

HETEROGENEOUS TRAJECTORIES OF DEPRESSIVE SYMPTOMS FROM
ADOLESCENCE TO YOUNG ADULTHOOD:
NON-COGNITIVE RISK FACTORS AND LABOR MARKET OUTCOMES

Jessica Nicole Gladstone

A DISSERTATION

in

Education

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2020

Supervisor of Dissertation

Robert F. Boruch, Professor of Education

Graduate Group Chairperson

J. Matthew Hartley, Professor of Education

Dissertation Committee

Robert F. Boruch, Professor of Education

Paul A. McDermott, Professor of Education

Michael J. Rovine, Senior Fellow

HETEROGENEOUS TRAJECTORIES OF DEPRESSIVE SYMPTOMS FROM
ADOLESCENCE TO YOUNG ADULthood:
NON-COGNITIVE RISK FACTORS AND LABOR MARKET OUTCOMES

COPYRIGHT

2020

Jessica Nicole Gladstone

ACKNOWLEDGMENTS

Bob, thank you for believing in me since my first day at interview weekend. Throughout my experiences with you as a student, researcher, teaching assistant, and instructor, you always encouraged me to think like “an engineer”, while still emphasizing the importance of retaining the humanity in my work. In particular, I thank you for encouraging me to explore new projects and join research opportunities which, in effect, expanded my comfort zone and demonstrated what I am truly capable of.

Paul, thank you for your time, generosity and enthusiasm when sharing your invaluable experiences and wisdom. To you and Andrea, thank you for kindly opening up your home (and phone), thus allowing me to learn more about you outside of the classroom. After the first summer session on latent growth mixture modeling, I immediately knew that I wanted to apply this for my dissertation. I am confident when I say that without you, I would not be where I am today.

Mike, I am grateful for you, without hesitation, sharing your knowledge, resources, and impressive quick wit. I always looked forward to our meetings, not only to pick your brain, but also because it was always a pleasure speaking with you. Perhaps my enjoyment was due to our affinities for baseball statistics, reading, and stochastic modeling. Or, perhaps writing such a list serves no good purpose at all. I’d like to think of it as something better left unexplained.

To my friends and colleagues in the QM department, Wharton School, and beyond: thank you for being there. My journey had its share of painful challenges and circumstances, so I rapidly learned how strong my foundation was at Penn. I know many friendships will last a lifetime. In particular, I would like to thank Christine Lee for being my “second mom” for every peak and valley in my journey. Christine, you were great as our program coordinator, but honestly, you were so much more than that. Ann Tiao, thank you for being my best advocate and cheerleader. You both mean so much to me.

Mom and Dad, I’d like to thank you both for staying patient and interested in my research and new ideas. You both have fostered a love for learning and instilled my strong work ethic. I can proudly say that I love what I do and can’t wait to do this for the rest of my life. From the bottom of my heart, I am so lucky to be your daughter.

Danielle, you have been patient with me while I grew as a writer and thinker. Thank you for supporting me without hesitation. Rachel, thank you for challenging me to fight for what I believe in, to be humble, and more open-minded. To the both of you, I love you immensely, and am so honored and lucky to call myself your big sister.

And finally, thank you Dan. When I struggled, you encouraged and supported me to keep pushing through. When I got hesitant, you encouraged me still to try. And when I doubted myself, you were the first person to believe in me. Thank you for showing me the true meaning of sacrifice, dedication, hard work, and of course, love.

ABSTRACT

HETEROGENEOUS TRAJECTORIES OF DEPRESSIVE SYMPTOMS FROM ADOLESCENCE TO
YOUNG ADULthood:
NON-COGNITIVE RISK FACTORS AND LABOR MARKET OUTCOMES

Jessica Gladstone

Robert F. Boruch

Depression commonly emerges during adolescence and is conservatively estimated to affect up to 12.5% of 12- to 17-year-olds annually (Clayborne, Varin & Colman, 2019). Prior longitudinal analyses have identified significant heterogeneity in the level and growth of depressive symptoms during the transitional period from adolescence to young adulthood. The purpose of this study was to follow one representative cohort during this transition to identify non-cognitive, in-school risk factors for atypical depression trajectories and contextualizing them using impactful labor market outcomes. Latent growth mixture modeling (GMM) was used to assess and classify depressive symptom change trajectories using four occasions of measurement from 1994 to 2008. The study used the public-use dataset from the National Study of Adolescent to Adult Health (Add Health). Two distinct change trajectories were identified using a latent basis model and classified as being either Normative (82.2%) or Elevated (17.8%) in its symptom level and shape. The adolescents in the Elevated class exhibited elevated and increasing depressive symptoms, while the Normative class showed consistently lower and decreasing depressive symptoms. Several demographic factors—being female, Black, or Native American—were risk factors for membership in the Elevated class. In addition, four non-cognitive, within-school indicators were associated with a significantly higher risk for an Elevated classification. The strongest non-cognitive risk factor was low levels of school connection, followed by high delinquency, low self-perceived likelihood for college admission, and retention in grade. Lastly, adults who were classified as Elevated in their depressive symptoms reported significantly lower socioeconomic outcomes across *all* eleven labor market indicators, including measures of employment benefits, job satisfaction, income, and public assistance.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iii
ABSTRACT	iv
LIST OF TABLES	viii
LIST OF ILLUSTRATIONS	ix
CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW	1
Understanding Depression	1
Depressive Symptoms	1
Depression in the United States	2
Depression in Adolescence	3
Depression, Adolescence and School	5
Depression Trends from Adolescence to Adulthood	8
Heterogeneity of Depression Trajectories	10
Demographics and Depression	12
Adjustment Behaviors and Depression	14
Labor Market Participation	16
Current Study	19
Purpose of Study	19
Primary Research Questions	22
CHAPTER 2: METHODS	24
Data	24

Sample.....	24
Measures	27
Longitudinal Missing Data	37
Case-Level Missingness.....	37
Item-Level Missingness – Depression	39
Item-Level Missingness – Antecedent Variables	40
Imputation Procedure	41
Dimensionality	43
Longitudinal Measurement Invariance	44
Equating and Scaling.....	45
Latent Growth Mixture Models	46
Risk Factors Models	48
Distal Outcomes Models.....	50
CHAPTER 3: RESULTS	51
Antecedent Variables – School Disengagement and Delinquency	51
Depressive Symptoms.....	52
Dimensionality	54
Longitudinal Measurement Invariance	56
Equating and Scaling	56
Latent Growth Mixture Models	57
Multinomial Logistic Regression.....	69
Logistic Regression.....	73
CHAPTER 4: DISCUSSION	78

Methodological Challenges	78
Review of Research Questions	80
Limitations and Future Directions	86
Missing Data	86
Using Wave vs. Chronological Age	87
Self-Report Items	87
The Great Recession (2007-2009)	88
Other Revisions	90
Research Contributions and Implications	92
BIBLIOGRAPHY	95

LIST OF TABLES

Table 1. <i>Sample Demographic and Student Characteristics</i>	27
Table 2. <i>School Disengagement Items</i>	31
Table 3. <i>Delinquency Items</i>	35
Table 4. <i>Non-Missing Cases of Depression Items by Wave</i>	39
Table 5. <i>Complete Observations by Wave for Depression Items</i>	40
Table 6. <i>Item Level Characteristics of Depressive Symptoms by Wave</i>	52
Table 7. <i>Initial Factor Structure for CES-D (9-item) Scale</i>	55
Table 8. <i>Properties, Fit Statistics, and Parameter Estimates for Latent Basis, Latent Growth Mixture Models</i>	59
Table 9. <i>Properties, Fit Statistics, and Parameter Estimates for Fixed Basis, Latent Growth Mixture Models</i>	62
Table 10. <i>Relationship between Demographic and Non-Cognitive Adolescent Characteristics, and Latent Classes</i>	71

LIST OF ILLUSTRATIONS

<i>Figure 1.</i> Visual depiction of a latent growth mixture model.....	47
<i>Figure 2.</i> Depiction of a latent growth mixture model with antecedent covariates.....	49
<i>Figure 3.</i> Depiction of a latent growth mixture model with distal outcomes	50
<i>Figure 4.</i> Estimated mean latent growth trajectories for depressive symptoms	65
<i>Figure 5.</i> Estimated mean latent growth trajectories for non-missing data.....	67
<i>Figure 6.</i> Distribution of depressive symptom scores by latent class and wave.	69
<i>Figure 7.</i> Predicted mean probability (and 95% confidence bands) of labor market outcomes (A-D) associated with class membership	75
<i>Figure 8.</i> Predicted mean probability (and 95% confidence bands) of labor market outcomes (E-H) associated with class membership.....	76
<i>Figure 9.</i> Predicted mean probability (and 95% confidence bands) of labor market outcomes (I-K) associated with class membership.....	77

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

Understanding Depression

The World Health Organization (WHO) ranked depression as one of the leading causes of disability worldwide and projected that it will rank second—only to heart disease—by the year 2020 (World Health Organization, 2017). Depression is a risk factor for mortality across numerous causes of death, with risks on par with the associations between smoking and mortality (Mykletun et al., 2007). Living with depression has been linked to significant reductions in personal functioning, including decreased performance within occupational roles, impaired social relationships (Spijker et al., 2004), and an elevated risk for physical comorbidities, such as cardiovascular disease, arthritis, and obesity (Chapman, Perry & Strine, 2005; Ferrari et al., 2010). Empirical evidence also suggests that the persistence of depressive symptoms leads to a progressive decline in daily functioning – there may be a positive feedback loop between depression and functional disability (Spijker et al., 2004).

Depressive Symptoms

Depression is a common and major public health problem in adolescents and adults and negatively impacts how you behave, feel, and think (American Psychiatric Association, 2013; Torres, 2020). Depression frequently results in a loss of personal interest in occasions and activities which were previously enjoyable. As outlined in the DSM-V (American Psychiatric Association, 2013), depressive symptoms vary in severity, with diagnostic indicators including trouble sleeping or sleeping too much, changing appetite, feeling sad, having depressed mood, feeling guilty or worthless,

decreasing concentration, thinking, and/or making decisions, and displaying thoughts of death or suicide (Torres, 2020). Not all episodes of depression manifest with all possible symptoms. The DSM-V provides instructions on diagnostic criteria for a major depressive episode (i.e., clinical depression) and requires that depressive symptoms persist for at least two weeks and hinders one's previous level of functioning (American Psychiatric Association, 2013).

Furthermore, depression can result in a multitude of problems at work and at home by decreasing the ability to create and maintain healthy relationships (Torres, 2020). Depression can be triggered from a single negative belief or occurrence where an individual perceives as a lessening of his/her own self-worth (Hammen, 2009). However, depression also affects individuals who appear to live in somewhat "ideal circumstances" (Torres, 2020). Several key factors have been shown to play a role in depression susceptibility and include personality, genetics, biochemistry, and environmental characteristics (Torres, 2020).

Depression in the United States

Depression is one of the most common mental disorders in the United States (National Institute of Mental Health [NIMH], 2017). An estimated 17.3 million American adults reported having at least one episode of major depressive disorder in a single year (NIMH, 2017). Among American adults, the incidence of depression is most common among young adults between 18 and 25 years old (13.1%). However, no echelon of American adults has negligible rates of major depression, with 7.7% and 4.7% Americans aged 26-49 and 50+ estimated to experience at least one episode of clinical

depression within a single year (NIMH, 2017). In 2017, an estimated 11 million American adults aged 18 or older had at least one major depression episode with severe functional impairment (NIMH, 2017), which are described as clear-cut, manifest losses in normative daily functioning across several domains (Üstün & Kennedy, 2009). Such impairments may include a reduction in routine hygienic care, failure to form and maintain interpersonal relationships, and significant loss in productivity at work or school (American Psychiatric Association, 2013; Hammer-Helmich et al., 2018; Üstün & Kennedy, 2009). Within one year, depression classified with severe functional impairment was estimated to make up 64% of major depressive episodes among U.S. adults (NIMH, 2017). However, despite its prevalence and significant effects on normative functioning, over 35% of adults with clinical depression never receive adequate treatment (NIMH, 2017).

Depression in Adolescence

Within a single year, more than three million U.S. adolescents (ages 12-17) were estimated to have at least one major depressive episode (NIMH, 2017). In addition, over two million adolescents had at least one episode with severe functional impairment, which represents 9.4% of the American adolescent population aged between 12 and 17 years (NIMH, 2017). Substantial evidence suggests that depressive episodes are most common in adolescents than in any other age group (Rao, Hammen & Poland, 2010; Rutter, 1991), with rates for major and minor depression estimated at 12.4% and 7.1%, respectively (Kessler & Walters, 1998), which nearly doubles the prevalence of depression in adults across the United States. During adolescence, depression becomes

relatively common, with prevalence rates significantly higher than during childhood (Lewinsohn, Clarke, Seeley & Rohde, 1994). Among all adolescent psychological disorders, clinical depression is most frequently diagnosed condition (Hammen, 2009).

The transitional period of adolescence is marked with increased psychological stress and life-altering developmental processes that, to some degree, contribute to a rise in depressive symptoms. Some of these changes include puberty-related hormone transformations (Ge, Conger & Elder, 2001), new and changing relationships with parents, peers, and romantic partners (Hankin, Mermelstein & Roesch, 2007), and growing a capacity for personal reflection and abstract thought (Costello, Swendsen, Rose & Dierker, 2008; Nolen-Hoeksema, 1994). Depression during adolescence may affect the usual development of healthy competencies across several domains that increase the risk of significant consequences in adulthood (Costello et al., 2008; Kessler & Walters, 1998). For example, depressed youth may not possess sufficient normative interpersonal interactions that are needed to forge and maintain healthy familial, platonic, and romantic relationships (Wickrama & Wickrama, 2010).

Like other internalizing disorders (i.e., anxiety), depression is a significant predictor of later sickness and disability, with evidence indicating a stronger association when depression is experienced during adolescence rather than adulthood (Mykletun et al., 2007). In a meta-analysis conducted by Melkevik et al. (2016), the authors reported that regardless of research design, most (six out of seven) studies found significant associations between depression and early school dropout, which is important to consider when studying potential impacts on later socioeconomic outcomes (i.e., higher risk for poor financial health and low employment attainment [Campbell, 2015]).

Adolescent depression transpires during a critical point in life that strongly impacts a person's growth, self-confidence, and lifestyle (Zhu, 2018). A single episode of depression during adolescence has been significantly linked to higher rates of adjustment problems in occupational, educational, and financial domains later in life (Costello et al., 2008; Fergusson et al., 2007; Holsen & Birkeland, 2017; Zhu, 2018). Furthermore, having at least one episode of clinical depression during adolescence is significantly associated with a higher risk for additional depressive episodes later during adolescence and adulthood (Fergusson et al., 2007; Zhu, 2018).

Depression, Adolescence and School

School can be viewed as a location where the macro-levels of society and culture meet middle- and micro-levels of community, school and classroom. Adolescents spend more time in school than any other setting, with the exception of their home. In school, personal affect, temperament, and experience exhibit strong influences on intellectual development, school engagement, and psychological well-being (Eccles & Roeser, 2011; Levitt, Saka, Romanelli & Hoagwood, 2007). Parents and peers frequently portray school and its outcomes as high stakes, “make or break” moments. While some students flourish in a school setting, many others grapple with increased levels of anxiety and unhappiness, while others struggle with social alienation and heightened emotional pain (Eccles & Roeser, 2011).

School experiences, whether with school leaders, teachers or peers, can shape adolescents' learning, motivation, and development (Eccles & Roeser, 2010). The design, content, and scope of academic curriculum not only impacts what students learn,

but also how students perceive what they learn. Of course, many non-cognitive qualities such as grit, integrity, and motivation play significant roles in learning, but researchers across both cross-sectional and longitudinal domains support the premise that meaningful academic curriculum can increase school connectedness and promote motivation to learn (Burchinal, Roberts, Zeisel & Rowley, 2008; Eccles & Roeser, 2010; Roeser, Eccles & Sameroff, 2000; Shocet et al., 2006).

School disengagement. Adolescent depression has been associated with reduced academic performance and elevated school disengagement, both of which are often cited as risk factors for early school dropout (Jaycox et al., 2009; Resnick, 1997). School (dis)engagement broadly relates to adolescents' affective experiences within the school setting and its interrelated constructs, with an emphasis on student perceptions of connections within their school community (Ozer, Wolf & Kong, 2008). Most research into schooling and depression includes some construct of measured school engagement, or closeness. This construct often consists of self-reported of school belonging, safety, happiness, closeness with others, and perceived relationships with peers and teachers (Ozer, Wolf & Kong, 2008).

Both cognitive and non-cognitive factors are related to school disengagement (Ozer, Wolf & Kong, 2008; Resnick, 1997). School disengagement has been linked to unhealthy school behaviors and its consequences, such as skipping/cutting class, receiving in and/or out of school suspension, plagiarism, frequent absenteeism, and permanent expulsion (Levitt, Saka, Romanelli & Hoagwood, 2007; Ozer, Wolf & Kong, 2008). School disengagement is a risk factor for numerous psychiatric ailments, such as depression, conduct disorder, and ADHD (Vaughn et al., 2010). Therefore, there may be

relevant indicators of adolescent school disengagement worth considering when studying later psychological and socioeconomic outcomes (Ozer, Wolf & Kong, 2008; Vaughn et al., 2010).

Delinquency. Comorbidities between internalizing and externalizing behaviors demonstrate significant, positive, and reciprocal trends across their co-development through adolescence, suggesting a link between atypical adolescent development and behavior which may be mutually and/or causally associated over time (Gilliom & Shaw, 2004; Loeber & Burke, 2011; Reynolds & Crea, 2015). For example, one explanation for such findings is that both internalizing and externalizing adolescent behaviors are driven by the same, or similar, temperamental factors, such as high emotionality (Gilliom & Shaw, 2004; Weeks et al., 2016). Both Gilliom & Shaw (2004) and Weeks et al. (2016) highlighted that externalizing behaviors during early adolescence (ages 12-13) should be of particular focus when studying associations between externalizing behaviors and adolescent depression, because this age coincides with a new transition to secondary school and the typical onset of puberty, both of which promote stressful changes that may negatively affect mental adjustment and later development.

Poor mental health during adolescent development not only weakens the healthy development of psychosocial competencies, but also contributes to other physical and mental health problems (Wickrama & Wickrama, 2010). For example, depressed youth often lack appropriate faculties for strong and normative interpersonal interactions with others, thereby eroding positive, healthy social relationships with their family and friends (Elder & Caspi, 1988). The loss and/or dysfunction of an adolescent's healthy

friendships and family relationships is significantly associated with low levels of positive influence and puts them at higher risk for criminality and other risky behavior(s) (Hammen, 2009; Wickrama & Wickrama, 2010). Some longitudinal researchers found adolescent depression was significantly correlated to dangerous conduct and significant outcomes in young adulthood, which included greater abuse of illicit substances, higher levels of incarceration, and more frequent sexual encounters with multiple partners (Wickrama & Wickrama, 2010).

Based on the general theory of self-control, individuals with a reduced or lack of self-control make choices based on simple gratification of desires, an inability to consider subsequent consequences, and a lack of planning for long-term goals (Gottfredson & Hirschi, 1990). Within this framework, adolescent depression may contribute to personal degradation in self-esteem and loss of the sense of control, which puts them at greater risk for actions for present fulfilment of wants and a neglect for future consequences (Reynolds & Crea, 2015). Furthermore, adolescents in a depressed state often suffer from lack of aspiration(s). Therefore, they have been associated to be at a higher risk of making poor choices which undermine preservation of a healthy, normative lifestyle (Ng & Jeffery, 2003).

Depression Trends from Adolescence to Adulthood

Personal sensitivities in one's mental health can quickly manifest during life's unforeseen changes and circumstances. For some individuals, such situations may act as a turning point in their mental health (Holsen & Birkeland, 2017; Schulenberg & Zarrett, 2006). The manner in which individuals react and cope during life changes may be

reflected in individual patterns of change in measures of mental health, including depressive symptoms, and also may be revealed across measures of normative functioning during later life (Arnett, 2000).

When averaged across all Americans from adolescence to young adulthood, the mean change trajectory of depressive symptoms is often characterized with an increase in average symptoms from early through mid-adolescence, peaking around mid-adolescence, which is followed by a steady decline in depressive symptoms during late-adolescence through young adulthood (Adkins et al., 2009; Barr, 2018; Cole et al., 2002; Ge, Lorenz, Conger, Elder & Simons, 1994; Holsen & Birkeland, 2017; Yaroslavsky et al., 2013).

Dissimilar from the average trajectory described above, the modal trajectory of depressive symptoms during this period starts negligible or slightly elevated during early and mid-adolescence, with most, if not all, depressive symptoms subsiding by young adulthood (Ames & Leadbeater, 2018; Costello et al., 2008; Stoolmiller, Kim & Capaldi, 2005; Wickrama & Wickrama, 2010). Although adolescents have the highest prevalence rate of depression and associated symptoms, the majority of adolescents are characterized by few depressive symptoms, which remain stable and low into adulthood (Holsen & Birkeland, 2017). Based on the literature, this subgroup, or class, contains the largest proportion of adolescents and is typically classified as “normative” in depressive symptoms from adolescence to young adulthood (Holsen & Birkeland, 2017; Wickrama & Wickrama, 2010; Yaroslavsky et al., 2013).

Heterogeneity of Depression Trajectories

A growing body of evidence suggests there is significant heterogeneity in both the level and shape of depressive symptom trajectories within the target population (i.e., Ames & Leadbeater, 2018; Barr, 2018; Stoolmiller, Kim & Capaldi, 2005; Yaroslavsky et al., 2013; Zhu, 2018). Most identified trajectories are not described by the average nor modal class trajectory. Despite robust evidence supporting the typical level and shape of change in depressive symptoms, there still exists considerable support for additional latent classes in the population with significantly different levels and/or shapes in their change trajectories. Symptoms measures, such as the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977), measure depression as a “continuum of severity” (Costello et al., 2008, p. 174), which recognizes clinical and subclinical findings (Brendgen et al., 2005). By using symptom scales, researchers have uncovered heterogeneity in depressive symptoms that “are likely to eventuate through a variety of developmental pathways” (Cicchetti & Toth, 1998, p. 221).

Consistent with developmental taxonomic theories, several studies examining the heterogeneity of depressive symptoms have found evidence of non-normative latent classes that follow distinct patterns of change of depressive symptoms from adolescence to young adulthood (Barr, 2018; Costello et al., 2008; Moffit & Caspi, 2002; Repetto, Caldwell & Zimmerman, 2004; Stoolmiller, Kim & Capaldi, 2005; Yaroslavsky et al., 2013). Although these studies had varied by methodology (i.e., latent growth mixture modeling, latent class analysis, hierarchical linear modeling), sample characteristics, scale development, model selection criteria, and measure of depression, most studies had

settled on either a 3- or 4-class solution (Ames & Leadbeater, 2018; Brendgen et al., 2005; Costello et al., 2008; Stoolmiller, Kim & Capaldi, 2005; Wickrama & Wickrama, 2010).

There are several approaches one may take when modeling longitudinal data, including methods such as latent growth curve analysis (Willett & Sayer, 1994), hierarchical linear modeling (Bryk & Raudenbush, 1992), latent class analysis (Muthén, 2004; Nagin, 2005), latent growth curve modeling (Duncan, Duncan & Strycker, 2009) and latent growth mixture modeling (Muthén & Muthén, 2000). Model choice should be determined based on the purpose of the study (Frankfurt, Frazier, Syed & Jung, 2016), since some models are interested in examining an average pattern of change, while other models were developed and used to identify distinct patterns of change based on empirical likenesses (Frankfurt et al., 2016). One such method, latent growth mixture modeling, tests whether there are latent classes within the target population, where each class follows a qualitatively distinct developmental change trajectory for a particular outcome, such as depressive symptoms (Muthén & Muthén, 2000; Ram & Grimm, 2009; Frankfurt et al., 2016).

Using latent growth mixture modeling, researchers can address traditional inquiries in psychological research, including whether an intervention works similarly for all individuals on average, and addressing the extent to which individuals may vary in their response to a new school program or treatment. In addition, latent growth mixture modeling allows researchers to examine each latent class' associations with relevant antecedent variables and distal outcomes, providing further support for the meaningfulness of latent subgroups (Brendgen et al., 2005; Muthén & Muthén, 2000;

Ram & Grimm, 2009), and offering a different perspective when viewing psychological change over time (Frankfurt et al., 2016).

By classifying temporal change patterns of depressive symptoms via latent growth mixture modeling, researchers, policymakers and clinicians can better understand who is likely to develop non-normative depressive symptoms and the manner(s) in which symptoms manifest over time (Barr, 2018; Frankfurt et al., 2016; Stoolmiller, Kim & Capaldi, 2005). Since the onset of depressive symptoms generally starts during early to mid-adolescence (Schubert et al., 2017), studying this development period is of particular importance to identify distinct patterns of change and establish psychological profiles for at-risk youth within schools.

Demographics and Depression

Gender. In previous growth mixture modeling studies, the most common risk factor for classification in a non-normative depression trajectory was being female (Barr, 2018; Brendgen et al., 2005; Costello et al., 2008; Meadows, Brown & Elder, 2006; Yaroslavsky et al., 2013). Starting around age 12-13, females begin to consistently report higher average levels of depression than their male counterparts (Campbell, Byrne & Baron, 1992; Holsen & Birkeland, 2017; Twenge & Nolen-Hoeksema, 2002). In addition, females tend to exhibit a steeper increase in depressive symptoms during early-to mid-adolescence, but also with a steeper decline into young adulthood, relative to males (Holsen & Birkeland, 2017). Relative to adolescent females, trajectories of depressive symptoms among males were relatively steady and lower, on average (Cole et al., 2002; Garber, 2006; Stoolmiller, Kim & Capaldi, 2005). Even so, there is still

supporting evidence of significant heterogeneity among male adolescents, as reported in the all-male study conducted by Stoolmiller, Kim & Capaldi (2005) in Oregon.

Females are theorized to be more susceptible to peer-related disruptions than males (Hankin & Abramson, 2001). Some studies found that in both adolescence and adulthood, women were not only more likely to report exposure to stressful life events but also were more vulnerable to stress having larger, more damaging effects on their reported mental health (Ge et al., 1994). However, other studies on adolescent depression found no support for differential stress exposure as the explanation for gender discrepancies in average depressive symptoms (Avison & McAlpine, 1992; Gore, Aseltine & Colten, 1993).

Race/Ethnicity. Among all American adolescents, the prevalence of clinical depression was highest among those who identified with more than one race/ethnicity, with approximately 17% of multiracial/ethnic adolescents having at least one major depressive episode within one year (NIMH, 2017). Significant demographic factors from other studies that used latent growth mixture modeling identified that being Black, Hispanic, Native American, Pacific Islander and Asian American (Wight, Aneshensel, Botticello & Sepúlveda, 2005) were associated with higher risk of non-normative depressive symptom trajectories. Furthermore, Wight et al. (2005) found that non-White Americans were more likely to have significantly higher depressive symptoms during adolescence, but these differences diminished by young adulthood.

Adjustment Behaviors and Depression

Behavioral comorbidities, such as delinquency, were more common in people with depression than those who were not (Brière, Janosz, Fallu & Morizot, 2015; Hammen, 2009). Academic problems, like early dropout, poor school bonding and high absenteeism were also related to elevated depressive symptoms during adolescence (Costello et al., 2008; Shocet et al., 2006). Temperamental characteristics are believed to play a role in both the development of depressive symptoms and behavioral patterns (Brendgen et al., 2005). Temperament is roughly described as one's emotional and behavioral style that remains stable over time and situations, which has a biological basis that can be modified by the environment (Prior, 1992; Rothbart & Bates, 2007). The primary temperamental characteristic that appears to play a key role in the development of depressive symptoms is reactivity/negative emotionality, which is theoretically linked to neuroticism, and a tendency to feel sadness, discomfort, and fear (Brendgen et al., 2005; Compas, Connor-Smith & Jaser, 2004; Sanson, Hemphill & Smart, 2004).

Delinquent behavior. There is evidence that externalizing disorders, such as conduct disorder and ADHD, and related behaviors act as risk factors for elevated depressive symptoms during both adolescence and later adulthood (Brière et al., 2015; Costello et al., 2008; McLeod, Uemura & Rohrman, 2012; Wallin et al., 2018). During adolescence, researchers found co-occurring patterns of delinquent behaviors/conduct problems and depressive symptoms (Capaldi & Stoolmiller, 1999; Wiesner & Kim, 2006). Studies using both clinical and community samples found considerable support that adolescents with high levels of delinquency were at significantly greater risk for

elevated depressive symptoms, and vice versa (Brière et al., 2015; Wiesner & Kim, 2006).

School disengagement. School connectedness, as defined by Goodenow (1993) is “the extent to which students feel personally accepted, respected, included, and supported by others in the school social environment” and consistently has been associated with adolescent depression (Goodenow, 1993; Zhu, 2018). When adolescents perceive care and feel appreciated by teachers and school faculty, they are significantly less likely to engage in risky behaviors, including participating in violent acts and use of illicit substances (McNeely, Nonnemaker & Blum, 2002). Social support from teachers, parents, and peers was found to act as a protective factor for both depression and early school dropout (Costello et al., 2008; Lewinsohn et al., 1994). Likewise, some researchers found that measures of school connectedness were among the strongest predictors of adolescent depression (Costello et al., 2008; McNeely, Nonnemaker & Blum, 2002; Millings et al., 2012; Shochet, Dadds, Ham & Montague, 2006; Zhu, 2018).

One theory for the significant associations among school disengagement, adolescent depression, and early school leaving, suggests that teachers prefer students who complete their work with positive attitudes, stay organized, and are not disruptive in the classroom (Henricsson & Rydell, 2004; McLeod, Uemura & Rohrman, 2012). Teachers were found to give heavier weight to non-cognitive skills and behaviors while evaluating school performance and academic merit, even when done so subconsciously (Farkas, 1996; McLeod, Uemura & Rohrman, 2012). Out of the classroom, but still within schools, students whose personal temperament and behaviors work to maintain social order are often rewarded, while schools repeatedly reprimand and punish students

with behaviors that are viewed as disruptive, threatening and/or not conducive for learning (Farkas, 1996; McLeod, Uemura & Rohrman, 2012). It is theorized that students who drop out of school tend to do so through a process of “withdrawal” and “disengagement” (Vaughn et al., 2010).

Labor Market Participation

Veldman et al. (2015) used latent growth mixture modeling to demonstrate that mental health problems in adolescence alone was associated with significantly lower expectations in educational and employment attainment during adulthood (Veldman et al., 2015). Furthermore, Kessler’s (2012) meta-analysis on the cost of depression cited findings from several studies that found having a history of one or more depressive episodes during childhood or adolescence was associated with significant reductions in expected per capita and household income during adulthood (Goodman, Joyce & Smith, 2011; Kessler, 2012; Smith & Smith, 2010). As mentioned previously, the onset of mental health problems, including depression, during adolescence is a prominent risk factor for early school dropout, which can directly limit career opportunities and standard of living (Lee et al., 2009).

When studied cross-sectionally, the average personal and household earnings were significantly lower for those reportedly suffering from depression, relative to those who were not (Kessler, 2012). In one quasi-experimental study that focused on job loss and depressive symptoms (Dooley, Fielding & Levi, 1996), found that the relationship between low income and depressive symptom severity significantly operated in a positive feedback loop.

Numerous cross-sectional studies have noted that elevated depressive mood was associated with significant deficiencies in personal employment (Fried & Nesse, 2014; Judd, Paulus, Wells & Rapaport, 1997). For example, it was reported that adults with elevated depressive symptoms were more likely to avoid challenges and opportunities in their occupation (Holsen & Birkeland, 2017). Workers who are depressed were at a significantly higher risk for permanent job termination, relative to those who are not depressed (Kessler, 2012). Based on these findings, people with elevated depressive symptoms find it more difficult to maintain a stability at work (i.e., more frequent change in positions, job termination, and employment dissatisfaction).

Researchers have also reported that depressed individuals may have substantial (and at times, permanent) deficiencies in multiple functioning domains, with severe functioning impairments on par with a disability caused by a chronic physical ailment (Buist-Bouwman, de Graff, Volleberg & Ormel, 2005; Druss et al., 2008; Hays, 1995; Ormel et al., 1998). Prior studies indicated that the mostly expensive conditions incurred at the societal level were those that were most common (Stewart, Ricci & Chee, 2003). In 2000, the two diseases with the highest overall burden on the United States' population were both mental disorders: major depression and alcoholism (Druss et al., 2008). In 2003, Stewart, Ricci & Chee estimated that the symptoms associated with depression alone cost \$40 billion dollars per year in the United States. Economic studies have estimated that approximately one-third of work absences due to sickness were caused by mental, not physical, ailments (Merikangas et al., 2007). Furthermore, several studies found there was no significant difference between the number of sick days caused by mental or physical suffering (Buist-Bouwman et al., 2005; Druss et al., 2008; Kessler,

Ormel, Demler & Stang, 2003). Depression, while quite devastating for the struggling employee, also directly impacts employers who are affected by work-related losses, such as lower productivity and increased sick leave. When considered separately or together, physical and mental illness cause significantly more absences from work (Buist-Bouwman et al., 2005) and fewer hours worked per day (Kessler et al., 2003). In the early 2000s, U.S. estimates of the salary-equivalent capital value of losses due to depression ranged from \$30.1 billion (Stewart, Ricci & Chee, 2003) to \$51.5 billion (Greenberg et al., 2003).

In addition to losses in work productivity, there are additional costs for the treatment and support of persons with mental health problems. Individuals suffering from depression significantly rely on increased use of publicly funded social and healthcare services (Stewart, Ricci & Chee, 2003; Wilson & Drury, 1984). Depression is also a significant risk factor for both temporary and permanent disability for American workers (Kessler, 2012). In fact, about one-third of all Americans currently on disability cited permanent functional impairments caused by a mental disorder alone (Druss et al., 2008). Approximately 8% of Americans on disability reported major depression as the reason they are unfit for work (Druss et al., 2008). Additionally, Fergusson, Boden & Horwood (2007) reported that adults with at least ten depressive episodes from the beginning of adolescence to present-day were more than twice as likely to be fully reliant on welfare program(s) (Fergusson, Boden & Horwood, 2007). Despite direct and indirect effects at the societal level, depression is not always viewed as a prominent public health crisis. The single greatest source of lost productivity was due to health-related reductions in work performance, not work absence (Stewart, Ricci & Chee, 2003). Research

indicates that depressed people operate a sub-standard level, relative to their ideal, fully realized potential.

Current Study

Purpose of Study

The primary purpose of this study was to discover whether there were latent depressive mood change trajectories in a large, nationally representative sample measured from adolescence to young adulthood. Since a growing number of studies have identified significant and distinct latent trajectories of depressive symptom change, traditional longitudinal methods which assume a single-population model may not be suitable for research used to inform comprehensive policy and targeted prevention (Brière et al., 2015; Frankfurt et al., 2016; Veldman et al., 2015). Numerous findings of heterogeneity in depressive symptoms were consistent with developmental taxonomic theories (Wickrama & Wickrama, 2010). However, the number of clear, identifiable depressive change trajectories in the United States remained somewhat unclear. While there is well-documented evidence about harm that depression may cause during a single point in time (i.e., an episode), research on the scope and depth of long-term impacts from non-normative developmental change of depressive symptoms during the transition from adolescence to adulthood remains still somewhat unexplored.

More than 60% of adolescents with a major depressive episode do not receive any form of treatment (NIMH, 2017). The aim of in-school mental health programs is to help adolescents attain their academic and emotional best, while also reducing the stigma of mental health struggles. Traditional mental health services and programs attempt to

reach this aim by offering additional psychological testing, time-restricted individual and group counseling, evaluation and placement in special education programs, and consultations with teachers, staff and family members (Levitt, Saka, Romanelli & Hoagwood, 2007). However, most of these traditional offerings offer minimal support to adolescents struggling with depression. Recent interest has been focused on evaluating the expansion of within-school mental health services by fully integrating programs into school structures (Adelman & Taylor, 2000; Levitt et al., 2007).

From a public health perspective, schools are an ideal place for screening adolescents for latent mental health problems because schools offer unparalleled access to a large number of youth (Levitt et al., 2007). From a research perspective, better identification and management of mental health problems among adolescent students is important. There is better evidence and recognition that adolescents' mental health functioning in school may encourage, or discourage, learning and normative development. In addition, consequential outcomes associated with untreated mental health ailments are somewhat avoidable with proper treatment and de-stigmatization (Levitt et al., 2007). This can happen through changes in school curricula and a much-needed change in mental health stigmatization.

Studies revealing the high prevalence of unidentified and untreated adolescent depression have inspired many researchers to make improvements for early and time-sensitive identification of at-risk youth and subsequent treatment (Levitt et al., 2007). Current research supports the claim that targeted psychological interventions for specific adolescents are more effective than generalized programs for adolescent mental health (Costello et al., 2008; Horowitz & Garber, 2006; Levitt et al., 2007). By identifying and

gauging the individual impacts that risk and protective factors have on developmental trajectories of depressive symptoms, researchers and policy makers have the ability to make more informed decisions when planning intervention designs for at-risk youth (Costello et al., 2008; Yaroslavsky et al., 2013).

One of this study's aims was to reveal whether adolescent school disengagement was associated with latent class membership of depressive symptoms. Within-school behavioral and academic problems are characterized as both influenced by and influencers of depressive symptoms (Brière et al., 2015; Stoolmiller, Kim & Capaldi, 2005). A better awareness of non-cognitive, school-related risk factors may reveal the type(s) of students to target for intervention, and lead to more effective policies on mental health and treatments within schools. In addition, unhealthy non-cognitive behaviors, habits, and thoughts can be directly addressed through interventions and new school programs, which would be far harder to implement if the study's focus was cognitive risk factors (Costello et al., 2008; Yaroslavsky et al., 2013). School-based mental health services offer provide regular access to all students. For certain students, especially those who are under-resourced or lack family support, school is the only place where adolescents may have access to mental health care, which makes school a distinct and special provider for mental health services (Levitt et al., 2007).

Another study aim was to test whether adverse adolescent psychosocial behavior (e.g., adolescent delinquency) was associated with latent class trajectories, and if so, what extent was its risk. The dual failure model maintains that conduct problems are risk factors for mood disorders, including depression (Patterson, DeBaryshe & Ramsey, 2017). This theory has been backed by empirical evidence of the positive association

between externalizing behaviors, like delinquency, and internalizing behaviors, like depression (Davies et al., 2019; Willner, Gatzke-Kopp & Bray, 2016). Engaging in delinquency has been associated with elevated depressive symptoms in adolescents for both males and females at roughly the same magnitude (Costello et al., 2008; Kandel & Davies, 1982; Stoolmiller, Kim & Capaldi, 2005).

Finally, the study intended to address a gap in the literature concerning the extent to which latent class trajectories forecast labor market outcomes in American young adults. An enhancement of such knowledge may better demonstrate tangible consequences, if any, of long-term, non-normative patterns of change in depressive mood using specific indicators of employment and financial well-being. Few studies that have studied depressive symptom trajectories and socioeconomic outcomes have found that adults with persistent low depressive mood report higher salary and life satisfaction but were limited by their study's scope and/or sample (Salmela-Aro, Aunola & Nurmi, 2008). For example, Holsen & Birkeland (2017) used latent growth mixture modeling on a sample of Norwegian adolescents measured from age 13 to 30 and found that a trajectory with an increased level of depressive symptoms was related to lower income, poor employment status, and fewer meeting adulthood milestones in age 30.

Primary Research Questions

From the reviewed literature, this study was devised to address three research questions using latent growth mixture modeling:

- 1) Are there multiple subpopulations (latent classes) of developmental change trajectories for depression from adolescence to early adulthood?
- 2) Do pre-existing demographic factors and non-cognitive behaviors and perceptions relate to membership in those subpopulations?
- 3) Does class membership in those subpopulations predict relevant distal outcomes, including labor market participation, financial well-being, and job satisfaction?

CHAPTER 2: METHODS

Data

Sample

The data for this study were derived from the public-use observations from the National Longitudinal Study of Adolescent to Adult Health (Add Health), a nationally representative, probability-based survey of adolescents between Grades 7 to 12 (Harris & Udry, 1995). Add Health study's primary purpose was to assess a variety of health-related attitudes and behaviors of American adolescents, along with changes in health measures. The same cohort of individuals have been followed up with and measured since 1994 (Harris et al., 2009). Add Health's public data is a representative sample of the restricted use, full data set. The primary differences between the public use and restricted use data sets were considered insignificant for the purposes of this analysis, as the restricted-use data set contained identifiable and confidential information about individuals' physical health, including genetic information and numerous medical test results, which were not relevant for this study.

Add Health's school sampling frame was generated from the Quality Education Database (Harris, 2013). Participants were selected from a multistage, stratified, school-based, cluster sample of secondary schools (Harris et al., 2009; Zhu, 2018). First, high schools (defined as schools with an 11th grade and at least 30 students) from 80 school communities were sampled based on several strata: enrollment size, region, urbanicity, school type (public, private, parochial), and ethnic composition (Costello et al., 2008; Harris, 2013). Next, when appropriate, a feeder middle school was identified and

contacted for the study to select eligible 7th and 8th grade students. Some schools did not have a middle school feeder pair because the selected school served Grades 7 through 12 (Harris, 2013). In total, 79% of the 80 school communities responded and agreed to participate. The sample constitutes student data from 132 different schools from 80 nationally representative school communities.

Among the 132 schools, over 90,000 students completed an in-school survey (~45-60 minutes), with a response rate of 98.5% (Harris et al., 2009). More importantly, for the context of this study, an in-home interview was later conducted on a sample drawn from a sampling frame made from the union of all enrolled students from each school roster and students who completed the in-school questionnaire (Harris, 2013). The students selected for the in-home interview formed the ‘core’ cohort for Wave I, who would be followed up during future measurement occasions. When constructing the core sample, researchers stratified students in each school community by grade and sex, randomly selecting approximately 17 students from each stratum (Harris, 2013). Therefore, about 200 adolescents from Grades 7 to 12 were selected from each of the 80 school communities. The core sample consisted of 20,745 adolescents in Wave I, with a response rate of about 79%. Four waves of longitudinal data were publicly available, and spanned approximately 14 years, with two measurements when the sample were adolescents and another two when they were considered young, but legal adults. Although sampling weights were provided by Harris (2013), they lacked relevant context for this study and were not used.

As stated previously, this study used the observations from the public use Add Health data set, which was reported by Harris et al. (1990) as a representative sample of

the full, restricted use data set. At baseline (1994-1995) the Wave I cohort contained data from 6504 middle and high school students between 7th and 12th grade. The following year (1996), the same core participants were followed up in Wave II, with the exception of all Wave I seniors ($N = 993$), because they were no longer considered adolescents. Relative to the public-use cohort from Wave I, which included seniors, the response rate for Wave II was 75.4% ($N = 4834$). All Wave I participants were followed up for Wave III (2001-2002), with a response rate of 76.2% ($N = 4882$). The final wave of public data, Wave IV (2008), followed up on all Wave I participants and had a response rate of 79.8% ($N = 5114$). More than half of all participants completed all four waves of data, with more than three-quarters completing three or more waves (Harris et al., 2009).

This study put limited, yet necessary inclusion criteria, when selecting eligible participants. Study participants included all Wave I adolescents who were enrolled in school within the past year, had completed at least one item used to measure depressive symptoms during Wave I, and had completed at least one antecedent variable. After exploring the original data from the 6504 students, 138 adolescents were excluded from the study, because they either had full-missingness on all antecedents ($N = 125$) or all depression-related items ($N = 13$) during Wave I. Thus, the final sample size was made up of 6366 adolescents. Sample demographic characteristics are shown in Table 1.

Since the United States consists of many people with diverse identities and ethnic/racial backgrounds, the six race/ethnicity indicators were not forced to be mutually exclusive. About 84.3% of adolescents ($N = 5369$) identified with one single race/ethnicity, with the remaining having more than one identifier selected.

Table 1. *Sample Demographic and Student Characteristics*

Descriptive Statistics	<i>n</i>	Percent ^a
Student Sex		
Male	3090	48.5
Female	3276	51.4
Student race/ethnicity ^b		
White	4211	66.1
Hispanic/Latino	717	10.8
Black	1587	24.9
Native American	228	3.6
Asian/Pacific Islander	267	4.2
Other	411	6.5
Grade at Wave 1		
7 th grade	979	15.4
8 th grade	990	15.6
9 th grade	1128	17.7
10 th grade	1154	18.1
11 th grade	1122	17.6
12 th grade	993	15.6

^aOut of *N* = 6366.^bStudents allowed to identify with more than one ethnic/racial indicator.

Measures

Depressive symptoms. Depressive symptoms were assessed using items from the Center for Epidemiological Studies-Depression Scale (CES-D; Radloff, 1977). The original scale consisted of twenty items that measured the incidence and frequency of specific symptoms associated with depression. The CES-D was initially developed to screen for depressive symptoms in the general population of adults (Radloff, 1977)

depressive symptoms among adults. However, it has since been validated in several studies for use with adolescents (Garrison et al., 1991; Stoolmiller, Kim & Capaldi, 2005). The CES-D is one of the most popular screening scales for assessing depression. However, its only appropriate use is for screening, not diagnosing depression (Radloff, 1977; Vilagut, Forero, Barbaglia & Alonso, 2016).

All CES-D items assessed the frequency of respondents' depressive symptoms over the past week and were scored on the same four-point Likert scale: zero – rarely or none of the time; one – some or a little of the time; two – occasionally or a moderate amount of the time (three to four days); three – most or all of the time (five to seven days). Therefore, on the nine-item short-version scale, raw sum scores could range from zero to twenty-seven (Radloff, 1977). Two items were re-valenced to reflect the uniform directionality of the construct, where higher scores reflected stronger depressive symptoms. Prior studies with factor analytic techniques have found that the full twenty-item CES-D was composed of four subfactors—somatic-retarded activity (being bothered by things, unable to shake the blues, having trouble keeping focused, and feeling too tired to do things), depressed affect (feeling depressed, feeling sad, and crying frequently), positive affect (feeling that you're as good as others, and enjoying life), and interpersonal relationships (feeling that people dislike you) (Meadows, Brown & Elder, 2006).

Since the nine-item version of the CES-D had little to no information about its psychometric properties and dimensional stability/structure using public Add Health data, an analysis of its dimensionality and scale invariance was required to address the three research questions in this study. Add Health used a common group of nine CES-D items across all four waves. Shortened versions of the CES-D have been used and validated in

past studies (e.g., Zhang et al., 2012). In this study, however, there was not sufficient evidence to warrant haphazard use of the short-form of the CES-D instrument, so a thorough and independent examination of the nine-item's qualitative and quantitative structure both within and across timepoints was conducted. Rather than reporting sum scores, scaled scores were produced using vertical equating with IRT-based calibration. None of the prior Add Health studies followed this approach and merely reported internal consistency values for each wave. For example, Meadows, Brown & Elder (2006) used the three available waves of the Add Health's CES-D data (restricted use) and reported that the nine-item scale was consistently reliable ($\alpha = .80-.82$; Meadows, Brown & Elder, 2006) and used the raw sum score at each wave when examining latent trajectories via growth mixture modeling.

Antecedent variables. All antecedent variables were measured during the Wave I in-home interview. Antecedent variables were used to identify adolescent characteristics that were more (i.e., a risk factor) or less (i.e., a protective factor) associated with having an atypical depressive symptom trajectory during the period from adolescence to young adulthood. Most antecedent variables, like biological sex, were easily constructed via single-item indicators, but two latent variables, school disengagement and delinquency, required a more comprehensive statistical analysis prior to producing latent factor scores. Both latent constructs were confirmed to each hold a unidimensional factor structure. Typically, exploratory and confirmatory factor analyses would be properly conducted. However, each latent construct was measured non-consecutively with large measurement

lapses between their positions within clearly defined and labeled testlets, and all other antecedent and depressive symptom variables used in this study.

Given the structure and presentation of this instrument, confirmatory factor analysis was run to confirm unidimensionality within the two separate testlets. Neither face validity nor item location within a testlet provided sufficient statistical evidence to create unidimensional factor scores straightaway. Therefore, it was essential to assess the reliability and unidimensionality for each antecedent latent construct: school disengagement and delinquency.

Demographic variables. All demographic variables were assessed using dichotomous indicators for each demographic category. Two demographic characteristics—biological sex and race/ethnicity—were tested using seven different dichotomous indicators, with each indicator set to one if an adolescent self-identified as a specific racial/ethnic category and zero if they did not. The six race/ethnicity categories were: White, Hispanic/Latino, Black, Native American, Asian/Pacific Islander, and Other. These measures relied on students' self-reports from Wave I only.

School disengagement. School disengagement was assessed with thirteen items from a distinct testlet labeled “Academics/Education”. The proposed items measuring this construct are shown in Table 2.

Table 2. *School Disengagement Items*

Item	Scale
Skipped school more than twice	Dichotomous (1 = yes)
Has received out of school suspension	Dichotomous (1 = yes)
Has received expulsion	Dichotomous (1 = yes)
Trouble with teachers (freq. per week)	Ordinal (0 – 4) (Never – Every Day)
Trouble paying attention (freq. per week)	Ordinal (0 – 4) (Never – Every Day)
Trouble finishing homework (freq. per week)	Ordinal (0 – 4) (Never – Every Day)
Trouble with other students (freq. per week)	Ordinal (0 – 4) (Never – Every Day)
Feel close to people at school	Ordinal (1 – 5) (Strongly Agree – Strongly Disagree)
Feel part of school	Ordinal (1 – 5) (Strongly Agree – Strongly Disagree)
Students at school are prejudiced	Ordinal (1 – 5) (Strongly Disagree – Strongly Agree)
Happy at school	Ordinal (1 – 5) (Strongly Agree – Strongly Disagree)
Teachers treat students fairly	Ordinal (1 – 5) (Strongly Agree – Strongly Disagree)
Feel safe at school	Ordinal (1 – 5) (Strongly Agree – Strongly Disagree)

Data from this section included self-reported information about an adolescent's conduct and trouble in school, and perceived engagement with peers, teachers, and school overall. These items were kept together because they were in non-contiguous and separately presented from other items used in this analysis.

Therefore, due to the presentation and isolation of these items within their own labeled testlet related to academics and education, a confirmatory factor analysis was used to test whether these thirteen items plausibly reflected a single construct: school disengagement.

Based on face validity, these items revealed adolescents' perceived behaviors and thoughts about themselves in school but did not directly assess objective measures of school closeness, such as engagement with club activities and/or membership on a sport team. Items ranged in scale, with some items scored dichotomously, others scored on a four- or five-point Likert scale, and one item on a continuum. The continuous item, frequency of skipping school without an excused absence, ranged from 0 to 99, and was rescaled as a dichotomous indicator and was set to 1 if the student skipped more than one time and 0 otherwise. This rescaling was needed because the original scale had an unacceptably high skewness and kurtosis. After rescaling this variable, the skewness and kurtosis changed from 8.11 to 1.94 and 83.96 to 1.75, respectively.

All items were properly valanced such that higher scores reflected higher disengagement from school. Due to the ordinal nature of the thirteen items' scales, a smoothed, polychoric correlation matrix was used (instead of a traditional Pearson correlation matrix) for the confirmatory analyses. Polychoric correlation matrices treat item scales as ordinal, rather than continuous, which was appropriate for this scale because some items had properties that would have grossly violated assumptions of normality. In addition, since items were scored

using different scales, latent factor scoring coefficients were used to generate factor scores.

Retained in grade. The only item from the Academics/Education testlet which was separately evaluated was whether the student had ever repeated a grade. This decision to evaluate this item separately was initially based on the item's face validity and was further confirmed after noting its observed relationships with the other thirteen school disengagement items. Based on face validity alone, it is unknown whether cognitive and/or non-cognitive factors were used when deciding to hold a student back a grade at least once prior to Wave I. Therefore, being retained in a grade may reflect academic competencies which were not of direct interest in this study and were thought to be poorly associated with the remaining disengagement items. Furthermore, initial analyses identified its item-total (remaining) correlation was near-zero, indicating that it poorly correlated with the total remaining score for the thirteen items. This item was still preserved because it still may reflect non-cognitive school factors, but was kept alone as its own dichotomous indicator, where 1 was assigned if the student ever repeated a grade, and 0 otherwise.

Delinquency. Thirteen delinquency items, like school disengagement, were obtained from their own distinct testlet labeled 'Delinquency' (Harris & Udry, 1995), which was presented distinctly during the in-home interview as items measuring delinquent behaviors. The items measuring delinquency were related to property, violent and disrespectful behaviors. Some behaviors were considered legal, while many others were not. However, unlike school disengagement, all

delinquency items were scored on the same four-point Likert scale which all started with the stem, “In the past 12 months, how often...?”, followed by scale options that ranged from: “Never (0)”, “1 or 2 times (1)”, “3 or 4 times (2)”, and “5 or more times (3)”. Each item is listed in Table 3. As with school disengagement, since these items were presented in a non-consecutive distinct testlet, these thirteen items were initially tested to confirm it held a unidimensional structure (i.e., overall delinquency) by using a smoothed polychoric correlation matrix to adjust for the categorical nature of the items’ scale.

Table 3. *Delinquency Items*

Item Description
(Stem: Over past 12 months, how often did you...)
Paint graffiti on someone else's property or in public
Deliberately damage property that didn't belong to you
Drive car without owner's permission
Take something from store without paying for it
Steal something worth more than \$50
Go into house or building to steal something
Get into serious fight
Hurt someone badly enough to need bandages or care from doctor or nurse
Use or threaten to use a weapon to get something from someone
Take part in fight where a group of your friends was against another group
Lie to parents/guardian about where you have been or whom you were with
Run away from home
Were loud, rowdy, or unruly in public place

Distal outcomes. All distal outcomes were scaled to represent dichotomous labor markers, with all measured during Wave IV, where study participants were aged 24 to 32 years old (Harris & Udry, 2018). These distal outcomes were used to contextualize and provide external validity to depressive symptom latent trajectory classes with respect to their socioeconomic outcomes. The eleven indicators were associated with an individual's financial well-being and employment status prior to and during 2008.

Financial health/security. Financial security was assessed using several items that measured aspects of monetary well-being over the past year. Three dichotomous indicators were used to separately indicate if the respondent worried about not having enough money to pay their rent/mortgage, covering utilities, or, restocking food/groceries. Individuals were also asked whether they had ever received a form of federal/state assistance based on financial need (i.e., food stamps or welfare payments) between Waves II and IV. Monetary health also was explored by asking individuals what their net status would be if their assets were liquefied, with 1 indicating they would still be in debt and 0 if they were even or had more assets than debt. Finally, the median household income in 2008 was used to create a dichotomous indicator to assess a respondent's current household income, where 1 indicated that their household income was below the median, or 50th percentile, of US household incomes in 2008, and 0 if their household made an income at or above the US median household income.

Employment history and status. Three employment indicators were used to determine the benefits offered by the respondent's current employer. Namely, separate indicators determined whether their current employer offered work-provided healthcare, some form of retirement benefits, or offered employees paid vacation/sick leave. In addition, one variable measuring job satisfaction was dichotomized where 1 indicated that the respondent's current job satisfaction was either "Not Satisfied" or "Extremely Not Satisfied", and set to 0 for the other three responses categories (e.g., "Neither Satisfied nor Dissatisfied", "Satisfied", and "Extremely Satisfied"). Finally, an additional dichotomous indicator was

used to determine whether the individual had ever worked in a supervisory role, where 1 indicated that they ever had, and 0 otherwise.

Longitudinal Missing Data

It is not uncommon to encounter missing data in most research studies, especially for longitudinal research, where the tracking and compliance of participants over multiple years and locations can become difficult to maintain. Missing data occurred through one of two processes: (1) attrition (respondent missed an entire occasion measuring the outcome of interest), or (2) incomplete/skipped items within the outcome for a given measurement occasion. In this study, both types of missing data were identified and required different analytical methods to deal with them.

Case-Level Missingness

An individual was classified as case-level missing if they had no depressive symptom items completed in a given wave, and so, case-level missingness was computed separately for each wave. All study participants were case-level present during Wave I, because part of the inclusion criteria included a requirement that at least one CES-D item had to be completed in Wave I. As long as one out of nine CES-D items was complete, the individual would not be considered case-level missing for a wave (i.e., they would be considered case-level *present* during that wave), with case-level present tallies shown by wave in Table 4.

After Wave I, there was a fair drop in the number of participants who were present for Wave II, dropping from 6366 to 4745 adolescents. The systemic loss of Wave I seniors ($N = 993$) during Wave II did not fully account for this drop, which was

an attrition of over 1600 Wave I participants. After Wave II, the number of participants marginally increased during Wave III ($N = 4789$), which was the first wave where youngest participant was at least 18 years old, making Wave III the first wave during young adulthood and the first wave where all Wave I participants were followed up.

During Waves II, III, and IV, the proportion of case-level missingness settled around 25%. The highest attrition occurred during Wave II. This was thought to be moderately affected from losing the Wave I seniors ($N = 993$) who had graduated, which likely inflated the proportion of attrition in depressive symptom data. Excluding Wave I, the lowest case-level missingness occurred during Wave IV, which was somewhat interesting, but could be due to either random fluctuations in response patterns or helped from study participants becoming more settled and easier to follow-up with and track down. The time intervals between waves were not equal. Compared to the one-year difference between Wave I and II, there was a five-year difference between Wave II and III, and a seven-year difference between Wave III and IV. Time can make longitudinal studies somewhat difficult, as people relocate, lose interest in participation, change contact information, etc.

Table 4. *Non-Missing Cases of Depression Items by Wave*

Wave (Year)	<i>n</i>	Percent ^a
Wave I (1994-1995)	6366	100
Wave II (1996)	4745	74.5
Wave III (2001-2002)	4756	74.7
Wave IV (2008)	5018	78.8

^aOut of *N* = 6366.

More than half of the sample had zero case-level missingness at every wave (*N* = 3285, 51.6%). In addition, another 30% of participants had one full wave of depression data missing (*N* = 1911, 30.0%), with another 10% having two waves fully missing (*N* = 875, 13.7%). Slightly less than 5% of the original cohort case-level present at Wave I missed all depression items for all three remaining waves (*N* = 295, 4.6%).

Item-Level Missingness – Depression

The proportion (out of the full cohort) with full item completion of the nine depression items ranged from 74.2% (Wave II) to 99.7% (Wave I). As stated previously, Wave I seniors were systematically absent from Wave II. If Wave II's item-level completion rate was calculated out of *eligible* adolescents who were followed up during Wave II, the completion rate of item-level depression data would jump up to 88.0%.

In addition to calculating the full item-level completion rates out of the original Wave I cohort, the rates were also computed separately by wave, out of those who were present for at least one item in each given wave (i.e., they were case-level present at that

wave). Therefore, these rates measured the proportion of those who completed all nine items, given that they already completed at least one. The proportion of full depression data ranged from 99.6% to 100%. Therefore, almost all non-missing cases for a given wave had little to no missing items.

Table 5. *Complete Observations by Wave for Depression Items*

Wave (Year)	<i>n</i>	Percent ^a	Percent of Non-Missing ^b
Wave I (1994-1995)	6347	99.7	99.7
Wave II (1996)	4726	74.2	99.6
Wave III (2001-2002)	4789	75.3	100
Wave IV (2008)	5010	78.7	99.8

^aOut of *N* = 6366.

^bOut of number of non-missing cases in Table 4.

Item-Level Missingness – Antecedent Variables

All antecedent variables were measured at Wave I, so they had no patterns of case-level missingness (antecedents were measured during one occasion). The rates of non-missing ancillary items were fairly high, with data recorded at a rate approximately 99% or above for all antecedent items. There was some variation in the rate of item-level missingness among antecedent variables, but these rates of item-level missingness were at or below 1% for all ancillary items. The highest missingness rates across all antecedent items were for those measuring delinquency, ranging from 50 (“How often did you run away from home?”) to 62 missing item responses (“How often did you lie to

your parents or guardians about where you have been or whom you were with?”, “How often did you hurt someone badly enough to need bandages or care from a doctor or nurse?”).

Imputation Procedure

Both case- and item-level missingness were observed in this study’s longitudinal data set. Listwise deletion should be avoided in nearly all studies with missing data (Allison, 2009). Here, the loss of information from a considerably smaller sample size (6366 vs. 2057) was considered more impactful than the benefit(s) that listwise deletion could provide. Implementing appropriate imputation procedures required separate attention and the appropriate time to handle case- and item-level missingness. Item-level missingness was handled prior to running any confirmatory analyses of factor dimensionality or testing of measurement invariance.

All item-level missingness was handled using multiple imputation (MI) under the Markov Chain Monte Carlo (MCMC) procedure (Rubin, 1987; Schafer, 1999). Under MI, missing values were assumed to be missing at random, and is outlined in greater detail in Rubin (1987) and Schafer (1999). Classical applications of MI assume that it operates missing data with an underlying multivariate, normal distribution (Rubin, 1987). However, recent studies have demonstrated that this assumption can be somewhat relaxed under certain conditions. For example, Lee and Carlin (2010) found that under fully conditional specification methods, MI generally yield similar estimates when using either continuous or ordinal/categorical data. Furthermore, Leite & Beretvas (2010) found that using MI procedures on data with five or more ordinal categories produced acceptable

estimates with up to 30% of missing data, which was far greater than the rate of item-level missingness in this study (no greater than 1%). However, the validity of using normally based MI and naively rounding imputed non-integer values with ordinal variables is still being considered and debated among researchers (i.e., Allison, 2005; Lee, Galati, Simpson & Carlin, 2012; Rhemtulla, Brousseau-Liard & Savalei, 2012; Xia & Yang, 2016), but the current consensus is that appropriate use of MI on ordinal data depends on model characteristics and imputation's purpose (Chao, 2017). Model characteristics include sample size, number of categories, symmetry of item distributions, proportion of missingness, and purpose/consequences of imputation (Allison, 2005; Chao, 2017).

The alternative methods to MCMC MI, such as MI without rounding, two-stage calibration, and adaptive rounding were either too cumbersome or not suitable for this study, which required items to be analyzed at the ordinal level (Allison, 2005; Chao, 2017; Lee et al., 2012). Other strategies designed for polytomous data, such as multinomial logistic regression and proportional odds methods performed more poorly when naively rounding in ordinal/categorical models than in normal, continuous models (Chao, 2017; Wu, Jia & Enders, 2015).

Despite its shortcomings, MCMC MI with naïve rounding was selected to best handle the item-level missing data within each wave of depression and among antecedent items. Antecedent variables were also imputed in this manner, because group-based trajectory methods, like growth mixture modeling, cannot accommodate observations with missing antecedents (Costello et al., 2008). Selecting this imputation strategy was primarily based on the low rates of item-level missingness among ancillary variables in

Wave I and among depression items within each timepoint. MI with naïve rounding was also selected based on inappropriate and/or cumbersome nature of alternative imputation approaches.

The purpose of this study was to identify and assess latent class trajectories, so the missing data at the case/wave-level was maintained. Therefore, MI was only used for individuals present at the case level (i.e., those who had at least one depression item completed). To circumvent autocorrelation of items across measurement occasions, MI was applied on each wave separately. As suggested by Bodner (2008), each dataset had been imputed using fifty rounds with SAS's PROC MI. To maintain the original ordinal scale of each CES-D item, each item's fifty imputations were first averaged, with this pooled value rounded to the nearest integer and whenever appropriate, knocked within the scale's original bounded values of zero and three (Chao, 2017). The effects on all correlation matrices were considered adequate, with all items' efficiency values greater than .99.

Dimensionality

The unidimensionality of depression, school disengagement, and delinquency were each analyzed using *R*'s lavaan (Rosseel, 2012) and psych (Revelle, 2020) packages. Since the two antecedent constructs were measured non-consecutively in their own distinct, labeled testlet, exploratory analyses were not appropriate. Since all item-level data was ordinal, confirmatory factor analyses were performed and adjusted using weighted least squares means and variance (WLSMV) estimation. Analyses with

WLSMV estimation used a robust, diagonally weighted, least squares method specifically devised for categorical data with no distributional assumptions (Brown & Little, 2015).

Acceptable model fit criteria were outlined using thresholds detailed in Marsh, Liem, Martin, Morin & Nagengast (2011), where acceptable model fit is determined with a Root Mean Squared Error of Approximation (RMSEA) $\leq .08$ and Comparative Fit Index (CFI) $\geq .90$. In addition to meeting the acceptable model fit criteria, unidimensionality for each factor structure was ensured, with all items holding salient loadings (loadings $\geq .35$) on their respective factors.

Longitudinal Measurement Invariance

Measurement invariance was necessary to allow the use of the nine-item scale to measure and validly interpret depressive symptoms scores across every timepoint. Nested levels of measurement invariance for the four unidimensional factor structures of depression were evaluated within an SEM framework. The first level, configural invariance, established that all items were related to the same construct (i.e., all factors were salient). Since this study involved measuring the nine-item scale over time, it was necessary to also validate that each item measured depressive symptoms in the same manner over time. Once the first level (i.e., configural invariance) of measurement invariance was met, the factor structures were tested at the second level (i.e., weak/factorial invariance) by placing a constraint that forced item-level factor loadings to be equal across time.

Models were estimated using robust maximum likelihood, which was considered to be an acceptable estimation procedure for non-normal data based on literature

published by Browne (1984), Li (2015), and Rhemtulla, Brosseau-Liard & Savalei (2012). In studies with a small sample size, measurement invariance may be formally tested using a scaled chi-square difference test (Satorra & Bentler, 2001). However, since this statistic has a strong dependence on sample size, it is almost always statistically significant when working with large samples, as in this study. Therefore, model fit was assessed primarily using alternative fit statistics, such as the RMSEA and CFI (Cheung & Rensvold, 2002; Marsh et al., 2011). Acceptable fit criteria for the SEM model with factorial/weak invariance had been outlined in the previous section (Marsh et al., 2011).

Equating and Scaling

Once weak invariance was confirmed across each wave of depression items, the individual wave scores were calibrated and vertically equated using an Item Response Theory (IRT) approach that placed longitudinal scores on a single scale using linking items between two adjacent timepoints (connecting Wave I to Wave II, Wave II to Wave III, etc.). All depressive symptom items from each factor were calibrated and equated using the IRT software *flexMIRT* (Cai, 2013). In longitudinal research, simulations have established that equating procedures via IRT provide better unbiased estimates relative to equating under methods derived from True Score Theory (Barr, 2018; Gortler, Fox & Twisk, 2015).

Since all CES-D items were polytomous, the graded response model was selected during the calibration and equating processes. First, linking items were identified through a multiple-group (wave) IRT analysis to reveal any CES-D items with differential item functioning (DIF) between waves. Each pair of consecutive waves were

separately analyzed to detect items that exhibited non-significant DIF and, thus, could be plausibly used as a linking item. Significant DIF was inferentially tested using each item's χ^2 test of the residuals, which was based on the expected similarities of item difficulty parameters between consecutive waves. Within each wave pair, any items displaying significant DIF were eliminated as possible linking items.

Two to three non-DIF items were identified and used as linking items between each wave gap, with no linking item was used more than once. Vertically equated scores were calculated using Bayesian Expected a Posteriori (EAP) estimation, where Wave I was used as the reference group/wave, with scaled scores located at $M = 50$ and $SD = 10$.

Latent Growth Mixture Models

Latent growth mixture modeling (Duncan, Duncan & Strycker, 2009; Ram & Grimm, 2009) was applied to identify unobserved sub-groups (latent classes) of longitudinal change of depressive symptoms. Using the standard analytic approach, observations were included in the growth mixture model when data for at least one measurement occasion was observed, which at minimum, was during Wave I (Muthén & Muthén, 2019; Kim, Thompson, Walsh & Schepp, 2015). Wave/case-level missing data was imputed and handled using full-information maximum-likelihood (FIML) estimation.

Models were estimated by applying both latent and fixed (linear and polynomial) basis estimation procedures across four waves. *Mplus* version 8.4 (Muthén & Muthén, 2019) was used to estimate all latent growth mixture models. A visual depiction of a latent growth mixture model is shown in Figure 1, with squares labeled representing

manifest (observed) variables and circles representing latent variables, which were estimated in *Mplus*.

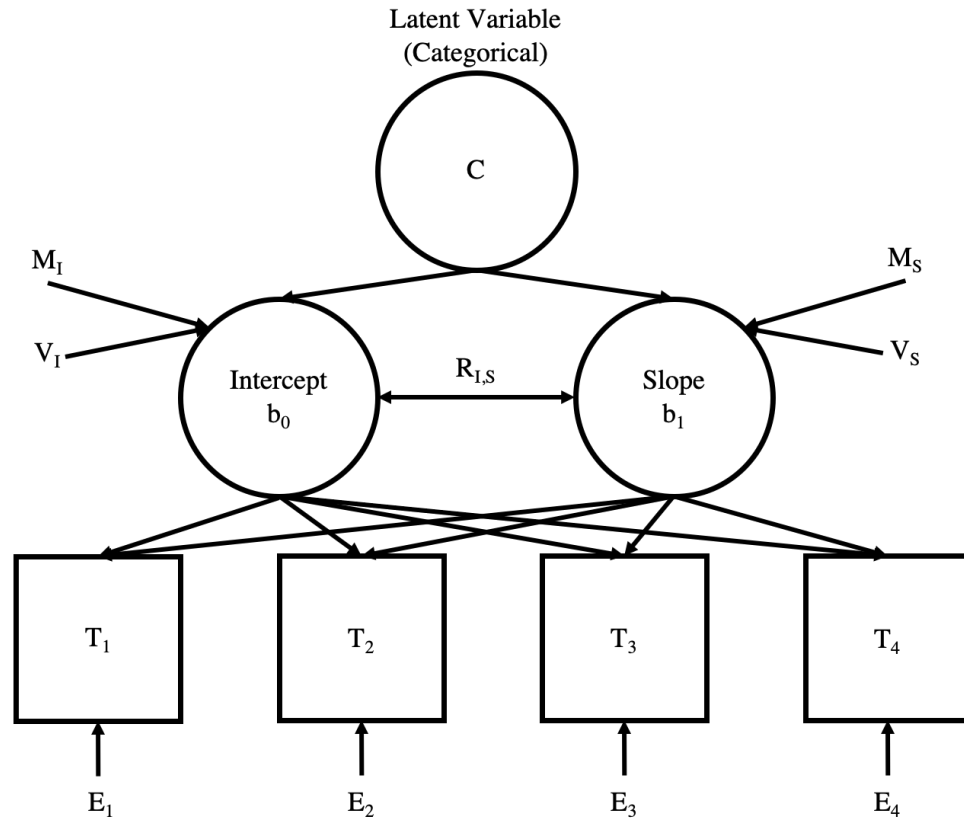


Figure 1. Visual depiction of a latent growth mixture model.

Relative to less complex models, ideal growth models will produce better fit statistics (including lower values for Akaike's Information Criterion [AIC], Adjusted BIC [ABIC], Schwarz's Bayesian Information Criterion [BIC], and minimal values via the Integrated Classification Likelihood with Bayesian Approximation [ICL-BIC; MacLachlan & Peel, 2000; Nylund, Asparouhov & Muthén, 2007]), relative to less complex models. Out of all alternative fit indices, the ICL-BIC statistic highly favored,

because this statistic tends to lead to a more parsimonious model and avoids overfitting, whereas other fit statistics like BIC tend to favor more complex models that may be prone to overfitting (Fruhwirth-Schnatter, 2006).

In addition, ideal growth models should produce higher (ideally, maximal) values for entropy and average posterior classification accuracy (Chao, 2017; Greenbaum et al., 2005; Nagin, 1999), and a significant improvement in model fit when compared to a model with one less latent class (as per the Vuong-Lo-Mendell-Rubin, Lo-Mendell-Rubin, and parameter bootstrap [using 100 draws] likelihood ratio tests) (Nylund, Asparouhov & Muthén, 2007). Optimal models should produce stable configural profiles across each class when testing imputation effects in a stability test. And lastly, superior models must produce classes with nontrivial membership size (>5%) and retain theoretical meaning, backed up by the literature (Ram & Grimm, 2009).

Risk Factors Models

All proposed antecedents reflected demographic and non-cognitive adolescent characteristics which were supported in the literature. Antecedent variables were used to quantify the relative probabilities of latent class membership in a non-normative vs. normative class for each significant risk/protective factor. All antecedents were placed on a binary scale, so re-coding was only necessary for the latent factor scores of school disengagement and delinquency. All individuals with scores in the upper quintile for each latent factor (i.e., students with the greatest behavior maladjustment) were coded as 1, with scores in the lower three quintiles coded as 0.

The initial model, with all antecedents included, used a multinomial logistic regression model and applied the general logit link function to assess possible significant risk/protective factors. Backwards elimination was used to trim and eliminate antecedents based on the variable with the highest non-significant p -value ($\alpha = .05$). A series of models were produced to identify and trim non-significant antecedents one at a time, which continued until only significant variables remained. The purpose of using a backwards elimination approach was to identify variables associated with a significant risk of being classified in a non-normative latent class, after controlling for the effects of all other antecedents in the model. Figure 2 depicts a representation of a structural equation model with antecedents, X , predicting latent class membership, C .

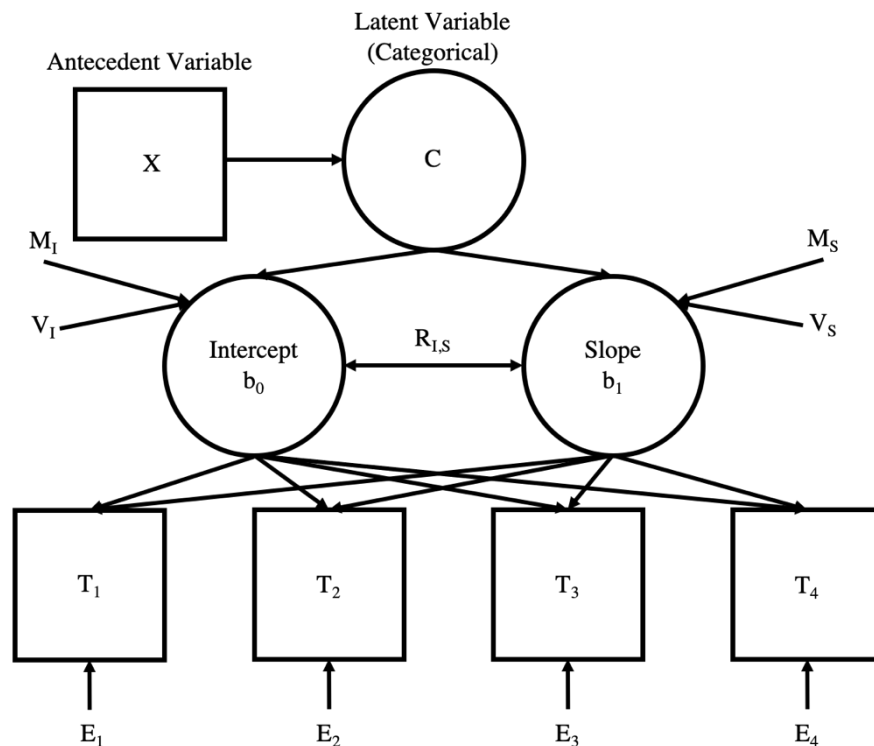


Figure 2. Depiction of a latent growth mixture model with antecedent covariates.

Distal Outcomes Models

Dichotomous labor market outcomes which assessed different attributes of financial security and employment status during young adulthood were regressed on the latent classification variables. Eleven socioeconomic outcomes were used to calculate the relative probability of a given outcome being true (vs. not) as a function of latent class trajectory membership. Each class probability of a favorable vs. non-favorable socioeconomic outcome was obtained via the *Mplus*'s DCAT function. Figure 3 shows a structural equation model where a categorical variable indicating latent class membership, C , is regressed on some distal outcome, U , which can be used to generate the probability of some outcome, given an individual's latent class membership, relative to if they were categorized in a different latent class.

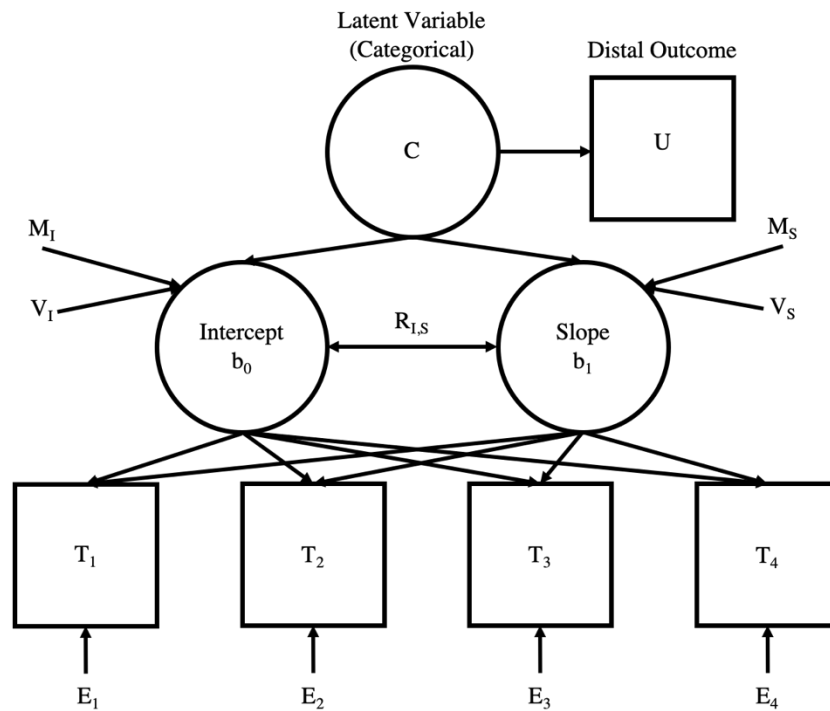


Figure 3. Depiction of a latent growth mixture model with distal outcomes.

CHAPTER 3: RESULTS

Antecedent Variables – School Disengagement and Delinquency

All items behaved in accordance with the other items on their respective scale.

School disengagement's item-total correlation adjusted for item overlap, ranged from .32 (ever received expulsion) to .64 (feel part of school). Delinquency's item-total correlation, adjusted for item overlap, ranged from .46 (frequency over past year of running away from home) to .64 (frequency over past year of taking something without paying for it). No antecedents had to be trimmed based on initial item analyses.

School disengagement's and delinquency's unidimensional factor structures were both confirmed using *R*'s SEM package, 'lavaan'. Both latent variables were modeled using a smoothed, polychoric correlation matrix and WLSMV estimator to account for the ordinal nature of both scales. The confirmatory model (configural) with the two factors demonstrated an acceptable fit based on recommendations in Marsh et al. (2011), based on relative and absolute fit indices (CFI = .924, TLI = .920, RMSEA = .061 (90% CI = [.060, .061]), SRMR = .089). The correlation between school disengagement and delinquency was estimated and found to be statistically significant, with $r_{school,delin.} = .51$ ($p < .001$). Scores on each factor were estimated using Empirical Bayes to obtain factor scoring coefficients. Once factor scores were produced, the highest quartile of each variable's scores was set to 1, with all other scores set to 0.

Depressive Symptoms

Initial item-level descriptive statistics like the means, standard deviations, and distribution statistics are shown in Table 6. Summary statistics are reported for each item separately by wave.

Table 6. *Item-Level Characteristics of Depressive Symptoms by Wave*

	Wave	Mean	SD	Skew	Kurtosis
1. Bothered by things that usually don't bother you	1	0.49	0.69	1.39	1.73
	2	0.54	0.70	1.26	1.35
	3	0.53	0.69	1.21	1.25
	4	0.53	0.71	1.37	1.76
2. Felt that you could not shake off the blues, even with help from family/friends	1	0.38	0.70	1.93	3.33
	2	0.40	0.70	1.85	3.03
	3	0.32	0.64	2.25	5.09
	4	0.32	0.71	2.24	5.05
3. Felt that you were as good as other people	1	1.07	1.00	0.52	-0.87
	2	1.02	0.99	0.61	-0.72
	3	0.68	0.92	2.25	0.13
	4	0.79	0.88	0.85	-0.23
4. Had trouble keeping mind on what you were doing	1	0.81	0.81	0.82	0.21
	2	0.82	0.79	0.77	0.17
	3	0.62	0.75	1.16	1.11
	4	0.82	0.79	0.84	0.41
5. Felt depressed	1	0.51	0.75	1.49	1.80
	2	0.50	0.73	1.52	2.00
	3	0.33	0.64	2.12	4.47
	4	0.38	0.67	1.95	3.78

	1	0.72	0.74	0.87	0.55
6. Felt too tired to do things	2	0.74	0.74	0.79	0.37
	3	0.63	0.73	1.08	0.99
	4	0.87	0.79	0.79	0.37
	1	0.76	0.86	0.86	-0.20
7. You enjoyed life	2	0.75	0.84	0.85	-0.16
	3	0.61	0.81	1.06	0.07
	4	0.68	0.79	0.81	-0.34
	1	0.56	0.68	1.15	1.43
8. You felt sad	2	0.55	0.67	1.15	1.45
	3	0.49	0.67	1.34	1.78
	4	0.56	0.66	1.07	1.26
	1	0.42	0.65	1.63	2.70
9. You felt that people disliked you	2	0.37	0.60	1.71	3.18
	3	0.26	0.56	2.40	6.29
	4	0.29	0.58	2.22	5.51

As reported in Kim (2013), values of skewness and kurtosis in large samples ($N > 300$) deviate from normality when they exceed 2 and 7, respectively (Hoyle, 2000; Kim, 2013). Since the CES-D items were designed to measure atypical behaviors (i.e., symptoms of depression), a positive skew was expected, since the presence of symptoms should be relatively uncommon in the target population of a community sample. Four items exceed a skewness of 2: CES-D 2 (Waves III and IV), CES-D 3 (Wave III only), CES-D 5 (Wave III only) and CES-D 9 (Waves III and IV). There were no items that had leptokurtic distributions (kurtosis > 7), so no item had markedly constrained

variance. All item-total correlations were within appropriate range. Based on these descriptive statistics, no items needed to be dropped from the analysis.

Dimensionality

Using *R*'s lavaan package, confirmatory analysis using an WLSMV estimator verified that each wave's depressive symptoms were best modeled using one dimension, with all items maintaining simple structure at each timepoint. Higher-order factor structures were tested as a counterfactual to proposed unidimensional model but were not reliable nor held simple structure (Tucker, 1955). Table 7 includes the factor structure of depressive symptoms by wave. Initial estimates of internal consistency (Cronbach's α) ranged from .80 (Wave I) to .82 (Wave III and IV). Due to the asymmetry of the CES-D item distributions, estimates of internal consistency were not fully appropriate (Yuan, Guarnaccia & Hayslip, 2003). However, an alternative reliability index, the composite reliability, will be calculated further on after calculating EAP scores and their standard errors derived after calibrating and equating via IRT (du Toit, 2003).

Table 7. *Initial Factor Structure for CES-D (9-item) Scale*

Item Description	Wave I	Wave II	Wave III	Wave IV
Bothered by things that usually don't bother you	.65 (.42)	.65 (.42)	.65 (.42)	.65 (.42)
Felt that you could not shake off the blues, even with help from family/friends	.78 (.61)	.82 (.67)	.82 (.67)	.83 (.72)
Felt that you were as good as other people	.44 (.19)	.46 (.21)	.56 (.31)	.56 (.31)
Had trouble keeping mind on what you were doing	.55 (.31)	.55 (.30)	.56 (.31)	.54 (.29)
Felt depressed	.86 (.74)	.88 (.77)	.92 (.84)	.90 (.81)
Felt too tired to do things	.52 (.25)	.56 (.31)	.50 (.25)	.49 (.24)
You enjoyed life	.56 (.31)	.61 (.37)	.67 (.45)	.67 (.45)
You felt sad	.78 (.60)	.83 (.69)	.82 (.67)	.81 (.67)
You felt that people disliked you	.61 (.37)	.60 (.36)	.55 (.30)	.54 (.29)

Note. $N = 6366$.

¹Final communality estimates are italicized and indicated in parenthesis underneath final factor loadings.

Longitudinal Measurement Invariance

Using the *R* package ‘lavaan’, all CES-D items loaded saliently on a single factor at each wave and demonstrated an acceptable fit under guidelines outlined in Marsh et al. (2011), with $CFI \geq .90$ and $RMSEA \leq .08$. Anderson & Gerbing’s (1988) indicated that when using large samples, it is likely the chi-square statistic will be significant. As anticipated, given the study’s sample size, all scaled chi-square difference tests were statistically significant, so model fit was evaluated using relative and absolute fit indices (Cheung & Rensvold, 2002; Marsh et al., 2011).

The model with weak/factorial invariance was found to have an acceptable fit ($\chi^2_{scaled}(562) = 7259.7$, $CFI = .943$, $TLI = .936$, $RMSEA = .07$ [90% CI = .069-071], $SRMR = .07$). Since the data were treated as ordinal using a WLSMV estimator, the Satorra-Bentler (2001) scaled chi-square statistic was reported instead of the normal chi-square statistic. Based on suggested fit statistics reported by Marsh et al. (2011), all fit statistics fell within the range of an appropriately fitting model. Therefore, there was sufficient evidence to support using the 9-item version of the CES-D scale to measure depressive symptoms across all four waves.

Equating and Scaling

The item scales from each wave needed to be linked together to create one scale that spanned across all four waves. First, DIF analyses were performed on each of the three consecutive wave pairs (Wave I and II, II and III, and III and IV) to identify potential linking items. Linking items (i.e., non-DIF) were identified using a graded response model with difficulty parameters constrained equal between consecutive waves.

Two items were used to link scores between waves: Items 2 and 7 between Wave I and II, Items 5 and 6 between Waves II and III, and Items 8 and 9 between Waves III and IV. Linking items were then used to vertically equate scores across the four waves, with Wave I used as the reference group/occasion.

The EAP scores were centered at $M = 50$ and $SD = 10$ at Wave I. Scale location was selected based on simply easing the interpretation for reading and reporting. Composite reliabilities were calculated based on the IRT-derived EAP scores and their standard errors (du Toit, 2003), and were acceptable at each wave: .75 (Wave I), .77 (Wave II), .74 (Wave III) and .76 (Wave IV). The resulting T -scores for each depressive symptom wave were used to generate latent growth models.

Latent Growth Mixture Models

Neither estimation method produced uniformly better fitting models, so models derived using both latent and fixed basis estimation were analyzed, compared, and reported in Tables 8 and 9, respectively. Both estimation methods were able to produce up to a 4-class model, but models higher in complexity did not reach convergence. All fixed basis models were best modeled using a quadratic polynomial where at least one latent class had significant quadratic curvature. No growth model was ideal based on all specified selection criteria.

Out of all candidate models, the 2-class, latent basis model was selected and determined to fit best, as it had the lowest ICL-BIC and retained class sizes with sufficient statistical power ($> 5\%$ membership). Model selection was not obvious, however, since the 2-class, latent basis model had neither the highest average posterior

probability (.747) nor entropy (.614). Compared to the selected 2-class latent basis model, the 2-class, fixed basis polynomial had the highest average posterior probability (.804) but had a lower entropy (.542), while the 4-class fixed basis (polynomial) model had the highest entropy (.662) but had a lower posterior probability (.700). No model's entropy was considered to be good. Aiming for parsimony, although the likelihood ratio tests (LRTs) suggested that higher-order models provided significant improvement in fit when compared to the model with one less class, there was no model greater in complexity (i.e., more latent classes) that had sufficient evidence to justify its selection.

All residuals by timepoint were statistically significant, which indicated there was a substantial degree of within-class variance among scaled scores. The variation within classes was also reflected in the relatively low entropy, which was a measure of how well the model was able to distinguish individuals into distinct classes. The posterior classification probabilities of Normative and Elevated latent classes were .967 and .524, respectively, which suggested that the growth mixture model's low entropy and poor average classification probability was mostly impaired by the low classification accuracy of the Elevated class.

Table 8. *Properties, Fit Statistics, and Parameter Estimates for Latent Basis, Latent Growth Mixture Models*

	1-Class model	2-Class model	3-Class model	4-Class model
Sample size				
Class 1, N_{C1}	6366.00	5232.28	4577.30	4119.00
Class 2, N_{C2}		1133.72	952.25	1021.60
Class 3, N_{C3}			836.45	340.92
Class 4, N_{C4}				884.48
Fit statistics				
# Free parameters	11	14	17	24
Akaike's Information Criterion (AIC)	148800	148560	148464	148214
Schwarz's Bayesian Information Criterion (BIC)	148874	148654	148580	148376
Sample size adjusted BIC (ABIC)	148839	148610	148526	148300
Integrated Classification Likelihood (ICL-BIC)		152061	154357	155136
Entropy		.614	.587	.617
Average class membership posterior probability		.747	.646	.614

Table 8 (continued)

Vuong-Lo-Mendell-Rubin LRT, p		<.0001	<.0001	<.001
Lo-Mendell-Rubin adjusted LRT, p		<.0001	<.0001	<.01
Parametric bootstrap LRT (1000 draws), p		<.0001	<.0001	<.0001
Latent variable means				
Class 1 intercept, γ_{01}	50.09 (0.11)	49.15 (0.25)	47.38 (0.39)	46.40 (0.30)
Class 1 slope, γ_{11}	-2.45 (0.12)	-3.49 (0.21)	-2.35 (0.30)	0.06 (0.28) [†]
Class 2 intercept, γ_{02}		54.48 (0.69)	54.58 (0.71)	60.00 (0.64)
Class 2 slope, γ_{12}		2.68 (0.73)	3.15 (0.72)	-11.68 (0.69)
Class 3 intercept, γ_{03}			60.17 (0.90)	45.28 (0.99)
Class 3 slope, γ_{13}			-9.52 (0.99)	9.80 (0.99)
Class 4 intercept, γ_{04}				57.07 (0.71)
Class 4 slope, γ_{14}				-3.67 (1.17)

Table 8 (continued)

Latent variable variances and covariance

Intercept, $\sigma_{\gamma_0}^2$	44.76 (1.48)	40.32 (2.17)	24.05 (2.65)	17.42 (1.18)
Slope, $\sigma_{\gamma_1}^2$	14.38 (2.07)	8.22 (1.77)	4.21 (1.42)	0.00 [fixed]
Intercept by slope, $\sigma_{\gamma_0}^2 \sigma_{\gamma_1}^2$	-12.96 (1.66)	-16.85 (1.92)	-7.83 (1.71)	0.00 [fixed]

Slope loadings, A_1

Wave I	0.00 [fixed]	0.00 [fixed]	0.00 [fixed]	0.00 [fixed]
Wave II	-0.02 (0.05) [†]	-0.01 (0.05) [†]	0.04 (0.05) [†]	-0.48 (0.36) [†]
Wave III	1.46 (0.09)	1.54 (0.11)	1.49 (0.10)	-0.45 (0.39) [†]
Wave IV	1.00 [fixed]	1.00 [fixed]	1.00 [fixed]	1.00 [fixed]

Residual variances

Wave I, σ_{e1}^2	36.14 (1.30)	32.89 (2.06)	37.70 (3.49)	30.40 (2.32)
Wave II, σ_{e2}^2	39.29 (1.39)	26.70 (2.25)	19.93 (4.41)	42.83 (2.45)
Wave III, σ_{e3}^2	53.81 (2.59)	38.79 (4.85)	39.28 (4.72)	38.91 (3.23)
Wave IV, σ_{e4}^2	52.18 (1.61)	36.00 (4.27)	36.28 (4.20)	53.56 (1.93)

Note. LRT = Likelihood Ratio Test. All parameter estimates are statistically significant unless indicated by the [†] symbol. Parenthetical values are estimated standard errors.

Table 9. *Properties, Fit Statistics, and Parameter Estimates for Fixed Basis, Latent Growth Mixture Models*

	1-Class model	2-Class model	3-Class model	4-Class model
Sample size				
Class 1, N_{C1}	6366.00	1727.87	4628.55	2323.84
Class 2, N_{C2}		4638.13	752.27	398.89
Class 3, N_{C3}			985.19	3291.61
Class 4, N_{C4}				351.67
Fit statistics				
# Free parameters	10	14	18	22
Akaike's Information Criterion (AIC)	149113	148934	148777	148679
Schwarz's Bayesian Information Criterion (BIC)	149200	149028	148898	148828
Sample size adjusted BIC (ABIC)	149159	148984	148841	148758
Integrated Classification Likelihood (ICL-BIC)		153070	154381	154794
Entropy		.542	.608	.662
Average class membership posterior probability	s	.804	.657	.700

Table 9 (continued)

Vuong-Lo-Mendell-Rubin LRT, p		<.0001	<.001	<.05
Lo-Mendell-Rubin adjusted LRT, p		<.0001	<.001	<.05
Parametric bootstrap LRT (1000 draws), p		<.0001	<.0001	<.0001
Latent variable means				
Class 1 intercept, γ_{01}	50.12 (0.12)	59.65 (0.61)	47.12 (0.56)	55.37 (0.42)
Class 1 slope, γ_{11}	-1.62 (0.15)	-6.39 (1.52)	-0.44 (0.29) [†]	-4.85 (0.43)
Class 1 quadratic slope, γ_{21}	0.30 (0.05)	1.30 (0.49)	-0.06 (0.08) [†]	0.82 (0.15)
Class 2 intercept, γ_{02}		46.52 (0.47)	53.85 (1.01)	43.71 (0.26)
Class 2 slope, γ_{12}		0.10 (0.37) [†]	2.04 (1.56) [†]	1.10 (0.36)
Class 2 quadratic slope, γ_{22}		-0.06 (0.13) [†]	-0.06 (0.49) [†]	-0.94 (0.76) [†]
Class 3 intercept, γ_{03}			61.50 (1.21)	43.71 (0.26)
Class 3 slope, γ_{13}			-10.38 (2.15)	1.10 (0.36)
Class 3 quadratic slope, γ_{23}			2.26 (0.67)	-0.17 (0.10) [†]
Class 4 intercept, γ_{04}				68.55 (0.69)
Class 4 slope, γ_{14}				-13.79 (1.89)
Class 4 quadratic slope, γ_{24}				2.93 (0.59)

Table 9 (continued)

Latent variable variances and covariance

Intercept, $\sigma_{\gamma_0}^2$	77.35 (5.00)	33.10 (6.23)	31.51 (4.01)	15.32 (1.75)
Slope, $\sigma_{\gamma_1}^2$	34.53 (6.27)	5.79 (0.98)	0.00 [fixed]	0.00 [fixed]
Quadratic slope, $\sigma_{\gamma_2}^2$	1.92 (0.49)	0.00 [fixed]	0.49 (0.07)	0.00 [fixed]
Intercept by linear slope, $\sigma_{\gamma_0}^2 \sigma_{\gamma_1}^2$	-36.80 (5.66)	-8.12 (2.22)	0.00 [fixed]	0.00 [fixed]
Intercept by quadratic slope, $\sigma_{\gamma_0}^2 \sigma_{\gamma_2}^2$	6.40 (1.39)	0.00 [fixed]	-2.63 (0.46)	-1.12 (0.25)
Linear slope by quadratic slope, $\sigma_{\gamma_1}^2 \sigma_{\gamma_2}^2$	-7.31 (1.60)	0.00 [fixed]	0.00 [fixed]	0.00 [fixed]

Residual variances

Wave I, $\sigma_{e_1}^2$	3.14 (4.82) [†]	23.67 (1.55)	20.88 (1.34)	13.44 (1.78)
Wave II, $\sigma_{e_2}^2$	51.08 (1.97)	47.46 (1.35)	47.80 (1.59)	51.25 (1.57)
Wave III, $\sigma_{e_3}^2$	65.49 (1.88)	70.35 (1.62)	67.73 (1.85)	69.19 (1.89)
Wave IV, $\sigma_{e_4}^2$	35.50 (5.27)	40.62 (1.94)	29.96 (3.89)	28.07 (3.36)

Note. LRT = Likelihood Ratio Test. All parameter estimates are statistically significant unless indicated by the [†] symbol. Parenthetical values are estimated standard errors.

The 2-class latent basis model's classes were named Normative (82.2%) and Elevated (17.8%) in their depressive symptom change trajectory from adolescence to young adulthood. The estimated mean class trajectories for both groups by wave are displayed in Figure 4. In this figure, the x -axis (wave number) was spaced to reflect the linear change in time between measurement occasions.

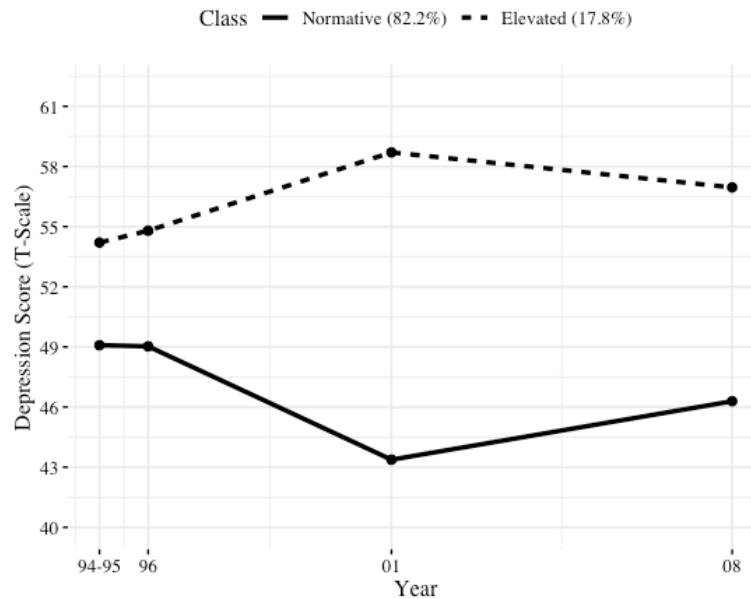


Figure 4. Estimated mean latent growth trajectories for depressive symptoms ($N = 6366$).

The mean trajectories appeared shaped like a tuning fork, with the Elevated class' mean scaled scores always higher than that of the Normative class. Mean depressive symptom trajectories start approximately $\frac{1}{2}$ SD from one another, with this difference statistically significant ($\alpha = .05$). Based on the non-significant slope loading parameter for Wave II, there were no discernable differences between Wave I and II's mean scaled scores for both classes. Between Wave II and III, both classes had a significant change in

their average depressive symptoms, with the Normative class having a decrease steeper than the Elevated class had increase. Wave III measured the greatest difference in mean symptoms between classes, at about 1 and $\frac{1}{2}$ *SD* difference. Between Wave III and IV, both waves settle somewhat towards each other, still with a mean difference of over 1 *SD* by Wave IV.

The stability of the 2-class, latent basis model was tested using a strict sub-sample of observations who had no item- or case-level missingness across all depression waves, all antecedents and distal outcomes ($N = 2057$). The mean change trajectories of the sub-sample are shown in Figure 5. The 2-class, latent basis model with non-missing data had a similar configural profile and distribution among classes relative to the model with fully imputed data ($N = 6366$). This evidence supports the use of imputation methods by demonstrating that the estimation of model parameters was not overly impacted by the imputation process.

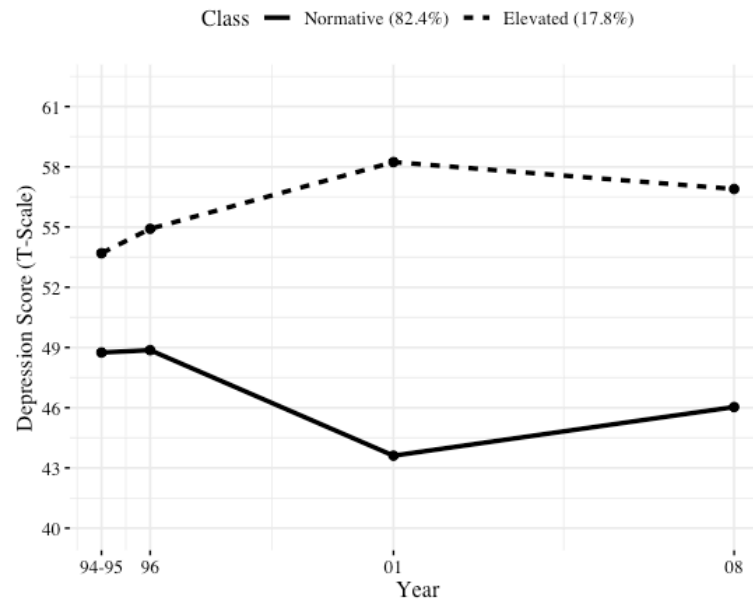


Figure 5. Estimated mean latent growth trajectories for non-missing data's depressive symptoms ($N = 2057$).

The distributions of depressive symptoms scaled scores by latent class and wave number were shown with boxplots in Figure 6. During adolescence in Waves I and II, the interquartile range (IQR) of both classes largely overlapped, which was no longer the case during measures taken in young adulthood. In addition, the difference between median scaled scores in Waves I and II were both much smaller than the differences in Waves III and IV. High levels of depressive mood were considered highly unusual (e.g., outliers) for the Normative class, but such outliers were observed during each timepoint, most notably in Waves I, II and IV. In the Elevated class, the prevalence of low depressed mood was somewhat balanced by high depressive scaled scores, as indicated by a lower prevalence of outliers and higher span and symmetry in the box and whiskers, relative to the Normative class. The Elevated class had fewer outliers than the Normative

class but had a few negative outliers during Waves III and IV, when the Normative class had none.

These distributional characteristics gave possible insight as to why the entropy and average posterior probability was quite poor. Two out of four timepoints had a large amount of score overlap between the classes, which may have made classification somewhat reliant on scores measured later during young adulthood. The overlap between scaled scores may have made classification more sensitive to idiosyncrasies and random chance. The difference between average scores by class grew after the first two waves in adolescence. This suggested that a low depressed mood measured during young adulthood was more telling of Normative membership than when measured during adolescence. Likewise, a high level of depressed mood may be more telling if observed during Waves III and IV, relative to if it were measured during Waves I and II.

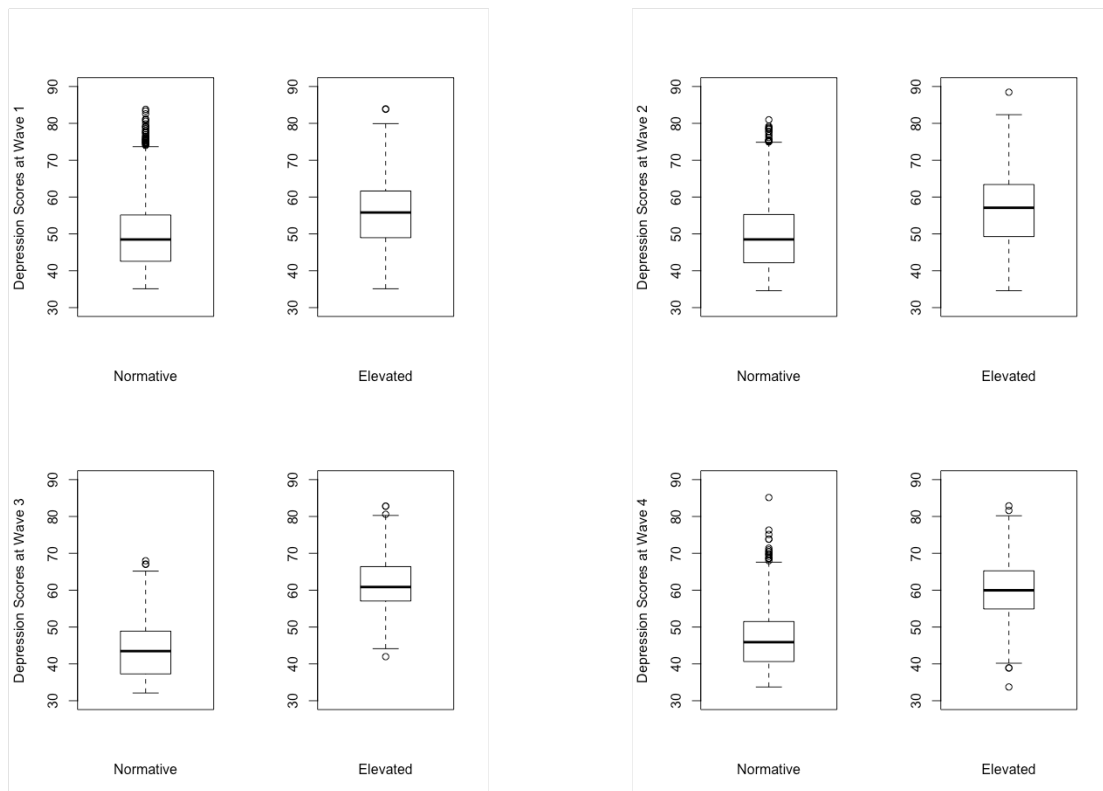


Figure 6. Distribution of depressive symptom scores by latent class and wave.

Multinomial Logistic Regression

Antecedent variables were evaluated using backwards elimination to identify variables with significant partial associations with latent class membership of depressive symptoms. Odds ratios derived from logistic regression estimated the relative likelihood of class membership, after controlling for all other antecedent variables in the model. Backwards elimination was used to incrementally remove non-significant variables until the only remaining antecedents were statistically significant ($\alpha = .05$). Seven antecedent variables were determined to share significant associations with class membership, with

results shown in Table 10. Each antecedent was shown to reflect its relative risk associated with an Elevated classification (vs. Normative).

In addition, two interaction variables, sex and school disengagement, and sex and delinquency, were tested in this model to determine if there were significant differences in trajectory classification by sex. Neither interaction was found to be statistically significant.

Table 10. *Relationship between Demographic and Non-Cognitive Adolescent Characteristics, and Latent Classes of Change in Depressive Symptoms*

Explanatory variable	Odds ratio (95% confidence limits)	% Risk increment ^a
Odds for classification as Elevated (latent class 2) vs. Normative (latent class 1)		
Child is female (vs. male)	3.40 (2.57/4.50)	240.1
Child is African American (vs. other ethnicity)	1.45 (1.10/1.90)	44.8
Child is Native American (vs. other ethnicity)	2.29 (1.36/3.84)	128.6
Child ever retained in grade (vs. never retained)	1.66 (1.25/2.20)	65.7
Child perceives a low likelihood of college (vs. not)	1.73 (1.18/2.54)	73.0
Child in highest quartile of school disengagement (vs. not)	2.64 (2.01/3.47)	164.1
Child in highest quartile of delinquency (vs. not)	1.76 (1.32/2.34)	75.6

Note. Values are estimated through multinomial logistic regression applying the generalized logit link function, where the latent growth classes are regressed simultaneously on antecedent variables.

^aEntries equal odds ratio - 1 (100).

Significant demographic risk factors for being classified as Elevated in average depressive symptom change trajectory, rather than Normative, represents the associated risk after controlling for all other factors in the model. Demographic risk factors included being (a) female, with a 240.1% greater risk ($OR = 3.40$, 95% $CI = [2.57, 4.50]$) than males, (b) being Native American, with a 128.6% greater risk ($OR = 2.29$, 95% $CI = [1.36, 3.84]$) than not identifying as Native American, and (c) being Black, with a 44.8% greater risk ($OR = 1.45$, 95% $CI = [1.10, 1.90]$) than not identifying as Black. Relative to males, females were nearly 3.5 times more likely to have an elevated depressive change trajectory.

Significant non-cognitive risk factors for being classified as Elevated, relative to Normative, represented the associated risk after controlling for all other factors in the model. Non-cognitive risk factors included (a) having high school disengagement, with a 164.1% higher risk ($OR = 2.64$, 95% $CI = [2.01, 3.47]$) relative to those with lower school disengagement, (b) having high delinquency, with a 75.6% higher risk ($OR = 1.76$, 95% $CI = [1.32, 2.34]$) than those with lower delinquency behaviors, (c) self-perceptions of a low likelihood of attending college, with a 73.0% higher risk ($OR = 1.73$, 95% $CI = [1.18, 2.54]$) relative to adolescents who believe there is at least some likelihood of attending college, and (d) being retained in a grade at least once, with a 65.7% higher risk ($OR = 1.66$, 95% $CI = [1.25, 2.20]$) relative to adolescents who had never been held back in a school grade.

Logistic Regression

Relative probabilities of all eleven distal outcomes were significantly associated with latent class membership of Depressive Symptoms and illustrated using boxplots by outcome in Figures 7-9.

A point estimates for each class represents the estimated probability of a given socioeconomic outcome being true, with estimated uncertainty denoted by the confidence intervals (vertical lines) surrounding each point estimate. Non-significant distal outcomes would coincide in boxplot wherever confidence intervals (i.e., the bars around point estimates in Figure 7-9) overlap between the two classes, which did not occur for any outcome in this study. Therefore, all outcome probabilities were significantly associated with latent class membership.

Relative to the Normative class, the Elevated class had a significantly lower likelihood of receiving every employer-provided benefit (e.g., paid sick leave/vacation, employer-provided healthcare, and employer provided retirement benefits), in addition to a lower likelihood of ever working in a supervisory role, and a lower probability of having some degree of job satisfaction.

The Elevated class was also significantly more likely to report a household income below the median income (defined as US median household income in 2008), have more debt than assets, and were more likely to have worried about having enough money to pay their rent/mortgage, utilities, and food over the past year. In addition, the Elevated class was more likely to have received some form of federal assistance during young adulthood (at least once between Wave II and IV). For example, the probability of

individuals in the Normative class had a 18% (95% CI = [16%, 20%] likelihood of receiving federal aid between Wave II and IV, while the Elevated class had a 48% estimated probability (95% CI = [43%, 53%]).

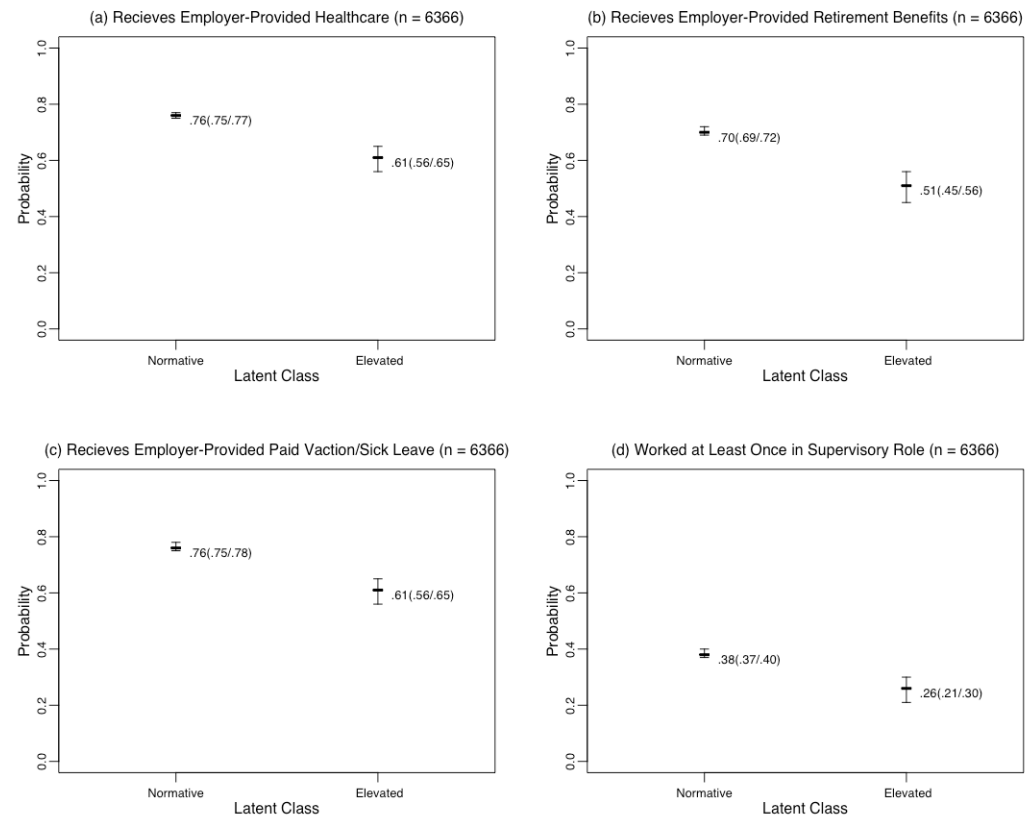


Figure 7. Predicted mean probability (and 95% confidence bands) of labor market outcomes (A-D) associated with class membership.

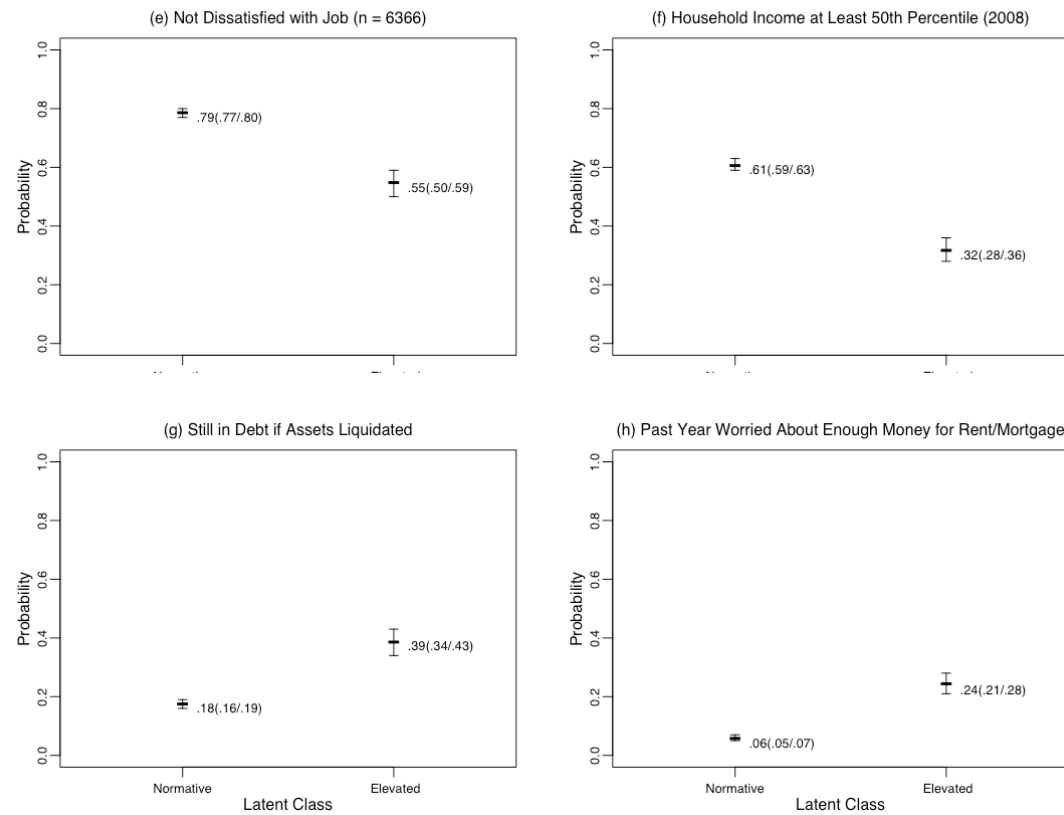


Figure 8. Predicted mean probability (and 95% confidence bands) of labor market outcomes (E-H) associated with class membership.

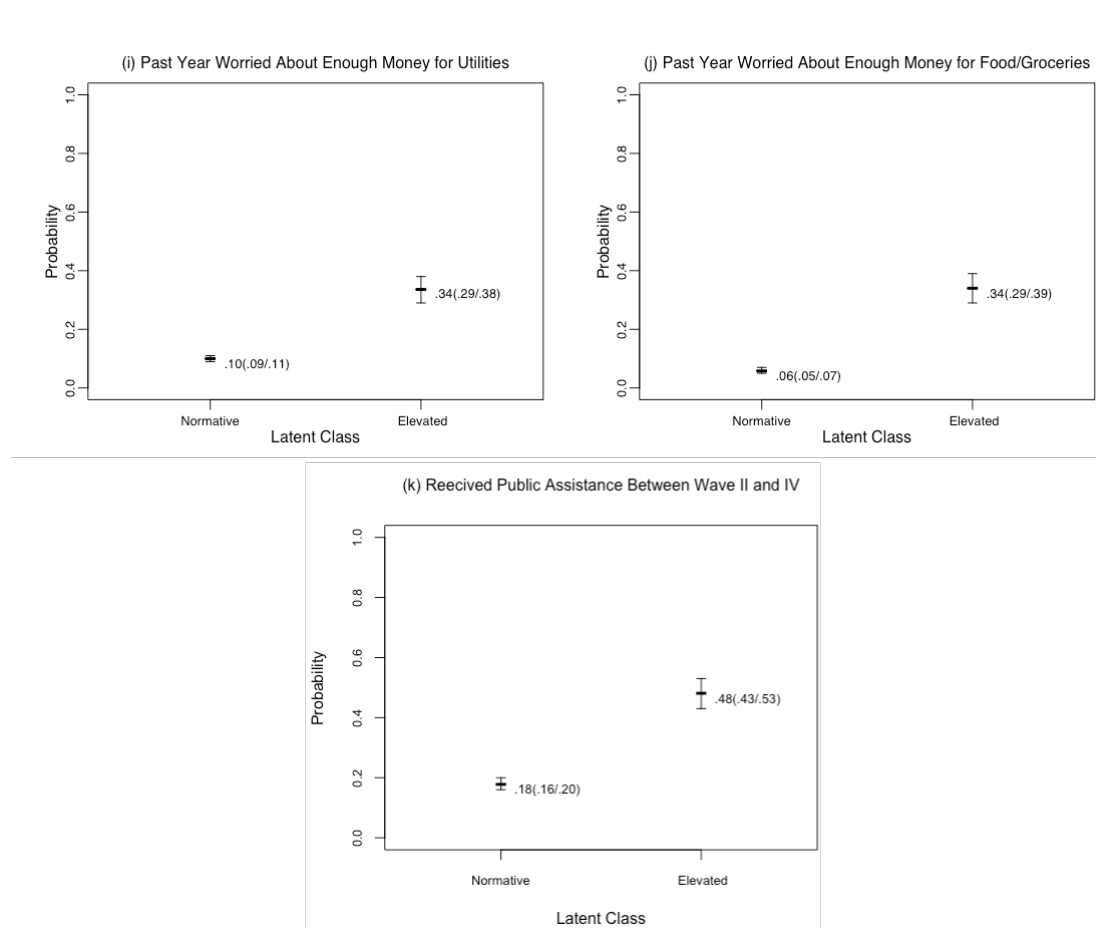


Figure 9. Predicted mean probability (and 95% confidence bands) of labor market outcomes (I-K) associated with class membership.

CHAPTER 4: DISCUSSION

Methodological Challenges

The purpose of this study was to apply latent growth mixture modeling on a nine-item measure of depressive mood/symptoms to evaluate its latent class trajectories between adolescence to young adulthood. Before producing the final latent classes, there were several methodological challenges that had to be addressed.

For instance, the missingness patterns of the data was fairly complex, such that there were missing data at both the item- and wave-level of depressive symptoms, as well as patterns of item-level missingness in the antecedent variables, which could not be handled on *Mplus*. The complexity of missingness required two different imputation approaches and application at different points during the analyses. When handling item-level missingness, MCMC MI was best suited, despite shortcomings and imputing missing case-level depressive scores using FIML. For both imputation methods, missing data were assumed to be missing at random (Allison, 2001). The stability of the 2-class model demonstrated that the estimation of model parameters was not unduly affected by the imputation process.

Another methodological challenge occurred when constructing scales for the both latent constructs (school disengagement and delinquency) which were both measured with multiple items. In particular, special modeling techniques were required to account for the items' ordinal scales. For example, a smoothed polychoric correlation was used instead of the traditional Pearson correlation matrix, a Satorra-Bentler's (2001) χ^2 was reported instead of the normal/standard χ^2 , and CFA models were estimated using

WLSMV which is specifically designed to handle ordinal data and makes no distributional assumptions.

In addition, depressive symptoms were measured using a subset of nine items from the CES-D instrument (Radloff, 1977), which were validated for configural and weak/factorial invariance, so scores could be calibrated, equated and placed on the same scale. This short form of the CES-D highlights the potential “sins”, or validity issues, that may be produced from using a shortened form of the parent instrument (Smith, McCarthy & Anderson, 2000). For example, the transfer of validity from the parent form from the short form may not contain full coverage of the content domain, as seen in this nine-item CES-D instrument. For example, one of the key facets of depression is recurrent suicidal ideation (Association of Psychiatric Association, 2013), which was measured in the original CES-D scale, but not in this short form. Smith, McCarthy & Anderson (2000) argue that although short forms may be validly used in appropriate settings, a thorough analysis must be completed to ensure adequate transfer of validity and an application of appropriate psychometric principles when validating the short form from the parent.

These methodological challenges were able to be handled. However, they did illustrate a few of the challenges which may arise when working with a public use, longitudinal data set. Addressing all issues required a thorough awareness of the available quantitative methods and strategies, given the underlying scales of measurement (i.e., ordinal vs. continuous).

Review of Research Questions

1) Are there multiple subpopulations (latent classes) of developmental change trajectories for depression from adolescence to early adulthood?

Using latent growth mixture modeling and aiming for parsimony, latent change trajectories were best represented with a 2-class, latent basis model. The average trajectories of both classes never crossed one another, where the Elevated class was consistently higher in symptom scores than the Normative class, on average. When compared to initial symptom severity during Wave I, the Normative class exhibited fewer depressive symptoms over time, while the Elevated class had more, on average. In addition, the configural profiles of both classes were maintained when using a sub-sample of data with no item- or wave-level missingness, which suggested that missing patterns were not related to the level and change of depressive symptoms (Little & Rubin, 2002).

These findings were somewhat consistent with the literature, as there is emerging evidence that there are latent sub-groups with distinct depressive symptom change trajectories between adolescence and young adulthood. Researchers commonly found a high stable and a low stable trajectory between adolescence and young adulthood (Yaroslavsky et al., 2013; Brière et al., 2015). Davies et al. (2019) applied latent growth mixture modeling on adolescents aged between 11 to 17 and identified a 2-class model with a normative class (84.1%) and an elevated class (15.9%), which was a finding that was quite similar to this study, although they did not model trajectories beyond adolescence. Several other studies also identified a moderately high and

increasing class, which had a similar configural profile to the Elevated class in this study (Stoolmiller, Kim & Capaldi, 2005; Yaroslavsky et al., 2013; Costello et al., 2008). Furthermore, Brière et al. (2015) reported that in nearly all longitudinal studies of depression trajectory heterogeneity, the greatest proportion of adolescents were represented by a stably low class, which is consistent with the Normative class.

However, unlike this study, most researchers reported three or more latent class trajectories (i.e., Schubert et al., 2017; Brière et al., 2015; Ames & Leadbeater, 2018; Costello et al., 2008; Brendgen et al., 2005; Meadows, Brown & Elder, 2006; Yaroslavsky et al., 2013). Typical patterns of change were still steadily low and steadily high (Yaroslavsky et al., 2013; Schubert et al., 2017), but other typical patterns included a low and increasing trajectory, and high and decreasing trajectory (Costello et al., 2008; Brendgen et al., 2005; Repetto, Caldwell & Zimmerman, 2004; Stoolmiller, Kim & Capaldi, 2005).

Studies that used the restricted-use Add Health data set primarily settled on a 3- or 4-class solution (Yaroslavsky et al., 2013; Wickrama & Wickrama, 2010; Costello et al., 2008). Despite the similarities in their data, no known Add Health study was equivalent in their modeling strategy and selection criteria. For example, several studies (Costello et al., 2008; Wickrama & Wickrama, 2010) selected models based on fit indices like the BIC, ABIC, and LRTs. This may indicate that their solution was due to overfitting, since they made no use nor mention of the ICL-BIC.

In addition, models produced by Yaroslavsky et al. (2013) used a different measure of depressive symptoms that only used three items from the Depressed Affect subscale of the all CES-D items. Wickrama & Wickrama (2010) used listwise

deletion to handle item- and case-level missingness and applied latent class analysis on an eight-item summed score of depression. Costello et al. (2008) modeled a three-item summed score from Waves I-III as a function of timepoint (there was no Wave IV), rather than wave number. They found that two trajectories (stably low, and initially low and escalating) started with the same scores during adolescence, but had diverged by late adolescence (~16 years) and young adulthood, while the other two trajectories (stably moderate, and early high but decreasing) started at significantly different scores but converged by early adulthood (~20 years) (Costello et al., 2008). No studies reported analyses that used any form of score equating of depressive symptom scores over time.

Studies using a short form of the CES-D to measure depressive symptoms often referred to the construct being measured as depression, rather than depressive symptoms. This highlights one of the key concerns when using short forms from the parent instrument and may be one of the causes of discrepancies between this study's results and other solutions (Smith, McCarthy & Anderson, 2000). This study's final model was unlike the results from other studies using Add Health data, which could be due to a variety of differences including timespan, sample characteristics, statistical methodology and/or handling of missing data. Of all available studies using Add Health data, none used vertical equating nor mentioned any modifications in methodology due to the ordinal nature of item data.'

2) *Do pre-existing demographic factors and non-cognitive behaviors and thoughts relate to membership in those subpopulations?*

There were seven significant risk factors for classification in the Elevated subgroup (vs. Normative) and comprised of three demographic variables (being female, Black, and Native American) and four non-cognitive variables (highest school disengagement [upper quartile], highest delinquency [upper quartile], perceived low likelihood of attending college, and ever repeating a grade level).

Demographic factors were backed up by literature which suggested females and certain racial/ethnic minorities were at greater risk for elevated depressive symptom trajectories (Wickrama & Wickrama, 2010; Barr, 2018; Meadows, Brown & Elder, 2006; Yaroslavsky et al., 2013). Starting in adolescence, girls begin to consistently report higher average depressive symptoms than boys, with this difference increasing over the span of adolescence and is sustained into adulthood (Wickrama & Wickrama, 2010; Ge et al., 1994). Furthermore, demographic factors related to occupying a lower status, including socioeconomic status, race/ethnicity, and sex, were associated with a higher level and growth rate of depressive symptoms over time (Barr, 2018). Childhood disadvantage, family disruption, and abuse have all been associated with later life depression (Barr, 2018; Gilman, Kawachi, Fitzmaurice & Buka, 2003; Goosby, 2013), which may offer plausible explanations as to why Black and Native American adolescents were found at a higher risk for Elevated trajectory class membership. Since a measure of household SES at Wave I was not included, it is possible that there was some degree of omitted variable bias. African Americans and Native Americans are both substantially more likely to be living under

the poverty line (Wight et al., 2005; Farley, 1995), and life in disadvantaged circumstances has been associated with greater emotional distress among adolescents (Wight et al., 2005).

There was considerable support from the literature that found adolescent depression was associated with non-cognitive factors, including delinquency, school disengagement, retention in grade level, and low perceptions of future aspirations (Brière et al., 2015; Joyce & Early, 2014; Schulte-Körne, 2016; Stoolmiller, Kim & Capaldi, 2005; Yaroslavsky et al., 2013). Shochet & Smith (2014) found that the classroom environment and school connectedness accounted for approximately 41 to 45% of the variance in concomitant depressive symptoms and 14% of later depressive symptoms after controlling for previous symptoms. Wickrama & Wickrama (2010) reported that adolescent depression moderately contributes to risky and delinquent behavior in young adulthood. The authors also found that a sudden onset in depressive symptom intensity was associated with new and unsafe lifestyle behaviors, including delinquency, substance abuse, and crime (Wickrama & Wickrama, 2010).

Both being retained in a school grade and a perceived low likelihood of attending college are indirectly related to self-esteem, which has been associated with depressive symptoms in the past (Allgood-Merten, Lewinsohn & Hops, 1990; Costello et al., 2008). For being retained in a grade, despite well-intentioned attempts to remediate children/adolescents with poor behavior and/or academic competencies, many researchers have found evidence that retaining a child/adolescent may have additional adverse effects on their educational achievement and socioeconomic

development (Freiburger, 2015; Jimerson et al., 1997; Jimerson, Anderson & Whipple, 2002; Pagani et al., 2001).

To examine gender differences in their risk for depression, two interactions were tested (e.g., sex by school disengagement and sex by delinquency). There was no significant interaction effect between gender and school disengagement nor gender and delinquency, which was not backed up by the literature. Several studies identified significant interactions between sex and non-cognitive behaviors, and depressive symptoms (Costello et al., 2008; Wickrama & Wickrama, 2010). For example, when compared to males, females who smoked and/or had multiple sex partners in adolescence had a significantly higher risk for membership in the elevated and increasing depression group (Wickrama & Wickrama, 2010). One possible reason for this study's findings is that scales for school disengagement and delinquency were both too general in scope to pick up on any gender-specific differences in coping strategies (Nolen-Hoeksema, 1994) and/or how depressive symptoms typically manifest (i.e., internalizing vs. externalizing behaviors; Kandel & Davies, 1982).

3) *Does class membership in those subpopulations predict relevant distal outcomes, including labor market participation, financial well-being, and job satisfaction?*

Latent class membership was significantly associated with outcomes for all eleven labor market outcomes. This demonstrated that those classified as Elevated were at a lower likelihood of positive employment conditions and a higher risk of financial instability in Wave IV, relative to those classified as Normative. These findings were

consistent with findings in the literature that young adults who struggle with elevated or high depressed mood are more likely to report lower annual income and lower job satisfaction than those with low depressed mood (Salmela-Aro, Aunola & Nurmi, 2008; Holsen & Birkeland, 2017; Yaroslavsky et al., 2013).

Limitations and Future Directions

Missing Data

Missing data required careful consideration at the item and case level. Like many longitudinal studies, the sample had evidence of attrition and occasional dropout. Before applying factor analysis and scaling, all item-level missingness was handled using MCMC MI independently by wave for those who were wave-level present. Multiple imputation was required for antecedent variables and all waves of depressive items.

For depression, although the best predictors of a timepoint variable are thought to be the values immediately preceding and succeeding it, each depression wave was independently imputed without using aid from the other waves in the imputation process. This was done to avoid autocorrelation between wave sets. After handling applying multiple imputation, the missing wave-level data was imputed using FIML within the growth mixture modeling stage. Appropriate use of FIML was evaluated by using the sub-sample of those with no missingness to estimate the same 2-class, latent basis model. The configural profiles of both models were quite similar, which supported that the estimation of latent growth mixture model parameters was not overly affected by imputation.

Using Wave vs. Chronological Age

The current study examined depressive symptoms as a function of study wave. However, from a developmental perspective, it may be valuable to examine the same outcome as a function of chronological age. During Wave I, the age of participants spanned the entire period of adolescence, which ranged from 12 to 17 years old. Based on prior literature and theory, depressive symptoms have distinct latent trajectories when studied in adolescence alone (Davies et al., 2019; Stoolmiller, Kim & Capaldi, 2005). In the future, modeling depressive symptoms as a function of chronological time may provide better data coverage and reveal more nuances in the developmental processes. Implementing this change may or may not produce models with a similar pattern of latent class trajectories found here. This should be looked into more thoroughly in the future. In addition, as done in Costello et al. (2008), chronological age may be modeled using different age intervals (i.e., bin sizes).

Self-Report Items

The study relied on self-reported measures which have the potential to produce biased and unreliable results (Garcia & Gustavson, 1997). Results from this study may not be generalizable until more work is done to ensure the validity of the self-report scales (Chao, 2017). Several of the latent traits in this study were relative by necessity of their definition (e.g., depression, school disengagement), with such items reliant wholly on an individual's current mood, temperament and perspective. When self-reporting may be unavoidable, efforts should still be made to improve self-report items as much as possible, such as adding additional items to improve reliability and data coverage, or by

mixing the order of items and removing presentation of items in testlets. Whenever appropriate, objective measures should be used to reduce self-report biases which may improve reliability and reduce bias (Garcia & Gustavson, 1997; Wickrama & Wickrama, 2010).

In addition, a new symptom scale for depressive symptoms may be used instead of the CES-D (Radloff, 1977) such as the Mood and Feelings Questionnaire (MFQ; Kent, Vostanis & Feehan, 1997) or the Beck Depression Inventory II (BDI-II; Beck, Steer, Ball & Ranieri, 1996). Both the MFQ and BDI-II evaluate the prevalence and severity of depressive symptoms over the last two weeks, which is in line with current diagnostic criteria in the DSM-V (American Psychiatric Association, 2013). However, the depressive symptoms measured in the CES-D scale evaluated the prevalence and intensity of symptoms over the last week. Although these items were appropriate for quick screening of recent depressive symptom severity, they relied on a much smaller screening period to assess symptoms, relative to the time elapsed between measurement occasions (waves). Increased reliance on the timing of a respondent's mood introduced an element of randomness to the data which may have diminished the reliability of the results and limited interpretations. A replication of this study may benefit from using a different symptom scale, such as the MFQ or BDI-II, which both span a longer period of time.

The Great Recession (2007-2009)

The Normative class' uptick in depressive symptoms at Wave IV (2008) may be somewhat influenced by the timing of the Great Recession (2007-2009). During this

time, over 30 million American jobs were lost and the average net loss in household income dropped by approximately 18 percent (Kalleberg & von Wachter, 2017). As mentioned in previous sections, there is strong evidence that normative depressive symptom change between adolescence and young adulthood can be characterized by a trajectory that is either consistently low or low and decreasing, which was not observed in this study's Normative class. As shown by the boxplot in Figure 6, high-scoring outliers became more prevalent in Wave IV relative to Wave III, which was not supported in the literature. The Great Recession offers a plausible, logical explanation as to why the frequency of unusually high elevated symptom scores increased rather than remained stable or decreased in 2008. Extreme outliers exhibit a stronger pull on the average, which may explain the observed increase in average symptoms in Wave IV.

In addition, Wave IV measurement timing may have contributed to a larger disparity in distal outcomes for the Normative and Elevated classes. Those in the Elevated class may have been more likely to experience stronger fallout from the recession (i.e., layoffs, evictions, bankruptcy, etc.). The Great Depression did not affect all Americans equally and disproportionately impacted those who were male, Black and less educated (Kalleberg & von Wachter, 2017). In this study, depressive symptoms were also associated with being Black and less engaged in school, so it is possible that the labor market outcomes were more disparate than they would have been without the Great Recession. Significant distal outcomes can be evaluated by replicating this study, with socioeconomic outcomes measured during a period with no economic recession.

Other Revisions

This study offers several possibilities for a follow-up. For example, this study should be replicated on a new, national sample to determine its robustness over time. Wherever possible, efforts should be made to minimize missingness both within and across time, which can be quite difficult in a longitudinal study. Although this current study identified a stable, 2-class latent basis model, it is possible that missingness impacted the precision and ability to detect smaller latent change trajectories. Most studies that used Add Health data (restricted use) identified three or four distinct trajectories, not two found here. Nevertheless, these studies varied by modeling approach and handling of missing data (Wickrama & Wickrama, 2010; Costello et al., 2008).

In addition, this study can be replicated separately by gender. This is supported by study results which found that gender shared the single greatest association with latent class membership. Furthermore, there is considerable support from gender-specific depression studies which revealed that the number, level, and shape of distinct trajectories varied by gender (i.e., Allgood-Merten, Lewinsohn & Hops, 1990; Ames & Leadbeater, 2018; Brendgen et al., 2005; Costello et al., 2008; Holsen & Birkeland, 2017; Stoolmiller, Kim & Capaldi, 2005). By separating analyses by gender, a more thorough understanding of distinct etiological pathways for males and females may be revealed.

The two scales of school disengagement and delinquency were both formed using confirmatory analyses because of their location within clearly defined testlets. No exploratory analyses were performed on either item set. It is possible there are sub-factors within the testlet(s) which may reveal a more detailed insight into the specific

thoughts and behaviors of school disengagement and delinquency which are associated with atypical change in depressive symptoms. In addition, without the presence of testlets, there may be certain school disengagement items that belong with delinquency, and vice versa. This premise is backed up by strong correlations between the two latent factors ($r_{school,diseng} = .51, p < .001$). For example, a history of expulsion, out-of-school suspension, and skipping school without an excused absence were more strongly correlated to more delinquency items than they were to school disengagement items.

There is also possibly an unidentified variable(s) which caused a spurious relationship between trajectory class and identified risk factor(s). Using appropriate exploratory and confirmatory analyses, additional non-cognitive, in-school variables should be used to provide a fuller illustration of the partial risk/protective factors associated with atypical, elevated depressive change trajectories. New variables should still be backed up by relevant literature and psychological theory. Constructs such as self-esteem, abstract thinking, and self-reflection have been previously identified as developmental markers for normative adolescent development (Costello et al., 2008; Nolen-Hoeksema, 1994) and may serve to function as protective factors against atypical change in depressive symptoms (Allgood-Merten, Lewinsohn & Hops, 1990). These findings may be of importance when developing school policies and mental health interventions.

In line with sociometer theory (Leary, Haupt, Strausser & Chokel, 1998), school disengagement may reflect a student's perceived relational value within the school setting (Shochet et al., 2006; Shochet & Smith, 2014), where a sociometer 'alerts' for possible

rejection and may result in a negative affect state. Sociometer theory also states that individuals have different calibrations of their own sociometers, indicating that there are varied responses to the same stimuli in school environments (Leary et al., 1998; Shochet & Smith, 2014). In the future, this hypothesis can be tested by including items that measure adolescent temperament as a latent construct, with interaction effects also tested between temperament and school disengagement, and temperament and delinquency.

There were only two antecedent interactions tested in this model. Future analyses should include additional interaction effects informed by the literature. Based on developmental research and theory, interactions other than gender should be evaluated, such as the interaction between race/ethnicity and school disengagement, and/or delinquency. Several studies have revealed that certain minorities exhibited elevated risks for an atypical depressive symptom trajectory classification (Ames & Leadbeater, 2018; Costello et al., 2008; Wickrama, Wickrama & Lott, 2009), so a new investigation into how associations differ by ethnicity/race may be worthwhile.

Research Contributions and Implications

Positive psychological adjustment during the developmental period from adolescence to young adulthood relies on, in part, the interaction between each person and his/her environment (Shochet & Smith, 2014). During adolescence, depression becomes one of the most common mental health disorders (Schubert et al., 2017). Poor mental health in adolescence has the potential to impede healthy development of cognitive, social, and psychological abilities, with effects that may be carried into adulthood (Elder & Caspi, 1988; Masten, 2007; Wickrama & Wickrama, 2010).

Impairments in the developmental domains increases the propensity for unregulated, dangerous, and uncontrolled behaviors, including early school leaving, delinquency and substance use (Birmaher et al., 1996; Fergusson & Elliott, 2002; Wickrama & Wickrama, 2010). There is strong evidence that depression can have significant unfavorable impacts on later outcomes of life (Carr, 2004; Shochet & Smith, 2014).

Adding to the current understanding of longitudinal development of depressive symptoms, factors relating to school disengagement and delinquency were used to provide empirical evidence which can be used to develop and improve school-based interventions which address depressive symptoms (Shochet & Smith, 2014). In addition, a gap in the research was addressed using distal outcomes that contextualized and revealed the cost of long-term elevated depressive symptoms using a variety of labor market outcomes assessed in young adulthood.

This study identified two stable latent growth trajectories with distinct levels and growth rates. Compared with the Normative class, the Elevated class exhibited higher initial symptoms and significant positive growth in symptom severity during the transition from adolescence to young adulthood. Both classes had a significant change in average symptoms between Wave II and III, which were the last period of adolescence and first period of young adulthood, respectively. The shift out of adolescence into young adulthood somewhat coincided with significant, distinct change in depressive symptoms, with significantly lower and higher symptoms in the Normative and Elevated classes, respectively. This transition may mark a significant junction of interest for stakeholders who are interested and responsible for developing targeted mental health interventions and related school policy. Possible in-school programs may include

increased presence of mentors, more diverse opportunities for adolescent engagement within the school community and supporting the development of new skills and interests (O'Connell, Boat & Warner, 2009).

The Elevated class was associated with being female, Black, and Native American, in addition to displaying high delinquency, high school disengagement, low perceived likelihood of attending college, and ever being held back in a school grade. This initial study offers several opportunities for future research into how these personal characteristics were associated with depressive symptoms. Finally, latent class membership was associated with all socioeconomic markers in Wave IV, including eleven indicators that measured an individual's employment conditions and financial health during Wave IV. The findings from distal outcome analysis painted a clearer picture of the potential long-term socioeconomic consequences of elevated depressive symptoms and offer additional evidence that depression is a major health problem in the United States.

BIBLIOGRAPHY

- Adelman, H., & Taylor, L. (2000). Moving Prevention From the Fringes Into the Fabric of School Improvement. *Journal of Educational and Psychological Consultation*, 11(1), 7-36. doi:10.1207/s1532768xjepc1101_3
- Adkins, D. E., Wang, V., Dupre, M. E., E. J. C. G. Van Den Oord, & Elder, G. H. (2009). Structure and Stress: Trajectories of Depressive Symptoms across Adolescence and Young Adulthood. *Social Forces*, 88(1), 31-60. doi:10.1353/sof.0.0238
- Allgood-Merten, B., Lewinsohn, P. M., & Hops, H. (1990). Sex differences and adolescent depression. *Journal of Abnormal Psychology*, 99(1), 55-63. doi:10.1037/0021-843x.99.1.55
- Allison, P. D. (2005). Imputation of categorical variables with PROC MI. Retrieved 2020, from <http://www2.sas.com/proceedings/sugi30/113-30.pdf>
- Allison, P. D. (2009). *Missing data*. Thousand Oaks, CA, CA: Sage Publications.
- Ames, M. E., & Leadbeater, B. J. (2018). Depressive symptom trajectories and physical health: Persistence of problems from adolescence to young adulthood. *Journal of Affective Disorders*, 240, 121-129. doi:10.1016/j.jad.2018.07.001
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103(3), 411-423. doi:10.1037/0033-2909.103.3.411
- Arnett, J. J. (2000). Emerging adulthood: A theory of development from the late teens through the twenties. *American Psychologist*, 55(5), 469-480. doi:10.1037/0003-066x.55.5.469
- Avison, W. R., & McAlpine, D. D. (1992). Gender Differences in Symptoms of Depression Among Adolescents. *Journal of Health and Social Behavior*, 33(2), 77. doi:10.2307/2137248
- Barr, P. B. (2018). Early neighborhood conditions and trajectories of depressive symptoms across adolescence and into adulthood. *Advances in Life Course Research*, 35, 57-68. doi:10.1016/j.alcr.2018.01.005
- Beck, A. T., Steer, R. A., Ball, R., & Ranieri, W. F. (1996). Comparison of Beck Depression Inventories-IA and-II in Psychiatric Outpatients. *Journal of Personality Assessment*, 67(3), 588-597. doi:10.1207/s15327752jpa6703_13

- Birmaher, B., Ryan, N. D., Williamson, D. E., Brent, D. A., & Kaufman, J. (1996). Childhood and Adolescent Depression: A Review of the Past 10 Years. *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(12), 1575-1583. doi:10.1097/00004583-199612000-00008
- Bodner, T. E. (2008). What Improves with Increased Missing Data Imputations? *Structural Equation Modeling: A Multidisciplinary Journal*, 15(4), 651-675. doi:10.1080/10705510802339072
- Brendgen, M., Wanner, B., Morin, A. J., & Vitaro, F. (2005). Relations with Parents and with Peers, Temperament, and Trajectories of Depressed Mood During Early Adolescence. *Journal of Abnormal Child Psychology*, 33(5), 579-594. doi:10.1007/s10802-005-6739-2
- Brière, F. N., Janosz, M., Fallu, J., & Morizot, J. (2015). Adolescent Trajectories of Depressive Symptoms: Codevelopment of Behavioral and Academic Problems. *Journal of Adolescent Health*, 57(3), 313-319. doi:10.1016/j.jadohealth.2015.05.012
- Brown, T. A., & Little, T. D. (2015). *Confirmatory factor analysis for applied research*. New York, NY: The Guilford Press.
- Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37(1), 62-83. doi:10.1111/j.2044-8317.1984.tb00789.x
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park: Sage Publ.
- Buist-Bouwman, M. A., Graaf, R., Vollebergh, W. A., & Ormel, J. (2005). Comorbidity of physical and mental disorders and the effect on work-loss days. *Acta Psychiatrica Scandinavica*, 111(6), 436-443. doi:10.1111/j.1600-0447.2005.00513.x
- Burchinal, M. R., Roberts, J. E., Zeisel, S. A., & Rowley, S. J. (2008). Social risk and protective factors for African American childrens academic achievement and adjustment during the transition to middle school. *Developmental Psychology*, 44(1), 286-292. doi:10.1037/0012-1649.44.1.286
- Cai, L. (2013). FlexMIRT (Version 2) [Computer software]. Chapel Hill: Vector Psychometric Group.
- Campbell, C. (2015). The socioeconomic consequences of dropping out of high school: Evidence from an analysis of siblings. *Social Science Research*, 51, 108-118. doi:10.1016/j.ssresearch.2014.12.011

- Campbell, T. L., Byrne, B. M., & Baron, P. (1992). Gender Differences in the Expression of Depressive Symptoms in Early Adolescents. *The Journal of Early Adolescence*, 12(3), 326-338. doi:10.1177/0272431692012003007
- Capaldi, D. M., & Stoolmiller, M. (1999). Co-occurrence of conduct problems and depressive symptoms in early adolescent boys: III. Prediction to young-adult adjustment. *Development and Psychopathology*, 11(1), 59-84. doi:10.1017/s0954579499001959
- Carr, A. (2004). *Depression and attempted suicide in adolescence*. Camberwell, Victoria: ACER Press.
- Chao, J. L. (2017). The Impact of Attitudes toward Science and Core Self-Evaluation on Science Achievement and Career Outcomes: A Trajectory-Based Approach. Retrieved September 28, 2020, from <https://repository.upenn.edu/dissertations/AAI10273669>
- Chapman, D. P., Perry, G. S., & Strine, T. W. (2005). The vital link between chronic disease and depressive disorders. Preventing chronic disease, 2(1), A14. *Preventing Chronic Disease*, 2(1).
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 233-255. doi:10.1207/s15328007sem0902_5
- Cicchetti, D., & Toth, S. L. (1998). The development of depression in children and adolescents. *American Psychologist*, 53(2), 221-241. doi:10.1037//0003-066x.53.2.221
- Clayborne, Z. M., Varin, M., & Colman, I. (2019). Systematic Review and Meta-Analysis: Adolescent Depression and Long-Term Psychosocial Outcomes. *Journal of the American Academy of Child & Adolescent Psychiatry*, 58(1), 72-79. doi:10.1016/j.jaac.2018.07.896
- Cole, D. A., Tram, J. M., Martin, J. M., Hoffman, K. B., Ruiz, M. D., Jacquez, F. M., & Maschman, T. L. (2002). Individual differences in the emergence of depressive symptoms in children and adolescents: A longitudinal investigation of parent and child reports. *Journal of Abnormal Psychology*, 111(1), 156-165. doi:10.1037/0021-843x.111.1.156
- Compas, B. E., Connor-Smith, J., & Jaser, S. S. (2004). Temperament, Stress Reactivity, and Coping: Implications for Depression in Childhood and Adolescence. *Journal of Clinical Child & Adolescent Psychology*, 33(1), 21-31. doi:10.1207/s15374424jccp3301_3

- Costello, D. M., Swendsen, J., Rose, J. S., & Dierker, L. C. (2008). Risk and protective factors associated with trajectories of depressed mood from adolescence to early adulthood. *Journal of Consulting and Clinical Psychology, 76*(2), 173-183. doi:10.1037/0022-006x.76.2.173
- Davies, S. E., Neufeld, S. A., Sprang, E., Schweren, L., Keivit, R., Fonagy, P., Goodyer, I. M. (2019). Trajectories of depression symptom change during and following treatment in adolescents with unipolar major depression. *Journal of Child Psychology and Psychiatry, 61*(5), 565-574. doi:10.1111/jcpp.13145
- Deković, M., Buist, K. L., & Reitz, E. (2004). Stability and Changes in Problem Behavior During Adolescence: Latent Growth Analysis. *Journal of Youth and Adolescence, 33*(1), 1-12. doi:10.1023/a:1027305312204
- Diamond, G., Russon, J., & Levy, S. (2016). Attachment-Based Family Therapy: A Review of the Empirical Support. *Family Process, 55*(3), 595-610. doi:10.1111/famp.12241
- Diego, M. A., Field, T. M., & Sanders, C. E. (2003, April/May). Academic performance, popularity, and depression predict adolescent substance use. *Adolescence, 38*(149), 35-42.
- Dooley, D., Fielding, J., & Levi, L. (1996). Health and Unemployment. *Annual Review of Public Health, 17*(1), 449-465. doi:10.1146/annurev.pu.17.050196.002313
- Druss, B. G., Hwang, I., Petukhova, M., Sampson, N. A., Wang, P. S., & Kessler, R. C. (2008). Impairment in role functioning in mental and chronic medical disorders in the United States: Results from the National Comorbidity Survey Replication. *Molecular Psychiatry, 14*(7), 728-737. doi:10.1038/mp.2008.13
- Duerden, M. (2009, April). *What is...? series*. Reading presented at Hayward Medical Communications in United Kingdom, London.
- Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2009). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.
- du Toit, M. (2003). *IRT from SSI: BILOG-MG MULTILOG PARSCALE TESTFACT*. Lincolnwood, IL: Scientific Software International. Retrieved 2020, from https://books.google.com/books?id=fbH0RgiZmoAC&dq=du+toit+2003&lr=&source=gbs_navlinks_s
- Eccles, J. S., & Roeser, R. W. (2010). An Ecological View of Schools and Development. *Handbook of Research on Schools, Schooling and Human Development*. doi:10.4324/9780203874844.ch2

- Eccles, J. S., & Roeser, R. W. (2011). Schools as Developmental Contexts During Adolescence. *Journal of Research on Adolescence*, 21(1), 225-241. doi:10.1111/j.1532-7795.2010.00725.x
- Ehrenberg, M. F., Cox, D. N., & Koopman, R. F. (1991). The relationship between self-efficacy and depression in adolescents. *Adolescence*, 26(102).
- Elder, G. H., & Caspi, A. (1988). Economic Stress in Lives: Developmental Perspectives. *Journal of Social Issues*, 44(4), 25-45. doi:10.1111/j.1540-4560.1988.tb02090.x
- Elder, G. H., & Caspi, A. (1988). Economic Stress in Lives: Developmental Perspectives. *Journal of Social Issues*, 44(4), 25-45. doi:10.1111/j.1540-4560.1988.tb02090.x
- Ende, J. V., Verhulst, F. C., & Tiemeier, H. (2016). The bidirectional pathways between internalizing and externalizing problems and academic performance from 6 to 18 years. *Development and Psychopathology*, 28(3), 855-867. doi:10.1017/s0954579416000353
- Farkas, G. (1996). *Human capital or cultural capital? Ethnicity and poverty groups in an urban school district*. New York, NY: Aldine De Gruyter.
- Farley, R. (1995). *State of the union: America in the 1990s*. New York, NY: Russell Sage Foundation.
- Fergusson, D. M., & Woodward, L. J. (2002). Mental Health, Educational, and Social Role Outcomes of Adolescents With Depression. *Archives of General Psychiatry*, 59(3), 225. doi:10.1001/archpsyc.59.3.225
- Fergusson, D. M., Boden, J. M., & Horwood, L. J. (2007). Recurrence of major depression in adolescence and early adulthood, and later mental health, educational and economic outcomes. *British Journal of Psychiatry*, 191(4), 335-342. doi:10.1192/bjp.bp.107.036079
- Fergusson, D. M., Boden, J. M., & Horwood, L. J. (2007). Recurrence of major depression in adolescence and early adulthood, and later mental health, educational and economic outcomes. *British Journal of Psychiatry*, 191(4), 335-342. doi:10.1192/bjp.bp.107.036079
- Ferrari, A. J., Charlson, F. J., Norman, R. E., Patten, S. B., Freedman, G., Murray, C. J., . . . Whiteford, H. A. (2013). Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010. *PLoS Medicine*, 10(11). doi:10.1371/journal.pmed.1001547

- Freeman, H., & Brown, B. B. (2001). Primary Attachment to Parents and Peers during Adolescence: Differences by Attachment Style. *Journal of Youth and Adolescence*, 30(6), 653-674. doi:10.1023/a:1012200511045
- Freiburger, M. B. (ga). *Relationship between grade retention and psychological outcomes during adolescence* (Master's thesis, University of Georgia, 2015) (pp. 1-78). Athens: University of Georgia.
- Fried, E. I., & Nesse, R. M. (2014). The Impact of Individual Depressive Symptoms on Impairment of Psychosocial Functioning. *PLoS ONE*, 9(2). doi:10.1371/journal.pone.0090311
- Frühwirth-Schnatter, S. (2006). *Finite mixture and Markov switching models*. New York, NY: Springer.
- Garber, J. (2006). Depression in Children and Adolescents. *American Journal of Preventive Medicine*, 31(6), 104-125. doi:10.1016/j.amepre.2006.07.007
- Garcia, J., & Gustavson, A. R. (1997, January 01). The Science of Self-Report. Retrieved 2020, from <https://www.psychologicalscience.org/observer/the-science-of-self-report>
- Garrison, C. Z., Addy, C. L., Jackson, K. L., Mckeown, R. E., & Waller, J. L. (1991). The CES-D as a Screen for Depression and Other Psychiatric Disorders in Adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, 30(4), 636-641. doi:10.1097/00004583-199107000-00017
- Ge, X., Conger, R. D., & Elder, G. H. (2001). Pubertal transition, stressful life events, and the emergence of gender differences in adolescent depressive symptoms. *Developmental Psychology*, 37(3), 404-417. doi:10.1037//0012-1649.37.3.404
- Ge, X., Lorenz, F. O., Conger, R. D., Elder, G. H., & Simons, R. L. (1994). Trajectories of stressful life events and depressive symptoms during adolescence. *Developmental Psychology*, 30(4), 467-483. doi:10.1037//0012-1649.30.4.467
- Gilliom, M., & Shaw, D. S. (2004). Codevelopment of externalizing and internalizing problems in early childhood. *Development and Psychopathology*, 16(02). doi:10.1017/s0954579404044530
- Gilman, S. E., Kawachi, I., Fitzmaurice, G. M., & Buka, S. L. (2003). Family Disruption in Childhood and Risk of Adult Depression. *American Journal of Psychiatry*, 160(5), 939-946. doi:10.1176/appi.ajp.160.5.939

- Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychology in the Schools, 30*(1), 79-90. doi:10.1002/1520-6807(199301)30:13.0.co;2-x
- Goodman, A., Joyce, R., & Smith, J. P. (2011). The long shadow cast by childhood physical and mental problems on adult life. *Proceedings of the National Academy of Sciences, 108*(15), 6032-6037. doi:10.1073/pnas.1016970108
- Goosby, B. J. (2013). Early Life Course Pathways of Adult Depression and Chronic Pain. *Journal of Health and Social Behavior, 54*(1), 75-91. doi:10.1177/0022146512475089
- Gore, S., Jr., R. H., Aseltine & Colten, M. E. (1993). Gender, Social-Relationship Involvement, and Depression. *Journal of Research on Adolescence, 3*(2), 101-125. doi:10.1207/s15327795jra0302_1
- Gorter, R., Fox, J., & Twisk, J. W. (2015). Why item response theory should be used for longitudinal questionnaire data analysis in medical research. *BMC Medical Research Methodology, 15*(1). doi:10.1186/s12874-015-0050-x
- Greenbaum, P. E., Del Boca, F. K., Darkes, J., Wang, C., & Goldman, M. S. (2005). Variation in the drinking trajectories of freshmen college students. *Journal of Consulting and Clinical Psychology, 73*(2). doi:https://doi.org/10.1037/0022-006x.73.2.229
- Greenberg, P. E., Kessler, R. C., Birnbaum, H. G., Leong, S. A., Lowe, S. W., Berglund, P. A., & Corey-Lisle, P. K. (2003). The Economic Burden of Depression in the United States. *The Journal of Clinical Psychiatry, 64*(12), 1465-1475. doi:10.4088/jcp.v64n1211
- Hammen, C. (2009). Adolescent Depression. *Current Directions in Psychological Science, 18*(4), 200-204. doi:10.1111/j.1467-8721.2009.01636.x
- Hammer-Helmich, L., Haro, J. M., Jönsson, B., Melac, A. T., Nicola, S. D., Chollet, J., Milea, D., Rive, B. & Saragoussi, D. (2018). Functional impairment in patients with major depressive disorder: The 2-year PERFORM study. *Neuropsychiatric Disease and Treatment, Volume 14*, 239-249. doi:10.2147/ndt.s146098
- Hankin, B. L., & Abramson, L. Y. (2001). Development of gender differences in depression: An elaborated cognitive vulnerability-transactional stress theory. *Psychological Bulletin, 127*(6), 773-796. doi:10.1037//0033-2909.127.6.773
- Hankin, B. L., Mermelstein, R., & Roesch, L. (2007). Sex Differences in Adolescent Depression: Stress Exposure and Reactivity Models. *Child Development, 78*(1), 279-295. doi:10.1111/j.1467-8624.2007.00997.x

- Harris, K. M. (2013). *The Add Health Study: Design and Accomplishments*. Retrieved 2020, from Add Health. doi:<https://doi.org/10.17615/C6TW87>
- Harris, K. M., & Udry, J. R. (1995). National Longitudinal Study of Adolescent to Adult Health, Public Use Contextual Database, Wave I. Retrieved from <http://www.thearda.com/Archive/Files/Descriptions/ADDHW1C.asp>
- Harris, K. M., & Udry, R. J. (2018). National Longitudinal Study of Adolescent to Adult Health (Add Health), 1994-2008 [Public Use]. Retrieved September 28, 2020, from <https://doi.org/10.3886/ICPSR21600.v21>
- Harris, K. M., Halpern, C. T., Whitsel, E., Hussey, J., Tabor, J., Entzel, P., & Udry, J. R. (2009). The National Longitudinal Study of Adolescent to Adult Health: Research Design. Retrieved from <http://www.cpc.unc.edu/projects/addhealth/design>
- Hays, R. D. (1995). Functioning and Well-being Outcomes of Patients With Depression Compared With Chronic General Medical Illnesses. *Archives of General Psychiatry*, 52(1), 11. doi:10.1001/archpsyc.1995.03950130011002
- Henricsson, L., & Rydell, A. (2004). Elementary School Children with Behavior Problems: Teacher-Child Relations and Self-Perception. A Prospective Study. *Merrill-Palmer Quarterly*, 50(2), 111-138. doi:10.1353/mpq.2004.0012
- Holsen, I., & Birkeland, M. S. (2017). Course, Predictors, and Outcomes of Depressed Mood in a Cohort of Norwegian Adolescents Followed From Age 13 Into Adulthood Age 30. *Emerging Adulthood*, 5(1), 3-15. doi:10.1177/2167696816635810
- Horowitz, J. L., & Garber, J. (2006). The prevention of depressive symptoms in children and adolescents: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 74(3), 401-415. doi:10.1037/0022-006x.74.3.401
- Hoyle, R. H. (2000). *Structural equation modeling: Concepts, issues, and applications*. Sage Publ.
- Jaycox, L. H., Stein, B. D., Paddock, S., Miles, J. N., Chandra, A., Meredith, L. S., . . . Burnam, M. A. (2009). Impact of Teen Depression on Academic, Social, and Physical Functioning. *Pediatrics*, 124(4). doi:10.1542/peds.2008-3348
- Jimerson, S. R., Anderson, G. E., & Whipple, A. D. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools*, 39(4), 441-457. doi:10.1002/pits.10046

- Jimerson, S., Carlson, E., Rotert, M., Egeland, B., & Sroufe, L. (1997). A prospective, longitudinal study of the correlates and consequences of early grade retention. *Journal of School Psychology, 35*(1), 3-25. doi:10.1016/s0022-4405(96)00033-7
- Joyce, H. D., & Early, T. J. (2014). The impact of school connectedness and teacher support on depressive symptoms in adolescents: A multilevel analysis. *Children and Youth Services Review, 39*, 101-107. doi:10.1016/j.childyouth.2014.02.005
- Judd, L. L., Paulus, M. P., Wells, K. B., & Rapaport, M. H. (1997). Functional Impairment Associated With Subsyndromal Depression. *PsycEXTRA Dataset*. doi:10.1037/e331142004-011
- Kalleberg, A. L., & von Wachter, T. M. (2017). The U.S. Labor Market During and After the Great Recession: Continuities and Transformations. *RSF: The Russell Sage Foundation Journal of the Social Sciences, 3*(3), 1. doi:10.7758/rsf.2017.3.3.01
- Kandel, D. B., & Davies, M. (1982). Epidemiology of Depressive Mood in Adolescents. *Archives of General Psychiatry, 39*(10), 1205. doi:10.1001/archpsyc.1982.04290100065011
- Kent, L., Vostanis, P., & Feehan, C. (1997). Detection of Major and Minor Depression in Children and Adolescents: Evaluation of the Mood and Feelings Questionnaire. *Journal of Child Psychology and Psychiatry, 38*(5), 565-573. doi:10.1111/j.1469-7610.1997.tb01543.x
- Kessler, R. C. (2012). The Costs of Depression. *Psychiatric Clinics of North America, 35*(1), 1-14. doi:10.1016/j.psc.2011.11.005
- Kessler, R. C., & Walters, E. E. (1998). Epidemiology of DSM-III-R major depression and minor depression among adolescents and young adults in the national comorbidity survey. *Depression and Anxiety, 7*(1), 3-14. doi:10.1002/(sici)1520-6394(1998)7:13.0.co;2-f
- Kessler, R. C., Ormel, J., Demler, O., & Stang, P. E. (2003). Comorbid Mental Disorders Account for the Role Impairment of Commonly Occurring Chronic Physical Disorders: Results From the National Comorbidity Survey. *Journal of Occupational and Environmental Medicine, 45*(12), 1257-1266. doi:10.1097/01.jom.0000100000.70011.bb
- Kim, H. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics, 38*(1), 52. doi:10.5395/rde.2013.38.1.52
- Kim, J., Thompson, E. A., Walsh, E. M., & Schepp, K. G. (2015). Trajectories of Parent-Adolescent Relationship Quality Among At-Risk Youth: Parental Depression and

- Adolescent Developmental Outcomes. *Archives of Psychiatric Nursing*, 29(6), 434-440. doi:10.1016/j.apnu.2015.07.001
- Leary, M. R., Haupt, A. L., Strausser, K. S., & Chokel, J. T. (1998). Calibrating the sociometer: The relationship between interpersonal appraisals and the state self-esteem. *Journal of Personality and Social Psychology*, 74(5), 1290-1299. doi:10.1037/0022-3514.74.5.1290
- Lee, K. J., & Carlin, J. B. (2010). Multiple Imputation for Missing Data: Fully Conditional Specification Versus Multivariate Normal Imputation. *American Journal of Epidemiology*, 171(5), 624-632. doi:10.1093/aje/kwp425
- Lee, K. J., Galati, J. C., Simpson, J. A., & Carlin, J. B. (2012). Comparison of methods for imputing ordinal data using multivariate normal imputation: A case study of non-linear effects in a large cohort study. *Statistics in Medicine*, 31(30), 4164-4174. doi:10.1002/sim.5445
- Lee, S., Tsang, A., Breslau, J., Aguilar-Gaxiola, S., Angermeyer, M., Borges, G., Kessler, R. C. (2009). Mental disorders and termination of education in high-income and low- and middle-income countries: Epidemiological study. *British Journal of Psychiatry*, 194(5), 411-417. doi:10.1192/bjp.bp.108.054841
- Leite, W., & Beretvas, S. N. (2010). The Performance of Multiple Imputation for Likert-type Items with Missing Data. *Journal of Modern Applied Statistical Methods*, 9(1), 64-74. doi:10.22237/jmasm/1272686820
- Levitt, J. M., Saka, N., Romanelli, L. H., & Hoagwood, K. (2007). Early identification of mental health problems in schools: The status of instrumentation. *Journal of School Psychology*, 45(2), 163-191. doi:10.1016/j.jsp.2006.11.005
- Lewinsohn, P. M., Clarke, G. N., Seeley, J. R., & Rohde, P. (1994). Major Depression in Community Adolescents: Age at Onset, Episode Duration, and Time to Recurrence. *Journal of the American Academy of Child & Adolescent Psychiatry*, 33(6), 809-818. doi:10.1097/00004583-199407000-00006
- Li, C. (2015). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936-949. doi:10.3758/s13428-015-0619-7
- Little, R. J., & Rubin, D. B. (2002). Statistical Analysis with Missing Data. doi:10.1002/9781119013563
- Loeber, R., & Burke, J. D. (2011). Developmental Pathways in Juvenile Externalizing and Internalizing Problems. *Journal of Research on Adolescence*, 21(1), 34-46. doi:10.1111/j.1532-7795.2010.00713.x

- MacLachlan, G. J., & Peel, D. (2000). *Finite mixture models*. John Wiley & Sons.
- Marsh, H. W., Liem, G. A., Martin, A. J., Morin, A. J., & Nagengast, B. (2011). Methodological Measurement Fruitfulness of Exploratory Structural Equation Modeling (ESEM): New Approaches to Key Substantive Issues in Motivation and Engagement. *Journal of Psychoeducational Assessment*, 29(4), 322-346. doi:10.1177/0734282911406657
- Masten, A. S. (2007). Competence, Resilience, and Development in Adolescence. *Adolescent Psychopathology and the Developing Brain*, 31-52. doi:10.1093/acprof:oso/9780195306255.003.0002
- McLeod, J. D., Uemura, R., & Rohrman, S. (2012). Adolescent Mental Health, Behavior Problems, and Academic Achievement. *Journal of Health and Social Behavior*, 53(4), 482-497. doi:10.1177/0022146512462888
- McNeely, C. A., Nonnemaker, J. M., & Blum, R. W. (2002). Promoting School Connectedness: Evidence from the National Longitudinal Study of Adolescent Health. *Journal of School Health*, 72(4), 138-146. doi:10.1111/j.1746-1561.2002.tb06533.x
- Meadows, S. O., Brown, J. S., & Elder, G. H. (2006). Depressive Symptoms, Stress, and Support: Gendered Trajectories From Adolescence to Young Adulthood. *Journal of Youth and Adolescence*, 35(1), 89-99. doi:10.1007/s10964-005-9021-6
- Melkevik, O., Nilsen, W., Evensen, M., Reneflot, A., & Mykletun, A. (2016). Internalizing Disorders as Risk Factors for Early School Leaving: A Systematic Review. *Adolescent Research Review*, 1(3), 245-255. doi:10.1007/s40894-016-0024-1
- Merikangas, K. R., Ames, M., Cui, L., Stang, P. E., Ustun, T. B., Korff, M. V., & Kessler, R. C. (2007). The Impact of Comorbidity of Mental and Physical Conditions on Role Disability in the US Adult Household Population. *Archives of General Psychiatry*, 64(10), 1180. doi:10.1001/archpsyc.64.10.1180
- Millings, A., Buck, R., Montgomery, A., Spears, M., & Stallard, P. (2012). School connectedness, peer attachment, and self-esteem as predictors of adolescent depression. *Journal of Adolescence*, 35(4), 1061-1067. doi:10.1016/j.adolescence.2012.02.015
- Muthén, L. K., & Muthén, B. O. (2019). Mplus (Version 8.4) [Computer software]. Retrieved 2020, from <http://www.statmodel.com/ugexcerpts.shtml>

- Muthén, B. (2004). Latent Variable Analysis: Growth Mixture Modeling and Related Techniques for Longitudinal Data. *The SAGE Handbook of Quantitative Methodology for the Social Sciences*, 346-369. doi:10.4135/9781412986311.n19
- Muthén, B., & Muthén, L. K. (2000). Integrating Person-Centered and Variable-Centered Analyses: Growth Mixture Modeling With Latent Trajectory Classes. *Alcoholism: Clinical and Experimental Research*, 24(6), 882-891. doi:10.1097/00000374-200006000-00020
- Mykletun, A., Bjerkeset, O., Dewey, M., Prince, M., Overland, S., & Stewart, R. (2007). Anxiety, depression and cause - specific mortality. The HUNT historical cohort study. *European Psychiatry*, 22. doi:10.1016/j.eurpsy.2007.01.212
- Nagin, D. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychological Methods*, 4(2), 139-157.
- Nagin, D. (2005). *Group-based modeling of development: A semiparametric, group-based approach*. Harvard University Press.
- National Institute of Mental Health [NIMH]. (2017). Retrieved from <https://www.nimh.nih.gov/health/statistics/major-depression.shtml>
- Ng, D. M., & Jeffery, R. W. (2003). Relationships Between Perceived Stress and Health Behaviors in a Sample of Working Adults. *Health Psychology*, 22(6), 638-642. doi:10.1037/0278-6133.22.6.638
- Nolen-Hoeksema, S. (1994). An Interactive Model for the Emergence of Gender Differences in Depression in Adolescence. *Journal of Research on Adolescence*, 4(4), 519-534. doi:10.1207/s15327795jra0404_5
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the Number of Classes in Latent Class Analysis and Growth Mixture Modeling: A Monte Carlo Simulation Study. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(4), 535-569. doi:10.1080/10705510701575396
- O'Connell, M. E., Boat, T. F., & Warner, K. E. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC, Washington, DC: National Academies Press.
- Ormel, J., Kempen, G. I., Deeg, D. J., Brilman, E. I., Sonderen, E. V., & Relyveld, J. (1998). Functioning, Well-Being, and Health Perception in Late Middle-Aged and Older People: Comparing the Effects of Depressive Symptoms and Chronic Medical Conditions. *Journal of the American Geriatrics Society*, 46(1), 39-48. doi:10.1111/j.1532-5415.1998.tb01011.x

- Ozer, E. J., Wolf, J. P., & Kong, C. (2008). Sources of Perceived School Connection Among Ethnically-Diverse Urban Adolescents. *Journal of Adolescent Research*, 23(4), 438-470. doi:10.1177/0743558408316725z
- Pagani, L., Tremblay, R. E., Vitaro, F., Boulerice, B., & McDuff, P. (2001). Effects of grade retention on academic performance and behavioral development. *Development and Psychopathology*, 13(2), 297-315. doi:10.1017/s0954579401002061
- Patterson, G. R., Debaryshe, B. D., & Ramsey, E. (2017). A Developmental Perspective on Antisocial Behavior. *Developmental and Life-course Criminological Theories*, 29-35. doi:10.4324/9781315094908-2
- Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends in Cognitive Sciences*, 9(2), 60-68. doi:10.1016/j.tics.2004.12.008
- Quiroga, C. V., Janosz, M., Lyons, J. S., & Morin, A. J. (2012). Grade Retention and Seventh-Grade Depression Symptoms in the Course of School Dropout among High-Risk Adolescents. *Psychology*, 03(09), 749-755. doi:10.4236/psych.2012.329113
- Radloff, L. S. (1977). The CES-D Scale. *Applied Psychological Measurement*, 1(3), 385-401. doi:10.1177/014662167700100306
- Ram, N. J., & Grimm, K. J. (2009). Growth mixture modeling: A method for identifying differences in longitudinal change among unobserved groups. *International Journal of Behavioral Development*, 33, 565-576. doi:https://doi.org/10.1177/0165025409343765
- Repetto, P., Caldwell, C., & Zimmerman, M. (2004). Trajectories of depressive symptoms among high risk African-American adolescents. *Journal of Adolescent Health*, 35(6), 468-477. doi:10.1016/s1054-139x(04)00095-3
- Resnick, M. D. (1997). Protecting adolescents from harm. Findings from the National Longitudinal Study on Adolescent Health. *JAMA: The Journal of the American Medical Association*, 278(10), 823-832. doi:10.1001/jama.278.10.823
- Revelle, W. (2020, September 04). Procedures for Psychological, Psychometric, and Personality Research [R package psych version 2.0.8]. Retrieved September 28, 2020, from <https://CRAN.R-project.org/package=psych>
- Reynolds, A. D., & Crea, T. M. (2015). Peer influence processes for youth delinquency and depression. *Journal of Adolescence*, 43, 83-95. doi:10.1016/j.adolescence.2015.05.013

- Rhemtulla, M., Brosseau-Liard, P. É., & Savalei, V. (2012). When can categorical variables be treated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17*(3), 354-373. doi:10.1037/a0029315
- Roeser, R. W., Eccles, J. S., & Sameroff, A. J. (2000). School as a Context of Early Adolescents Academic and Social-Emotional Development: A Summary of Research Findings. *The Elementary School Journal, 100*(5), 443-471. doi:10.1086/499650
- Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2). doi:10.18637/jss.v048.i02
- Rothbart, M. K., & Bates, J. E. (2007). Temperament. *Handbook of Child Psychology*. doi:10.1002/9780470147658.chpsy0303
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: John Wiley & Sons. <http://dx.doi.org/10.1002/9780470316696>
- Rutter, M. (1991). Age changes in depressive disorders: Some developmental considerations. *The Development of Emotion Regulation and Dysregulation, 273-300*. doi:10.1017/cbo9780511663963.013
- Salmela-Aro, K., Aunola, K., & Nurmi, J. (2008). Trajectories of depressive symptoms during emerging adulthood: Antecedents and consequences. *European Journal of Developmental Psychology, 5*(4), 439-465. doi:10.1080/17405620600867014
- Sanson, A., Hemphill, S. A., & Smart, D. (2004). Connections between Temperament and Social Development: A Review. *Social Development, 13*(1), 142-170. doi:10.1046/j.1467-9507.2004.00261.x
- Satorra, A., & Bentler, P. M. (2001). A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika, 66*(4), 507-514. doi:10.1007/bf02296192
- Schafer, J. L. (1999). *Analysis of incomplete multivariate data*. Boca Raton, FL: Chapman & Hall/CRC.
- Schulenberg, J. E., & Zarrett, N. R. (2006). Mental Health During Emerging Adulthood: Continuity and Discontinuity in Courses, Causes, and Functions. *Emerging Adults in America: Coming of Age in the 21st Century.*, 135-172. doi:10.1037/11381-006
- Schulte-Körne, G. (2016). Mental Health Problems in a School Setting in Children and Adolescents. *Deutsches Ärzteblatt Online*. doi:10.3238/arztebl.2016.0183

- Shochet, I. M., & Smith, C. L. (2014). A Prospective Study Investigating The Links Among Classroom Environment, School Connectedness, And Depressive Symptoms In Adolescents. *Psychology in the Schools*, 51(5), 480-492. doi:10.1002/pits.21759
- Shochet, I. M., Dadds, M. R., Ham, D., & Montague, R. (2006). School Connectedness Is an Underemphasized Parameter in Adolescent Mental Health: Results of a Community Prediction Study. *Journal of Clinical Child & Adolescent Psychology*, 35(2), 170-179. doi:10.1207/s15374424jccp3502_1
- Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological assessment*, 12(1), 102–111. <https://doi.org/10.1037//1040-3590.12.1.102>
- Smith, J. P., & Smith, G. C. (2010). Long-term economic costs of psychological problems during childhood. *Social Science & Medicine*, 71(1), 110-115. doi:10.1016/j.socscimed.2010.02.046
- Spijker, J., Graaf, R., Bijl, R. V., Beekman, A. T., Ormel, J., & Nolen, W. A. (2004). Functional disability and depression in the general population. Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *Acta Psychiatrica Scandinavica*, 110(3), 208-214. doi:10.1111/j.1600-0447.2004.00335.x
- Stewart, W. F., Ricci, J. A., & Chee, E. (2003). Cost of Lost Productive Work Time Among US Workers With Depression—Correction. *Jama*, 290(16), 2128. doi:10.1001/jama.290.16.2128
- Stoolmiller, M., Kim, H. K., & Capaldi, D. M. (2005). The Course of Depressive Symptoms in Men From Early Adolescence to Young Adulthood: Identifying Latent Trajectories and Early Predictors. *Journal of Abnormal Psychology*, 114(3), 331-345. doi:10.1037/0021-843x.114.3.331
- Torres, F. (2020). What Is Depression? Retrieved 2020, from <https://www.psychiatry.org/patients-families/depression/what-is-depression>
- Tucker, L. R. (1955). The objective definition of simple structure in linear factor analysis. *Psychometrika*, 20(3), 209-225. doi:10.1007/bf02289018
- Twenge, J. M., & Nolen-Hoeksema, S. (2002). Age, gender, race, socioeconomic status, and birth cohort difference on the children's depression inventory: A meta-analysis. *Journal of Abnormal Psychology*, 111(4), 578-588. doi:10.1037/0021-843x.111.4.578

- Üstün, B., & Kennedy, C. (2009). What is “functional impairment”? Disentangling disability from clinical significance. *World Psychiatry*, 8(2), 82-85. doi:10.1002/j.2051-5545.2009.tb00219.x
- Vaughn, M. G., Wexler, J., Beaver, K. M., Perron, B. E., Roberts, G., & Fu, Q. (2010). Psychiatric Correlates of Behavioral Indicators of School Disengagement in the United States. *Psychiatric Quarterly*, 82(3), 191-206. doi:10.1007/s11126-010-9160-0
- Veldman, K., Reijneveld, S. A., Ortiz, J. A., Verhulst, F. C., & Bültmann, U. (2015). Mental health trajectories from childhood to young adulthood affect the educational and employment status of young adults: Results from the TRAILS study. *Journal of Epidemiology and Community Health*, 69(6), 588-593. doi:10.1136/jech-2014-204421
- Vilagut, G., Forero, C. G., Barbaglia, G., & Alonso, J. (2016). Screening for Depression in the General Population with the Center for Epidemiologic Studies Depression (CES-D): A Systematic Review with Meta-Analysis. *Plos One*, 11(5). doi:10.1371/journal.pone.0155431
- Wallin, A. S., Koupil, I., Gustafsson, J., Zammit, S., Allebeck, P., & Falkstedt, D. (2018). Academic performance and depression: 26 000 adolescents followed into adulthood. *European Journal of Public Health*, 28(Suppl_4). doi:10.1093/eurpub/cky212.360
- Weeks, M., Ploubidis, G. B., Cairney, J., Wild, T. C., Naicker, K., & Colman, I. (2016). Developmental pathways linking childhood and adolescent internalizing, externalizing, academic competence, and adolescent depression. *Journal of Adolescence*, 51, 30-40. doi:10.1016/j.adolescence.2016.05.009
- Wickrama, K., Wickrama, T., & Lott, R. (2009). Heterogeneity in Youth Depressive Symptom Trajectories: Social Stratification and Implications for Young Adult Physical Health. *Journal of Adolescent Health*, 45(4), 335-343. doi:10.1016/j.jadohealth.2009.04.018
- Wickrama, T., & Wickrama, K. (2010). Heterogeneity in Adolescent Depressive Symptom Trajectories: Implications for Young Adults Risky Lifestyle. *Journal of Adolescent Health*, 47(4), 407-413. doi:10.1016/j.jadohealth.2010.02.013
- Wight, R. G., Aneshensel, C. S., Botticello, A. L., & Sepúlveda, J. E. (2005). A multilevel analysis of ethnic variation in depressive symptoms among adolescents in the United States. *Social Science & Medicine*, 60(9), 2073-2084. doi:10.1016/j.socscimed.2004.08.065

- Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin*, 116(2), 363-381. doi:10.1037//0033-2909.116.2.363
- Willner, C. J., Gatzke-Kopp, L. M., & Bray, B. C. (2016). The dynamics of internalizing and externalizing comorbidity across the early school years. *Development and Psychopathology*, 28(4pt1), 1033-1052. doi:10.1017/s0954579416000687
- Wilson, R. W., & Drury, T. F. (1984). Interpreting Trends in Illness and Disability: Health Statistics and Health Status. *Annual Review of Public Health*, 5(1), 83-106. doi:10.1146/annurev.pu.05.050184.000503
- World Health Organization. (2017). Retrieved from <https://www.who.int/en/news-room/fact-sheets/detail/depression>
- Wu, W., Jia, F., & Enders, C. (2015). A Comparison of Imputation Strategies for Ordinal Missing Data on Likert Scale Variables. *Multivariate Behavioral Research*, 50(5), 484-503. doi:10.1080/00273171.2015.1022644
- Xia, Y., & Yang, Y. (2016). Bias Introduced by Rounding in Multiple Imputation for Ordered Categorical Variables. *The American Statistician*, 70(4), 358-364. doi:10.1080/00031305.2016.1200486
- Yaroslavsky, I., Pettit, J. W., Lewinsohn, P. M., Seeley, J. R., & Roberts, R. E. (2013). Heterogeneous trajectories of depressive symptoms: Adolescent predictors and adult outcomes. *Journal of Affective Disorders*, 148(2-3), 391-399. doi:10.1016/j.jad.2012.06.028
- Yuan, K., Guarnaccia, C. A., & Hayslip, B. (2003). A Study Of The Distribution Of Sample Coefficient Alpha With The Hopkins Symptom Checklist: Bootstrap Versus Asymptotics. *Educational and Psychological Measurement*, 63(1), 5-23. doi:10.1177/0013164402239314
- Zhang, W., O'Brien, N., Forrest, J. I., Salters, K. A., Patterson, T. L., Montaner, J. S., Lima, V. D. (2012). Validating a Shortened Depression Scale (10 Item CES-D) among HIV-Positive People in British Columbia, Canada. *PLoS ONE*, 7(7). doi:10.1371/journal.pone.0040793
- Zhu, L. (2018, April 25). A Multi-Level Analysis on School Connectedness, Family Support, and Adolescent Depression: Evidence from the National Longitudinal Study of Adolescent Health, 1995–1996. Retrieved from <https://www.mdpi.com/2076-0760/7/5/72>