=FALSE,
    label.cex= 4.0,
    legend.cex = 0.4,
    title= "Wave 2",
    maximum=Max)
plot(Wave3NetworkIF,
    layout=L,
    cut=0,
    theme="colorblind",
    groups=Traits,
    #nodeNames=Names,
    #edge.color="black",
    label.scale.equal=TRUE,
    label.cex= 4.0,
    legend.cex = 0.4,
    legend=FALSE,
    title= "Wave 3",
    maximum=Max)

NCTW2vsW3<- NCT(Wave2NetworkIF, Wave3NetworkIF, test.edges=TRUE,
            it=100)


#difference in global strength between the networks of the observed data sets
```

```
NCTW2vsW3$glstrinv.real
#0.7016259

#global strength values of individual networks
NCTW2vsW3$glstrinv.sep
#59.29835 vs 59.99997

NCTW2vsW3$glstrinv.perm
#there's 100 since we did 100 permutations

#Difference in global strength p-value
NCTW2vsW3$glstrinv.pval
#0.7524752 - so not significantly different from one another in regard to global strength

#Value of the max difference in edge weights of observed networks
NCTW2vsW3$nwinv.real
#1.322886

NCTW2vsW3$nwinv.perm
#there's 100 since we did 100 permutations

#Maximum difference in edge weights
NCTW2vsW3$nwinv.pval
#0.2277228 - so not significantly different from one another in regard to number of edge weights

#Which edges significantly differ?
NCTW2vsW3$einv.pvals[which(NCTW2vsW3$einv.pvals[,3]<0.05),]
#            Var1      Var2 p-value
#155         ECIG      PTSD 0.02970297
#207 Dual CIG + ECIG Listening 0.01980198
#209      Marijuana Listening 0.01980198
#251      Listening    Fights 0.01980198
#256          CIG  Restless 0.01980198
#262      Depressed  Restless 0.01980198
#265         PTSD  Restless 0.03960396
#276       Alcohol  Answered 0.02970297
#286         Bully  Answered 0.04950495
```

R File name: W1W2W3 Merged Network

```
#PATH WAVE 1, WAVE 2, WAVE 3 - SA 3
#Merge all data and develop network


############################
# Read in merged dataset  #
############################

setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files")
getwd()
master<-read.csv("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files/w1w2w3master.csv", header=T, sep=',')
names(master)
```

```
#how many complete data
master_complete_cases <- master[complete.cases(master),]
## MASTER TOTAL          = 33106
## MASTER COMPLETE CASES = 21353
## MASTER MISSING        = 11753


#select vars to keep for network modeling
myvars <- c("W1_ACUR_CIG" , "W1_ACUR_ECI" , "W1_ACUR_DUA" , "W1_ACUR_ALC" ,
        "W1_ACUR_MAR" , "W1_ACUR_PAI" ,
        "W1_DEPRESS" ,  "W1_SLEEPING" , "W1_ANXIOUS"  , "W1_PTSD"  ,
        "W1_LIED"    ,   "W1_ATTENTIO" , "W1_LISTENING" , "W1_BULLY" ,
        "W1_FIGHTS"  ,  "W1_RESTLESS" , "W1_ANSWERED",

        "W2_ACUR_CIG" , "W2_ACUR_ECI" , "W2_ACUR_DUA" , "W2_ACUR_ALC" ,
        "W2_ACUR_MAR" , "W2_ACUR_PAI" ,
        "W2_DEPRESS" ,  "W2_SLEEPING" ,  "W2_ANXIOUS" , "W2_PTSD" ,
        "W2_LIED"     ,  "W2_ATTENTIO" , "W2_LISTENIN" , "W2_BULLY"    ,
        "W2_FIGHTS" ,   "W2_RESTLESS",  "W2_ANSWERED",

        "W3_ACUR_CIG" , "W3_ACUR_ECI" , "W3_ACUR_DUA", "W3_ACUR_ALC" ,
        "W3_ACUR_MAR" , "W3_ACUR_PAI" ,
        "W3_DEPRESS",   "W3_SLEEPING" , "W3_ANXIOUS" ,  "W3_PTSD",
        "W3_LIED" ,     "W3_ATTENTIO" , "W3_LISTENIN" , "W3_BULLY" ,
        "W3_FIGHTS"  ,  "W3_RESTLESS"  ,"W3_ANSWERED")
new_master <- master[myvars]


#check distributions
table(new_master$W1_ACUR_PAI)

#rename variables so they look nice on the network
names(new_master)[names(new_master) == "W1_ACUR_CIG"] <- "W1 CIG"
names(new_master)[names(new_master) == "W1_ACUR_ECI"] <- "W1 ECIG"
names(new_master)[names(new_master) == "W1_ACUR_DUA"] <- "W1 Dual CIG + ECIG"
names(new_master)[names(new_master) == "W1_ACUR_ALC"] <- "W1 Alcohol"
names(new_master)[names(new_master) == "W1_ACUR_MAR"] <- "W1 Marijuana"
names(new_master)[names(new_master) == "W1_ACUR_PAI"] <- "W1 PDNP"
names(new_master)[names(new_master) == "W1_DEPRESS"] <- "W1 Depressed"
names(new_master)[names(new_master) == "W1_SLEEPING"] <- "W1 Sleeping"
names(new_master)[names(new_master) == "W1_ANXIOUS"] <- "W1 Anxious"
names(new_master)[names(new_master) == "W1_PTSD"] <- "W1 Distressed/Past"
names(new_master)[names(new_master) == "W1_LIED"] <- "W1 Lied"
names(new_master)[names(new_master) == "W1_ATTENTIO"] <- "W1 Attention"
names(new_master)[names(new_master) == "W1_LISTENING"] <- "W1 Listening"
names(new_master)[names(new_master) == "W1_BULLY"] <- "W1 Bully"
names(new_master)[names(new_master) == "W1_FIGHTS"] <- "W1 Fights"
names(new_master)[names(new_master) == "W1_RESTLESS"] <- "W1 Restless"
names(new_master)[names(new_master) == "W1_ANSWERED"] <- "W1 Answered"

names(new_master)[names(new_master) == "W2_ACUR_CIG"] <- "W2 CIG"
names(new_master)[names(new_master) == "W2_ACUR_ECI"] <- "W2 ECIG"
names(new_master)[names(new_master) == "W2_ACUR_DUA"] <- "W2 Dual CIG + ECIG"
names(new_master)[names(new_master) == "W2_ACUR_ALC"] <- "W2 Alcohol"
names(new_master)[names(new_master) == "W2_ACUR_MAR"] <- "W2 Marijuana"
names(new_master)[names(new_master) == "W2_ACUR_PAI"] <- "W2 PDNP"
```

```
names(new_master)[names(new_master) == "W2_DEPRESS"] <- "W2 Depressed"
names(new_master)[names(new_master) == "W2_SLEEPING"] <- "W2 Sleeping"
names(new_master)[names(new_master) == "W2_ANXIOUS"] <- "W2 Anxious"
names(new_master)[names(new_master) == "W2_PTSD"] <- "W2 Distressed/Past"
names(new_master)[names(new_master) == "W2_LIED"] <- "W2 Lied"
names(new_master)[names(new_master) == "W2_ATTENTIO"] <- "W2 Attention"
names(new_master)[names(new_master) == "W2_LISTENIN"] <- "W2 Listening"
names(new_master)[names(new_master) == "W2_BULLY"] <- "W2 Bully"
names(new_master)[names(new_master) == "W2_FIGHTS"] <- "W2 Fights"
names(new_master)[names(new_master) == "W2_RESTLESS"] <- "W2 Restless"
names(new_master)[names(new_master) == "W2_ANSWERED"] <- "W2 Answered"

names(new_master)[names(new_master) == "W3_ACUR_CIG"] <- "W3 CIG"
names(new_master)[names(new_master) == "W3_ACUR_ECI"] <- "W3 ECIG"
names(new_master)[names(new_master) == "W3_ACUR_DUA"] <- "W3 Dual CIG + ECIG"
names(new_master)[names(new_master) == "W3_ACUR_ALC"] <- "W3 Alcohol"
names(new_master)[names(new_master) == "W3_ACUR_MAR"] <- "W3 Marijuana"
names(new_master)[names(new_master) == "W3_ACUR_PAI"] <- "W3 PDNP"
names(new_master)[names(new_master) == "W3_DEPRESS"] <- "W3 Depressed"
names(new_master)[names(new_master) == "W3_SLEEPING"] <- "W3 Sleeping"
names(new_master)[names(new_master) == "W3_ANXIOUS"] <- "W3 Anxious"
names(new_master)[names(new_master) == "W3_PTSD"] <- "W3 Distressed/Past"
names(new_master)[names(new_master) == "W3_LIED"] <- "W3 Lied"
names(new_master)[names(new_master) == "W3_ATTENTIO"] <- "W3 Attention"
names(new_master)[names(new_master) == "W3_LISTENIN"] <- "W3 Listening"
names(new_master)[names(new_master) == "W3_BULLY"] <- "W3 Bully"
names(new_master)[names(new_master) == "W3_FIGHTS"] <- "W3 Fights"
names(new_master)[names(new_master) == "W3_RESTLESS"] <- "W3 Restless"
names(new_master)[names(new_master) == "W3_ANSWERED"] <- "W3 Answered"


names(new_master)


require(ggplot2)
require(bootnet)
require(IsingFit)
require(IsingSampler)
require(qgraph)


#########
#IsingFit
MasterNetworkIF <-estimateNetwork(new_master, default="IsingFit", missing="listwise")
plot(MasterNetworkIF, layout = "spring", vsize = 10, cex=8)

MasterTraits <- rep(c(
  'Substance Use',
  'Negative Affect',
  'Externalizing',
  'Substance Use',
  'Negative Affect',
  'Externalizing',
  'Substance Use',
  'Negative Affect',
  'Externalizing'
```

```
), times=c(6,4,7,6,4,7,6,4,7))

#COLORED EDGES
layout(t(1))
plot(MasterNetworkIF,
    layout="spring",
    cut=0,
    theme="colorblind",
    groups=MasterTraits,
    #labels=Names,
    #nodeNames=Names,
    #edge.color="black",
    label.scale.equal=TRUE,
    label.cex= 3,
    legend.cex = 0.4,
    title= "Wave 1, Wave 2, and Wave 3")

MasterNetworkIF

#edges
MasterEdges <- MasterNetworkIF$graph
write.csv(MasterEdges, file="MasterEdges.csv")
#this worked!
#(do this for all other edge matrices)

#Very little overlap across the waves, edges within the waves are weaker
```

R File name: Checking for missing data from network

```
## Checking for missing data from network analyses ##


################
#OVERALL WAVE 1#
#setwd("C:\Users\blondinoct\\Documents\\PNASS\\Data Management\\CSVs to use in R")
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files")
getwd()
overall<-read.csv("/Users/courtneyblondino/Library/Mobile
Documents/com~apple~CloudDocs/Dissertation Files/overallwave1.csv", header=T, sep=',')


#rename variables so they look nice on the network
names(overall)[names(overall) == "ACUR_CIG"] <- "CC"
names(overall)[names(overall) == "ACUR_ECI"] <- "EC"
names(overall)[names(overall) == "ACUR_DUA"] <- "Dual CC + EC"
names(overall)[names(overall) == "ACUR_ALC"] <- "Alcohol"
names(overall)[names(overall) == "ACUR_MAR"] <- "Marijuana"
names(overall)[names(overall) == "ACUR_PAI"] <- "PDNP"
names(overall)[names(overall) == "DEPRESS"] <- "Depressed"
names(overall)[names(overall) == "SLEEPING"] <- "Sleeping"
names(overall)[names(overall) == "ANXIOUS"] <- "Anxious"
names(overall)[names(overall) == "PTSD"] <- "PTSD"
names(overall)[names(overall) == "LIED"] <- "Lied"
names(overall)[names(overall) == "ATTENTIO"] <- "Attention"
names(overall)[names(overall) == "LISTENING"] <- "Listening"
```

```
names(overall)[names(overall) == "BULLY"] <- "Bully"
names(overall)[names(overall) == "FIGHTS"] <- "Fights"
names(overall)[names(overall) == "RESTLESS"] <- "Restless"
names(overall)[names(overall) == "ANSWERED"] <- "Answered"


#delete obs with missing data
overall_complete_cases <- overall[complete.cases(overall),]

## W1 OVERALL TOTAL  = 32,320
## W1 COMPLETE CASES = 15,299
## W1 MISSING        = 17,021


######################
##WAVE 1 - MALE ONLY##
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files")
getwd()
male<-read.csv("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation
Files/malewave1.csv", header=T, sep=',')
dim(male)
names(male)

#rename variables so they look nice on the network
names(male)[names(male) == "ACUR_CIG"] <- "CC"
names(male)[names(male) == "ACUR_ECI"] <- "EC"
names(male)[names(male) == "ACUR_DUA"] <- "Dual CC + EC"
names(male)[names(male) == "ACUR_ALC"] <- "Alcohol"
names(male)[names(male) == "ACUR_MAR"] <- "Marijuana"
names(male)[names(male) == "ACUR_PAI"] <- "PDNP"
names(male)[names(male) == "DEPRESS"] <- "Depressed"
names(male)[names(male) == "SLEEPING"] <- "Sleeping"
names(male)[names(male) == "ANXIOUS"] <- "Anxious"
names(male)[names(male) == "PTSD"] <- "PTSD"
names(male)[names(male) == "LIED"] <- "Lied"
names(male)[names(male) == "ATTENTIO"] <- "Attention"
names(male)[names(male) == "LISTENING"] <- "Listening"
names(male)[names(male) == "BULLY"] <- "Bully"
names(male)[names(male) == "FIGHTS"] <- "Fights"
names(male)[names(male) == "RESTLESS"] <- "Restless"
names(male)[names(male) == "ANSWERED"] <- "Answered"


#delete obs with missing data
malew1_complete_cases <- male[complete.cases(male),]

## W1 MALE OVERALL TOTAL  = 16,306
## W1 MALE COMPLETE CASES = 8,406
## W1 MALE MISSING        = 7,900


########################
##WAVE 1 - FEMALE ONLY##
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files")
getwd()
```

```
female<-read.csv("/Users/courtneyblondino/Library/Mobile
Documents/com~apple~CloudDocs/Dissertation Files/femalewave1.csv", header=T, sep=',')
dim(female)
names(female)

#rename variables so they look nice on the network
names(female)[names(female) == "ACUR_CIG"] <- "CC"
names(female)[names(female) == "ACUR_ECI"] <- "EC"
names(female)[names(female) == "ACUR_DUA"] <- "Dual CC + EC"
names(female)[names(female) == "ACUR_ALC"] <- "Alcohol"
names(female)[names(female) == "ACUR_MAR"] <- "Marijuana"
names(female)[names(female) == "ACUR_PAI"] <- "PDNP"
names(female)[names(female) == "DEPRESS"] <- "Depressed"
names(female)[names(female) == "SLEEPING"] <- "Sleeping"
names(female)[names(female) == "ANXIOUS"] <- "Anxious"
names(female)[names(female) == "PTSD"] <- "PTSD"
names(female)[names(female) == "LIED"] <- "Lied"
names(female)[names(female) == "ATTENTIO"] <- "Attention"
names(female)[names(female) == "LISTENING"] <- "Listening"
names(female)[names(female) == "BULLY"] <- "Bully"
names(female)[names(female) == "FIGHTS"] <- "Fights"
names(female)[names(female) == "RESTLESS"] <- "Restless"
names(female)[names(female) == "ANSWERED"] <- "Answered"

#delete obs with missing data
femalew1_complete_cases <- female[complete.cases(female),]


## W1 FEMALE OVERALL TOTAL  = 15,980
## W1 FEMALE COMPLETE CASES = 6,888
## W1 FEMALE MISSING        = 9,092


########
## W2 ##
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files")
getwd()
Wave2<-read.csv("/Users/courtneyblondino/Library/Mobile
Documents/com~apple~CloudDocs/Dissertation Files/w2datafornetcomp232021.csv", header=T, sep=',')
dim(Wave2)
names(Wave2)

#Have to tell R what missing means
Wave2$acur_cignew[Wave2$acur_cignew==-99999] <- NA
Wave2$acur_ecignew[Wave2$acur_ecignew==-99999] <- NA
Wave2$acur_dualnew[Wave2$acur_dualnew==-99999] <- NA
Wave2$acur_alc[Wave2$acur_alc==-99999] <- NA
Wave2$acur_marijuana[Wave2$acur_marijuana==-99999] <- NA
Wave2$acur_painkiller[Wave2$acur_painkiller==-99999] <- NA
Wave2$depressed[Wave2$depressed==-99999] <- NA
Wave2$sleeping[Wave2$sleeping==-99999] <- NA
Wave2$anxious[Wave2$anxious==-99999] <- NA
Wave2$ptsd[Wave2$ptsd==-99999] <- NA
Wave2$lied[Wave2$lied==-99999] <- NA
Wave2$attention[Wave2$attention==-99999] <- NA
Wave2$listening[Wave2$listening==-99999] <- NA
```

```
Wave2$bully[Wave2$bully==-99999] <- NA
Wave2$fights[Wave2$fights==-99999] <- NA
Wave2$restless[Wave2$restless==-99999] <- NA
Wave2$answered[Wave2$answered==-99999] <- NA

#rename variables so they look nice on the network
names(Wave2)[names(Wave2) == "acur_cignew"] <- "CC"
names(Wave2)[names(Wave2) == "acur_ecignew"] <- "EC"
names(Wave2)[names(Wave2) == "acur_dualnew"] <- "Dual CC + EC"
names(Wave2)[names(Wave2) == "acur_alc"] <- "Alcohol"
names(Wave2)[names(Wave2) == "acur_marijuana"] <- "Marijuana"
names(Wave2)[names(Wave2) == "acur_painkiller"] <- "PDNP"
names(Wave2)[names(Wave2) == "depressed"] <- "Depressed"
names(Wave2)[names(Wave2) == "sleeping"] <- "Sleeping"
names(Wave2)[names(Wave2) == "anxious"] <- "Anxious"
names(Wave2)[names(Wave2) == "ptsd"] <- "PTSD"
names(Wave2)[names(Wave2) == "lied"] <- "Lied"
names(Wave2)[names(Wave2) == "attention"] <- "Attention"
names(Wave2)[names(Wave2) == "listening"] <- "Listening"
names(Wave2)[names(Wave2) == "bully"] <- "Bully"
names(Wave2)[names(Wave2) == "fights"] <- "Fights"
names(Wave2)[names(Wave2) == "restless"] <- "Restless"
names(Wave2)[names(Wave2) == "answered"] <- "Answered"

#delete obs with missing data
w2_complete_cases <- Wave2[complete.cases(Wave2),]


## W2 TOTAL          = 26,444
## W2 COMPLETE CASES = 25,592
## W2 MISSING        = 852


########
## W3 ##
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation Files")
getwd()
Wave3<-read.csv("/Users/courtneyblondino/Library/Mobile
Documents/com~apple~CloudDocs/Dissertation Files/w3datafornetcomp332021.csv", header=T, sep=',')
dim(Wave3)
names(Wave3)

#Have to tell R what missing means
Wave3$acur_cignew[Wave3$acur_cignew==-99999] <- NA
Wave3$acur_ecignew[Wave3$acur_ecignew==-99999] <- NA
Wave3$acur_dualnew[Wave3$acur_dualnew==-99999] <- NA
Wave3$acur_alc[Wave3$acur_alc==-99999] <- NA
Wave3$acur_marijuana[Wave3$acur_marijuana==-99999] <- NA
Wave3$acur_painkiller[Wave3$acur_painkiller==-99999] <- NA
Wave3$depressed[Wave3$depressed==-99999] <- NA
Wave3$sleeping[Wave3$sleeping==-99999] <- NA
Wave3$anxious[Wave3$anxious==-99999] <- NA
Wave3$ptsd[Wave3$ptsd==-99999] <- NA
Wave3$lied[Wave3$lied==-99999] <- NA
Wave3$attention[Wave3$attention==-99999] <- NA
Wave3$listening[Wave3$listening==-99999] <- NA
```

```
Wave3$bully[Wave3$bully==-99999] <- NA
Wave3$fights[Wave3$fights==-99999] <- NA
Wave3$restless[Wave3$restless==-99999] <- NA
Wave3$answered[Wave3$answered==-99999] <- NA

#rename variables so they look nice on the network
names(Wave3)[names(Wave3) == "acur_cignew"] <- "CC"
names(Wave3)[names(Wave3) == "acur_ecignew"] <- "EC"
names(Wave3)[names(Wave3) == "acur_dualnew"] <- "Dual CC + EC"
names(Wave3)[names(Wave3) == "acur_alc"] <- "Alcohol"
names(Wave3)[names(Wave3) == "acur_marijuana"] <- "Marijuana"
names(Wave3)[names(Wave3) == "acur_painkiller"] <- "PDNP"
names(Wave3)[names(Wave3) == "depressed"] <- "Depressed"
names(Wave3)[names(Wave3) == "sleeping"] <- "Sleeping"
names(Wave3)[names(Wave3) == "anxious"] <- "Anxious"
names(Wave3)[names(Wave3) == "ptsd"] <- "PTSD"
names(Wave3)[names(Wave3) == "lied"] <- "Lied"
names(Wave3)[names(Wave3) == "attention"] <- "Attention"
names(Wave3)[names(Wave3) == "listening"] <- "Listening"
names(Wave3)[names(Wave3) == "bully"] <- "Bully"
names(Wave3)[names(Wave3) == "fights"] <- "Fights"
names(Wave3)[names(Wave3) == "restless"] <- "Restless"
names(Wave3)[names(Wave3) == "answered"] <- "Answered"


#delete obs with missing data
w3_complete_cases <- Wave3[complete.cases(Wave3),]


## W3 TOTAL          = 26,239
## W3 COMPLETE CASES = 25,359
## W3 MISSING        = 880




###NEW DATA
###APRIL 14
###############
#OVERALL WAVE 1#
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation
Files/Network Chapter")
getwd()
overall<-read.csv("/Users/courtneyblondino/Library/Mobile
Documents/com~apple~CloudDocs/Dissertation Files/Network Chapter/OverallWave1-4142021.csv",
header=T, sep=',')


#rename variables so they look nice on the network
names(overall)[names(overall) == "ACUR_CIG"] <- "CIG"
names(overall)[names(overall) == "ACUR_ECI"] <- "ECIG"
names(overall)[names(overall) == "ACUR_DUA"] <- "Dual CIG + ECIG"
names(overall)[names(overall) == "ACUR_ALC"] <- "Alcohol"
names(overall)[names(overall) == "ACUR_MAR"] <- "Marijuana"
names(overall)[names(overall) == "ACUR_PAI"] <- "PDNP"
names(overall)[names(overall) == "DEPRESS"] <- "Depressed"
```

```
names(overall)[names(overall) == "SLEEPING"] <- "Sleeping"
names(overall)[names(overall) == "ANXIOUS"] <- "Anxious"
names(overall)[names(overall) == "PTSD"] <- "PTSD"
names(overall)[names(overall) == "LIED"] <- "Lied"
names(overall)[names(overall) == "ATTENTIO"] <- "Attention"
names(overall)[names(overall) == "LISTENING"] <- "Listening"
names(overall)[names(overall) == "BULLY"] <- "Bully"
names(overall)[names(overall) == "FIGHTS"] <- "Fights"
names(overall)[names(overall) == "RESTLESS"] <- "Restless"
names(overall)[names(overall) == "ANSWERED"] <- "Answered"


#delete obs with missing data
overall_complete_cases <- overall[complete.cases(overall),]

## W1 OVERALL TOTAL  = 32,320
## W1 COMPLETE CASES = 30,211
## W1 MISSING        = 2,109


######################
##WAVE 1 - MALE ONLY##
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation
Files/Network Chapter")
getwd()
male<-read.csv("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation
Files/Network Chapter/MaleWave1-4142021.csv", header=T, sep=',')
dim(male)
names(male)

#rename variables so they look nice on the network
names(male)[names(male) == "ACUR_CIG"] <- "CIG"
names(male)[names(male) == "ACUR_ECI"] <- "ECIG"
names(male)[names(male) == "ACUR_DUA"] <- "Dual CIG + ECIG"
names(male)[names(male) == "ACUR_ALC"] <- "Alcohol"
names(male)[names(male) == "ACUR_MAR"] <- "Marijuana"
names(male)[names(male) == "ACUR_PAI"] <- "PDNP"
names(male)[names(male) == "DEPRESS"] <- "Depressed"
names(male)[names(male) == "SLEEPING"] <- "Sleeping"
names(male)[names(male) == "ANXIOUS"] <- "Anxious"
names(male)[names(male) == "PTSD"] <- "PTSD"
names(male)[names(male) == "LIED"] <- "Lied"
names(male)[names(male) == "ATTENTIO"] <- "Attention"
names(male)[names(male) == "LISTENING"] <- "Listening"
names(male)[names(male) == "BULLY"] <- "Bully"
names(male)[names(male) == "FIGHTS"] <- "Fights"
names(male)[names(male) == "RESTLESS"] <- "Restless"
names(male)[names(male) == "ANSWERED"] <- "Answered"


#delete obs with missing data
malew1_complete_cases <- male[complete.cases(male),]

## W1 MALE OVERALL TOTAL  = 16,306
## W1 MALE COMPLETE CASES = 15,268
## W1 MALE MISSING        = 1,038
```

```
########################
##WAVE 1 - FEMALE ONLY##
setwd("/Users/courtneyblondino/Library/Mobile Documents/com~apple~CloudDocs/Dissertation
Files/Network Chapter")
getwd()
female<-read.csv("/Users/courtneyblondino/Library/Mobile
Documents/com~apple~CloudDocs/Dissertation Files/Network Chapter/FemaleWave1-4142021.csv",
header=T, sep=',')
dim(female)
names(female)

#rename variables so they look nice on the network
names(female)[names(female) == "ACUR_CIG"] <- "CIG"
names(female)[names(female) == "ACUR_ECI"] <- "ECIG"
names(female)[names(female) == "ACUR_DUA"] <- "Dual CIG + ECIG"
names(female)[names(female) == "ACUR_ALC"] <- "Alcohol"
names(female)[names(female) == "ACUR_MAR"] <- "Marijuana"
names(female)[names(female) == "ACUR_PAI"] <- "PDNP"
names(female)[names(female) == "DEPRESS"] <- "Depressed"
names(female)[names(female) == "SLEEPING"] <- "Sleeping"
names(female)[names(female) == "ANXIOUS"] <- "Anxious"
names(female)[names(female) == "PTSD"] <- "PTSD"
names(female)[names(female) == "LIED"] <- "Lied"
names(female)[names(female) == "ATTENTIO"] <- "Attention"
names(female)[names(female) == "LISTENING"] <- "Listening"
names(female)[names(female) == "BULLY"] <- "Bully"
names(female)[names(female) == "FIGHTS"] <- "Fights"
names(female)[names(female) == "RESTLESS"] <- "Restless"
names(female)[names(female) == "ANSWERED"] <- "Answered"

#delete obs with missing data
femalew1_complete_cases <- female[complete.cases(female),]


## W1 FEMALE OVERALL TOTAL  = 15,980
## W1 FEMALE COMPLETE CASES = 14,925
## W1 FEMALE MISSING        = 1,055
```

VITA

Courtney Taylor Blondino was born September 24, 1993 in Richmond, Virginia and is an American citizen. She graduated from Mills E. Godwin High School, Richmond, VA in 2011. She received her Bachelor of Science in Human Nutrition, Foods, and Exercise with a minor in Medicine and Society from Virginia Tech, Blacksburg, VA in 2015. She received a Master of Public Health in Epidemiology from the University of Kentucky, Lexington, KY in 2017.

Professional and Research Positions

2020-2021 Epidemiology Consultant, Venebio Group LLC

2018-2021 Graduate Student Researcher/Writer, American Medical Association

2018-2021 Graduate Research Assistant/Writer, Virginia Commonwealth University, Office of Assessment, Evaluation and Scholarship

2017-2021 Graduate Research Assistant, Virginia Commonwealth University, Department of Family Medicine and Population Health, Division of Epidemiology

2016-2017 Research Assistant, University of Kentucky, College of Public Health

2016 Public Health Intern, Virginia Department of Health, Office of the Chief Medical Examiner

Publications

Blondino, C., Clifford, J., Lu, J., & Prom-Wormley, E.C. The Association between Internalizing and Externalizing Severity with Current Use of Cigarettes, E-cigarettes, and Alcohol in Adults: Wave 1 (2013-2014) of the Population Assessment of Tobacco and Health (PATH) Study. Accepted to *Addictive Behaviors.*

Usidame, B., Gibson, E., Diallo, A., **Blondino, C.,** Clifford, J., Zanjani, F., Sargent, L., Price, E., Slattum, P., Parsons, P., Prom-Wormley, E.C. Understanding the Preference for Receiving Mental Health and Substance Use Support in African Americans 50 and Older. Accepted to *Journal of Prevention and Intervention in the Community.*

Daly, N., Parsons, M., **Blondino, C.,** Clifford, J., Prom-Wormley, E.C. Association between caregiver depression and child after-school program participation. Accepted to *Journal of Family Social Work.*

Blondino, C.*, Gormley, M.*, Taylor, D.H., Lowery, E., Clifford, J., Burkart, B., Graves, W., Lu, J., & Prom-Wormley, E.C. The Influence of Co-Occurring Substance Use on the

Effectiveness of Opiate Treatment Programs According to Intervention Type. Accepted to *Epidemiologic Reviews.*

Gormley, M.*, **Blondino C.***, Taylor, D., Lowery, E., Graves, W., Clifford, J.S., Prom-Wormley, E.C., & Lu, J. Assessment of polysubstance use during opioid use disorder treatment in the United States: A systematic review. Accepted to *Epidemiologic Reviews.*

Park Y, Ryan MH, Santen SA, Sabo R, **Blondino C**, Magee ML. Nurturing the Student, Sustaining the Mission: 20 Years of the International/Inner-City/Rural Preceptorship Program. Fam Med. 2019;51(10):823-829. https://doi.org/10.22454/FamMed.2019.358223.

Prom-Wormley, E.C., Clifford, J.S., Bourdon, J., Barr, P., **Blondino, C.**, Ball, K. M., Montgomery, J., Davis, J. K., Real, J.E., Edwards, A., Thiselton, D., Wilson, D., Creighton, G.C., & Newbille, C. (2019). Developing community-based strategies with family health history education: Assessing the association between community resident family history and interest in health education. *Social Science and Medicine*. doi: 10.1016/j.socscimed.2019.02.011

Santen, S.A., Feldman, M., Weir, S., **Blondino, C.,** Rawls, M., & DiGiovanni, S. (2018). Developing Comprehensive Strategies to Evaluate Medical School Curricula. *Medical Science Educator*. doi:10.1007/s40670-018-00640-x

*Denotes co-first author.

Presentations

Chartier, K. G., Prom-Wormley, E., Bares, C. B., **Blondino, C. T.**, Mulroy, T., Miles, K., Lee, A. G., Sankoh, M., & Karriker-Jaffe, K. J. When genomics are unavailable: Prioritizing prevention using family history of alcohol problems for medically-underserved adults. Oral presentation at the Society for Prevention Research Annual Meeting, Virtual Program. June, 2021.

Wilson, T.L., **Blondino, C.T.**, Prom-Wormley, E.C. Exploring the Role of E-cigarettes across Cessation Behaviors among US Adults. 27th Annual Meeting of the Society for Research on Nicotine and Tobacco, Virtual, February 24-27, 2021.

Blondino, C., Perera, R., Prom-Wormley, E.C. Using Latent Class Analysis to Understand Comorbidity of Substance Use and Internalizing and Externalizing Symptoms in U.S. Adults. 53rd Annual Society for Epidemiologic Research, Virtual, December 16-18, 2020.

Wilson, T.L., Clifford, J.S., **Blondino, C.T.**, Prom-Wormley, E.C. Examining the Association between Race and Mental Health on Lifetime Frequency of E-nicotine Use in U.S. Adults. 2020 Virginia Public Health Association, Virtual, May, 2020.

Clifford, J.S., Wilson, T., **Blondino, C.**, Prom-Wormley, E.C. The effect of electronic and conventional cigarette coupon receipt on the relationship between income level and past 12-month use in adults in PATH. 26th Annual Meeting of the Society for Research on Nicotine and Tobacco, New Orleans, LA, March 11-14, 2020.

Blondino, C., Perera, R., Prom-Wormley, E.C. Using Latent Class Analysis to Characterize Comorbidity of Substance Use and Internalizing and Externalizing Symptoms in U.S. Adults. Virginia Commonwealth University Division of Epidemiology Seminar Series, Richmond, VA, February 18, 2020.

Lee, J., **Blondino, C.**, Chapman, D. Evaluation of the Autoregressive Cross Lagged Effects among Adolescents' Smartphone Addiction and Online/Offline Delinquency. 52nd Annual Society for Epidemiologic Research, Minneapolis, MN, June 18-21, 2019.

Blondino, C., Gormley, M., Taylor, D.D.H., Lowery, E., Clifford, J.S., Burkart, B., Graves, W.C., Lu, J., Prom-Wormley E.C. The Impact of Polysubstance Use on the Effectiveness of Opioid Use Disorder Therapy by Treatment Type: A Systematic Review. 52nd Annual Society for Epidemiologic Research, Minneapolis, MN, June 18-21, 2019.

Gormley, M., **Blondino, C.**, Taylor, D.D.H., Lowery, E., Clifford, J.S., Burkart, B., Graves, W.C., Prom-Wormley, E.C., Lu, J. Assessment of Polysubstance Use During Opioid Use Disorder Treatment in the United States: A Systematic Review. 52nd Annual Society for Epidemiologic Research, Minneapolis, MN, June 18-21, 2019.

MacDonald, J., Usidame, B., Forr, T., **Blondino, C.**, Prom-Wormley, E.C. Smoking, Drinking, and Aging: Tobacco and Alcohol Effects on Mental Health in Older American Adults. 11th Annual Virginia Commonwealth University Annual Poster Symposium for Undergraduate Research and Creativity, Richmond, VA, April 24, 2019.

Blondino, C., Prom-Wormley, E.C. Using Social Network Analysis to Evaluate Collaboration Among Community-Serving Organizations in the East End of Richmond, Virginia. Virginia Commonwealth University Division of Epidemiology Seminar Series, Richmond, VA, March 19, 2019.

Blondino, C., Perera, R., Prom-Wormley, E.C. Confirmatory Factor Analysis of Internalizing, Externalizing, and Substance Use Problem Symptoms in PATH. Virginia Commonwealth University Division of Epidemiology Seminar Series, Richmond, VA, May 1, 2018.

Paulson, L., Carlyle, K., **Blondino, C.**, Prom-Wormley, E.C. Association Between Threat, Attempt, and Occurrence of Intimate Partner Violence and Chronic Alcohol Use, 14th Annual VCU Women’s Health Research Day, Richmond, VA, April 12, 2018.

Clifford, J.S., Ball, K.M., **Blondino, C**., Prom-Wormley, E.C. The associations between conventional and electronic cigarette use with tobacco messaging. Annual conference of the Virginia Public Health Association, Lynchburg, VA, April, 2018.

Ball, K.M., Clifford, J., **Blondino, C.**, Do, E., Maes, H., Prom-Wormley, E.C. Understanding the associations between opinions towards tobacco and youth current poly-tobacco use. 6th triennial conference of Virginia Conference on Youth Tobacco Use, Richmond, VA, March, 2018.

Blondino, C., Lu, J., Prom-Wormley, E.C. Mental health among current tobacco and alcohol dual users. 24th Annual Meeting of the Society for Research on Nicotine and Tobacco, Baltimore, MD, February 21-24, 2018.

Blondino, C., Seals, J., Brown, S. Do numbers matter? Comparing single homicide followed by suicide and multiple homicide followed by suicide using the National Violent Death Reporting System, 2003-2012. 12th Annual University of Kentucky Center for Clinical and Translation Science, Lexington, KY, March 30, 2017.

Blondino, C., Seals, J., Brown, S. Do numbers matter? Comparing single homicide followed by suicide and multiple homicide followed by suicide using the National Violent Death Reporting System, 2003-2012. Kentucky Violent Death Reporting System Advisory Board Meeting, Frankfort, KY, 2017.