

LONGITUDINAL MODELING OF STUDENT SELF-REPORTED PURPOSE IN
URBAN MIDDLE SCHOOLS

By

DANIELLE RYAN HATCHIMONJI

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Psychology

Written under the direction of

Maurice J. Elias, Ph.D.

And approved by

New Brunswick, New Jersey

October 2019

ABSTRACT OF THE DISSERTATION

Longitudinal Modeling of Student Self-Reported Purpose in Urban Middle Schools

By DANIELLE RYAN HATCHIMONJI

Dissertation Director:

Maurice J. Elias, Ph.D.

Purpose is “a stable and generalized intention to accomplish something that is at once meaningful to the self and of consequence to the world beyond the self” (p. 121, Damon, Menon, & Bronk, 2003). Purpose incorporates three dimensions: Intention toward purpose, Engagement in purpose, and the Beyond-the-self quality of purpose. Purpose is considered an essential developmental asset and galvanizing force for adolescents of all backgrounds, but most of the literature on adolescent purpose development has investigated the construct in majority White and middle to upper income settings. This study sought to address gaps in the purpose literature by quantitatively modeling purpose among students in the setting of largely low-socioeconomic status, predominately racial/ethnic minority, urban middle schools. Using a Developmental Systems Theory framework, the study investigated the ability of relevant contextual factors, including race, gender, “failing” school status, and mental health, to predict the initial status and change in purpose.

Participants were middle school students ($n = 2629$) attending six low-resourced urban middle schools in the 2015-2016 and 2016-2017 school years. Hierarchical Linear Modeling was used to model purpose, and the three purpose dimensions, over time. Model building was guided by the procedures outlined by Singer and Willett (2003).

Results were unexpected in that purpose was found to decline over the course of the study. While many of the hypothesized predictors were related to the initial status of purpose, few predictors were related to change in purpose over time. The overall pattern of purpose decline and the lack of significant predictors of purpose change held for the three purpose dimensions as well.

Overall, the results of this study failed to provide clear answers to the research questions. Instead, results pointed to challenges with conceptual clarity and measurement of purpose, difficulty measuring purpose in middle school students, and the mismatch between Developmental Systems Theory and contemporary options for quantitative analysis. Understanding youth purpose development, particularly in the context of low-resourced urban schools, remains a critical priority. This study highlights the need to develop innovative procedures for statistically modeling the complex nature of youth development in context.

ACKNOWLEDGMENTS

I would like to thank my mentor, Maurice Elias, for his support in preparing the dissertation and in helping me fine tune my own purpose in life. I would also like to thank my committee for their thoughtful input and feedback, which has allowed the dissertation to become more clearly grounded in Developmental Systems Theory. I would also like to thank my friends and family who have dealt with a grueling five years, most notably the last several months. This dissertation would not have been possible without the material and emotional support from my husband, Justin Hatchimonji, and patience from our son, Evan! Finally, I would like to extend a note of gratitude to the Social-Emotional and Character Development Lab. My fellow labmates, Gwyne White, Cesalie Stepney, Arielle Linsky, and Sam Nayman have been integral to my development as a researcher and a person. The research coordinators (Esha Vaid, Sarah DeMarchena, Sarah Kim) and research assistants in the lab (too many to list!) have made my work not only possible but also enjoyable.

TABLE OF CONTENTS

Abstract of the Dissertation	ii
Acknowledgments.....	iv
List of Tables	vi
List of Figures.....	vii
Introduction.....	1
Methods.....	22
Results.....	28
Discussion.....	45
References.....	56
Tables.....	61
Figures.....	73
List of Appendices.....	79

LIST OF TABLES

Table 1. Purpose Scale	61
Table 2. Strengths and Difficulties Questionnaire.....	62
Table 3. Description of Hypothesized Predictors of Purpose Initial Status and Change..	63
Table 4. Sample Characteristics.....	64
Table 5. Percentage of Cohort with Complete Purpose Scale.....	65
Table 6. Correlations of Purpose Subscales.....	66
Table 7. Correlations of Purpose Total Scores	67
Table 8. Correlations of Purpose and SDQ Scores in SDQ Subsample.....	68
Table 9. Building Multilevel Model for Change in Student Self-Reported Purpose Scores.....	69
Table 10. Final Models Predicting Purpose Total Score and Purpose Dimensions Over Time.....	71
Table 11. Mental Health Predicting Purpose Total Score Over Time.....	72

LIST OF FIGURES

Figure 1. Mean Purpose Scores by Total Score and Subscales.....	73
Figure 2. Mean Purpose Scores by (a) Race/Ethnicity, (b) Gender, and (c) School.....	74
Figure 3. Randomly Selected 30 Cases Fitted with Regression Lines.....	75
Figure 4. Graphical Representations of Final Model Predicting Self-Reported Purpose Over Time.....	76
Figure 5. Individual Plots for Randomly Selected Cases after Analysis.....	77
Figure 6. Hierarchical Linear Model with Baseline High Purpose Raters as a Dichotomous Grouping Variable.....	78

Introduction

The study of purpose in adolescence is part of the larger positive youth development movement, which—in contrast to more traditional models of deficit-focused development—conceptualizes development in a strength-based framework, (Damon, Menon, Bronk, 2003; Lerner, Dowling, & Anderson, 2003; Seligman & Csikszentmihalyi, 2000). Purpose has been related to several positive aspects of mental health, including increased life satisfaction (Bronk, Hill, Lapsley, Talib, & Finch, 2009) and positive affect (Burrow, O'Dell, & Hill, 2010; Hill, Edmonds, Peterson, Luyckx, & Andrews, 2016) and experiencing less stress (Hill et al., 2016). While purpose has been studied across the lifespan, purpose is thought to be particularly salient for adolescents because of normative changes in brain structure and cognitive skills, including improved executive function, social cognition (Blakemore & Mills, 2014), and empathy (Hoffman, 2000). These emerging abilities to understand the emotional states of others and engage in meta-cognitive reasoning allow adolescents to consider and identify a life purpose.

Because of the benefits associated with purpose, there has been an increasing focus on understanding purpose development in adolescence; however, there has been limited longitudinal research on purpose to provide evidence of developmental pathways. Instead, the existing conclusions drawn about purpose development stem from cross-sectional qualitative or quantitative studies (e.g., Bronk et al., 2009; Hill, Sumner, & Burrow, 2014). The few longitudinal studies of purpose in adolescence have relied on interview data, with convenience-based and small samples (e.g., Bronk, 2012; Malin, Reilly, Quinn, & Moran, 2013; Quinn, 2016). While these qualitative studies have provided important theoretical contributions to the study of purpose development, there is

a need to test the theories of purpose development quantitatively so that larger samples can be used. Based on literature review to date, no study has quantitatively modeled purpose over time.

The existing literature on purpose in adolescence has not only been limited by its qualitative research methodology and small samples but it has also focused largely on mid- to late-adolescence, with an overrepresentation of White middle class individuals. This focus is particularly concerning because there has been little effort to theoretically or empirically account for the role of context in adolescent purpose development. The goal of this study is to address these concerns by quantitatively modeling the development of purpose over the course of middle school among students in the setting of largely low-socioeconomic status (SES), predominately racial/ethnic minority, urban middle schools. Better understanding the course of purpose in urban middle schools may point to suggestions for fostering purpose development for early adolescents in this setting.

Defining Purpose

Prior to Damon et al.'s (2003) seminal article on adolescent purpose development, purpose was inconsistently operationalized in the developmental and psychological literature. The terms “purpose” and “meaning” were vaguely defined and were often used interchangeably. Although some researchers defined purpose more specifically—relating it to identity, goal-directedness, or achievement—few researchers connected purpose directly to a beyond-the-self intention (Bronk, 2014; Damon et al., 2003). Since Damon et al. (2003)'s presentation of a precise and multidimensional definition of purpose, the construct has been used more consistently. Purpose is considered a “stable

and generalized intention to accomplish something that is at once meaningful to the self and of consequence to the world beyond the self” (p. 121, Damon et al., 2003).

The Damon et al. (2003) conceptualization of purpose includes three dimensions: intention, engagement, and a beyond-the-self orientation (Moran, 2009). The intention dimension of purpose describes a stable, future-oriented goal (Malin et al., 2013). The engagement dimension suggests that an individual must be meaningfully and actively working toward their intention (Malin et al., 2013). The beyond-the-self dimension is critical for two reasons. First, this dimension distinguishes purpose from other life goals. Second, this beyond-the-self dimension separates the Damon et al. (2003) definition of purpose from previously articulated definitions of purpose. According to the Damon et al. (2003) framework, a beyond-the-self goal is one that aims to contribute to something larger than the self. While a beyond-the-self orientation often takes the form of a prosocial purpose, these constructs are not necessarily synonymous (Quinn, 2016). For example, a prosocial purpose aims to benefit others; thus, serving others through a helping profession, such as a doctor or teacher, would be considered prosocial purpose. Yet, it is possible for a beyond-the-self purpose to both contribute and be of consequence to the world without the primary goal of directly benefiting others. Examples of such beyond-the-self purposes that would not be considered prosocial include making a theoretical contribution to science or creating a masterful work of art. Despite these distinctions, some purpose researchers use the terms “beyond-the-self” and “prosocial” interchangeably (e.g., Mariano, Going, Schrock, & Sweeting, 2011). For the current study, a “beyond-the-self” purpose is conceptualized in its broadest sense, including both

prosocial purposes and those purposes that contribute to larger society without necessarily directly benefiting individuals.

Another essential characteristic of the Damon et al. conceptualization of purpose is that the specific content of an individual's life purpose is of less consequence than the directionality created by the purpose. Accordingly, an individual's purposeful goal "may be material or nonmaterial, external or internal, reachable or nonreachable" (p. 121, Damon et al., 2003). For example, a purpose may be aspirational, meaning that it is "nonreachable," such as eradicating racism or poverty, or a purpose may be non-material, such as spreading joy or love to others. What makes a purpose powerful is that it serves as a galvanizing and motivating force that organizes an individual's energy in its service, regardless of the specific content of the goal (Damon, 2008). Damon et al. (2003) also distinguish between noble and ignoble purposes, noting that noble purposes serve humanity, whereas ignoble purposes aim to destroy. For this study and consistent with previous research, purpose is viewed as positive, or "noble," in nature, making the assumption that purpose is necessarily connected to seeking to serve and contribute, rather than to destroy.

The Damon et al. (2003) definition of purpose differentiates purpose from similar constructs, such as identity or meaning. Whereas identity answers the question, "*who* am I?", purpose answers the questions "*why* am I?" (Yeager & Bundick, 2009) and "*what* will I accomplish?" (Bronk, 2012). These questions about purpose are connected to a long-term life goal, versus a momentary or short-term focus. While the construct of meaning can also answer the "why" question, unlike purpose, meaning is an internal

experience that may provide a reason for existence but is not necessarily directed toward a goal (Damon et al., 2003).

Theoretical Underpinnings of Purpose Development

Victor Frankl (1959) brought the concepts of meaning and purpose (terms he used interchangeably) to the attention of the psychology field in his *Man's Search for Meaning*. Frankl, a survivor of several years in a Nazi concentration camp, saw purpose as something that could defend against the effects of inhumanity and cruelty and provide a healing power (Damon et al., 2003). In describing his school of psychotherapy, which he called logotherapy, Frankl further asserted that the search for meaning was the primary motivating force in a person's life; in the concentration camp, he had observed that loss of purpose was a form of death no less cruel than physical torture. Frankl's work laid the groundwork for contemporary study of purpose.

Frankl's focus on purpose highlighted the importance of purpose in the face of adverse life experiences. From the contemporary perspective of the positive youth development and positive psychology movements, purpose is also viewed as a galvanizing force of positive development and one of many important developmental assets (Benson et al., 2006; Damon et al., 2003). Thus, in contrast to Frankl's original focus, contemporary literature suggests that purpose development need not rely on reacting to a negative life experience. In fact, Hill et al. (2014) examined three different pathways of deliberation toward identifying a purpose and found that all three pathways (proactive, reactive, or social learning) were associated with positive well-being in emerging adults ($n = 179$; 77% female, 66% White) and adults ($n = 307$, 65% female, 75% White). Their findings suggest that identifying a purpose in life has positive

benefits, regardless of an individual's reasons for developing purpose or the pathway taken to determine this purpose.

Developmental Systems Theory

Developmental systems theory, the basis of the field of positive youth development, examines human development by considering the bidirectional developmental processes transacting between individuals and their contexts (Lerner et al., 2003). Developmental systems theory was largely influenced by Bronfenbrenner's (1979) ecological theory of development, which posits that human development involves mutual accommodation between a human being and the immediate and larger settings in which the individual is embedded. Contemporary conceptualizations of developmental systems theory emphasize the ongoing and dynamic impact of relational processes and contextual co-influences on development (Osher, Cantor, Berg, Steyer, & Rose, 2018; Overton, 2015).

Malin et al. (2013) used developmental systems theory to investigate how youth purpose changed over time and across life contexts. The authors analyzed interview data from a sample of 146 adolescents from four different age groups (middle school to college). The adolescents were interviewed twice over a two-year interval. Interview data were coded to identify the three dimensions of purpose: intention, engagement, and beyond-the-self orientation. Adolescents were considered to exhibit purpose if they described all three purpose dimensions to a high degree. If they described only one or two purpose dimensions, they were considered to exhibit a "precursor of purpose." Precursors of purpose included 1) dreaming (having a strong future intention without engagement or acting toward it), 2) dabbling (engaging in a beyond-the-self-oriented

activity without a clear future intention), and 3) having a self-oriented goal (engaged in working toward a future intention without beyond-the-self reasons for pursuing the goal). Adolescents with no evidence of any of the three purpose dimensions were not considered to be purposeful.

Despite coding interviews from 146 participants, only a small number of adolescents were found to exhibit purpose within each age cohort. Most relevant to the current study are the findings from the early adolescent sample. These students were in 6th grade at the first interview and in 8th grade at the second interview ($n = 46$, 57% female; 30% Asian, 24% White, 15% Mixed race, 13% Hispanic, 9% Filipino). Of the 46 early adolescents interviewed, only eight individuals were purposeful at the first timepoint and only two of these individuals maintained purpose over two years. Of those who were no longer purposeful, five out of six had changed their goal to be self-oriented. Three participants in the early adolescent sample were not purposeful in 6th grade but became purposeful over the two-year period; they did so by becoming more specific about how to act out their purpose.

From their coding of the interview data, Malin and colleagues suggested that purpose in adolescence develops through 1) orienting toward empathy, 2) envisioning a societal role, 3) reevaluating values and priorities, and 4) developing a pathway toward actualizing the societal role. Although the small number of students who were considered purposeful makes it difficult to interpret the authors' conclusions about purpose development, this study is useful for its framing of purpose development within the developmental systems approach. Rather than developing from precursors, Malin et

al. (2013) suggest that purpose develops according to one's life-stage and is influenced by life experiences, context, and individual factors.

Identity Development Model

Although he did not use the term purpose consistently, Erikson (1968) highlighted purpose as a means to resolve the primary developmental task of adolescence: developing a sense of identity. He writes that adolescents, in their search for identity, look for the “opportunity to decide [...] on one of the available or unavoidable avenues of duty and service” (p.129, Erikson, 1968). Because of this link between purpose and identity, purpose development has been conceptualized within an identity development framework. Marcia's (1966) identity development model delineates four identity statuses (achieved, moratorium, foreclosed, diffused) that are related to high or low levels of identity exploration and commitment. This model suggests that individuals undergo distinct exploration and commitment processes while forging an identity.

Burrow et al. (2010) proposed that the pathway toward establishing purpose can be explained by a developmental process that is similar to Marcia's identity development model. In a cross-sectional sample of 318 high school adolescents from three midwestern suburban high schools (55% female, 76.3% White, non-Hispanic), Burrow et al. (2010) found empirical support for a four-cluster model that parallels Marcia's identity model. They used the Bundick et al. (2006) 20-item *Youth Purpose Scale*, which includes “commitment” and “exploration” subscales, to measure purpose. They found the following four categories of purpose development in their sample: Foreclosed (high commitment, low search), Diffused (low commitment, low search), Moratorium/Uncommitted (low commitments, high search), and Achieved (high

commitment, high search). The authors found no statistically significant differences with respect to gender, year in school, or ethnicity between the unique purpose profiles. These findings suggest that purpose development may follow a process that is similar to identity development, meaning that searching for and committing to a purpose are distinct pathways.

Blattner, Liang, Lund, and Spencer (2013) also found support for the same four-cluster model in a sample of 207 adolescent females attending selective private schools in affluent suburbs (85% White). Unlike Burrow et al. (2010), Blattner et al. (2013) found significant differences among sixth, eighth, and tenth graders for both search for purpose, and commitment to purpose (using the same self-report scale). Tenth grade students had lower levels of commitment to purpose and higher levels of search for purpose than their sixth- and eighth-grade counterparts. Although these results were cross-sectional, they suggest that searching for purpose may increase while commitment to purpose may decrease over the course of early adolescence.

The work of Blattner et al. (2013) and Burrow et al. (2010) suggests that commitment to and search for purpose are distinct processes. Using the same self-report scale, these cross-sectional studies suggested different trajectories for the development of purpose over time, with Burrow et al. (2010) identifying no differences across grade levels but Blattner et al. (2013) suggesting increasing search and decreasing commitment from 6th to 10th grade. Of note, studies relied on purpose subscales representing “search” and “commitment,” rather than investigating the three dimensions of purpose (intention, engagement, and beyond-the-self), so it is not clear how the commitment and search processes align with these three dimensions. The different developmental patterns found

in these studies could be attributable to contextual factors within the sampled schools and surrounding area or to the different developmental periods examined.

Prosocial Development

Prosocial behavior is linked to the beyond-the-self dimension of purpose (Quinn, 2016) and shares several qualities with purpose. Both purpose and prosocial behavior are thought to be guided by the development of sociocognitive skills in adolescence that allow for increased perspective taking, empathy, and sympathy (Eisenberg, Carlo, Murphy, & van Court, 1995; Hoffman, 2000). Given the limited longitudinal research on adolescent purpose development, it is useful to consider literature on prosocial development.

While some research points to prosociality increasing over adolescence (e.g., Eisenberg et al., 1995), Kanacri, Pastorelli, Zuffianò, and Eisenberg (2013) found that Italian adolescents ($n = 573$) decreased in their overall level of prosociality from age 13 to 17. These conflicting findings could be explained by developmental systems theory, in that these outcomes could be the result of contextual differences. A study by Padilla-Walker, Dyer, Yorgason, Fraser, and Coyne (2015) underscores the necessity of considering the context of prosocial behavior. The authors found that prosocial behavior toward family decreased slightly or stabilized over adolescence, whereas prosocial behavior toward friends increased. Thus, the literature on prosocial development reinforces the tenets of developmental systems theory: that purpose development in adolescence must take into account an individual's context.

Further support for the developmental systems theory perspective comes from the debate over the competing roles of individual personality factors and context-specific

situational factors in prosocial development. For instance, specific personality traits have been related to emotional reactions to moral decisions in adolescence (Krettenauer, Colasante, Buchman, & Malti, 2014; Malti & Buchmann, 2010). In addition, dispositional sympathy has been positively associated with prosocial behavior (Eisenberg, Fabes, & Spinrad, 2006; Padilla-Walker et al., 2015). Contextual variables, including cultural and family factors, have also been implicated in the development of prosocial behavior. Perspective taking has been found to partially mediate the relationship between familism values and prosocial tendencies in a sample of Mexican American early adolescents (Knight, Carlo, Basilio, & Jacobson, 2015). Further, maternal warmth has been associated with higher levels of prosocial behavior, particularly in adolescence (Padilla-Walker et al., 2015). In sum, research on prosocial development, like the limited research on purpose development, suggests that development through adolescence is not straightforward and likely depends on how specific settings and contexts interact with the individual dispositions of an adolescent.

Moral Development

Prosocial development is part of a larger literature on moral development; thus, the extensive literature on moral development may also contribute to a theory of purpose development. Theories of moral development explain why and how individuals develop the capacity to make moral decisions and act in moral ways. Kohlberg's (1984) stage model of moral development has been foundational in understanding the development of moral reasoning, but it has also long been recognized as an overgeneralization of a white male perspective on moral reasoning. Carol Gilligan (1982) noted that females tend to consider relationships and caring as the most important aspects of moral decision-

making, in contrast to the Kohlbergian male emphasis on justice. Gilligan explained that ignoring these distinct perspectives has led to erroneously relegating women to a “lower” developmental level when compared to males. Her work emphasizes the importance of examining implicit cultural assumptions embedded within developmental theory, and particularly in the development of moral reasoning.

More recent work also suggests that Kohlberg’s framework oversimplified moral development by not adequately considering how moral development occurs across multiple domains and contexts. In his *Education in the Moral Domain*, Larry Nucci (2001) argues that moral reasoning cannot be considered one overarching developmental system. Instead, he explains that individuals reason differently about right and wrong within three separate domains: morality, convention, and personal choice. Morality is independent of social norms and is thought to be consistent across contexts, whereas convention describes rules of a social system. The personal choice domain includes aspects of moral decisions that vary according to an individual’s disposition or experiences.

According to this social domain theory, individuals will develop these three systems at different rates; specifically, the theory predicts asynchrony between these three domains throughout childhood and adolescence. In the realm of purpose development, an asynchrony across domains could be related to the development of the three dimensions of purpose. For instance, adolescents might develop the engagement, intention, and beyond-the-self aspects of purpose at rates that differ by individual, domain, and experience. An inconsistent development of these different domains of purpose is in line with developmental systems theory and corroborates the findings of

Malin et al. (2013) that adolescents gained and lost specific dimensions of purpose at different life stages.

Correlates of Purpose

Purpose in adolescence has been related to several indicators of well-being and may also buffer against the developmentally normative risks of adolescence. For adolescents, identifying a purpose has been associated with increased life satisfaction (Bronk et al., 2009), higher levels of positive affect (Burrow et al., 2010), and higher levels of hope (Bronk et al., 2009). Identifying a purposeful career aspiration has been related to experiencing more meaning in life and in schoolwork, suggesting that identifying a purpose may buffer against disengagement from school in middle and high school (Yeager & Bundick, 2009). Further, adolescents' purpose in life has been found to mediate the relationship between higher levels of crystallized intelligence and lower levels of drug use (Minehan, Newcomb, & Galaif, 2000).

There appears to be a specific benefit to identifying a beyond-the-self purpose or career goal. Research on prosocial career goals shows that college students who endorsed a stronger desire to serve others in their future career were more optimistic about their career than those with a less of a desire to serve others (Duffy & Raque-Bogdan, 2010). Yeager, Bundick, and Johnson (2012) found that adolescents (6th, 9th, 12th grade students; $n = 99$; 60% female; 6% African American, 30% Asian American, 23% Latino, 41% White; 46% middle income, 20% high income, 3% low income) who held both self-oriented and beyond-the-self-oriented motives for career goals were more likely to experience higher levels of purpose and meaning over a two-year period than those who held neither type of motive.

Prevalence of Purpose in Adolescence

While purpose is considered to be a developmental task of adolescence, identifying a purpose appears to be a somewhat rare occurrence, particularly for early adolescents. In a sample of 64 high ability youth (students attending schools that required above average intelligence scores for admission), Bronk, Finch, and Talib (2010) found that 9% of the early adolescents and 34% of the late adolescents exhibited purpose (56% Female, 81% White). In a retrospective analysis of data from 1938-1942, Mariano and Vaillant (2012) found similar levels of beyond-the-self intention in college males (100% White): about 38% of cases endorsed a prosocial beyond-the-self intention ($n = 53$). Yeager and Bundick (2009) found that 30% of 148 adolescents had identified a purposeful work goal, which they defined as an occupational goal that aims to contribute to the world beyond the self (58% female; 4% African American, 30% Asian American, 25% Latino, 23% White). In the sample of 146 adolescents and emerging adults studied by Malin et al. (2013), only about 40% of youth were able to maintain or grow in their purpose development over a two-year period (32% White, 21% Asian American, 16% Hispanic; 11% Mixed race, 11% Other; 52% female). As reported above, only two of the 46 early adolescents in this study maintained purpose over two years.

The beyond-the-self dimension of purpose appears to be particularly rare for adolescents. Pointing to the Malin et al. (2013) finding that youth were less likely to name the beyond-the-self dimension of purpose than the engagement or intention dimensions, Quinn (2016) suggests that the beyond-the-self intention may not be a typical milestone of purpose development in adolescence. In her recoding of the interview data from the Malin et al. (2013) study, Quinn (2016) found that two

dimensions of purpose—engagement and intention—increased and stabilized over the course of adolescence and emerging adulthood. However, the beyond-the-self intention was a rare experience and did not clearly increase or stabilize with age. Only 12.3% of the sample gained a beyond-the-self orientation over the two years of the study, a pattern that did not vary across the middle school, high school, or college samples.

In sum, fully realized purpose in early adolescence appears to be a rare phenomenon. However, much of the research on the prevalence of purpose in adolescence stems from coding of interview data, meaning that participants must meet specific criteria to be considered purposeful. In these studies, participants are either purposeful or not purposeful—there is no gradation of identifying a purpose. On the other hand, studies that have used self-report scales as indicators of purpose allow for a continuum of being purposeful, but these studies have not attempted to categorize individuals as purposeful or not purposeful, making it difficult to report the prevalence of purpose. More work is needed to bring together the qualitative and quantitative strands of purpose research by measuring not only overall purpose with self-report scales but also investigating the three dimensions of purpose: intention, engagement, and beyond-the-self intention.

Contextual Considerations: The Urban Middle School Environment

Much of the extant research on youth purpose development has put forward broad conclusions about purpose development without theoretically or empirically accounting for the context of development. This context-neutral approach to understanding purpose development is not in line with developmental systems theory, which emphasizes the constant reciprocal interactions between individuals and environment (Osher et al., 2018;

Overton, 2015). More recently, purpose researchers have begun to examine purpose in more diverse settings and have started to explore the role context plays in shaping purpose.

The context of the urban middle school involves the intersecting impact of racial/ethnic background, poverty, and low-resourced and stressed school systems. Purpose may be particularly important for youth growing up in this setting because purpose may be able to mitigate the negative effects of poverty, discrimination, and exposure to violence and trauma (Hatchimonji, Linsky, & Elias, 2017).

Urban poverty. In an interview study of youth of color living in impoverished urban communities, Gutowski, White, Liang, Diamonti, and Berado (2017) found that the psychological stress experienced by these youth served as a barrier to purpose development because (1) youth believed it was impossible to achieve their goals and (2) youth were so overwhelmed that purpose was not a meaningful priority. However, the authors also found that psychological stress could serve as a purpose motivator in this setting because some youth felt pressure to succeed and a desire to escape from their difficult circumstances. An interview investigation of youth in Guatemala reinforced that poverty and discrimination experiences can be barriers to purpose development (Liang et al., 2017b). In addition to negative experiences of stigmatization and discrimination, the youth in this study had less access to opportunities to engage in activities or organizations that might foster the development of a life purpose. Positive childhood experiences and opportunities to engage in potentially purposeful activities, including extracurricular activities are important elements of purpose development (Bronk, 2014; Ishida & Okada, 2006).

Racial/ethnic background. There is limited research to clarify how youth from various racial and ethnic groups experience purpose. Most pertinent to the context of the current study is the intersection of marginalization based on racial/ethnic background, the experience of poverty, and the urban school setting. Racism affects the relationship between students of color and the education system at every level of the developmental system. Students of color experience disproportionate rates of discipline referrals, disparities in academic outcomes, and disparities in opportunities for learning and enrichment (Osher et al., 2018). The research on youth purpose in the context of poverty suggests that the psychological stressors based on racial/ethnic background would serve as barriers to purpose development for some students but motivation for others (Gutowski et al., 2017). Mariano et al. (2011) coded interviews of 46 African American middle school girls for forms of purpose that included drifter, dabbler, dreamer, self-life goal, and beyond-the-self purpose, a coding process that was similar to that used by Malin et al. (2013). They found that the forms of purpose they identified were similar to those found in other samples of students from diverse racial backgrounds. However, Mariano et al. found that the relationship of social support to purpose was unexpected: girls who reported more complete forms of purpose indicated that they experienced less social support in school when compared to those who described more diffuse forms of purpose. It may be that identifying a purpose in early adolescence is rare enough that the experience of working toward the purpose is isolating. An alternative perspective is that those with more complete purposes required less social support from their school environment. It may also be that school environments are not supportive of purposes

identified by early adolescents or the school may be unsupportive of purpose for African American girls.

Low-resourced urban school environment. Urban schools are plagued by lack of resources and high levels of need among students living in poverty. High rates of staff turnover in these schools and high rates of student mobility contribute to lower school climates and less sense of community and belonging (Esposito, 1999; Holme, Jabbar, Germain, & Dinning, 2018). In addition, many schools in impoverished urban communities perform poorly on standardized assessments and are labeled as “failing.” These schools are then pressured to increase test scores, leaving even fewer of the psychological and financial resources of the staff available for supporting positive youth development (Finnigan & Daly, 2012). The lack of social connections and community in the low-resourced urban school setting is troubling given that social support has been indicated as one of the most important conditions for youth purpose development (Bronk, 2014; Liang et al., 2017a; Liang et al., 2017b). Youth attending low-resourced urban schools may not experience adequate social support in the school environment to overcome the barriers to purpose development presented by poverty and marginalization based on racial/ethnic background.

An additional barrier to purpose development in the low-resourced urban school setting may be low expectations for future achievement held by students and staff alike (Ou & Reynolds, 2008). Opportunities to discuss and articulate purpose are critical to its development (Bundick, 2011; Dik et al., 2011; Pizzolato, Brown, & Kanny, 2011), yet those discussions and opportunities for self-reflection may be less likely to occur in

stressed school environments in which students and teachers hold lower expectations for student academic and vocational outcomes.

Academic achievement. Given the multitude of stressors in the low-resourced urban school environment, succeeding in this context requires significant resilience. Thus, students who perform well in this setting may be exceptionally resilient and may also demonstrate fuller forms of purpose. However, the directionality of the relationship between purpose and academic achievement is not clear. Purpose may serve as a motivator for academic achievement (Benson, 2008; Pizzolato et al., 2011), but purpose development may also be supported by academic successes. Performing well academically offers a host of benefits that support purpose development, such as enrichment opportunities, higher self-expectations, and higher expectations for success from teachers and parents.

Gender. Damon et al. (2003) summarize that early research on purpose found inconsistencies in whether males and females experienced different levels of purpose. However, there is reason to believe that middle school age females would experience higher levels of purpose than males: studies consistently find females report higher levels of developmental assets than males (Benson et al., 2006). In addition, Van der Graaf et al. (2014) found that adolescent girls showed steeper increases in perspective taking and higher levels of empathic concern than boys, skills that are theorized to be necessary for developing purpose. In the context of low-resourced urban middle schools, females are less likely to experience discipline referrals and tend to demonstrate higher academic achievement than males (Voyer & Voyer, 2014; Wallace, Goodkind, Wallace, & Bachman, 2008).

Mental Health. Much of the positive youth development literature supports the idea that mental wellness is critical to purpose development (Benson et al., 2006). Students in low-resourced urban schools are more likely to have been exposed to trauma and violence and to experience related emotional distress (Sacks & Murphey, 2018; Thompson et al., 2007). Psychological stress associated with poverty, marginalization, and discrimination likely impedes purpose development (Gutowski et al., 2017; Liang et al., 2017b). Purpose has also been related specifically to mental health indicators, including decreased risky behavior and antisocial behavior, and more positive mood (King, Hicks, Krull, & Del Gaiso, 2006; Machell, Disabato, & Kashdan, 2016; Sayles, 1995). Sayles (1995) found a negative association between purpose and risky behaviors in a sample of high school adolescents, a relationship that was consistent across gender and ethnicity. In a daily diary study, experiences of meaning and purpose were associated with increased positive mood (King et al., 2006). In a study of youth living in poverty, Machell et al. (2016) found that purpose in life mitigated the effects of poverty on antisocial behaviors.

The Current Study

Research on purpose in adolescence is now entering its second decade. It is clear that purpose is associated with several indicators of well-being. However, little is known empirically about how purpose develops, largely because the limited longitudinal research on purpose development has relied on interview data and small samples. Quantitative research on purpose development has been limited to cross-sectional studies using different cohorts to draw conclusions about development. While this research has been critical to informing a theory of purpose development, to move the field forward, it

is important to test the theory of purpose development with quantitative longitudinal models. The current study aims to address these gaps in the purpose literature by quantitatively modeling purpose over the course of middle school. Further, developmental systems theory suggests that individual and contextual factors must be considered when modeling developmental assets; thus, this study aims to investigate individual demographic characteristics and contextual factors as predictors of purpose development.

Purpose is considered an essential developmental asset and galvanizing force for adolescents of all backgrounds, but most of the literature on adolescent purpose development has investigated the construct in majority White and middle to upper income settings. This study explores the development of purpose in low-income, racially diverse, low-resourced urban middle schools to better understand purpose development in this context. Purpose may be particularly important to promoting resilience for youth in this setting, but these students may not have access to opportunities for purpose development in comparison to their more affluent peers (Bronk, 2014; Liang et al., 2017b). Thus, understanding purpose development in the context of low-resourced urban middle schools may offer important insights for cultivating youth purpose in this setting.

Hypotheses

Study hypotheses were informed by the review of the literature outlined above.

Hypothesis 1: Overall purpose is expected to increase over the course of the study.

Hypothesis 2: The “Intention” and “Engagement” dimensions of purpose are expected to increase over the course of the study. However, the “Beyond-the-self”

dimension of purpose, which is expected to be unstable over middle school, is not expected to clearly increase or decrease.

Hypothesis 3: In line with Developmental Systems Theory, individual demographic characteristics and school-level and individual-level contextual factors are expected to explain the initial status and trajectory of self-reported purpose. (Table 3 includes hypotheses.)

Hypothesis 4: Higher levels of baseline mental health symptoms are expected to predict lower levels of initial purpose and a negative slope in students' purpose trajectories.

Method

Participants

Participants were 6th, 7th, and 8th grade students attending six public urban middle schools in the 2015-2016 and 2016-2017 school years. The total number of students attending these six schools each school year was approximately 2100. Sample characteristics and the sample definition procedure are described in detail below. Survey return rates from the six schools varied widely over the course of the study, ranging from 0% to 87%. In the 2015-2016 school year, the average student self-report survey return rates were 75% (Fall 2015 Day 1), 64% (Fall 2015 Day 2), and 63% (Spring 2016). One school had a 0% return rate in Spring 2016. In the 2016-2017 school year, average return rates across schools were 64% (Fall 2016) and 72% (Spring 2017).

Measures

Purpose. Student purpose was measured by a 5-item self-report scale that was created for this study from two existing purpose scales for adolescents (Table 1). Six

items were initially selected based on face validity and readability in order to capture the three dimensions of purpose with the least overlap in items. Reliability analyses after the baseline data collection resulted in the removal of one reverse-scored item (Cronbach's $\alpha = .80$ with item removed). The resulting 5-item scale used in this study includes two items from the Lippman et al. (2014) *Purpose Scale*. These items appear to measure the "Beyond-the-self intention" ("My life will make a difference in the world") and "Engagement" ("I am doing things now that will help me to achieve my purpose in life") dimensions of purpose. Three items come from the *Revised Youth Purpose Survey* (Bundick et al., 2008). These items appear to capture the "Intention" ("My life has a clear sense of purpose" and "I have a purpose in my life that says a lot about who I am") and "Engagement" ("I am always working toward accomplishing my most important goals in life") dimensions. All five items are rated on a 5-point Likert scale ranging from "Disagree A LOT!" to "Agree A LOT!" Higher scores indicate higher levels of self-reported purpose. Cronbach's alphas for the current study indicated adequate reliability (range .81 to .88).

Mental Health. Mental health was measured by an adaptation of the well-validated *Strengths and Difficulties Questionnaire* (SDQ; Goodman, Meltzer, & Bailey, 1998). Factor analysis, correlations, and participant feedback led to reducing the original 25-item scale for time points beyond the baseline assessment (Table 2). The resulting scale used in this study includes selected items from three of the original five subscales: emotional problems (4 items), conduct problems (3 items), and peer problems (4 items). Items are rated on a 3-point scale, with the options: "Not True," "Somewhat True," and "Certainly True." Higher scores indicate higher levels of self-reported symptoms and

thus poorer mental health. Only one item is reverse scored (“I have one good friend or more”). The Cronbach’s alpha for the SDQ (11-item scale) in the current study was .76.

Potential predictors of purpose trajectories. Table 3 summarizes the contextual variables examined in this study as potential predictors of purpose initial status and change over time. Student demographics, including gender, race/ethnicity, grade level, special education accommodations, and lunch status were provided by the district. Two indicators of implementation were used to control for the effects of an ongoing intervention: Ambassador Status (whether an individual student participated in the MOSAIC student leadership program, known as the Ambassador Program) and a school-level indicator of high or low MOSAIC implementation. Student report card grades were also provided by the district. Academic report card grades were used to create an indicator of student achievement history by averaging all core subject grades reported by the district over the two years of the study.

Procedure

The study was part of a larger research project to implement, refine, and test a school-based social-emotional and character development intervention (Mastering Our Skills And Inspiring Character; MOSAIC). Six middle schools were recruited to take part in this study for the 2015-2016 and 2016-2017 school years. Schools were selected for their demographic diversity, representing the diverse array of public schools within the larger school district. The intervention included an additional leadership program for 1-2 students per MOSAIC classroom, called the Ambassador Program, in which students were elected by classmates to lead classroom discussions and represent their class in

conversations with school administrators about potential solutions to school problems. In the current sample, 247 students participated in the Ambassador Program.

At the beginning of each school year, students were consented to study participation through a passive consent process approved by the school district and the research institution's Institutional Review Board. Students were also provided an opportunity to decline participation through a passive assent process at the time of the survey. Students in sixth to eighth grade in the participating schools were asked to complete self-report surveys in the fall and spring of each school year. In 2015-2016 most students took surveys through an online survey platform. Due to inconsistent return rates across schools because of computer access difficulties, some students also took the 2015-2016 surveys by paper. In 2016-2017, all students completed the survey using scantrons and a paper survey. In total, the students were surveyed at four time points: Fall 2015, Spring 2016, Fall 2016, and Spring 2017.

Data Analysis Plan

Data were examined for missing data, normality, and outliers. Descriptive statistics and correlation analyses were examined for purpose (full scale and three dimensions) and mental health for the overall sample as well as by cohort, gender, race/ethnicity, and lunch status (as a proxy for socioeconomic status). Preliminary data analyses also investigated descriptive characteristics of the data by school to attempt to uncover any lack of equivalency across the schools included in the study. Data were visually examined to inform model specification, particularly to determine whether a linear, quadratic, or other function would best represent the data. Hierarchical Linear

Modeling (HLM) was used to test all study hypotheses using the HLM 7 software program (Raudenbush, Bryk, & Congdon, 2011).

HLM was selected to model purpose over time for several reasons. First, HLM is considered to be superior to traditional methods of analyzing change over time (such as multivariate repeated measures) because HLM is much more flexible in terms of its underlying assumptions (Raudenbush & Bryk, 2002). HLM can handle unequally spaced data, complex nonlinear trajectories, and multivariate growth processes (Curran, Obeidat, & Losardo, 2010). The hierarchical linear modeling approach to analyzing change over time uses time (Level 1) nested within persons (Level 2). For the current study, each hypothesis was tested using the models below:

Hypothesis 1: Modeling student self-reported purpose (total score) over time. The hypothesis that purpose would increase over the course of the study was tested using the unconditional growth model, with time as the only Level 1 variable. This model is represented by the following equations:

$$\text{Level 1: } PURPOSE = \pi_{0i} + \pi_{1i}(TIME)_{it} + e_{it}$$

$$\text{Level 2: } \pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

$$\text{Combined model: } PURPOSE_i = \beta_{00} + \beta_{10} * TIME_{ti} + r_{0i} + e_{ti}.$$

Hypothesis 2: Modeling dimensions of purpose over time. Three additional unconditional growth models, analogous to the equation tested in Hypothesis 1, were tested with each of the three dimensions of purpose (Intention, Engagement, Beyond-the-self) as the dependent variable.

Hypothesis 3: Effects of individual demographic characteristics and school-level and individual-level contextual factor on purpose over time. The contextual variables and covariates were added to Level 2 of the model predicting purpose to determine the influence of these factors on the initial status and trajectory of self-reported purpose. For example, adding gender to the Level-2 equation resulted in the following set of equations:

$$\text{Level 1: } PURPOSE_{it} = \pi_{0i} + \pi_{1i}(TIME)_{it} + e_{it}$$

$$\text{Level 2: } \pi_{0i} = \beta_{00} + \beta_{01}(GENDER_i) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11}(GENDER_i) + r_{1i}$$

$$\begin{aligned} \text{Combined model: } PURPOSE_{it} = & \beta_{00} + \beta_{01}(GENDER_i) + \beta_{10}(TIME_{it}) + \\ & \beta_{11}(GENDER_i * TIME_{it}) + r_{0i} + r_{1i}(TIME) + e_{it} \end{aligned}$$

Hypothesis 4: Effects of baseline mental health symptoms on purpose initial status and change over time. Baseline mental health status was a variable of particular interest and required creating a separate subsample. The models used to test the association of mental health symptoms with purpose initial status and change mirrored the equations outlined above. Mental health symptoms in Fall 2015 were added as a Level 2 predictor of purpose initial status and change.

Because there were only six schools represented in the study, there was not enough power to detect school-level effects. Thus, nesting by school (as a Level-3 variable) was not examined. Instead, school-level variables were included as person-level (Level-2) predictors. Model adequacy was examined using procedures delineated by Singer and Willett (2003). Statistical tests of model fit included the chi-square difference test and examination of goodness-of-fit statistics: deviance, Akaike

information criterion (AIC), and Bayesian information criterion (BIC). The deviance statistic can only be used to compare for models that are nested (the same outcomes and the same data are used with only specific predictors differing between the models). The chi-square difference test determines whether a model has made a significant improvement in the deviance statistic. The AIC and BIC statistics allow for comparison between non-nested models as long as the same data are used. These statistics include penalties for adding predictors so that parsimony is favored. The formula for the BIC places a larger penalty for complexity and is sometimes preferred to the AIC (Hox, Moerbeek, & van de Schoot, 2018). There is no clear criterion for how much change in these statistics signifies a better fit, which is why it is important to look at multiple methods for comparing models. Singer and Willett (2003) cite Rafferty (1995) as suggesting that a change in the BIC should follow these guidelines: a difference of 0-2 is weak, 2-6 is positive, 6-10 is strong, and over 10 is very strong. The pseudo- R^2 statistics for variance components, described by Singer and Willett (2003), were also calculated and examined to interpret model fit.

Results

Sample Definition

Over the 2015-2016 and 2016-2017 school years, 3623 sixth to eighth grade students were registered in the six schools involved in this study, according to the demographic information provided by the district. The district provided at least one academic course grade for 2975 of these students. Students were included in the sample if they had completed at least one self-report assessment (defined by completing at least 75% of the items on the purpose scale at any one timepoint) and had received at least one

academic grade (reading, writing, math, social studies, or science) in either school year. 183 students were removed from the sample because they were part of very small racial groups ($n = 36$), had incomplete demographic information reported by district ($n = 90$), moved schools between school years ($n = 44$), or had an atypical grade level pattern (same grade level for two years, skipped a grade, or no grade level reported; $n = 13$). In each case, these exclusions were intended to enhance interpretability of results, particularly for comparisons by racial group and by school. The characteristics of the resulting sample ($n = 2629$) are described in Table 4.

The sample was largely representative of the six schools in terms of the racial make-up, lunch status, and students with disabilities. An entire cohort of students from school 6 was lost due to inconsistencies in district academic grade reporting that disproportionately affected this group. Students who were not included in the sample were more likely to have been from school 5 or 6 ($\chi^2(5) = 287.92, p < .001$), to have qualified for full price lunch ($\chi^2(1) = 21.57, p < .001$), and to have received accommodations, as noted by having an IEP or 504 plan in place ($\chi^2(1) = 109.05, p < .001$). Students who were registered as American Indian, Pacific Islander, and Multiracial had been systematically removed from the sample. In addition to being more likely to come from those groups, students excluded from the sample were less likely to be registered as Asian, White, or Latino than those who were included and were more likely to be registered as Black ($\chi^2(6) = 184.13, p < .001$). Sample inclusion was unrelated to LEP status and gender. For the few students who were excluded from the sample for whom self-report data was available, there were no differences on purpose or mental health scores between the excluded students and those in the sample.

Missing Data

Missingness at item level. Examination of missingness on the individual items of the purpose scale (before sample definition) revealed that the most frequent missing data pattern was for a student to miss one of the five items. Examination of missingness at the item levels on the SDQ scale revealed a similar pattern. Once the sample was defined (students who had at least 75% of the purpose scale at one timepoint), missingness at the item level only affected a small proportion of the sample. When students were missing more than 25% of a scale, the data from that timepoint was deleted from the student's record but the student was retained in the sample. Findings from 17 administrations of purpose scales across the four timepoints and 19 SDQ scales from Fall 2015 were deleted because over 75% of data on the scale were missing. Mean substitution (substitution of the individual's mean on the reported scale data) was used to replace specific items for students who were missing 25% or less of a scale in order to preserve cases (.5% to 2.6% of cases were affected for each scale). The use of mean substitution on these items likely had a minimal effect on study outcomes because in large datasets, missingness under 5% is thought to have little impact on the data analytic outcomes (Tabachnik & Fidell, 2013, p. 97).

Wave missingness on purpose scale. An additional aspect of missing data was students who completed the survey at fewer than expected timepoints (depending on cohort, students were expected to complete two to four timepoints). At the timepoint level, total scores for purpose and mental health failed to meet Little's test of Missing Completely at Random. To use established missing data procedures, it was important that the missing data were Missing at Random (MAR). MAR assumes that the

“probability of missingness is unrelated to unobserved concurrent outcomes” (Singer & Willett, 2003, p. 159). In other words, to be considered MAR, the fact that a student is missing a wave of purpose data should be unrelated to their score on the purpose scale. Correlations between number of student surveys returned and purpose total scores were below $r = .07$. Because student response to surveys was largely dependent on adults (teachers and school administrators) or on their grade level (two cohorts of students were only present in the schools for one year), it is likely that missing data are not related to student characteristics. Fortunately, a benefit of multilevel modeling as an approach to longitudinal data analysis is that the method can incorporate cases with even a single wave of data collection. Thus, all cases were included in the model (Singer & Willett, 2003). Singer and Willett (2003) do caution that severely unbalanced data can lead to errors in modeling, including failure of the model to converge.

Missingness on SDQ scale. The SDQ subsample was defined as students who completed the purpose survey in Fall 2015 ($n = 1691$). Because the Fall 2015 survey was administered over two separate occasions, the purpose scale and the SDQ measure were given separately. Thus, there were significantly different return rates for those scales: Of the 1691 students who completed the purpose scale in Fall 2015, 478 of those students did not complete the SDQ scale. Completing the scale was defined as responding to over 75% of the items. Complete case analysis in this case would have resulted in the loss of those 478 cases and would have, therefore, resulted in significant bias (Garson, 2015). Multiple imputation is one of the preferred methods for handling missing data (Rubin, 1987), yet this method has been poorly described for multi-level models. In particular, little attention has been paid to missingness on Level 2 predictors (van Buuren, 2011).

While flat-file imputation (ignoring the hierarchical nature of the data) is generally discouraged, multilevel imputation is complex and therefore rarely implemented (Gottfredson, Sterba, & Jackson, 2017; van Buuren, 2011). In fact, van Buuren (2011) cites research suggesting that the flat-file imputation model is likely an acceptable procedure and certainly preferable to complete case analysis or single imputation. To address the missingness on the SDQ scale, the SPSS 24 multiple imputation feature was used to create 5 imputed datasets. The following variables were included as predictors in the imputation model: purpose total scores at all four timepoints, achievement history (average GPA), lunch status, accommodations, cohort membership, school, and the full scale and subscales of the SDQ. The five imputed files, a dataset averaging those imputed files, and complete cases were analyzed using the same model structures in HLM. The patterns of significance in coefficients were consistent across these datasets.

Preliminary Analysis

Normality. Examination of histograms, boxplots, normal probability plots, and detrended normal probability plots indicated significant negative skew in purpose total scores at all timepoints. Skewness ranged from -1.10 to -.93. Kurtosis ranged from .53 to 1.04. Both skew and kurtosis appeared to lessen over the course of the study, with the worst skew apparent in Fall 2015. Many students responded to the purpose scale items in a positive way; in fact, at the baseline assessment in Fall 2015, 362 students rated themselves at the top of the scale range (total = 25). Data transformation strategies were considered to address this violation; however, while transformations were able to mitigate the skew in the middle of the distribution, the overall negative skew could not be effectively addressed through any transformation because all 362 students who rated

themselves at the top of the range would still receive the same score in any transformed version of the dataset. Dropping students with extremely high scores was also considered, but even if students with scores of 24 or 25 (the top two possible scores) were removed from the sample, hundreds of students responded in the total score range of 20-23, which would retain the skew and further restrict the range. Therefore, it was decided to proceed with analysis on the original data. *Post hoc* analyses investigated the effect of the extremely high purpose reporters. The SDQ scale demonstrated less skew, but still revealed a bias toward positive reporting (skew = .32, kurtosis = .66).

Outliers. Univariate and multivariate outliers were examined for purpose total scores and SDQ total score for the SDQ subsample. Examination of *Z* scores for purpose total scores revealed between 12 and 17 cases at each timepoint had *Z* scores below -3. There were no *Z* scores above 2, highlighting the significant negative skew in the variables. In the SDQ subsample, 16 students' scores fell above a *Z* score of 3. These cases were thought to be a true part of the sample and were not transformed or deleted. Mahalanobis, Cook's distance, and standardized DFBeta statistics revealed no concerning multivariate outliers for the relationship of SDQ and Fall 2015 purpose.

Purpose subscales. The measurements of the three purpose dimensions (Engagement, Intention, and Beyond-the-self) were based on theory rather than past research on the scale. Thus, it was necessary to run factor analysis to determine whether the five items on the purpose scale loaded on the three dimensions. Principal component analysis with varimax rotation found support for the theorized subscales (Table 1) when the outcome was constrained to three factors. Correlations of the purpose subscales are provided in Table 6. Items on the purpose scales were correlated at appropriate levels (*r*s

.39 to .66) to suggest each item contributed to a related and distinct aspect of the purpose scale.

Descriptive statistics. Means and standard deviations of the purpose scale are found in Table 7 (SDQ subsample in Table 8). Examination of purpose means over time suggested two predominate patterns. Overall, mean scores declined over the four timepoints in the study. This pattern of decline was evident for the full scale and the subscales (Figure 1). An overall decline in purpose scores was evident in the overall sample and when broken into groups by school, gender, race/ethnicity, and cohort (Figure 2a-c). A second pattern emerged that appeared to be more pronounced in some subgroups (females, Black and Latino students, and students in Schools 1, 3, and 4): a cubic trend with higher scores in the fall and lower scores in the spring of each year. Means of the purpose subscales suggested a cubic trend for Engagement and Intention but a steady or slight decline for Beyond-the-self (Figure 1). Further visual examination of the full data revealed a more complex story, which is described further in the “Data Visualization” section below.

Relationships between study variables. Total scores for the purpose scale were positively associated between each of the timepoints (Table 7). For both the purpose subscales (Table 6) and the purpose full scale (Table 7), the associations between the scores became weaker over time. The Fall 2015 SDQ total score was negatively associated with purpose at all timepoints, but this relationship became weaker over time (Table 8). Of note, the SDQ subscales followed a distinct pattern from the total score in that both the peer problems and emotional problems subscales had no relationship with

purpose in the following year (Fall 2016 and Spring 2017). The conduct problems subscale appeared to be driving the correlation between the SDQ scale and purpose scale.

A series of independent t-tests and one-way ANOVAs were examined to inform model specification and to explore the relationships of individual (gender, race, free/reduced lunch, accommodations) and contextual variables (priority/focus school) with purpose and SDQ scores. Of note, all group differences in purpose scores occurred on the Engagement or Beyond-the-self subscale. No significant differences were found between groups on the Intention subscale. Females, students with accommodations, and Black students endorsed higher levels of mental health symptoms compared to other students.

Visual Examination of Data

In line with the recommendations of Singer and Willett (2003), a series of data visualization tactics were used to inform model specification. First, scatterplots of time and purpose total score were examined in the full sample and for demographic subgroups. Regression lines were graphed for the full sample and for the key demographic groups. Most subgroups were indistinguishable, with the exception of grouping by cohort. All regression lines demonstrated a slight decreasing trend. Next, two random subsamples of 30 students were selected from the full sample to examine individual student trajectories. Plots of these students' data were examined using both linear and curvilinear regression lines. Plots were visually examined on one graph as well as separated by demographic groups. The individual plots showed that while there were many students who rated a decline in purpose, there were also many students who reported increases in purpose (Figure 4). Visually comparing the graphs grouped by demographic subgroups did not

reveal any clear distinctions. Overall, visual examination indicated that there was an overall decreasing trend in the purpose scores over the course of the study. However, the random case analysis also demonstrated significant individual variation. Despite several curvilinear trajectories in the individual cases analysis, a linear trajectory appeared to be the most parsimonious and suitable way to model the course of purpose for the entire sample over the span of the study.

Models Predicting Self-Reported Purpose (Total Score)

Hierarchical linear models were built using the HLM 7.0 software program using full maximum likelihood estimation method. The total score on the purpose scale was the outcome. Full maximum likelihood (versus restricted maximum likelihood) estimates the fixed and random effects simultaneously and allows goodness of fit statistics to apply to the entire model (Singer & Willett, 2003, p. 88). For all models, the slope was allowed to vary randomly because slopes were expected to vary across groups and contexts. Table 9 shows the model building process that culminated in the final model predicting purpose over time (Table 10). The corresponding mathematical equations are listed in Appendix A. In addition to the final model predicting purpose, the models predicting the purpose subscales (Engagement, Intention, and Beyond-the-self) are listed in Table 10.

Model 1: Unconditional means model. The unconditional means model (or intercept only model) describes the means of students without regard to time. No predictors were included at Level 1 or Level 2. The amount of within-person and between-person variance suggested that there was significant variability in purpose scores, warranting further analysis ($\chi^2(2628) = 6478.63, p < .001$). The intraclass

correlation coefficient (ICC) of .40 demonstrated that 40% of the variation in purpose scores was attributable to differences among students.

Model 2: Unconditional growth model. The next model added time as a Level-1 predictor of purpose. The average change trajectory for purpose was found to have a non-zero intercept ($\beta_0 = 20.61$, $t(2628) = 245.22$, $p < .001$) and non-zero slope ($\beta_1 = 20.61$, $t(2628) = .8.06$, $p < .001$). In contrast to the study hypothesis, and in line with the preliminary analyses, purpose was found to decrease over the course of the study. Although the unconditional growth model did demonstrate significantly less deviance than Model 1 ($\chi^2(3) = 135.99$, $p < .001$), a large amount of variance in the intercept and slope remained unexplained, meaning that multilevel modeling was indicated to try to explain this heterogeneity. AIC and BIC fit statistics supported the deviance significance tests, showing that the unconditional growth model accounted for more variance in purpose scores than the unconditional means model.

Models 3-7: Effects of individual and contextual variables on purpose.

Several variables were considered as potential predictors of purpose (variables listed in Table 3). First, each predictor was tested as a single Level-2 predictor to determine whether the variable should be considered for inclusion in the final model. This analysis resulted in removing several variables from further consideration: receiving accommodations, free/reduced lunch status, and the gender and race interaction variables were not found to be significant predictors of the purpose initial status or change. In addition, school (dummy-coded school variables) and school implementation status (whether schools were high or low implementers of MOSAIC program), were not found to be significant predictors of purpose initial status or change. Overall, whereas several

variables demonstrated an association with the purpose initial status, few variables were found to be associated with purpose over time. A quadratic term for time was also added to Level-1 during this initial stage of model fitting and was found to be nonsignificant.

The next stage of model fitting involved building a model based on the primary study hypotheses, using only predictors found to be significant individual predictors of purpose initial status or change over time. Ambassador status (whether student was a MOSAIC Ambassador at any point in the two years) and student cohort were first included as covariates, and these factors were found to have a significant effect on initial status but not slope (Model 3). Though not a focus of this study, cohort effects played a significant role in predicting purpose. Cohorts who began the study in 7th and 8th grade (Cohorts 2 and 3) had significantly lower intercepts than the 6th graders from that same year (Cohort 1). Notably, 6th graders who began the study in 2016-2017 (the second year) did not significantly differ from the cohort who began the student as 6th graders in 2015-2016 (Cohort 1).

Gender was added as the first predictor to address a specific study hypothesis (Model 4). As expected, female students had significantly higher initial scores on purpose when compared to males ($t(2623) = 2.54, p = .011$). However, student gender was not associated with the purpose slope, when controlling for Ambassador status and cohort effects. Model 4 accounted for significantly more variance in purpose compared to Model 3 ($\chi^2(10) = 87.62, p < .001$); however, the modest decrease in the AIC (from 31857.72 to 31854.17) and the increase in the BIC statistic suggested that Model 4 was not an improvement over Model 3. In fact, the BIC statistic increased between Model 3 and 4 while the AIC statistic only decreased by a small amount. These statistics suggest

that Model 4 was roughly the same, or perhaps worse, in its ability to account for variance in purpose over time compared to Model 3 (the covariates-only model). Gender was retained in the model for the model building process but was removed in the final model.

Race (dummy-coded with White as reference group) was included as the next predictor of interest (Model 5). Black students were found to report higher levels of initial purpose compared to White students, controlling for the effects of gender, cohort, and Ambassador status ($t(2620) = 2.416, p = .016$). Latino students and Asian students did not demonstrate significant differences in initial status compared to White students. Black students also demonstrated steeper (more negative) slopes compared to White students ($t(2620) = -2.92, p = .003$). Latino students' slope was approaching significance ($t(2620) = -1.78, p = .08$). Model 5 was found to be an improvement in accounting for variance in purpose ($\chi^2(6) = 21.89, p = .002$), but again the BIC and AIC statistics suggested that a more parsimonious model would be a better fit. For Model 5, the AIC showed a modest decrease (31854.17 to 31844.28) while the BIC showed an increase (31948.16 to 31973.52).

Models 6 and 7 added school status (whether school was labeled “priority” or “focus”) and student achievement history (average GPA), respectively. Controlling for the other predictors entered previously, attending a focus or priority school was associated with lower initial purpose scores ($t(2619) = -3.25, p = .001$) but was not related to change in purpose. Higher GPA was associated with a higher initial purpose score ($t(2618) = 3.54, p < .001$) and was approaching significance in its relationship with a less negative slope ($t(2618) = 1.63, p = .10$). Of note, across all of the models described

in Table 9, predictors retained significance as subsequent predictors were added to the model. The exception was that gender lost significance when GPA was added to the model, suggesting that the effect of gender on purpose was captured by the effect of achievement history. Goodness-of-fit statistics demonstrated a statistically significant decrease in the deviance statistic with each model. However, as mentioned above, some models showed very small decreases, or even increases, in the AIC or BIC statistics. Thus, several variables were removed from the final model to improve model parsimony and enhance interpretability (Table 10). Gender was removed from the fixed effects models predicting intercept and slope of purpose. Cohort, school status, achievement history, and ambassador status were also removed from the Level-2 model predicting slope. In addition, a model was also examined with a fixed slope to determine whether this more parsimonious model would demonstrate a better fit (Singer & Willett, 2003). However, the deviance statistic increased for this model (increased from 31752.43 to 31805.75), which means that this model with a fixed slope was not a better fit for the data, so the final model retained the random slope.

Although the model building process led to some incremental improvements in predicting purpose trajectories as demonstrated by the statistically significant deviance statistics, the pseudo- R^2 statistics (Table 9) revealed that the models were only accounting for very small amounts of variance in purpose (Singer & Willett, 2003). The pseudo- R^2 statistic is not without interpretation difficulties, particularly when the statistic is negative (as happened in a few of the models, as seen in Table 9). However, the overall pattern of low amounts of explained variance is helpful in demonstrating that the model building process using the hypothesized predictors of purpose trajectories culminated in a model

that had limited ability to predict purpose over time. Figures 4a-e depict a graphical representation of the final model separated by predictors of interest.

Models Predicting Three Purpose Dimensions

The same procedure used to build a model predicting purpose total scores was used for each of the three purpose dimensions (final models given in Table 10).

Modeling the trajectories of these three dimensions showed several similarities to modeling of the overall purpose score. Unconditional growth models showed that students' self-reported Beyond-the-self, Engagement, and Intention dimensions of purpose declined significantly over the course of the study. Overall, the relationships of the predictors that were included in these final models demonstrated the same patterns found in the models predicting the purpose total score. Differences in the final models are highlighted here.

Beyond-the-self. The model building process for the Beyond-the-self dimension mirrored the model for the total score, meaning that the same deliberate steps taken with the total score as the outcome were taken with the Beyond-the-self item as the outcome. This process led to a final model including the same predictors of initial status and slope; however, in the final model for the Beyond-the-self model the intercept for the slope lost significance, indicating that in the full model tested here, the slope of the Beyond-the-self dimension was not significantly different from zero. In other words, the Beyond-the-self dimension did not demonstrate a clear decline once other predictors were included in the model.

Intention. The model building process for the Intention subscale differed from that of the purpose total score because preliminary analysis did not identify gender and

school status as independent predictors of Intention. Therefore, gender and school status were not considered for the models predicting Intention. Of the three final models predicting purpose dimensions, the Intention model included the fewest predictors of purpose initial status (race, cohort, and ambassador status) but the highest number of predictors of purpose slope (race, cohort, GPA). This pattern of results suggests that the change in the Intention dimension was more clearly linked to predictor variables.

Engagement. The final Engagement model differed from the other final models because race was not found to be a predictor of initial status. Thus, the finding that Black students had significantly higher initial reports of purpose was true for Beyond-the-self and Intention but not for the Engagement dimension. In the Engagement final model, Latino students did demonstrate a steeper (more negative slope) than White students ($t(2625) = -1.96, p = .05$).

Mental Health Predicting Purpose Change

A subsample was used to test the hypothesis about the relationship of mental health to purpose over time. This sample included all students from the larger sample who had completed the purpose scale in Fall 2015 ($n = 1691$). The SDQ full scale and subscales were examined alone as predictors in the Level-2 model and as additional predictors to the final model described above. As previously described, multiple imputation was used to impute SDQ total scores and subscale scores for 478 students. Five imputed datasets were created and each dataset was run separately in HLM. The same analyses were run using a complete case analysis and using a dataset that included the averages of the imputed values. Each of these methods led to a similar pattern of results; the results of the averaged imputed values are reported here. Overall, higher

levels of mental health symptoms were associated with lower initial scores on purpose ($t(1689) = -6.41, p < .001$). However, neither the full SDQ scale nor any of the individual SDQ subscales were related to the slope of purpose (full scale model reported in Table 11). These results echoed the findings reported above: few predictors were found to be related to the change in purpose over time.

Post-Analysis

Examining assumptions. Per the recommendations outlined by Singer and Willett (2003), after conducting the model specification procedures described above, several graphical and statistical procedures were used to examine the appropriateness of model assumptions. After running models in HLM, residual plots for the final model predicting purpose was examined to investigate the normality of the residuals. Q-Q plots of Level-1 residuals were close to linear, indicating that residuals were close to a normal distribution. Plots of residuals versus predicted values for the full model showed randomly scattered but somewhat more concentrated at higher levels of residuals and predicted value. This finding was not surprising given the negative skew in purpose scores.

To assess for functional form, a series of individual subject and subgroup graphs were examined using the HLM software program. Like the preliminary visual examination, the post-analysis graphs in HLM demonstrated that there was significant heterogeneity in student trajectories of purpose. While a linear model appeared to capture a summary of the data best, there were many clear departures from this trend, including quadratic and cubic trends as well as several students with flat or increasing trajectories (Figure 5). Overall, the model checking process reinforced weaknesses in the

data that were uncovered during preliminary analysis, which likely impacted results by making it difficult to accurately summarize an overall trajectory for purpose.

Post hoc analysis of skew. The extreme skew on the only outcome variable in this study created a restricted range and ceiling effect that was difficult to address. The extreme skew likely also resulted in a regression to the mean, making it difficult to interpret one of the study's primary findings: the decline in purpose over time. Possible reasons for this extreme skew included students not understanding the concept of purpose well enough to accurately reflect their true intentions and/or behaviors or students reacting to a positive response/social desirability bias. To further examine the impact of extremely high reporting, exploratory *post hoc* analyses were conducted. A dichotomous variable was created to account for students who may have inaccurately reported high levels of purpose at their first assessment (presumably because of social desirability bias or lack of understanding of the purpose construct). Students with extremely high ratings (total score of 24 or 25) at their first assessment (Fall 2015 for Cohorts 1, 2, 3 and Fall 2016 for Cohort 4) were coded as "high baseline purpose raters." Using this strategy, 708 students (27% of the sample) were identified. Using this dichotomous variable as a single level-2 predictor demonstrated that these students had an alarming impact on the results of the study. High purpose raters had an initial status of purpose that was 5 points higher than the remainder of the sample ($t(2627) = 445.99, p < .001$). High purpose raters also had a significantly more negative slope in purpose over time, which is most clear from the graphical display of this model (Figure 6). Accounting for those who rated themselves very high at their baseline assessment of purpose suggests that there may be a flat or slight increasing trend for the other remaining students.

Discussion

The current study sought to address several gaps in the youth purpose literature by quantitatively modeling the trajectory of self-reported purpose in the context of urban middle schools. In an attempt to understand purpose development from a Developmental Systems Theory perspective, relevant contextual factors were examined as potential predictors of initial purpose status and change over time. As hypothesized, higher academic achievement was related to higher initial purpose scores. Both higher (worse) mental health symptoms and attending a “failing” school were associated with lower initial purpose scores. Females had higher initial levels of purpose than males, but these effects disappeared when GPA was added to the model. Of note, the ongoing intervention (MOSAIC) had some impact on purpose in that participation in Ambassador Program was associated with higher initial purpose.

Many of the findings in this study were unexpected. Several hypothesized predictors of purpose were not significant, including free/reduced lunch status and level of MOSAIC implementation (high versus low implementing schools). Furthermore, it was surprising that few contextual factors were significant predictors of purpose change. As this is the first study to quantitatively model purpose over time, it is not clear whether this is a typical finding.

Several predictors were also associated with purpose in unexpected ways. For instance, Black students reported higher initial levels of purpose compared to White students. Again, because of the limited research in this area, this finding is difficult to interpret. Gutowski et al. (2017) found that some youth of color living in urban poverty experienced stress as a barrier to purpose development, while other youth in the same

setting viewed this stress as motivation to escape their environment. It is possible that Black students were more likely to feel motivated by their difficult surroundings than other groups. For three of the four cohorts (cohorts 1, 3, and 4), a higher percentage of Black students rated themselves with a perfect purpose (total score = 25) compared to the other racial/ethnic groups. In addition, Black students were more likely than other groups to fall into the category of a “high rater” at baseline. To examine possible general response bias patterns, independent sample *t*-tests compared SDQ scores for students who were high raters of purpose at baseline and those who were not. Overall, students who were high purpose raters at baseline had lower SDQ scores, indicating more positive reports of mental health ($t(1211) = 3.44, p = .001$). However, when these *t*-tests were examined separately by racial groups, this finding did not hold for Black or Latino students. Thus, the high ratings of Black students on the purpose scale do not seem to reflect an overall positive reporting bias. Black students who were high purpose raters at baseline did not report mental health symptoms at significantly different levels than those who were not high purpose raters.

It is possible that the steeper negative slope demonstrated by Black students is attributable to regression to the mean rather than any specific experience of the Black students. However, it also is quite possible that those Black students who initially saw themselves as purposeful were more affected by the harmful contextual factors over the course of the study. The fact that these differences were observed in the Black students may point to data patterns that were not recognized through the current approach. Given the great heterogeneity in individual trajectories seen in the visual examination of the

data, it is possible that the data analytic approach used in this study masked subgroups of students who respond to the urban middle school context in divergent ways.

Another surprising finding was that although females demonstrated higher initial purpose than males in some models, this relationship disappeared when academic achievement was taken into account. Thus, gender did not have an effect on purpose that was separate from the effect of academic achievement. The positive correlation between being female and academic achievement ($r = .26$ in this study) was not so high that this finding was likely. It could be that academic achievement mediates the link between gender and purpose. There is a small body of work on the association of academic achievement with purpose (e.g., Pizzolato et al., 2011), but no research has looked at the combined effects of academic achievement and contextual factors on purpose development. Future research should further examine the role of academic achievement in purpose development, particularly its relationship with gender and the additional contextual factors examined in this study.

Many of the findings regarding the three purpose dimensions (Engagement, Intention, and Beyond-the-self) were also unexpected. As with the models predicting the total purpose score, few predictors were able to account for change in the purpose dimensions. Most notable was that the contextual factors chosen to predict purpose change were most related to students' Intention toward purpose. In addition, it had been hypothesized that the Beyond-the-self dimension of purpose would be the least stable dimension; however, this dimension had the smallest relationship with time, indicating very little change over the course of the study. Another unexpected finding was that although Black students were likely to report higher initial levels of Beyond-the-self and

Intention dimensions of purpose, there were no significant differences by race/ethnicity in the initial status of Engagement in purpose.

The main unexpected finding was the decline in purpose over time. The original study hypothesis was that purpose would increase over the course of the study as a signifier of positive youth development. Not only did purpose total scores decline over the course of the study, but student cohorts had significantly different intercepts on purpose, with 6th graders demonstrating higher initial self-reported purpose than 7th or 8th graders. There are several explanations for self-reported purpose to have declined over the two years of the study. For one, as the *post hoc* analysis of the high purpose raters revealed, significant skew in the purpose variable may have created a regression to the mean effect that influenced the modeling process. The skew may have been influenced by students who had limited understanding of purpose at baseline but over time (and, perhaps, due to the ongoing MOSAIC intervention), students may have started to understand purpose better and rated themselves differently. Unfortunately, within the constraints of the data available for this study, it is not possible to separate out response bias, lack of understanding, and true high purpose scores. Thus, it is also worth considering additional explanations of this decline in purpose.

The relationship of cohort to purpose trajectory suggests that the decline in purpose found over the course of this study may be an expected developmental pattern. There is some precedent in the literature for decreases in purpose in adolescence and emerging adulthood. Crumbaugh and Maholick (1967) found that college freshman had higher reported purpose than seniors. Blattner et al. (2013) found that tenth grade students had lower levels of commitment to purpose and higher levels of search for

purpose than 6th and 8th graders. Some literature on other developmental assets has also identified declines over specific periods in adolescence. Kanacri et al. (2013) found decreases in prosociality from age 13 to 17, and a sense of school belonging has been found to decline over 6th and 7th grade (Anderman, 2003). It is important to note that one of the key studies cited in the literature review, the Malin et al. (2013) longitudinal interview study, found that only eight of 46 sixth graders were purposeful, and only two of these students maintained purpose at the follow-up interview two years later. While this study used a small sample, it does lend support to either a shifting understanding of purpose or a natural decline in purpose from sixth to eighth grade.

Reasons for this natural decline are also not clear. It is possible that declines in developmental assets, including purpose, represent iatrogenic effects of some school environments, particularly for students from marginalized racial groups and low socioeconomic status (Osher et al., 2018). The graphs of the mean scores (Figures 1 and 2) that show a slight cubic trend point to periodic higher fall scores and lower spring scores. This pattern could signify hopeful beginnings that slowly decline as the school year progresses and high-stakes testing pressure increases (testing is typically in the Spring). This kind of pattern could explain declines across the college years, as well. Further research is needed to uncover whether this cyclical pattern represents a contextual interaction of the school environment and aspects of student development.

Limitations

The study had several limitations, many of which are related to the difficulty of measuring youth purpose through self-report surveys. Most significant was the negative skew on the purpose scale, which was explored in detail in the *post hoc* exploratory

analysis. The skew in the purpose variable is related to another study limitation: relying on a single-informant approach to measure purpose. The single-informant self-report design of this study made it difficult to interpret the unexpected findings. Future research should combine self-report with other forms of reporting (teacher) and use more mixed methods (student interview or essay) to understand what students mean when they respond to self-report questions about their life purpose.

The difficulty of assessing character, particularly through self-report measures is well known (Card, 2017). A specific challenge of measuring purpose is the lack of conceptual clarity around the construct. Many of the widely used purpose instruments were developed from the original Frankl work on purpose, which means that these measures do not clearly differentiate “purpose” from the construct of “meaning” (see Bronk, 2014 for a review of purpose measures). The purpose scale used in this study was adapted from two existing measures that were designed to measure youth purpose. Both of these scales were developed after Damon et al. (2003) articulated a more precise definition of purpose. Although the study used existing items, these items were not originally intended to measure the three dimensions of purpose. The different patterns in the results of the purpose dimensions point to a need to further explore these dimensions quantitatively through deliberate scale development procedures. In particular, more research is needed to determine how to best reflect the Beyond-the-self dimension (Bronk 2014; Liang et al., 2017a), which was only assessed using one item in this study.

Part of the difficulty around purpose measurement lies in a lack of conceptual clarity in the field. Whereas the Damon group out of Stanford has explicitly labeled purpose as requiring a “Beyond-the-self” element, other researchers have not made this

distinction. Growing research supports that the Beyond-the-self dimension of purpose is specifically important to promoting mental health and positive outcomes. For example, Abramowski, Pierce, and Stoddard (2018) found substance use was inversely related to other-focused purpose, while a self-focused purpose was positively associated with substance use in high school students. However, the Beyond-the-self purpose may not be as critical for younger students. Future research should further clarify the role of the Beyond-the-self dimension of purpose and whether other forms of purpose (self-focused purpose) may be a developmental precursor or impediment to cultivating this dimension.

A further limitation in this study was the lack of measuring the search for purpose. Much of the purpose literature has differentiated between “searching” for and “committing” to purpose, but the scale used in this study did not include items that clearly measured search for purpose. The items used in this study intended to capture Intention and Engagement seem to incorporate the concept of commitment. The distinction between search and commitment is important because some researchers have found searching for purpose to be related to higher life satisfaction in adolescence (Bronk et al., 2009); others have found that while commitment to a purpose is associated with higher self-esteem, searching for a purpose is associated with lower self-esteem (Blattner et al., 2013). In line with the findings of Blattner et al. (2013), it may be that the students in this sample would have shown an increased search for purpose, even as their commitment to purpose declined.

An additional limitation was the unbalanced longitudinal data that resulted from both unexpected wave missingness and planned missingness for some cohorts (Cohorts 3 and 4 were only in the study for one of the two years). Singer and Willett (2003) and

Hox et al. (2018) indicate that unbalanced design is not a concern when using HLM for longitudinal analyses. However, Singer and Willett (2003) do warn that extreme unbalanced design could lead to errors in model specification. It is possible that the data in this study were unbalanced enough that this affected how well variables were able to predict purpose change.

The great heterogeneity in individual purpose trajectories that was observed in this study suggests that there were contextual factors that were not examined that may have accounted for purpose change. For example, analyses were not able to effectively control for school-level effects because of the small sample of schools. In addition, MOSAIC implementation data were not used at a classroom or student level. Being able to more effectively control for the intervention at the student level would clarify the role of the MOSAIC program in purpose development. Another approach to handling the heterogeneity in individual trajectories would be to use a person-centered approach, such as growth mixture modeling, to uncover categories of purpose growth patterns.

Future Directions

The theoretical tenets of Developmental Systems Theory are difficult to assess with current methodology. In his chapter describing the theory underlying developmental systems, Overton (2015) highlights the need to shift data analysis toward “creating categories, not [...] cutting nature at its joints” (p. 12). Current empirical procedures, such as those used in this study, represent an attempt to uncover relationships among variables to explain development on a broad scale. However, to take Overton’s (2015) recommendations into account, more nuanced approaches are needed. An attempt to describe development on a broad scale may be oversimplifying to the point of inaccuracy.

Current analytic strategies are not able to account for the dynamic relationships among contextual factors and development in youth. For example, each school included in this study represented a unique context, and this context was not a static collection of variables. Rather, the schools included in this study experienced important shifts in personnel, programming, and oversight over the two-year period. Yet, these contextual variables were not clearly captured by the variables in this study. Statistically accounting for school level effects would require including a larger sample of schools; for many studies, including this study, this requirement is not tenable. Therefore, it is essential that researchers develop innovative methods for investigating the dynamic role that context plays in youth development without requiring unrealistically large samples.

Future research should also consider a more complex developmental process for purpose in youth. Current models emphasize the identification of a single life purpose; however, it is possible that youth normatively identify and engage with several purposes over the course of adolescence. The literature on purpose development may then depend largely on timing: are students in the process of searching for a new purpose? This level of complexity has not been incorporated into the current models of purpose development and may account for some of the inconsistencies in the current study and in the literature overall.

Future research on purpose should also further account for the context of development by examining the potential impact of sociocultural influences at every level of the ecological system. For example, one of the findings in the current study was that Black students demonstrated a distinct relationship to purpose, both initially as well as over time. To make appropriate interpretations, it is imperative to consider how larger

national or community dialogues are interpreted by students and incorporated into their understanding of sense of purpose. The study took place from 2015 to 2017, during which a polarizing national election unfolded, national dialogues on race and police violence toward unarmed Black males waxed and waned, and increasing visibility of the Black Lives Matter movement shined a national spotlight on activism in Black communities. A mixed methods approach of interviewing students from a variety of racial/ethnic backgrounds with different purpose trajectories could uncover how such national events were understood by students and their community to clarify how the larger ecology may affect self-reported purpose. A case can be made for the same events to lead to either an enhanced or a diminished sense of positive purpose; knowing which of these stories is the case for an individual child would allow for interpretation of co-occurring intervention trajectories.

Purpose research in urban communities of color is limited; however, there are several bodies of literature investigating related constructs that may offer important theoretical contributions. For instance, personal agency and self-efficacy (Bandura, 2001) as well as critical consciousness (see Watts, Diemer, & Voight, 2011) support the ability of students of color to succeed in academic contexts (Weinstein, Gregory, & Strambler, 2004). Exactly how these terms are used and interpreted is important for understanding the relevance of this literature for purpose development. It may be that conceptualizations of purpose development need to be integrated into those of the development of critical consciousness and/or self-efficacy or agency; these dimensions may well turn out to have salience for White middle class students, as well. Future work should integrate literature from bodies of work that are specific to communities of color

to work toward a more inclusive and ecologically valid theory of purpose development for students from all backgrounds.

Conclusion

Purpose is considered an essential developmental asset and galvanizing force for adolescents of all backgrounds (Damon et al., 2003). However, most of the literature on adolescent purpose development has investigated the construct in majority White and middle to upper income settings, with little attention paid to the context of development. Using a Developmental Systems Theory Framework, this study sought to address gaps in the purpose literature by quantitatively modeling purpose among students in the setting of largely low-SES, predominately racial/ethnic minority, urban middle schools. Results were unexpected in that purpose was found to decline over the course of the study. While many of the hypothesized predictors were related to the initial status of purpose, few predictors were related to change in purpose over time. Overall, the results of this study failed to provide clear answers to the research questions. Instead, results pointed to challenges with conceptual clarity and measurement of purpose, difficulty measuring purpose in middle school students, and the mismatch between Developmental Systems Theory and current research methodology. Understanding youth purpose development, particularly in the context of low-resourced urban schools, remains a critical priority. This study highlights the need to develop innovative procedures for statistically modeling the complex nature of youth development in context.

References

- Abramoski, K., Pierce, J., Hauck, C., & Stoddard, S. A. (2018). Variations in adolescent purpose in life and their association with lifetime substance use. *The Journal of School Nursing, 34*(2), 114-120.
- Anderman, L. H. (2003). Academic and social perceptions as predictors of change in middle school students' sense of school belonging. *The Journal of Experimental Education, 72*(1), 5-22. doi:10.1080/00220970309600877
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1-26.
- Benson, P.L. (2008) *Sparks: How parents can help ignite the hidden strengths of teenagers*. San Francisco: Jossey Bass.
- Benson, P. L., Scales, P., Hamilton, S. F., & Sesman Jrs., A. (2006). Positive youth development: Theory, research, and applications. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Psychology: Vol 1: Theoretical Models of Human Development* (6th ed., pp. 894-941). New York, NY: Wiley.
- Blakemore, S. J., & Mills, K. L. (2014). Is adolescence a sensitive period for sociocultural processing? *Annual Review of Psychology, 65*, 187-207.
- Blattner, M. C. C., Liang, B., Lund, T., & Spencer, R. (2013). Searching for a sense of purpose: The role of parents and effects on self-esteem among female adolescents. *Journal of Adolescence, 36*(5), 839-848.
- Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.
- Bronk, K. C. (2012). A grounded theory of the development of noble youth purpose. *Journal of Adolescent Research, 27*(1), 78-109.
- Bronk, K. C. (2014). *Purpose in life: A critical component of optimal youth development*. New York, NY: Springer Science & Business Media.
- Bronk, K. C., Finch, W. H., & Talib, T. L. (2010). Purpose in life among high ability adolescents. *High Ability Studies, 21*(2), 133-145.
- Bronk, K. C., Hill, P. L., Lapsley, D. K., Talib, T. L., & Finch, H. (2009). Purpose, hope, and life satisfaction in three age groups. *Journal of Positive Psychology, 4*(6), 500-510.
- Bundick, M. (2011). The benefits of reflecting on and discussing purpose in life in emerging adulthood. *New Directions for Youth Development, 132*, 89-103.
- Bundick, M., Andrews, M., Jones, A., Mariano, J. M., Bronk, J. C., & Damon, W. (2006). *Revised Youth Purpose Survey*. Stanford, CA: Unpublished instrument, Stanford Center on Adolescence.
- Bundick, M.J., Andrews, M.C., Jones, A., Moran, S., Mariano, J.M., Bronk, K.C., & Damon, W. (2008). *Youth Purpose Survey Version 2008*. Stanford, CA: Unpublished instrument, Stanford Center on Adolescence.
- Burrow, A. L., O'Dell, A. C., & Hill, P. L. (2010). Profiles of a developmental asset: Youth purpose as a context for hope and well-being. *Journal of Youth & Adolescence, 39*(11), 1265-1273.
- Card, N. (2017). Methodological issues in measuring the development of character. *Journal of Character Education, 13*(2), 29-45.

- Crumbaugh, J.C., & Maholick, L.T. (1967). An experimental study in existentialism: The psychometric approach to Frankl's concept of noogenic neurosis. In V.E. Frankl (Ed.), *Psychotherapy and existentialism* (pp. 183-197). New York: Washington Square Press.
- Curran, P. J., Obeidat, K., & Losardo, D. (2010). Twelve frequently asked questions about growth curve modeling. *Journal of Cognitive Development, 11*(2), 121-136. doi:10.1080/15248371003699969
- Damon, W. (2008). *The path to purpose: Helping our children find their calling in life*. New York: The Free Press.
- Damon, W., Menon, J., & Bronk, K. C. (2003). The development of purpose during adolescence. *Applied Developmental Science, 7*, 119-128.
- Dik, B.J., Steger, M.F., Gibson, A., & Peisner, W. (2011). Make your work matter: Development and pilot evaluations of a purpose-centered career education intervention. *New Directions for Youth Development, 132*, 59-73.
- Duffy, R. D., & Raque-Bogdan, T. L. (2010). The motivation to serve others: Exploring relations to career development. *Journal of Career Assessment, 18*(3), 250-265.
- Eisenberg, N., Carlo, G., Murphy, B., & van Court, P. (1995). Prosocial development in late adolescence: A longitudinal study. *Child Development, 66*, 1179-1197.
- Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial behavior. In N. Eisenberg (Ed.), *Handbook of child psychology. Social, emotional, and personality development* (6th ed., Vol. 3, pp. 646-718). New York: Wiley.
- Erikson, E. H. (1968). *Identity: Youth and crisis*. New York: Norton.
- Esposito, C. (1999). Learning in urban blights: School climate and its effect on the school performance of urban, minority, low-income children. *School Psychology Review, 28*(3), 365
- Finnigan, K. S., & Daly, A. J. (2012). Mind the gap: Organizational learning and improvement in an underperforming urban system. *American Journal of Education, 119*(1), 41-71. doi:10.1086/667700
- Frankl, V. E. (1959). *Man's search for meaning: An introduction to logotherapy*. Boston: Beacon.
- Garson, G.D. (2015). *Missing values analysis and data imputation*. Asheboro, NC: Statistical Associates Publishing.
- Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge: Harvard University Press.
- Goodman, R., Meltzer, H., & Bailey, V. (1998). *The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version*. *European Child and Adolescent Psychiatry, 7*, 125-130.
- Gottfredson, N. C., Sterba, S. K., & Jackson, K. M. (2017). Explicating the conditions under which multilevel multiple imputation mitigates bias resulting from random coefficient-dependent missing longitudinal data. *Prevention Science, 18*(1), 12-19. doi:10.1007/s11121-016-0735-3
- Gutowski, E., White, A. E., Liang, B., Diamonti, A.-J., & Berado, D. (2017). How stress influences purpose development: The importance of social support. *Journal of Adolescent Research, 00*(0). doi:10.1177/0743558417737754
- Hatchimonji, D.R., Linsky, A.V., & Elias, M.J. (2017). Cultivating Noble Purpose in urban middle schools: A missing piece in school transformation. *Education, 138* (2) 162-178.

- Hill, P. L., Edmonds, G. W., Peterson, M., Luyckx, K., & Andrews, J. A. (2016). Purpose in life in emerging adulthood: Development and validation of a new brief measure. *Journal of Positive Psychology, 11*(3), 237-245.
- Hill, P. L., Sumner, R., & Burrow, A. L. (2014). Understanding the pathways to purpose: Examining personality and well-being correlates across adulthood. *Journal of Positive Psychology, 9*(3), 227-234.
- Hoffman, M. L. (2000). *Empathy and moral development: Implications for caring and justice*. New York: Cambridge University Press.
- Holme, J.J., Jabbar, H., Germain, E., & Dinning, J. (2018). New measures of teacher turnover can reveal deeper, underlying chronic staffing problems in schools. *PRC Research Brief 3*(4). DOI: 10.15781/T2VM43D9H.
- Hox, Moerbeek, & van de Schoot (2018). *Multilevel analysis: Techniques and applications, 3rd Edition*. New York, NY: Routledge.
- Ishida, R., & Okada, M. (2006). Effects of a firm purpose in life on anxiety and sympathetic nervous activity caused by emotional stress: Assessment by psychological method. *Stress and Health: Journal of the International Society for the Investigation of Stress, 22*(4), 275-281.
- Kanacri, B. P. L., Pastorelli, C., Zuffianò, A., Caprara, G. V., & Eisenberg, N. (2013). The development of prosociality from adolescence to early adulthood: the role of effortful control. *Journal of Personality, 81*(3), 302-312.
- King, L. A., Hicks, J. A., Krull, J., & Del Gaiso, A. K. (2006). Positive affect and the experience of meaning in life. *Journal of Personality and Social Psychology, 90*, 179-196.
- Knight, G. P., Basilio, C. D., Carlo, G., & Jacobson, R. P. (2015). Familism values, perspective taking, and prosocial moral reasoning: Predicting prosocial tendencies among Mexican American adolescents. *Journal of Research on Adolescence, 25*(4), 717-727.
- Kohlberg, L. (1984). *The psychology of moral development: The nature and validity of moral stages*. San Francisco: Harper & Row.
- Krettenauer, T., Colasante, T., Buchmann, M., & Malti, T. (2014). The development of moral emotions and decision-making from adolescence to early adulthood: A 6-year longitudinal study. *Journal of Youth and Adolescence, 43*(4), 583-596.
- Lerner, R. M., Dowling, E. M., & Anderson, P. M. (2003). Positive youth development: Thriving as a basis of personhood and civil society. *Applied Developmental Science, 7*, 172-180.
- Liang, B., Lund, T., Mousseau, A., White, A. E., Spencer, R., & Walsh, J. (2017a). Adolescent girls finding purpose: The role of parents and pro-sociality. *Youth & Society, 1*-17. doi: 10.1177/0044118x17697850
- Liang, B., White, A. E., Rhodes, H., Strodel, R., Gutowski, E., Mousseau, A. M. D., & Lund, T. J. (2017b). Pathways to purpose among impoverished youth from the Guatemala City Dump community. *Community Psychology in Global Perspectives, 3*(2), 1-21.
- Lippman, L.H., Moore, K.A., Guzman, L., Ryberg, R., McIntosh, H., Ramos, M.F., ...Kuhfeld, M. (2014) Pilot study and psychometric analyses. In *Flourishing Children: Defining and Testing Indicators of Positive Development*. (pp. 45-105). Springer US.

- Machell, K. A., Disabato, D. J., & Kashdan, T. B. (2016). Buffering the negative impact of poverty on youth: the power of purpose in life. *Social Indicators Research*, 126(2), 845-861. doi:https://doi.org/10.1007/s11205-015-0917-6
- Malin, H., Reilly, T. S., Quinn, B., & Moran, S. (2013). Adolescent purpose development: Exploring empathy, discovering roles, shifting priorities, and creating pathways. *Journal of Research on Adolescence*, 24(1), 186-199.
- Malti, T., & Buchmann, M. (2010). Socialization and individual antecedents of adolescents' and young adults' moral motivation. *Journal of Youth and Adolescence*, 39, 138-149.
- Marcia, J. E. (1966). Development and validation of ego identity status. *Journal of Personality and Social Psychology*, 3, 551-558.
- Mariano, J. M., Going, J., Schrock, K., & Sweeting, K. (2011). Youth purpose and the perception of social supports among African-American girls. *Journal of Youth Studies*, 14(8), 921-937.
- Mariano, J. M., & Vaillant, G. E. (2012). Youth purpose among the 'greatest generation'. *Journal of Positive Psychology*, 7(4), 281-293.
- Minehan, J. A., Newcomb, M. D., & Galaif, E. R. (2000). Predictors of adolescent drug use: Cognitive abilities, coping strategies, and purpose in life. *Journal of Child & Adolescent Substance Abuse*, 10, 33-52.
- Moran, S. (2009). Purpose: Giftedness in intrapersonal intelligence. *High Ability Studies*, 20, 143-159.
- Nucci, L. P. (2001). *Education in the moral domain*. Cambridge, UK: Cambridge University Press.
- Osher, D., Cantor, P., Berg, J., Steyer, L., & Rose, T. (2018). Drivers of human development: How relationships and context shape learning and development. *Applied Developmental Science*, 1-31. doi:10.1080/10888691.2017.1398650
- Ou, S. & Reynolds, A. (2008). Predictors of educational attainment in the Chicago longitudinal study. *School Psychology Quarterly*, 23 (2), 199-229.
- Overton, W. F. (2015). Processes, relations, and relational-developmental-systems. In W. F. Overton, & P. C. M. Molenaar (Eds.), *Handbook of Child Psychology and Developmental Science: Vol. 1. Theory and Method* (7th ed., pp. 9-62). Hoboken, NJ: Wiley.
- Padilla-Walker, L. M., Dyer, W. J., Yorgason, J. B., Fraser, A. M., & Coyne, S. M. (2015). Adolescents' prosocial behavior toward family, friends, and strangers: A person-centered approach. *Journal of Research on Adolescence*, 25(1), 135-150.
- Pizzolato, J.E., Brown, E.L., & Kanny, M.A. (2011). Purpose plus: Supporting youth purpose, control, and academic achievement. *New Directions for Youth Development*, 132, 75-88.
- Quinn, B. P. (2016). The beyond-the-self dimension of adolescent purpose: Absence and change. *The Journal of Positive Psychology*, 1-13.
- Raudenbush, S.W., & Bryk, A.S. (2002). *Hierarchical linear models: Applications and data analysis methods*, 2nd edition. Thousand Oaks: CA, Sage.
- Raudenbush, S.W., Bryk, A.S., & Congdon, R. (2011). HLM 7.00 for Windows [Computer software]. Lincolnwood, IL: Scientific Software International, Inc.
- Rubin, D.B. (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: Wiley.

- Sacks, V., & Murphey, D. (2018, February 12). Prevalence Adverse Childhood Experiences nationally by state and race and ethnicity. *ChildTrends.org*. Retrieved from <https://www.childtrends.org/publications/prevalence-adverse-childhood-experiences-nationally-state-race-ethnicity>
- Sayles, M. (1995). Adolescents' purpose in life and engagement in risky behaviors: Differences by gender and ethnicity. *Dissertation Abstracts International Section A: Humanities and Social Sciences, Vol 55(9-A)*, Mar 1995, 2727.
- Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*, 5–14.
- Singer, J.D., & Willett, J.B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York, NY: Oxford University Press.
- Tabachnik, B.G., & Fidell, L.S. (2013). *Using multivariate statistics, 6th edition*. Boston, MA: Pearson.
- Thompson, R., Proctor, L. J., Weisbart, C., Lewis, T. L., English, D. J., Hussey, J. M., & Runyan, D. K. (2007). Children's self-report about violence exposure: an examination of the Things I Have Seen and Heard Scale. *American Journal of Orthopsychiatry, 77*(3), 454-466. doi:10.1037/0002-9432.77.3.454
- van Buuren, S. (2011). Multiple imputation of multilevel data. In J.J. Hox and J.K. Roberts (Eds.) *The Handbook of Advanced Multilevel Analysis* (pp.173-196). New York: Routledge.
- Van der Graaff, J., Branje, S., De Wied, M., Hawk, S., Van Lier, P., & Meeus, W. (2014). Perspective taking and empathic concern in adolescence: Gender differences in developmental changes. *Developmental Psychology, 50*(3), 881-888.
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin, 140*(4), 1174.
- Wallace, J. M., Jr., Goodkind, S., Wallace, C. M., & Bachman, J. G. (2008). Racial, ethnic, and gender differences in school discipline among U.S. high school students: 1991-2005. *Negro Educational Review, 59*(1/2), 47-62.
- Watts, R. J., Diemer, M. A., & Voight, A. M. (2011). Critical consciousness: Current status and future directions. *New Directions in Child and Adolescent Development, 134*, 43-57. doi:10.1002/cd.310
- Weinstein, R. S., Gregory, A., & Strambler, M. J. (2004). Intractable self-fulfilling prophecies fifty years after Brown v. Board of Education. *American Psychologist, 59*(6), 511-520. doi:10.1037/0003-066X.59.6.511
- Yeager, D. S., & Bundick, M. J. (2009). The role of purposeful work goals in promoting meaning in life and in schoolwork during adolescence. *Journal of Adolescent Research, 24*(4), 423-452.
- Yeager, D. S., Bundick, M. J., & Johnson, R. (2012). The role of future work goal motives in adolescent identity development: A longitudinal mixed-methods investigation. *Contemporary Educational Psychology, 37*(3), 206-217.

Table 1
Purpose Scale

Item	Dimension	Source
My life will make a difference in the world.	Beyond the Self	Lippman et al., 2014
I am doing things now that will help me to achieve my purpose in life.	Engagement	Lippman et al., 2014
My life has a clear sense of purpose.	Intention	Bundick et al., 2008
I am always working toward accomplishing my most important goals in life.	Engagement	Bundick et al., 2008
I have a purpose in my life that says a lot about who I am.	Intention	Bundick et al., 2008

Table 2
Strengths Difficulties Questionnaire (Goodman et al., 1998)

Item	Subscale
I get a lot of headaches, stomach-aches or sickness.	Emotional Problems
I worry a lot.	Emotional Problems
I have many fears, I am easily scared.	Emotional Problems
I am nervous in new situations. I easily lose confidence.	Emotional Problems
I get very angry and often lose my temper.	Conduct Problems
I fight a lot. I can make other people do what I want.	Conduct Problems
I am often accused of lying or cheating.	Conduct Problems
I would rather be alone than with people my age.	Peer Problems
I have one good friend or more.	Peer Problems
I get along better with adults than with people my own age.	Peer Problems
Other children or young people pick on me or bully me.	Peer Problems

Table 3
Description of Hypothesized Predictors of Purpose Initial Status and Change

Predictor	Coding	Hypothesized relationship with purpose
Gender	0 = Male; 1 = Female	Intercept: Females higher than males Slope: Females more positive than males
Race	Dummy-coded with White as reference group (Black, Latino, Asian, White)	Intercept: Marginalized racial/ethnic groups in education setting, particularly Black and Latino students, lower initial status than White and Asian students Slope: Marginalized racial groups more negative slope
Free or Reduced Lunch	0 = Not eligible for Free or Reduced lunch; 1 = Eligible for Free or Reduced lunch	Intercept: Eligible for free or reduced lunch lower than those not eligible Slope: Eligible for free or reduced lunch more negative than those not eligible
Cohort	Dummy-coded with Cohort 1 (largest group) as reference Cohort 1: 6 th graders 2015-2016, 7 th graders 2016-2017 Cohort 2: 7 th graders 2015-2016, 8 th graders 2016-2017 Cohort 3: 8 th graders 2015-2016 Cohort 4: 6 th graders 2016-2017	Intercept: Expected Cohort 3 to have highest intercept and Cohorts 1 and 4 to have lowest intercepts (reflecting expected increase in purpose over time) Slope: Expected parallel slopes across cohorts (no differences)
Achievement history	Academic grades (reading, writing, math, science, social studies) reported by district in 2015-2016 and 2016-2017 averaged to create indicator of overall achievement	Intercept: Higher GPA associated with higher initial status Slope: Higher GPA associated with more positive slope
Focus or Priority School ^b	0 = School was neither Focus nor Priority (2, 4, 5); 1 = School was Focus or Priority (1, 3, 6)	Intercept: Focus/Priority schools lower than other schools Slope: Focus/Priority schools more negative than other schools
Mental Health	SDQ Scale	Intercept: Worse mental health symptoms, lower initial purpose status Slope: Worse mental health symptoms, more negative slope
Ambassador Status ^a	0 = Never MOSAIC Ambassador; 1 = MOSAIC Ambassador during 2015-2016 or 2016-2017	Control for intervention effects.
High or Low MOSAIC Implementer ^a	0 = Schools were low implementers of MOSAIC intervention (1, 4, 5); 1 = Schools were high implementers of MOSAIC intervention (2, 3, 6)	Control for intervention effects
Accommodations	0= Did not receive accommodations; 1= Student received 504 Plan or IEP during 2015-2016 or 2016-2017 school years	Potential control variable
Six schools	Dummy-coded with School 1 (largest school) as reference	Potential control variable

Note. ^a Variables were included to control for effects of ongoing intervention. ^b “Focus” and “Priority” labels are state designations indicating poor academic performance

Table 4
Sample Characteristics

School Number	1	2	3	4	5	6	Total
% Female	53.7	47.4	45.5	46.1	44.8	44.2	48.6
% White	10.7	28.1	17.49	9.6	7.1	4.9	13.3
% Black	28.5	22.4	44.0	9.6	36.8	58.0	29.9
% Hispanic	37.8	40.6	24.7	63.4	50.5	26.5	40.7
% Asian	23.0	8.9	13.9	17.3	5.7	10.6	16.1
% Free or Reduced Lunch ^a	73.1	83.3	75.6	75.0	62.7	88.9	75.8
% Received Accommodations ^a	9.1	13.5	14.7	13.4	15.1	27.9	13.5
% LEP ^b	4.0	24.7	17.1	10.3	.5	6.2	10.1
% Cohort 1 ^c	27.0	24.7	29.6	31.8	23.1	32.3	28.0
% Cohort 2 ^c	29.5	27.6	26.2	25.7	27.8	32.7	28.2
% Cohort 3 ^c	22.4	22.4	21.8	19.3	25.9	35.0	23.1
% Cohort 4 ^c	21.1	25.3	22.5	23.2	23.1	0	20.7
Total Students (<i>n</i>)	942	384	409	456	212	226	2629

Note. ^aFree or Reduced Lunch and Received Accommodations indicates that student qualified for free or reduced lunch or had an IEP or 504 accommodations in either one or both of the 15-16 or 16-17 school years. ^bLEP indicates student was designated with Limited English Proficiency at some point during their registration in the district. ^cCohort 1 = Fall 2015 6th grade, Cohort 2 = Fall 2015 7th grade, Cohort 3 = Fall 2015 8th grade, Cohort 4 = Fall 2016 6th grade.

Table 5
Percentage of Cohort with Complete Purpose Scale

	% Cohort Completed by Timepoint			
	Fall 2015	Spring 2016	Fall 2016	Spring 2017
Cohort 1 (<i>n</i> = 737)	72.7%	69.2%	73.0%	67.7%
Cohort 2 (<i>n</i> = 741)	81.5%	73.4%	52.1%	58.8%
Cohort 3 (<i>n</i> = 608)	90.6%	56.3%	Not Collected	Not Collected
Cohort 4 (<i>n</i> = 543)	Not Collected	Not Collected	82.7%	84.0%
Full Sample (<i>n</i> = 2629)	64.3%	53.1%	52.2%	52.9%
	% Cohort Completed by Number of Waves			
	1 Wave	2 Waves	3 Waves	4 Waves
Cohort 1 (<i>n</i> = 737)	12.9%	25.4%	28.0%	33.8%
Cohort 2 (<i>n</i> = 741)	12.6%	32.7%	31.2%	23.6%
Cohort 3 (<i>n</i> = 608)	53.1%	46.9%	Not Collected	Not Collected
Cohort 4 (<i>n</i> = 543)	33.3%	66.7%	Not Collected	Not Collected
Full Sample (<i>n</i> = 2629)	26.3%	40.9%	16.6%	16.1%

Note. Cohort 1 = Fall 2015 6th grade (Waves 1-4 expected), Cohort 2 = Fall 2015 7th grade (Waves 1-4 expected), Cohort 3 = Fall 2015 8th grade (Waves 1 and 2 only expected), Cohort 4 = Fall 2016 6th grade (Waves 3 and 4 only expected). Purpose Scale considered complete if students completed 75% of the scale.

Table 6
Correlations of Purpose Subscales

	1	2	3	4	5	6	7	8	9	10	11	12
1. Fall 2015 BTS ^a	-	.48** (n=1691)	.52** (n=1691)	.38** (n=1174)	.26** (n=1174)	.27** (n=1174)	.26** (n=714)	.21** (n=714)	.22** (n=714)	.26** (n=707)	.10** (n=707)	.11** (n=707)
2. Fall 2015 Eng. ^b		-	.64** (n=1691)	.34** (n=1174)	.39** (n=1174)	.36** (n=1174)	.25** (n=714)	.30** (n=714)	.25** (n=714)	.20** (n=707)	.21** (n=707)	.18** (n=707)
3. Fall 2015 Int. ^c			-	.33** (n=1174)	.31** (n=1174)	.40** (n=1174)	.22** (n=714)	.27** (n=714)	.30** (n=714)	.20** (n=707)	.21** (n=707)	.22** (n=707)
4. Spring 2016 BTS				-	.61** (n=1396)	.63** (n=1396)	.50** (n=662)	.36** (n=662)	.42** (n=662)	.31** (n=645)	.24** (n=645)	.26** (n=645)
5. Spring 2016 Eng.					-	.74** (n=1396)	.33** (n=662)	.40** (n=662)	.38** (n=662)	.26** (n=645)	.29** (n=645)	.30** (n=645)
6. Spring 2016 Int.						-	.34** (n=662)	.36** (n=662)	.43** (n=662)	.29** (n=645)	.28** (n=645)	.36** (n=645)
7. Fall 2016 BTS							-	.46** (n=1373)	.50** (n=1373)	.39** (n=1029)	.29** (n=1029)	.31** (n=1029)
8. Fall 2016 Eng.								-	.60** (n=1373)	.26** (n=1029)	.39** (n=1029)	.33** (n=1029)
9. Fall 2016 Int.									-	.32** (n=1029)	.35** (n=1029)	.39** (n=1029)
10. Spring 2017 BTS										-	.52** (n=1391)	.56** (n=1391)
11. Spring 2017 Eng.											-	.66** (n=1391)
12. Spring 2017 Int.												-

+ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

^aBTS= Beyond the Self ^bEng. = Engagement ^cInt. = Intention

=Note. Bold font indicates within timepoint correlations. Italic font indicates across timepoint comparisons for same subscale.

Table 7
Correlations of Purpose Total Scores

		M (SD)	1	2	3	4
1	Fall 2015 Purpose	21.08 (3.77)	1 (<i>n</i> =1691)	.456** (<i>n</i> = 1174)	.363** (<i>n</i> = 714)	.256** (<i>n</i> = 707)
2	Spring 2016 Purpose	20.50 (4.19)			.516** (<i>n</i> = 662)	.381** (<i>n</i> = 645)
3	Fall 2016 Purpose	20.79 (3.71)			1 (<i>n</i> =1373)	.477** (<i>n</i> = 1029)
4	Spring 2017 Purpose	20.08 (4.14)				1 (<i>n</i> =1391)

+ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 8
Correlations of Purpose and SDQ Scores in SDQ Subsample

		M (SD)	1	2	3	4	5	6	7	8
1	Fall 2015 Purpose	21.11 (3.74)	1	.45** (n=902)	.36** (n=583)	.26** (n=574)	-.14** (n=1213)	-.16** (n=1213)	-.119** (n=1213)	-.180** (n=1213)
2	Spring 2016 Purpose	20.67 (4.08)		1	.495** (n=473)	.359** (n=465)	-.160** (n = 902)	-.146** (n = 902)	-.156** (n = 902)	-.200** (n = 902)
3	Fall 2016 Purpose	20.55 (3.80)			1	.488** (n=432)	-.146** (n=583)	-.139** (n=583)	-0.058 (n=583)	-.148** (n=583)
4	Spring 2017 Purpose	19.733 (4.34)				1	-0.073 (n=574)	-.100* (n=574)	-0.034 (n=574)	-.089* (n=574)
5	SDQ: Emotional Problems	7.20 (2.03)					1	.402** (n=1213)	.382** (n=1213)	.805** (n=1213)
6	SDQ: Conduct Problems	4.53 (1.51)						1	.455** (n=1213)	.758** (n=1213)
7	SDQ: Peer Problems	6.07 (1.69)							1	.772** (n=1213)
8	Fall 2015 SDQ Total ^a	17.89 (4.08)								1

+ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Note. ^aSDQ = Strengths and Difficulties Questionnaire. Correlation analyses were carried out as part of preliminary analyses before running multiple imputation. Results reported here are from dataset using complete cases.

Table 9
Building Multilevel Model for Change in Student Self-Reported Purpose Scores

			Model 1 Unconditional Means Model Coeff. (s.e.)	Model 2 Unconditional Growth Model Coeff. (s.e.)	Model 3 Covariates Coeff. (s.e.)	Model 4 Gender Coeff. (s.e.)	Model 5 Race Coeff. (s.e.)	Model 6 School Status Coeff. (s.e.)	Model 7 GPA Coeff. (s.e.)
Fixed Effects	Intercept	β_{00}	20.61 (.06)***	21.09 (.09)***	21.45 (.14)***	21.25 (.17) ***	20.99 (.28)***	21.31 (.29)***	21.39 (.29)***
Initial status, π_{0i}	Cohort ^a								
	Cohort 2	β_{01}			-.78 (.21)***	-.79 (.21)***	-.78 (.21)***	-.77 (.21)***	-.78 (.21)***
	Cohort 3	β_{02}				-.94 (.22)***	-.98 (.22)***	-.97 (.22)***	-1.01(.22)***
	Cohort 4	β_{03}			-.94 (.22)*** 1.16 (.51) *	1.17 (.51)*	1.20 (.51)*	1.10 (.51)*	1.03 (.51)+
Race ^d	Ambassador ^b	β_{04}			.95 (.25)***	.93 (.25)***	.92 (.25)***	.92 (.25)***	.80 (.25)**
	Gender ^c	β_{05}				.43 (.17)*	.42 (.17)*	.43 (.17)*	.25 (.18) ns
	Black	β_{06}					.69 (.29) *	.79 (.29) **	.96 (.29) ***
	Latino	β_{07}					.03 (.28) ns	-.002 (.28) ns	-.06 (.27) ns
	Asian	β_{08}					.29 (.31) ns	.38 (.31) ns	.15(.31) ns
	School Status ^e	β_{09}						-.59 (.18) **	-.59 (.18) **
	Achieve History ^f	β_{010}							.05 (.01)***
Rate of Change, π_{1i}	Intercept	β_{10}		-.34 (.04)***	-.53 (.07)***	-.50 (.08) ***	-.26 (.14) +	-.31(.15) *	-.30(.15) *
Cohort	Cohort 2	β_{11}			.11 (.10) ns	.11 (.10) ns	.10 (.10) ns	.10 (.10) ns	.10 (.10) ns
	Cohort 3	β_{12}			.44 (.23)+	.44 (.23) +	.46 (.24) +	.45 (.23) +	.45 (.23) +
	Cohort 4	β_{13}			-.13 (.21) ns	-.13 (.21) ns	-.15 (.21) ns	-.11 (.21) ns	-.10 (.21) ns
	Ambassador	β_{14}			-.03 (.13) ns	-.02 (.13) ns	-.01 (.21) ns	-.04 (.13) ns	-.04 (.13) ns
Race	Gender	β_{15}				-.07 (.09) ns	-.07 (.08) ns	-.10 (.09) ns	-.10 (.09) ns
	Black	β_{16}					-.43 (.15)**	-.40 (.15)**	-.40 (.15)**
	Latino	β_{17}					-.25 (.14)+	-.22 (.14)	-.22 (.14) ns
	Asian	β_{18}					-.09 (.15) ns	-.13 (.15) ns	-.13 (.15) ns
	School Status	β_{19}						.07 (.09) ns	.08 (.09) ns
	Achieve History	β_{110}							.01 (.01) +

Variance Components									
Level 1	Within person	e	9.50	8.04	8.03	8.02	8.03	8.03	8.01
Level 2	In initial status	r_{0i}	6.23***	7.42***	7.15***	7.11***	7.04***	6.96***	6.89***
	In rate of change	r_{1i}		.78***	.76***	.76***	.74***	.74***	.74***
	Covariance			-.84	-.81	-.80	-.79	-.78	-.80
Pseudo R^2									
	R^2_e			0.15	.001	0.001	-0.001	0	0.002
	R^2_0			-0.19	.04	0.006	0.01	0.01	0.01
	R^2_l				.03	0	0.03	0	0
	R^2_{cov}				.04	0.01	0.01	0.01	-.03
Goodness of Fit Statistics									
	Deviance		32045.78	31909.79	31829.72	31822.17	31800.28	31787.05	31747.37
	AIC		32051.78	31921.79	31857.72	31854.17	31844.28	31835.05	31799.37
	BIC		32069.40	31957.04	31939.961	31948.16	31973.52	31976.04	31952.10

+ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Note. ^aCohort dummy coded: Dummy-coded with Cohort 1 (largest group) as reference (Cohort 1: 6th graders 2015-2016, 7th graders 2016-2017; Cohort 2: 7th graders 2015-2016, 8th graders 2016-2017; Cohort 3: 8th graders 2015-2016; Cohort 4: 6th graders 2016-2017. ^b Ambassador: 0 = Never MOSAIC Ambassador; 1 = MOSAIC Ambassador during 2015-2016 or 2016-2017; ^cGender coded: 0 = Male; 1 = Female; ^dRace dummy coded: Dummy-coded with White as reference group (Black, Latino, Asian, White) ^eSchool Status: 0 = School was neither Focus nor Priority (2, 4, 5); 1 = School was Focus or Priority (1, 3, 6) ^fAchievement History (grand-mean centered): Academic grades (reading, writing, math, science, social studies) reported by district in 2015-2016 and 2016-2017 averaged to create indicator of overall achievement

Table 10
Final Models Predicting Purpose Total Score and Purpose Dimensions Over Time

			Purpose Total Coeff. (s.e.)	Beyond-the-Self Coeff. (s.e.)	Intention Coeff. (s.e.)	Engagement Coeff. (s.e.)
Fixed Effects	Intercept	β_{00}	21.44 (.28)***	3.91 (.08)***	8.49 (.13)***	8.99 (.06)***
Initial status, π_{0i}						
Cohort ^a	Cohort 2	β_{01}	-.77 (.21)***	-.17 (.06) **	-.31 (.10)**	-.26 (.07)***
	Cohort 3	β_{02}	-1.02 (.29)***	-.19 (.06) **	-.50 (.10)***	-.29 (.08)***
	Cohort 4	β_{03}	1.03 (.51)*	.01 (.17) ns	.33 ns	.30 (.08)***
	Ambassador ^b	β_{04}	.75 (.20)***	.20 (.06)***	.29 (.09)**	.27 (.08)***
Race ^c	Black	β_{05}	1.01 (.29)***	.34 (.08) ***	.33 (.14)*	(not included)
	Latino	β_{06}	.09 (.27) ns	.04 (.08) ns	.06 (.13) ns	
	Asian	β_{07}	.05 (.31) ns	-.04 (.09) ns	.11 (.15) ns	
	School Status ^d	β_{08}	-.49 (.14)***	-.06 (.04) +	(not included)	-.22 (.05)***
	Achieve History ^e	β_{09}	.07 (.01)***	.02 (.003)***	(not included)	.04 (.004)***
Rate of Change, π_{1i}	Intercept	β_{10}	-.30(.13)*	-.01 (.04)ns	-.75 (.22) ***	-.12 (.04)***
Cohort ^a	Cohort 2	β_{11}	.09 (.10) ns	.04 (.03) ns	.03 (.05) ns	(not included)
	Cohort 3	β_{12}	.46 (.23)*	.15 (.06)*	.26 (.11)*	
	Cohort 4	β_{13}	-.10 (.21)ns	.04 (.07) ns	.02 (.11) ns	
Race ^c	Black	β_{14}	-.42 (.15)**	-.13 (.04)**	-.14(.07)+	-.05 (.04) ns
	Latino	β_{15}	-.24 (.14) +	-.08 (.04)+	-.09 (.07) ns	-.08 (.04)+
	Asian	β_{16}	-.07 (.15) ns	-.001 (.04) ns	-.06 (.07) ns	-.03 (.04) ns
	Achieve History ^e	β_{17}	(not included)	(not included)	.01 (.003) **	(not included)
Variance Components						
Level 1	Within person	e	8.01	.85	1.42	1.27
Level 2	In initial status	r_0	6.92 ***	.72 ***	1.16 ***	.96 ***
	In rate of change	r_i	.75 ***	.21 ***	.37 ***	.36 ***
	Covariance		-.82	-.07	-.14	-.12

+ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Note. ^aCohort dummy coded: Dummy-coded with Cohort 1 (largest group) as reference (Cohort 1: 6th graders 2015-2016, 7th graders 2016-2017; Cohort 2: 7th graders 2015-2016, 8th graders 2016-2017; Cohort 3: 8th graders 2015-2016; Cohort 4: 6th graders 2016-2017. ^b Ambassador: 0 = Never MOSAIC Ambassador; 1 = MOSAIC Ambassador during 2015-2016 or 2016-2017; ^c Race dummy-coded with White as reference group (Black, Latino, Asian, White ^dSchool Status: 0 = School was neither Focus nor Priority (2, 4, 5); 1 = School was Focus or Priority (1, 3, 6) ^eAchievement History (grand-mean centered): Academic grades (reading, writing, math, science, social studies) reported by district in 2015-2016 and 2016-2017 averaged to create indicator of overall achievement

Table 11
Mental Health Predicting Purpose Total Score Over Time

			Purpose Total Coeff. (s.e.)
Fixed Effects	Intercept	β_{00}	21.06 (.09)***
Initial status, π_{0i}	SDQ Total Score	β_{01}	-.17 (.03)***
Rate of Change, π_{1i}	Intercept	β_{10}	-.47 (.052)***
	SDQ Total Score	β_{11}	.02 (.02) ns
Variance Components			
Level 1	Within person	e	7.90
Level 2	In initial status	r_0	6.93***
	In rate of change	r_i	.89***
	Covariance		-.61

+ $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Note. Results reported are from a dataset comprised of averaged multiple imputation values for the SDQ scale. Complete case analysis and analyses of each of the five imputed datasets yielded similar results.

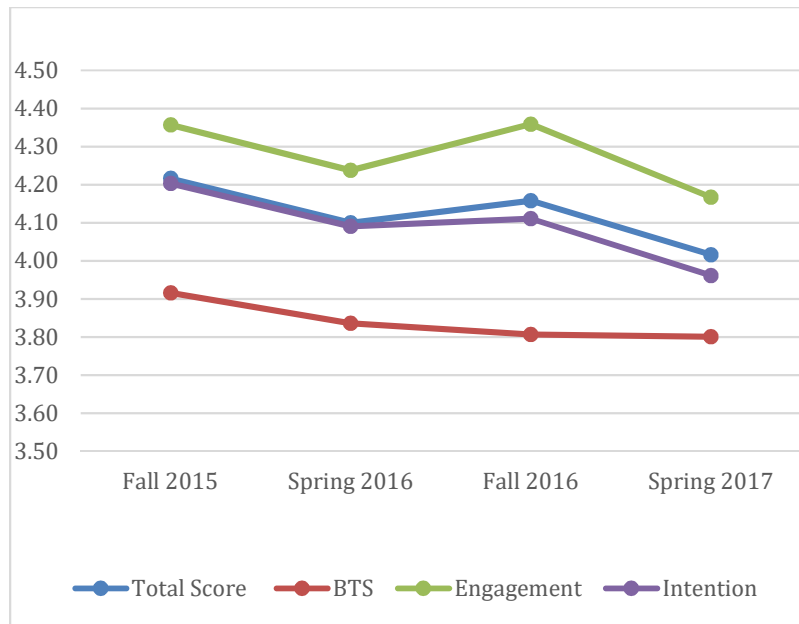
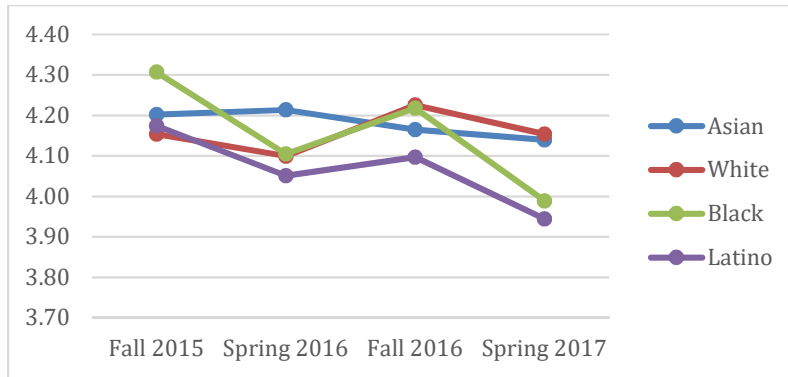
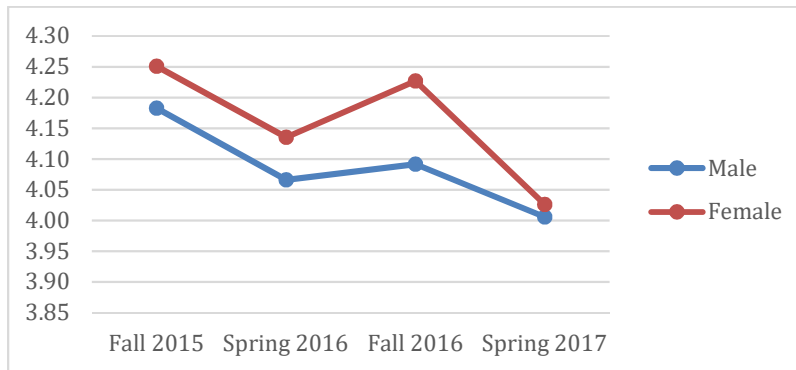


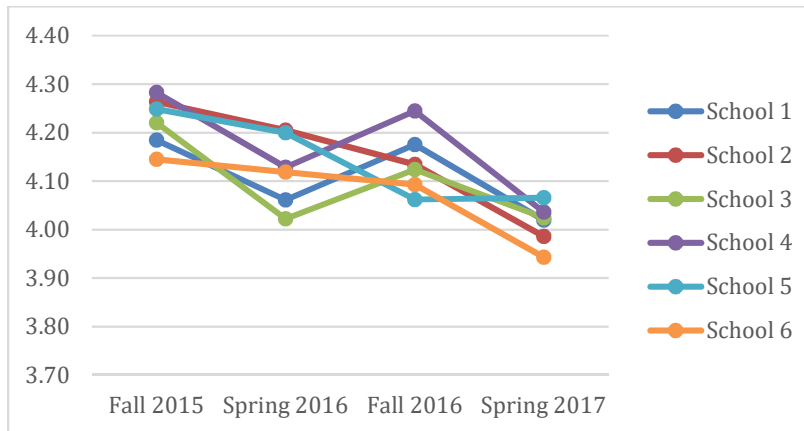
Figure 1. Mean Purpose Scores by Total Score and Subscales. Purpose mean scores reflect averages of items on the total scale or the subscales: Engagement, Intention, BTS (Beyond the Self).



2a



2b



2c.

Figure 2. Mean Purpose Scores by (a) Race/Ethnicity, (b) Gender, and (c) School.

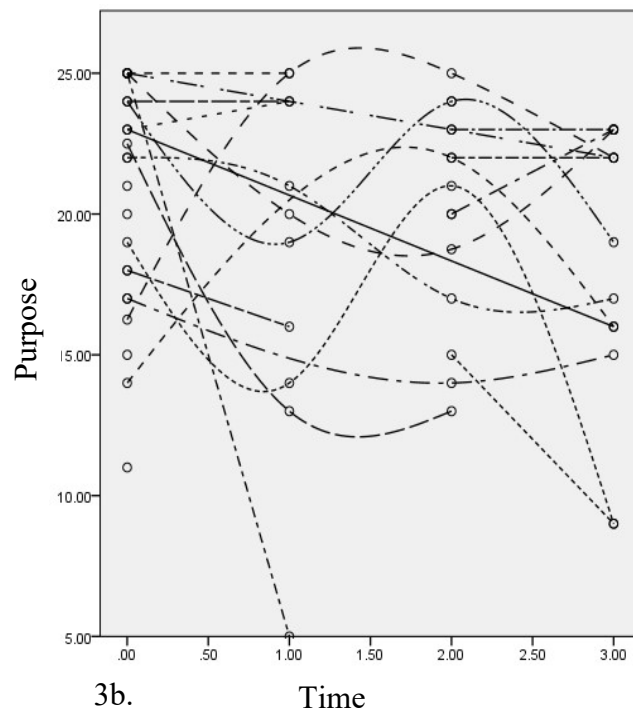
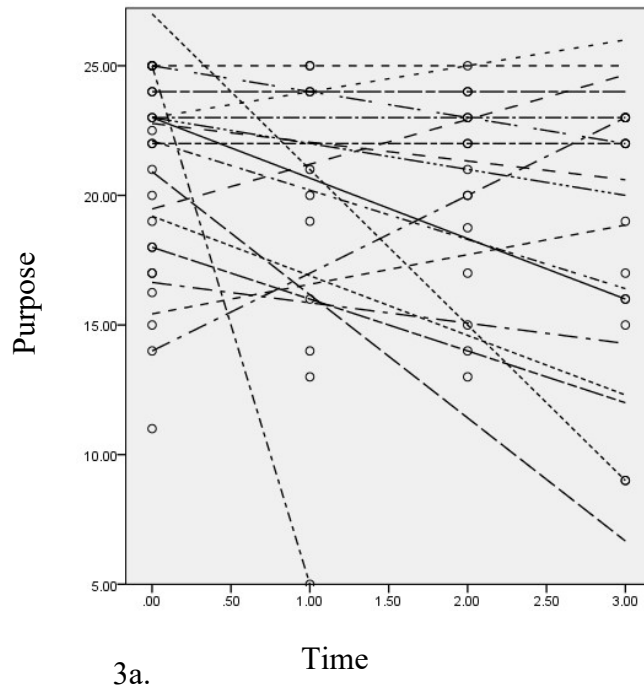
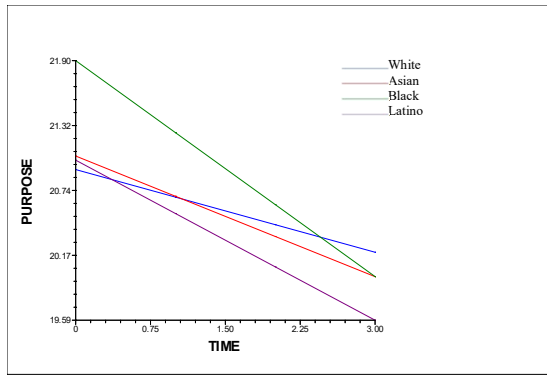
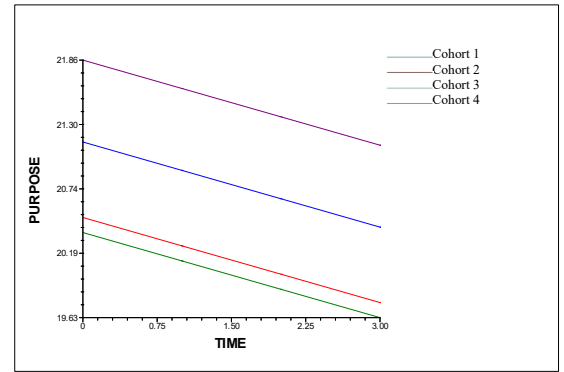


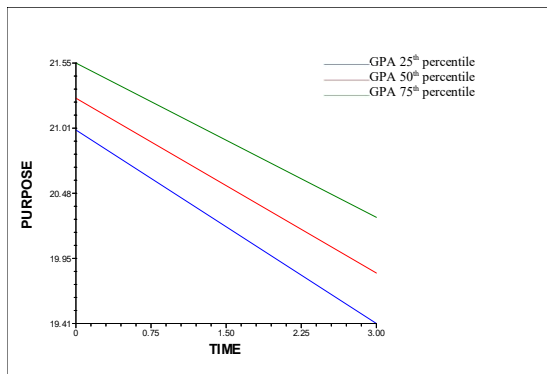
Figure 3. Randomly Selected 30 Cases Fitted with (a) Linear and (b) Curvilinear Regression Lines.



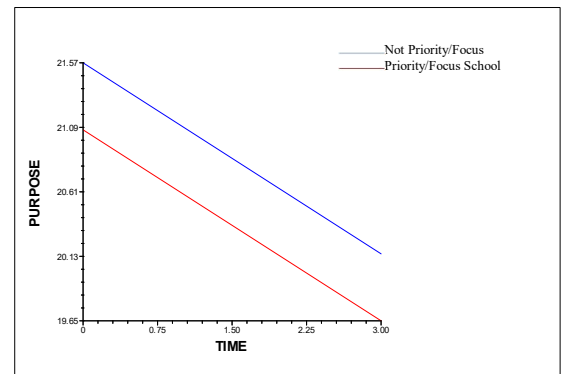
4a.



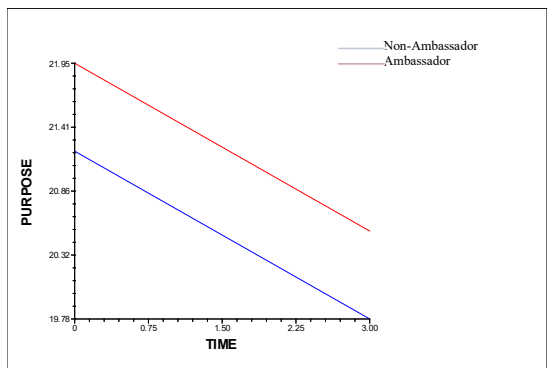
4b.



4c.



4d.



4e.

Figure 4. Graphical Representation of Final Model Predicting Self-Reported Purpose Over Time. Race/ethnicity (a), cohort (b), academic achievement history (c), school status (priority or focus school) (d), and Ambassador status (did student participate in the Ambassador program) (e) predicted initial status of purpose. Only race/ethnicity, specifically being a Black student, was related to the slope of purpose over time.

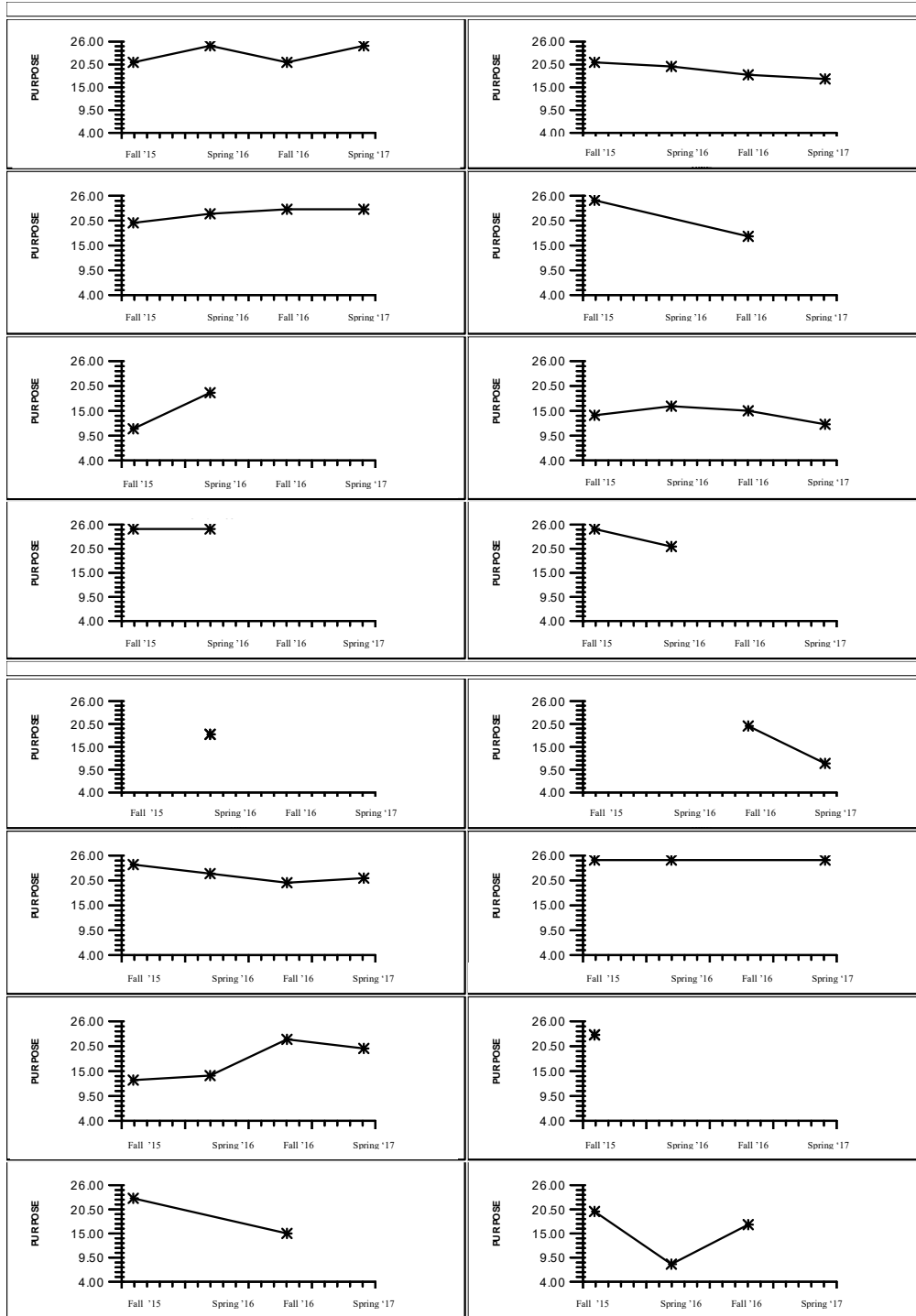


Figure 5. Individual Plots for Randomly Selected Cases after Analysis.

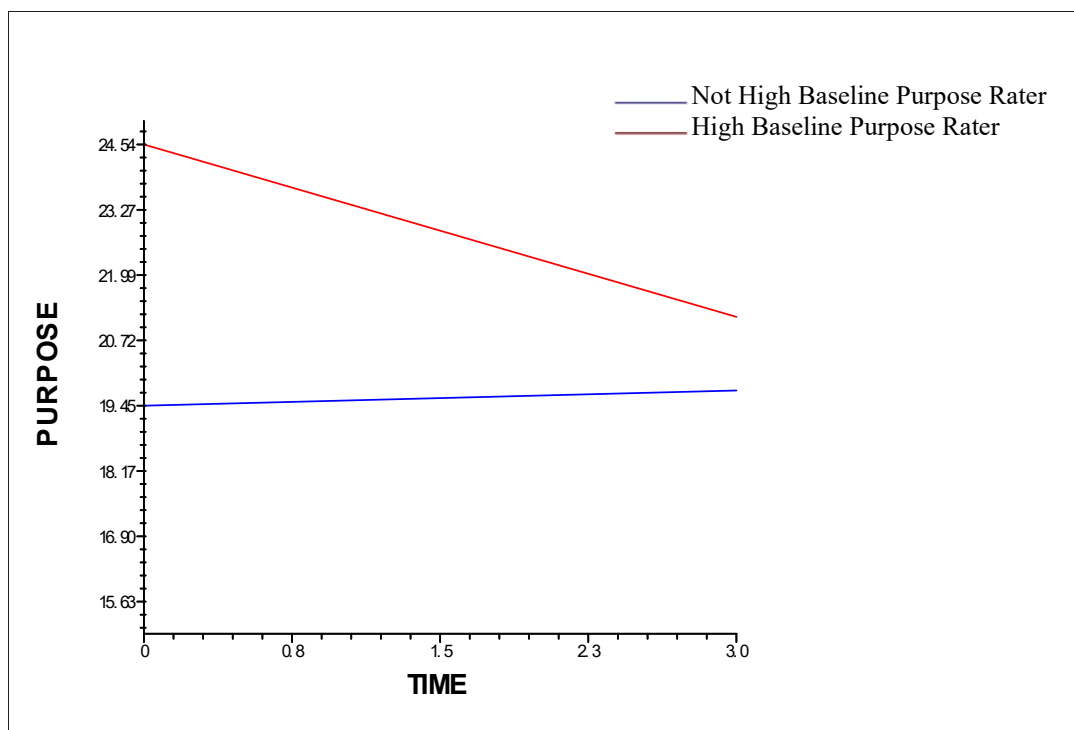


Figure 6. Hierarchical Linear Model with Baseline High Purpose Raters as a Dichotomous Grouping Variable.

LIST OF APPENDICES

Appendix A. Taxonomy of Multilevel Models for Purpose.....	79
--	----

Appendix A

Taxonomy of Multilevel Models for Purpose

Model	Level-1 model	Level-2 model(s)	Composite model
1	$PURPOSE = \pi_{0i} + e_{0i}$	$\pi_{0i} = \beta_{00} + r_{0i}$	$PURPOSE_i = \beta_{00} + r_{0i} + e_{0i}$
2	$PURPOSE = \pi_{0i} + \pi_{1i}*(TIME) + e_{it}$	$\pi_{0i} = \beta_{00} + r_{0i}$ $\pi_{1i} = \beta_{10} + r_{1i}$	$PURPOSE_{ti} = \beta_{00} + \beta_{10}*TIME_{ti} + r_{0i} + r_{1i}*(TIME) + e_{ti}$
3	$PURPOSE = \pi_{0i} + \pi_{1i}*(TIME) + e_{it}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(COH2_i) + \beta_{02}*(COH3_i) + \beta_{03}*(COH4_i) + \beta_{04}*(AMB_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(COH2_i) + \beta_{12}*(COH3_i) + \beta_{13}*(COH4_i) + \beta_{14}*(AMB_i) + r_{1i}$	$PURPOSE_{ti} = \beta_{00} + \beta_{01}*COH2_i + \beta_{02}*COH3_i + \beta_{03}*COH4_i + \beta_{04}*AMB_i + \beta_{10}*TIME_{ti} + \beta_{11}*COH2_i*TIME_{ti} + \beta_{12}*COH3_i*TIME_{ti} + \beta_{13}*COH4_i*TIME_{ti} + \beta_{14}*AMB_i*TIME_{ti} + r_{0i} + r_{1i}*TIME_{ti} + e_{ti}$
4	$PURPOSE = \pi_{0i} + \pi_{1i}*(TIME) + e_{it}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(COH2_i) + \beta_{02}*(COH3_i) + \beta_{03}*(COH4_i) + \beta_{04}*(AMB_i) + \beta_{05}*(GENDER_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(COH2_i) + \beta_{12}*(COH3_i) + \beta_{13}*(COH4_i) + \beta_{14}*(AMB_i) + \beta_{15}*(GENDER_i) + r_{1i}$	$PURPOSE_{ti} = \beta_{00} + \beta_{01}*COH2_i + \beta_{02}*COH3_i + \beta_{03}*COH4_i + \beta_{04}*AMB_i + \beta_{05}*GENDER_i + \beta_{10}*TIME_{ti} + \beta_{11}*COH2_i*TIME_{ti} + \beta_{12}*COH3_i*TIME_{ti} + \beta_{13}*COH4_i*TIME_{ti} + \beta_{14}*AMB_i*TIME_{ti} + \beta_{15}*GENDER_i*TIME_{ti} + r_{0i} + r_{1i}*TIME_{ti} + e_{ti}$
5	$PURPOSE = \pi_{0i} + \pi_{1i}*(TIME) + e_{it}$	$\pi_{0i} = \beta_{00} + \beta_{01}*(COH2_i) + \beta_{02}*(COH3_i) + \beta_{03}*(COH4_i) + \beta_{04}*(AMB_i) + \beta_{05}*(GENDER_i) + \beta_{06}*(BLACK_i) + \beta_{07}*(LATINO_i) + \beta_{08}*(ASIAN_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11}*(COH2_i) + \beta_{12}*(COH3_i) + \beta_{13}*(COH4_i) + \beta_{14}*(AMB_i) + \beta_{15}*(GENDER_i) + \beta_{16}*(BLACK_i) + \beta_{17}*(LATINO_i) + \beta_{18}*(ASIAN_i) + r_{1i}$	$PURPOSE_{ti} = \beta_{00} + \beta_{01}*(COH2_i) + \beta_{02}*(COH3_i) + \beta_{03}*(COH4_i) + \beta_{04}*(AMB_i) + \beta_{05}*(GENDER_i) + \beta_{06}*(BLACK_i) + \beta_{07}*(LATINO_i) + \beta_{08}*(ASIAN_i) + \beta_{10}*TIME_{ti} + \beta_{11}*COH2_i*TIME_{ti} + \beta_{12}*COH3_i*TIME_{ti} + \beta_{13}*COH4_i*TIME_{ti} + \beta_{14}*AMB_i*TIME_{ti} + \beta_{15}*GENDER_i*TIME_{ti} + \beta_{16}*BLACK_i*TIME_{ti} + \beta_{17}*LATINO_i*TIME_{ti} + \beta_{18}*ASIAN_i*TIME_{ti} + r_{0i} + r_{1i}*TIME_{ti} + e_{ti}$

6	$PURPOSE = \pi_{0i} + \pi_{1i} * (TIME) + e_{it}$	$\pi_{0i} = \beta_{00} + \beta_{01} * (COH2_i) + \beta_{02} * (COH3_i) + \beta_{03} * (COH4_i) + \beta_{04} * (AMB_i) + \beta_{05} * (GENDER_i) + \beta_{06} * (BLACK_i) + \beta_{07} * (LATINO_i) + \beta_{08} * (ASIAN_i) + \beta_{09} * (SCH_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11} * (COH2_i) + \beta_{12} * (COH3_i) + \beta_{13} * (COH4_i) + \beta_{14} * (AMB_i) + \beta_{15} * (GENDER_i) + \beta_{16} * (BLACK_i) + \beta_{17} * (LATINO_i) + \beta_{18} * (ASIAN_i) + \beta_{19} * (SCH_i) + r_{1i}$	$PURPOSE_{it} = \beta_{00} + \beta_{01} * (COH2_i) + \beta_{02} * (COH3_i) + \beta_{03} * (COH4_i) + \beta_{04} * (AMB_i) + \beta_{05} * (GENDER_i) + \beta_{06} * (BLACK_i) + \beta_{07} * (LATINO_i) + \beta_{08} * (ASIAN_i) + \beta_{09} * (SCH_i) + \beta_{10} * TIME_{it} + \beta_{11} * COH2_i * TIME_{it} + \beta_{12} * COH3_i * TIME_{it} + \beta_{13} * COH4_i * TIME_{it} + \beta_{14} * AMB_i * TIME_{it} + \beta_{15} * GENDER_i * TIME_{it} + \beta_{16} * BLACK_i * TIME_{it} + \beta_{17} * LATINO_i * TIME_{it} + \beta_{18} * ASIAN_i * TIME_{it} + \beta_{19} * (SCH_i) * TIME_{it} + r_{0i} + r_{1i} * TIME_{it} + e_{it}$
7	$PURPOSE = \pi_{0i} + \pi_{1i} * (TIME) + e_{it}$	$\pi_{0i} = \beta_{00} + \beta_{01} * (COH2_i) + \beta_{02} * (COH3_i) + \beta_{03} * (COH4_i) + \beta_{04} * (AMB_i) + \beta_{05} * (GENDER_i) + \beta_{06} * (BLACK_i) + \beta_{07} * (LATINO_i) + \beta_{08} * (ASIAN_i) + \beta_{09} * (SCH_i) + \beta_{010} * (ACH_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11} * (COH2_i) + \beta_{12} * (COH3_i) + \beta_{13} * (COH4_i) + \beta_{14} * (AMB_i) + \beta_{15} * (GENDER_i) + \beta_{16} * (BLACK_i) + \beta_{17} * (LATINO_i) + \beta_{18} * (ASIAN_i) + \beta_{19} * (SCH_i) + \beta_{110} * (ACH_i) + r_{1i}$	$PURPOSE_{it} = \beta_{00} + \beta_{01} * (COH2_i) + \beta_{02} * (COH3_i) + \beta_{03} * (COH4_i) + \beta_{04} * (AMB_i) + \beta_{05} * (GENDER_i) + \beta_{06} * (BLACK_i) + \beta_{07} * (LATINO_i) + \beta_{08} * (ASIAN_i) + \beta_{09} * (SCH_i) + \beta_{110} * (ACH_i) + \beta_{10} * TIME_{it} + \beta_{11} * COH2_i * TIME_{it} + \beta_{12} * COH3_i * TIME_{it} + \beta_{13} * COH4_i * TIME_{it} + \beta_{14} * AMB_i * TIME_{it} + \beta_{15} * GENDER_i * TIME_{it} + \beta_{16} * BLACK_i * TIME_{it} + \beta_{17} * LATINO_i * TIME_{it} + \beta_{18} * ASIAN_i * TIME_{it} + \beta_{19} * (SCH_i) * TIME_{it} + \beta_{110} * (ACH_i) * TIME_{it} + r_{0i} + r_{1i} * TIME_{it} + e_{it}$
Final Model	$PURPOSE = \pi_{0i} + \pi_{1i} * (TIME) + e_{it}$	$\pi_{0i} = \beta_{00} + \beta_{01} * (COH2_i) + \beta_{02} * (COH3_i) + \beta_{03} * (COH4_i) + \beta_{04} * (AMB_i) + \beta_{06} * (BLACK_i) + \beta_{07} * (LATINO_i) + \beta_{08} * (ASIAN_i) + \beta_{09} * (SCH_i) + \beta_{010} * (ACH_i) + r_{0i}$ $\pi_{1i} = \beta_{10} + \beta_{11} * (COH2_i) + \beta_{12} * (COH3_i) + \beta_{13} * (COH4_i) + \beta_{16} * (BLACK_i) + \beta_{17} * (LATINO_i) + \beta_{18} * (ASIAN_i) + \beta_{110} * (ACH_i) + r_{1i}$	$PURPOSE_{it} = \beta_{00} + \beta_{01} * (COH2_i) + \beta_{02} * (COH3_i) + \beta_{03} * (COH4_i) + \beta_{04} * (AMB_i) + \beta_{06} * (BLACK_i) + \beta_{07} * (LATINO_i) + \beta_{08} * (ASIAN_i) + \beta_{09} * (SCH_i) + \beta_{110} * (ACH_i) + \beta_{10} * TIME_{it} + \beta_{11} * COH2_i * TIME_{it} + \beta_{12} * COH3_i * TIME_{it} + \beta_{13} * COH4_i * TIME_{it} + \beta_{16} * BLACK_i * TIME_{it} + \beta_{17} * LATINO_i * TIME_{it} + \beta_{18} * ASIAN_i * TIME_{it} + \beta_{110} * (ACH_i) * TIME_{it} + r_{0i} + r_{1i} * TIME_{it} + e_{it}$

Note. COH2, COH3, COH4 = Cohort 2, Cohort 3, Cohort 4; AMB = Student Ambassador Program Status; SCH = School Status (Priority/Focus School); ACH = Academic Achievement