

THREE PERSPECTIVES ON POSTSECONDARY ATTAINMENT: STUDENTS,
TEACHERS, AND INSTITUTIONS

By

Michael D. Broda

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degrees of

Curriculum, Instruction, and Teacher Education - Doctor of Philosophy
Educational Policy - Doctor of Philosophy

2015

ABSTRACT

THREE PERSPECTIVES ON POSTSECONDARY ATTAINMENT: STUDENTS, TEACHERS, AND INSTITUTIONS

By

Michael D. Broda

This dissertation collects three independent but related empirical studies that explore different dimensions of postsecondary attainment in the United States. The first study uses multilevel models to examine predictors of nonresponse when using the experience sampling method (ESM), a form of momentary data collection that captures participants' situational thoughts, feelings, and emotions. Because ESM approaches often seek to generalize and compare participants' emotional states across days and times, it is important to understand how and when participants may miss response opportunities, and further to explore how this response bias may limit generalizability of findings. Results from this study, conducted in three mid-Michigan high schools in 2013-2014 with a sample of more than 140 students, find that time of day and day of week are significantly related to a given participant's odds of nonresponse. Specifically, ESM "beeps" occurring after school and over the weekend had much higher odds of being missed by participants, even after controlling for other covariates such as race/ethnicity, gender, and person-level emotional trends. These findings suggest if day and time contextual factors are significantly related to odds of nonresponse, then research approaches that seek to compare widely different time contexts should be mindful of possible generalizability concerns.

The second study reports on initial findings from the first year of a matched pair of social psychological interventions designed to improve students' postsecondary

persistence and academic success. These interventions, which target students' mindset regarding intelligence and their sense of social belonging, were implemented in a randomized sample of more than 6,000 incoming first-year students at a large university in mid-Michigan in summer 2014. Initial results suggest that the randomization achieved strong balance between treatment and control groups on all measured covariates, and rates of attrition (both overall and differential) were less than one percent in all cases, suggesting strong randomized properties. After one semester, the belonging intervention saw no positive main effects on GPA, credits attempted, or credits completed, while students in the mindset treatment group had slightly higher GPAs (0.15 points) compared to students in the control. These small but significant main effects after only one semester suggest that mindset interventions such as this may be a low-cost and minimally invasive approach to improving students' academic success.

Finally, a third study examines the impact of high school teachers' colleagues on their own beliefs and practices related to college going. Using two time-point survey data collected from 104 teachers in four mid-Michigan high schools in 2011-2012, this study employs cluster analysis to map collegial networks by school and influence models to examine how teachers are impacted by their closest colleagues. Results of the cluster analysis suggest that all four high schools show evidence of distinct clustering by subgroups. Results from influence models show that while positive in almost all cases, the impact of teachers' exposure to colleagues is not significantly related to a change in beliefs or practices. However, interaction with college-advising program staff is found to be related to a small positive change in some teacher practices.

ACKNOWLEDGEMENTS

Many outstanding people have contributed to both my thinking and the development of this work over the past six years. First, I must thank my dissertation committee: Ken Frank, Peter Youngs, John Yun, and of course my advisor, Barbara Schneider, who all provided essential feedback and pushed my development of these three essays. Barbara in particular has been invaluable to me both for her support of my work and her investment in my development as an intellectual- thank you so much!

I would also like to thank a few of the many friends and colleagues with whom I have had the pleasure of spending time while at MSU. Members of my book club- Andy Saultz, John Lane, Alan Hastings, Justina Judy Spicer, Ben Creed, and Alyssa Morley- have all been wonderful sounding boards and sources of motivation and inspiration, as have many colleagues in the Office of the Hannah Chair- Justin Bruner, Guan Saw, Heather Johnson, I-Chien Chen, Wei Li, and Kri Burkander.

Finally, I must thank my wonderful family, who have endured years of progress updates and sharing of highs and lows along the way. Your love and support have meant the world to me- I truly couldn't have accomplished this without you!

PREFACE

This research is supported by grants from the National Science Foundation (DRL-1316702 and DRL-1255807). The opinions expressed here are those of the author and do not necessarily reflect those of the National Science Foundation.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES	xi
 CHAPTER 1: USING MULTILEVEL MODELS TO EXPLORE PREDICTORS OF NONRESPONSE FROM EXPERIENCE SAMPLING METHOD (ESM) STUDIES	
Introduction.....	1
Related Literature.....	2
Nonresponse in Related Survey Research	2
Nonresponse in ESM Research.....	3
Missing Data and Generalizability.....	4
Research Questions	6
Methods.....	6
Data Collection	6
Sample.....	8
Measures	8
Study outcome: Missed beeps.....	8
Time-variant predictors.....	9
Time-invariant predictors.....	10
Analyses	10
Calculating Predicted Probabilities.....	13
Missing Data	14
Results.....	15
Model 1: Level-1 Time and Day Predictors	15
Model 2: Student Demographics and School Assignment	16
Model 3: Person-Level Emotional States	16
Model 4: Full Model, with Lagged Emotional States.....	17
Predicted Probabilities	18
Discussion	18
Limitations	20
Implications for Research and Practice.....	21
APPENDICES	23
Appendix A: Figures.....	24
Appendix B: Tables	27
Appendix C: Instrument.....	31
REFERENCES	32
 CHAPTER 2: A PRELIMINARY ANALYSIS OF THE EFFECTS OF RANDOMIZED MINDSET AND BELONGING INTERVENTIONS ON FRESHMEN STUDENTS’ ACADEMIC SUCCESS.....	
Introduction.....	35
Relevant Literature.....	36
Research Questions.....	39

Methods.....	39
Study Characteristics	39
Mindset Intervention Condition.....	40
Belonging Intervention Condition	41
Comparison Condition.....	41
Setting	42
Sample Formation.....	42
Outcome Measures.....	42
Analytic Approach.....	43
Model 1: Basic Treatment Model	45
Model 2: Treatment Plus Engagement.....	45
Model 3: Treatment, Engagement, and Demographic Characteristics	45
Model 4: Treatment, Engagement, Demographic Characteristics, and Achievement Pre-Measures.....	46
Model 5: Relevant Interaction Effects	46
Statistical Adjustments.....	46
Missing Data	47
Results.....	47
Equivalence of Treatment and Control Groups at Time 1.....	47
Equivalence of Treatment and Control Groups at Time 2.....	48
Attrition Rates.....	49
Initial Results- Mean Differences	50
Regression Models.....	50
Belonging Intervention	51
Mindset Intervention.....	52
Note on Survey Engagement Predictor.....	55
Discussion	55
Limitations	57
Conclusion	58
APPENDIX.....	60
REFERENCES	79

CHAPTER 3: TEACHERS' SOCIAL NETWORKS, COLLEGE ATTAINMENT, AND THE DIFFUSION OF A SCHOOL-BASED REFORM INITIATIVE.....	82
Introduction.....	82
College-going interventions and school-wide college culture.....	84
Research Questions.....	87
Theory and Hypothesis	88
Theoretical Framework.....	88
Methods.....	90
Sample.....	90
Development of Survey Instruments	90
Measures	91
Focal Covariates.....	93
Other covariates	94
Within-School Cluster Analysis	94

Modeling teacher network effects.....	95
Missing Data	97
Results.....	97
Within-School Cluster Analysis and Sociograms.....	97
Influence Model for Teacher Practices and Beliefs.....	98
Robustness Check- What Would it take to change the inference?	99
Discussion and Conclusion	100
Limitations	101
APPENDICES	103
Appendix A: Figures.....	104
Appendix B: Tables	106
Appendix C: Instrument.....	117
REFERENCES	119

LIST OF TABLES

Table 1.1 Sample Characteristics for High Schools in the Study.....	27
Table 1.2 Descriptive Statistics for the Analytical Sample.....	27
Table 1.3 Results of 2-Level HGLM Predicting a Missed Beep, Conditional (Unit-Specific) Model.....	28
Table 1.4 Results of 2-Level HGLM Predicting a Missed Beep, Marginal (Population-Average) Model.....	29
Table 1.5 Predicted Probabilities Based on Model 2 Log Likelihood Estimates.....	30
Table 2.1 Descriptive Statistics for and Outcomes and Controls used in Regression Analysis.....	60
Table 2.2 Descriptive Statistics for Pre-Randomization Questions at Time 1, Belonging Experiment.....	61
Table 2.3 Descriptive Statistics for Pre-Randomization Questions at Time 1, Mindset Experiment.....	62
Table 2.4 Descriptive Statistics for Pre-Randomization Questions at Time 2, Belonging Experiment.....	63
Table 2.5 Descriptive Statistics for Pre-Randomization Questions at Time 2, Mindset Experiment.....	64
Table 2.6 Sample Sizes and Attrition Rates for Mindset and Belonging Experiments...	65
Table 2.7 Estimated Treatment Effects, Belonging Experiment.....	66
Table 2.8 Estimated Treatment Effects, Mindset Experiment.....	66
Table 2.9 Regression Estimates for Impact of Belonging Treatment On Fall GPA.....	67
Table 2.10 Regression Estimates for Impact of Belonging Treatment on Credits Completed.....	69
Table 2.11 Regression Estimates for Impact of Belonging Treatment on Credits Attempted.....	71

Table 2.12 Regression Estimates for Impact of Mindset Treatment on Fall GPA.....	73
Table 2.13 Regression Estimates for Impact of Mindset Treatment on Credits Completed.....	75
Table 2.14 Regression Estimates for Impact of Mindset Treatment on Credits Attempted	77
Table 3.1 Relevant School Characteristics for Study Sample.....	106
Table 3.2 Descriptive Statistics for the Analytical Sample	107
Table 3.3 Regression Estimates for the Influence of Colleagues and CAP Staff on Teachers’ Practices (Sum)	109
Table 3.4 Regression Estimates for the Influence of Colleagues and CAP Staff on Teachers’ Practices (Mean).....	110
Table 3.5 Regression Estimates for the Influence of Colleagues and CAP Staff on Number of Recommendation Letters Written.....	111
Table 3.6 Regression Estimates for the Influence of Colleagues and CAP Staff on Teachers’ Perception of Students’ Capacity for College-Level Work.....	112
Table 3.7 Regression Estimates for the Influence of Colleagues and CAP Staff on Teachers’ Alignment with College-Going Priorities.....	113
Table 3.8 Regression Estimates for the Influence of Colleagues and CAP Staff on Teachers’ Familiarity with College-Going Resources.....	114
Table 3.9 Summary of All Regression Estimates, by Outcome.....	115

LIST OF FIGURES

Figure 1.1 Graphic Representation of Missed Beeps, 10% Missed, Random Selection Process.....	24
Figure 1.2 Graphic Representation of Missed Beeps, 30% Missed, Random Selection Process.....	25
Figure 1.3 Graphic Representation of Missed Beeps, 30% Missed, Nonrandom Selection Process.....	26
Figure 3.1 Sociogram of Collegial Ties at Middlebrook High School.....	104
Figure 3.2 Sociogram of Collegial Ties at Drew High School.....	104
Figure 3.3 Sociogram of Collegial Ties at Scherzer High School.....	105
Figure 3.4 Sociogram of Collegial Ties at Kelly High School.....	105

CHAPTER 1: USING MULTILEVEL MODELS TO EXPLORE PREDICTORS OF NONRESPONSE FROM EXPERIENCE SAMPLING METHOD (ESM) STUDIES

Introduction

Research designs using the experience sampling method- a form of time diary (Hektner, Schmidt, & Csikszentmihalyi, 2007; Mehl & Conner, 2011) have grown dramatically in popularity over the last decade, particularly in education and the social sciences (Zirkel, Garcia, & Murphy, 2015). The flexibility of the method, as well as its ability to capture in-the-moment data from participants across a variety of contexts, offer unique advantages to researchers seeking to understand phenomena embedded in daily life and social interactions. ESM approaches frequently provide rich subsamples of daily activity, but also present unique challenges in the areas of nonresponse and missing data (Jeong, 2005).

While the general psychometric properties of longitudinal data have been increasingly well established (Singer & Willett, 2003), less is known about the specific statistical properties of ESM designs (an *intensive* longitudinal method, see Bolger & Laurenceau, 2013) which usually compress longitudinal data collection into a period of days or weeks. Within this brief time period, participants can often take ESM surveys forty or fifty times. As a result, researchers often end up with rich data on participants' thoughts, feelings, and actions in multiple spatial and social contexts over the course of a week. Nevertheless, participants inevitably miss scheduled prompts (or "beeps"), and over the course of a week, these patterns of missed beeps need to be understood by researchers in order to make sound generalizations about their sample and population.

Related Literature

Nonresponse in Related Survey Research

While the body of research on nonresponse is less developed in the ESM context, the issue of nonresponse and potential biases resulting from nonresponse has been much more developed in other subfields of survey research and design. For example, researchers interested in time use have long considered the potential impacts of nonresponse on time use estimates. Mulligan, Schneider, & Wolfe (2005) analyzed ESM data from the Alfred P. Sloan Study of Youth and Social Development and compared respondents' estimates of time spent working to similar estimates in the American Time Use Survey (ATUS) and the National Educational Longitudinal Study NELS: 1988) obtained using a time diary approach, and found that differential rates of nonresponse by gender and time of day in the ESM sample significantly impacted population-level estimates of time use. However, this impacted could be corrected for using weighting techniques, similarly to approaches in other longitudinal studies that use imputation to adjust for missingness.

In addition to time use research, nonresponse bias is also a significant concern for researchers engaged in designing and implementing large-scale household surveys (Bethlehem, Cobben, & Schouten, 2011; Wagner, 2010; Kreuter, Couper, & Lyberg, 2010). Here, researchers usually build a probability sample by calling households and asking a series of survey questions. As rates of refusal or nonresponse for these types of calls has increased over the last fifteen years (Couper, 2005; Groves, 2006), survey statisticians have become much more interested in how factors like time of day and day of week might influence rates of participant nonresponse or compliance, and further, how

nonresponse bias might impact the generalizability of even large-scale household surveys.

Nonresponse in ESM Research

To date, limited research has been conducted on the predictors of nonresponse in ESM studies; however, results of this work suggest that demographic characteristics, such as gender, and contextual characteristics, such as time of day and day of week, may significantly impact the odds of nonresponse. For example, Silvia, Kwapil, Eddington, & Brown (2013) examined dispositional and situational predictors of nonresponse in an aggregate sample of 450 college students, and found that males were significantly more likely to miss responses than females. In addition, missed beeps were significantly more likely to occur later in the day (after 4pm) and later in the study week. The researchers attribute this effect to increased survey fatigue of participants, who tire of responding at the end of each day and, eventually, during the middle and end of the entire study period.

A second study by Messiah, Grondin, & Encrenaz (2011) examined factors influencing nonresponse in a single sample of 224 French university students, and again found that males were significantly more prone to nonresponse. Further, participants in this study had higher odds of nonresponse for beeps occurring in the middle of the week (on Tuesday, Wednesday, or Thursday) and first thing in the morning (from 8-11am). A third study by Courvoisier, Eid, & Lischetzke (2012) analyzed the factors influencing nonresponse in a sample of 318 university students in Switzerland and found that the odds of nonresponse were lowest between 5-7 pm in the evening and highest between 9 and 11 am in the morning. The researchers also found a slight increase in the odds of nonresponse as the study progressed, particularly between days four and seven.

Missing Data and Generalizability

As shown above, nonresponse poses challenges to survey-based research by potentially biasing estimates. In an ESM context, each missed beep represents a vector of missed data points. In general, quantitative researchers have developed different classifications for types of missing data (Allison, 2002; Rubin, 1976); these include *missing completely at random* (MCAR), *missing at random* (MAR), and *not missing at random* (NMAR). Each classification is associated with a series of assumptions about the underlying data. For example, missing data are considered to be missing completely at random (MCAR) if the likelihood of an item being missed is unrelated to any other observed variable, as well as unrelated to any unobserved confounders. In the ESM context, which frequently has higher rates of missing data and nonresponse than traditional survey methods, and also prompts participants to respond in a variety of different times and daily contexts, this standard is extremely difficult to meet. As seen above, in most cases, nonresponse seems to be influenced by one or more person-level or beep-level predictors.

Data are assumed to be missing at random (MAR), a less demanding assumption, if, after controlling for auxiliary variables, such as time, day, or gender, likelihood of missingness is not related to the outcome in question. In other words, if within each stratum, or level, of a particular control variable, the probability of a participant missing a beep is randomly distributed, then data can be assumed to be (conditionally) missing at random.

Figures 1.1-1.3 below demonstrate graphically what MAR and NMAR data may look like in the ESM context. All three figures display a 7 by 8 grid representing the total

possible sample of beeps one person may receive throughout the course of a week in an ESM study (typically participants receive about 8 prompts per day, over the course of 7 days, which total 56 possible beeps). Each cell in the grid represents one response; if the cell is clear, that particular beep was responded to, if the cell is red, that particular beep was missed.

Figure 1.1 illustrates a set of responses which would likely satisfy the MAR assumptions- overall, only 10% of beeps are missing, and the missingness appears to be occurring randomly across days and times (beeps). Likewise, Figure 1.2 reflects a largely MAR process- although in this case, rates of missingness are much higher, at 30%. Here, despite the higher overall presence of missed beeps, the probability of missingness does not appear to vary across days of times.

Figure 1.3 again displays a set of responses with 30% of beeps missed, however in this case the beeps are not missing at random. Beeps occurring later in the day and on the weekend appear to have much higher probabilities of nonresponse. This would be a situation that would not satisfy the assumptions of MAR, given the impact of day and time on response. This does not, however, automatically mean that data must be considered NMAR (not missing at random), which necessitates far more complex statistical approaches to reach generalizability (Allison, 2002). MAR assumptions could still be met by applying appropriate statistical controls to mitigate the differential probability of nonresponse by day and time. However, any outcome used in analysis would need to be carefully examined for its relationship to missingness, both before and after controlling for day and time.

These figures illustrate the complexity of understanding patterns of missing data and responses in ESM research. This study applies multilevel modeling, which can simultaneously examine patterns of missingness due to beep-level and person-level characteristics, to examine a sample of 7,961 beeps collected from 141 students in three high schools in mid-Michigan in 2013-2014. The relevant research questions are below:

Research Questions

- 1) How do factors of time and daily context impact participants' odds of missing an ESM response?
- 2) How do person-level factors (gender, race, school) impact participants' odds of missing an ESM response?
- 3) How do in-the-moment and person-averaged measures of student emotional states impact participants odds of missing an ESM response, after controlling for other relevant time- and person-level predictors?

Methods

Data Collection

Data were collected in this study using the Experience Sampling Method (ESM), which captures behaviors and subjective experiences across multiple contexts (Csikszentmihalyi & Larson, 1987; Csikszentmihalyi & Schneider, 2000; Hektner, Schmidt, & Csikszentmihalyi, 2007). ESM affords several advantages over one-time surveys. Most importantly, data are collected in the moment, thus reducing recall bias and the potential for providing socially desirable answers (Mulligan, Schneider, & Wolfe, 2005). Prior ESM studies have used programmed wristwatches, pagers, or PDAs that went off randomly throughout the day over a period of one or two weeks prompting

participants to answer a set of items such as, “What is the main activity you are doing?” and “How much did you enjoy what you were doing?” (See Appendix 1.C for the complete ESM instrument).

This study made use of a new smartphone application (app), Paco (www.pacoapp.com), developed by Google engineer Robert Evans, which is open source and specifically designed for the ESM. The ESM instrument was transferred to the smartphones using a web-based program that also uploads, stores, and protects participant data prior to analysis.

For a week in winter 2013 and early spring 2014, students in the study were given smartphones with Paco. Each time students were signaled during the day, they were asked to respond to a set of identical items within a fifteen-minute window. On average, it takes about sixty seconds to complete the items. All students received training on how to use Paco prior to data collection, including when not to respond to the beeps (e.g. when driving). Teachers also were briefed on study procedures. Students were asked to consent to the study; two students out of the sample of 143 declined to participate. All data were accumulated on a secure data server and de-identified to maintain participant confidentiality. After the data collection period, teachers were asked if the data collection was irregular or a distraction. Responses of teachers suggested this was not the case.

Each day over a week all students received eight beeps on their smartphones, which gave them 56 total response opportunities. The preprogrammed randomized beep schedule ran from 8am to 8pm daily, and guaranteed a minimum of 4 beeps in the morning and an additional four beeps between 1pm and 8pm. Beeps were scheduled to

be slightly more frequent in the morning in an effort to capture additional beeps in students' science classes, which occurred before noon.

Sample

The sample includes 141 students (62 males and 79 females) in three secondary schools in the mid-Michigan area. According to U.S. Census classifications, one school is urban, one is suburban, and one is rural. The schools and teachers were selected on a voluntary basis and all students in these science classes were asked to participate. About 19% of student in the sample self-identified as non-white. Table 1.1 presents descriptive statistics and characteristics for the three high schools in the study: Meridian High School, Northern High School, and Lowrey High School.

Measures

Table 1.2 presents descriptive statistics for the variables used in this study. Below, the measurement and operationalization of each variable are described. Variables are separated into two conceptual categories: 1) time-variant predictors, which vary within-student and across time, and 2) time-invariant predictors, which only vary between students. These distinctions will become clearer when contextualized by the statistical approach outlined in the analysis section.

Study outcome: Missed beeps

The outcome in this study, *MISSED_BEEP*, is measured as a 0 if a given beep was responded to, and 1 if the beep was not responded to within the fifteen minute window specified by the study design. The mean of this measure, 0.53, suggests overall, about 53% of scheduled beeps in the study were missed. The predictors described below

are included to help identify the factors that are most associated with the probability of missing a beep.

Time-variant predictors

Three predictors were used to identify effects of day and time on missed beeps. *MORNING* is a dichotomous measure that is equal to 1 when a beep occurred before noon on a given day, and 0 otherwise. This measure is especially useful in this case because students in the sample all attended science class in the morning. Since science teaching and learning were the driving motivation behind the initial research, and science teachers were active participants and collaborators in the work, one might expect that students miss fewer beeps in the morning.

A second time predictor, *SCHOOL_OUT*, is a dichotomous measure equal to 1 when a beep occurs after 4pm on a given day, and 0 otherwise. This variable allows for a useful distinction by contrasting students' behaviors during and after school. Since these two contexts are quite different, one would expect that patterns of missed beeps may vary accordingly. A final time predictor, *WEEKEND*, is a dichotomous measure equal to 1 when a beep occurred on the weekend, and 0 otherwise. In this study, students responded to beeps from Monday to Sunday. Like *SCHOOL_OUT*, *WEEKEND* represents a very different social context and may have much different missed beep patterns.

In addition to predictors of time and day, this study uses five predictors of momentary emotional states: *L.HAPPY*, *L.ACTIVE*, *L.ANXIOUS*, *L.BORED*, and *L.LONELY* are lagged effects of students' emotional states in the beep immediately preceding the beep that was missed. Each of these measures are interval-ratio in nature. Students were asked "How are you feeling?", and then for each emotion above, given a

range between 1 “Not at all” to 4 “Very much”. Lag measures in this case only refer to a students’ response in the immediately preceding time unit. Lags beyond one unit were not used as the gap in time between the initial response and the missed beep would be quite significant (i.e. many hours or even a full day).

Time-invariant predictors

In addition to time-variant predictors, this study uses several time-invariant predictors to examine person-level impacts on the probability of a missed beep. These include dichotomous predictors of gender (*MALE*), race/ethnicity (*NONWHITE*), and school assignment (*NORTHERN* and *LOWREY*). Further, average levels of students’ emotional states were included (*MEAN_HAPPY*, *MEAN_ACTIVE*, *MEAN_ANXIOUS*, *MEAN_BORED*, and *MEAN_LONELY*). These values represent the average response for each student over the course of the study, and were included to serve as proxies for person-level differences in average emotional levels.

Analyses

The ESM data are analyzed using two-level hierarchical generalized linear models (HGLM) (Raudenbush & Bryk, 2002; Rabe-Hesketh & Skrondal, 2008), with the individual student responses (level one) nested within each student (level two) with the outcome of missed beep. One key strength of multi-level modeling is its capacity to more accurately account for variation in nested phenomena, such as students in schools (organizational contexts) (Raudenbush & Bryk, 2002) or repeated measures within students (individual growth contexts) (Hong & Raudenbush, 2005). As opposed to a more traditional approach, such as OLS regression, HLM partitions variance by level; this leads to more accurately estimated standard errors. In this study, we use a two-level

model, which includes in-the-moment student responses and contextual measures (e.g. time, date, and prior emotional state) at level one and static student characteristics (e.g. gender, race/ethnicity, and average emotional states) at level two. To estimate the impact of time factors and person-level characteristics on the odds of a missed beep, we use the following two-level random-intercept logistic regression specified for time (instance) i nested within person j below:

Equation 1: The impact of beep-level and person-level predictors on the likelihood of nonresponse.

(1)

Level-1 Model (Beep Level)

$$\begin{aligned} \text{Prob}(\text{MISSED_BEEP}_{ij} = 1 | \mathbf{x}_{ij}, \zeta_j) &= \phi_{ij} \\ \log[\phi_{ij}/(1 - \phi_{ij})] &= \eta_{ij} \\ \eta_{ij} &= \beta_{0j} + \beta_{1j} * (\text{MORNING}_{ij}) + \beta_{2j} * (\text{SCHOOL_OUT}_{ij}) + \beta_{3j} * (\text{WEEKEND}_{ij}) + \\ &\beta_{4j} * (\text{L.HAPPY}_{ij}) + \beta_{5j} * (\text{L.ACTIVE}_{ij}) + \beta_{6j} * (\text{L.ANXTIOUS}_{ij}) + \beta_{7j} * (\text{L.BORED}_{ij}) + \\ &\beta_{8j} * (\text{L.LONELY}_{ij}) \end{aligned}$$

Level-2 Model (Person Level)

$$\begin{aligned} \beta_{0j} &= \gamma_{00} + \gamma_{01} * (\text{MALE}_j) + \gamma_{02} * (\text{NONWHITE}_j) + \gamma_{03} * (\text{NORTHERN}_j) + \gamma_{04} * (\text{LOWREY}_j) + \\ &\gamma_{05} * (\text{MEAN_HAPPY}_j) + \gamma_{06} * (\text{MEAN_ACTIVE}_j) + \gamma_{07} * (\text{MEAN_ANXTIOUS}_j) + \\ &\gamma_{08} * (\text{MEAN_BORED}_j) + \gamma_{09} * (\text{MEAN_LONELY}_j) + \zeta_{0j} \\ \beta_{pj} &= \gamma_{p0}, p = 1, 2, 3, 4, 5, 6, 7, 8 \end{aligned}$$

$$\text{Level-1 variance} = 1/[\phi_{it}(1-\phi_{it})]$$

$$\zeta_j \sim N(0, \psi)$$

Where:

$\text{MISSED_BEEP} = 1$ if student missed beep, 0 if beep is answered

$\text{MORNING} = 1$ if beep occurred before noon, 0 after noon

$\text{SCHOOL_OUT} = 1$ if beep occurred after 3pm, 0 before 3pm

WEEKEND = 1 if beep occurred on Saturday or Sunday, 0 if Monday-Friday

L.HAPPY = person *j*'s reported level of happiness at time (instance) *i*-1

L.ACTIVE = person *j*'s reported level of activity at time (instance) *i*-1

L.ANXIOUS = person *j*'s reported level of anxiousness at time (instance) *i*-1

L.BORED = person *j*'s reported level of boredom at time (instance) *i*-1

L.LONELY = person *j*'s reported level of loneliness at time (instance) *i*-1

MALE = 1 if student identifies as male

NONWHITE = 1 if student identifies as nonwhite

NORTHERN = 1 if student enrolled in Northern H.S.

LOWREY = 1 if student enrolled in Lowrey H.S.

MEAN.HAPPY = person *j*'s mean level of reported happiness

MEAN.ACTIVE = person *j*'s mean level of reported activity

MEAN.ANXIOUS = person *j*'s mean level of reported anxiousness

MEAN.BORED = person *j*'s mean level of reported boredom

MEAN.LONELY = person *j*'s mean level of reported loneliness

In order to better understand how groups of predictors relate to one another, the analysis was conducted following a typical model building approach (Hox, 2010) that includes 4 separate models. Model 1 includes only level one (within-person) predictors of time and day. Model 2 adds on level two predictors of gender, race/ethnicity, and school assignment. Model 3 includes person-level averages of emotional states, while Model 4, the full model, incorporates all the covariates above plus lagged momentary emotional states. The advantage of this approach is that it allows for examination of how

key covariates (time, place, gender) based on theory and previous research are impacted by the introduction of additional predictors.

Main results presented below come from Equation 1, applying conditional (or unit-specific) assumptions. Analysis was performed using the `-xtlogit-` package in STATA, which makes use of maximum likelihood estimation with adaptive Gauss-Hermite quadrature. This means that log-odds (η_{ij}) should be interpreted as conditional on all covariates as well as the level-2 random effect (ζ_j). This study's focus on within-person processes suggests that a conditional model, as opposed to a marginal, or population-average approach, is the best choice (Raudenbush & Bryk, 2002). However, Table 1.4 in Appendix 1.B presents population-averaged estimates. This estimation procedure employs a generalized estimation equation (GEE) to average coefficient estimates across all level-2 distributions (Rabe-Hesketh & Skrondal, 2008). In this case, however, estimates from the unit-specific and population-averaged approaches did not appear to differ significantly in terms of magnitude or significance.

Calculating Predicted Probabilities

Given the complexity of interpreting log-odds coefficients produced by HGLMs, often it is useful to convert the estimated log-odds (η_{ij}) to a predicted probability (φ_{ij}). This can be accomplished using a relatively straightforward conversion equation (Raudenbush & Bryk, 2002):

(2)

$$\varphi_{ij} = \frac{1}{1 + \exp\{-\eta_{ij}\}}$$

Calculating predicted probabilities provides an additional perspective on the relative impact of various covariates on rates of nonresponse.

Missing Data

Missing data is often an important consideration in ESM research, which can be susceptible to higher rates of missingness (a key motivation of this study). In this case, however, there was no missing data in the sample. This is largely due to the construction of the analysis, which relies on measures of dates and times that are automatically recorded by the ESM software, as well as person-level measures that can readily be obtained from school administrative data. Therefore, all scheduled beeps are included in the current sample.

One additional possibility exists for missing data; if cell phones are turned off, ESM beeps do not occur, and are likewise not registered as being missed. So, broken or powered-down phones may not record beeps properly. This is a somewhat difficult phenomenon to understand, but one approach is to estimate how many beeps *should* have occurred over the course of the week, and then compare that number to the number of beeps that *actually* occurred. In the case of this sample, with 141 total participants, the phones should have beeped a total of 7,896 times (the product of 141 and 56, the number of weekly beeps). In actuality, the phones in this sample beeped a total of 7,961 times, which is actually slightly higher than the expected value, but within 1% of the total beep range. Ideally, the expected and actual beep totals would be identical, but some small variation is expected due to technical features of the signaling program. In sum, there appears to be little evidence that participants turned their phones off during the day and

missed beeps, which suggests that the sample collected is an accurate snapshot of missed and responded beeps.

Results

Table 1.3 presents unit-specific results sequentially for all four statistical models. Results are summarized by model. Coefficients presented have been translated to odds ratios (O.R.) by exponentiating the log-odds (η_{ij}) coefficients ($O.R. = \exp(\eta_{ij})$). Thus, odds ratios greater than 1 suggest that particular times, states, etc. are associated with increased odds of missing beeps, and likewise, odds ratios less than 1 suggest that a predictor is associated with a decrease in odds of missing beeps. See Table 1.4 for corresponding population-average estimates, which did not change compared to the unit-specific results.

Model 1: Level-1 Time and Day Predictors

Fitting an HGLM with just beep-level covariates (see Table 1.3, Column 1), all three level-1 predictors were found to significantly impact the likelihood of nonresponse, after controlling for other covariates and the level-2 random effect. Beeps occurring in the morning ($O.R. = 0.59, p < .01$) had about half the odds of being missed compared to beeps in the afternoon. Beeps occurring after school ($O.R. = 1.32, p < .01$) were associated with significantly odds of nonresponse. In this case, after school beeps were about 1.3 times more likely to be missed than beeps occurring in the afternoon. Similarly, weekend beeps ($O.R. = 2.23, p < .01$) appear to be associated with much higher odds of nonresponse. This suggests that the odds of missing a beep over the weekend are more than twice the odds of missing a beep during the week.

Model 2: Student Demographics and School Assignment

Next, I fit an HGLM with the three level-1 predictors above along with several level-2 (person-level) predictors (see Table 1.3, Column 2). In this case, all three level-1 predictors were found to significantly impact the likelihood of nonresponse, after controlling for other covariates and the level-2 random effect. Beeps occurring in the morning (O.R. = 0.52, $p < .01$) had about half the odds of being missed compared to beeps in the afternoon. Beeps occurring after school (O.R. = 1.32, $p < .01$) were still associated with a significantly higher likelihood of nonresponse. Similarly, weekend beeps (O.R. = 2.22, $p < .01$) are still associated with a much higher likelihood of nonresponse.

At level-2, most person-level predictors do not appear to significantly impact the likelihood of nonresponse, after adjusting for other covariates. The impact of both gender and race/ethnicity were not significant. School assignment was found to be marginally significant, but only in the case of students at Northern High School (O.R. = 0.59, $p < .10$), who appear to be less likely to miss beeps compared to students at Meridian High School. Given the significantly smaller sample of students at Northern (17) compared to other schools, this borderline significance may be evidence of real contextual differences between Northern High School and other schools in the sample.

Model 3: Person-Level Emotional States

Model 3 (see Table 1.3, Column 3) adds to Models 1 and 2 by including person-level covariates that represent students' average reported emotional states. At level-1, all three predictors remain strongly significant and unchanged in magnitude. At level-2, the marginal effect for students at Northern High School (O.R. = 0.57, $p < .05$) continues to

hold, while only mean levels of boredom (O.R. = 1.500, $p < .05$) appear to significantly impact the odds of a missed beep, after controlling for all other covariates and the level-2 random effect. In this case, students who, on average, report one unit higher on boredom, tend to have odds of a missed beep that are 1.5 times higher than students who report less boredom. No effects appear for mean levels of other emotional states.

Model 4: Full Model, with Lagged Emotional States

Model 4 (see Table 1.3, Column 4), the final model in this progression, builds on Models 1-3 by adding lagged levels of students' momentary emotional states at level-1. In this model, beeps occurring in the morning still have significantly lower odds of missingness than other beeps (O.R. = 0.65, $p < .01$), while beeps occurring after school still have significantly higher odds of missingness than others (O.R. = 1.33, $p < .01$). The effect for weekend is no longer significant, and the lagged effect for boredom appears to be associated with a slight decrease in the odds of a missed beep (O.R. = 0.93, $p < .10$). At level-2, only the predictor for Northern High School is significant (O.R. = 0.55, $p < .01$). It is important to note, however, that Model 4 saw a significant decrease in sample size ($N=3,584$ beeps) relative to the other models ($N > 7,000$ beeps). This is due to the introduction of lagged effects, because lagged effects only can be calculated in the case that a student responded to the beep immediately prior to a missed beep. In other words, in situations where students missed multiple beeps in a row, lag values could not be used, causing a significant drop in sample size and power. Moreover, given the much higher odds of missed beeps over the weekend, this loss of sample could be most impactful for the estimate of the weekend predictor. This explains how an otherwise highly significant covariate might lose its significance with the inclusion of lagged values.

Predicted Probabilities

Another useful method for examining the impact of covariates is to use the intercept and covariate estimates to calculate predicted probabilities. Using the estimates obtained in Model 2 and Equation 2 above produced Table 1.5 which estimates the relative impact that time and day covariates have on the predicted probability of a missed beep. Later models that introduce continuous predictors complicate the calculation and interpretation of predicted probabilities in this context. For simplicity, and also because time and day predictors are the most impactful covariates on odds of missed beeps, only MORNING, SCHOOL_OUT, and WEEKEND are interpreted here.

As seen in Table 1.5, the predicted probability associated with the intercept is 0.58, which represents the probability of a missed beep when it occurs during a weekday afternoon. Weekday afternoons represent the reference category for the additional binary predictors of time and context. In other words, about 58% of beeps in this time window are likely to be missed. If a beep occurs in the morning, the predicted probability drops to 0.45, or about 45%. After school, the predicted probability increases to 65%, and over the weekend, the rate increases even more to 75%. This suggests that beeps occurring over the weekend have about a 75% of being missed by study participants.

Discussion

Results of the current ESM study suggest that several time-related factors impact the odds of nonresponse. In particular, participants were much less likely to miss a beep in the morning and much more likely to miss a beep after school. Previous research suggested a similar evening effect, however the lower odds of nonresponse in the current study are somewhat different than previous research, which suggested that participants

may be much more likely to miss a beep in the morning. One possible explanation for this discrepancy is the difference in study populations; the current study involved high school students, all of whom are in class early in the mornings and taking science classes with teachers who were aware and supportive of the research. This may have led more students to respond, perhaps due to peer/ normative pressure or encouragement from their teachers. In contrast, the previous ESM studies that found higher nonresponse odds in the morning all examined populations of college students, who might likely have less uniform and more disparate class and work schedules that might not be as conducive to responding to beeps.

In addition to time of day, day of week appears to be a strong predictor of nonresponse. Results from this analysis suggest that the odds of a missed beep on the weekend are double the odds of a missed beep during the week. Converting these odds to predicted probabilities, this means that the probability of missing a beep on the weekend could be as high as 0.75- in other words, three out of four beeps are going unanswered. Previous research found some evidence of elevated nonresponse at the end of the week, however the magnitude of the current findings is much larger. This may again be related to the difference in study populations. High school students, while more receptive to responding to beeps in school and class, might lack sufficient self-regulation or discipline to continue to respond to beeps on their own time, especially over the weekend. College students, however, may be more accustomed to managing their time and commitments independently and could more easily integrate ESM response into their out-of-school routine.

Finally, in contrast to several previous studies, this analysis finds no systematic difference in odds of nonresponse by gender. This includes models that control for beep-level covariates, all of which were found to be significant, as well as in models restricted to person-level covariates alone. Previous results would seem to suggest that males have higher odds of missing a beep than females, but in this study the difference was not found to be statistically significant. However, this study did find that higher average rates of reported boredom appeared to increase the odds of a missed beep, which suggests that person-level averages of emotional states may play a role in understanding nonresponse.

Limitations

While the findings regarding nonresponse and its predictors are certainly robust, it is important to underscore the limitations of this research. To begin, this study examined a purposive sample of high school students in three high schools; thus, one should not assume that these findings are generalizable to all ESM studies or even all ESM studies involving high schools students. While the trends uncovered in this work are strong, they should still be considered as an initial investigation into patterns and predictors of nonresponse in ESM designs. In fact, given the differences in findings between prior research above, which has relied largely on college students, and the current study, which focuses on high school students, one might expect predictors of nonresponse to vary widely between different populations of interest. Future research involving different samples and student populations might further flesh out these differences.

Another important limitation of this study is the operationalization of nonresponse. For the purposes of this analysis, an agnostic view of the causes of nonresponse was adopted- in other words, all missed responses were treated the same,

regardless of their cause or circumstance. A more nuanced approach might attempt to classify missed beeps into various categories- such as technical glitches or phones being turned off. As explained above in methods, the structure and pattern of the data collected in this study suggest that phone issues (both glitches and powered off) were minimal. In general, the difference between the “expected” number of beeps and the number of actual beeps was very small. However, additional research, maybe in the form of focus groups or follow-up interviews with participants, might shed more light on the various reasons why beeps were missed, which might in turn have implications for better classifying and understanding the larger phenomenon of nonresponse.

Implications for Research and Practice

This study has shown that in a sample of high school students, nonresponse appears to be significantly related to several time-related factors. Having established this claim, one must then consider potential implications this might have for future ESM research and practice. One key implication for research is to seriously interrogate the generalizability of findings that focus on ESM data collected over the weekend or outside of school. Our results suggest that beeps occurring in either of these contexts are susceptible to dramatically higher rates of nonresponse, and subsequently, increased risk of bias in ESM estimates. This is especially the case in this sample for beeps occurring over the weekend, where the nonresponse rate is so high as to raise real questions as to the viability of weekend-specific analyses.

One possibility is simply to discard observations from these more problematic time contexts and focus on research questions that can be explained by in-school and weekday responses. While viable, this strategy also eliminates a great deal of useful

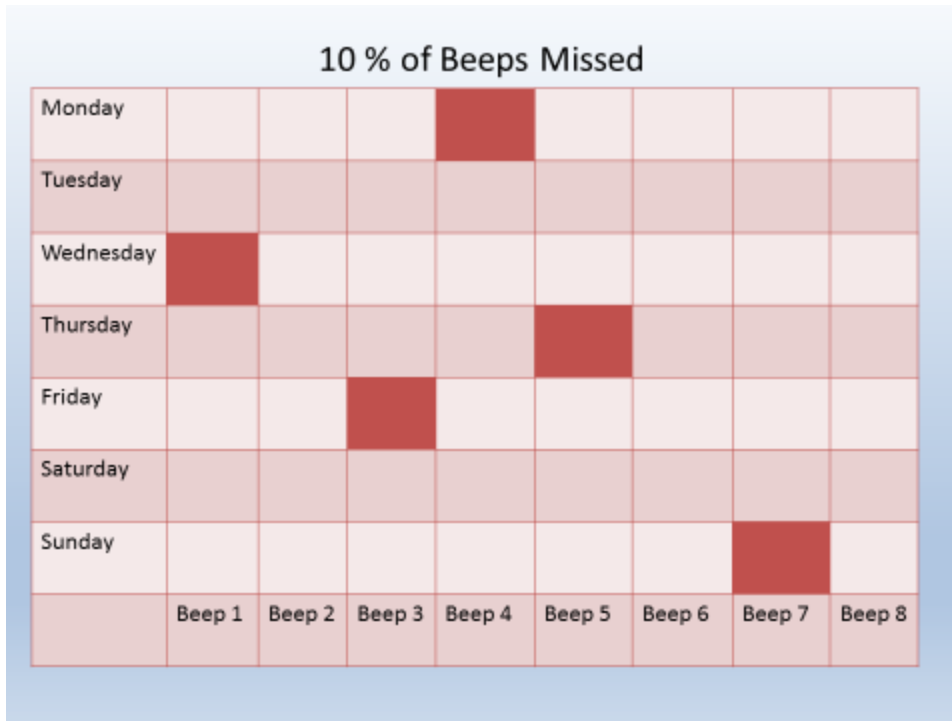
information found in the ESM responses participants do provide in non-school contexts. Another possible solution may be developing a weighting scheme, similar to that suggested by Mulligan, Schneider, & Wolfe (2005) and Jeong (2008) that would more heavily weight responses from time contexts that tend to be more prone to nonresponse. This could potentially allow the researcher to produce estimates that are more representative of the larger population under study. Similarly, one of several multiple imputation procedures (Black, Harel, & Matthews, 2011) could improve statistical power and allow for more complex analysis of weekend and evening contexts. In the case of either weighting or multiple imputation, however, drawbacks are inherent. For example, both techniques require participants to have some baseline level of response in all contexts; these responses can then be weighted-up or imputed accordingly. However, given the wide variability in response rates in the current sample, some participants may lack even the few observations necessary to establish a baseline, and would have to be excluded, thereby introducing the risk of additional bias.

Overall, much more needs to be known about trends and patterns of nonresponse in ESM research. One benefit of the growth of smartphone- and app-based ESM approaches is that these devices can more easily capture more contextual information (or *paradata*- see Kreuter, Couper, & Lyberg, 2010) related to when and where beeps occur, and if they are missed. This allows researchers to develop more sophisticated models for nonresponse and the factors that may influence it. As ESM designs continue to grow in their appeal and ease of use and implementation, the need for more information about nonresponse bias will only become more acute.

APPENDICES

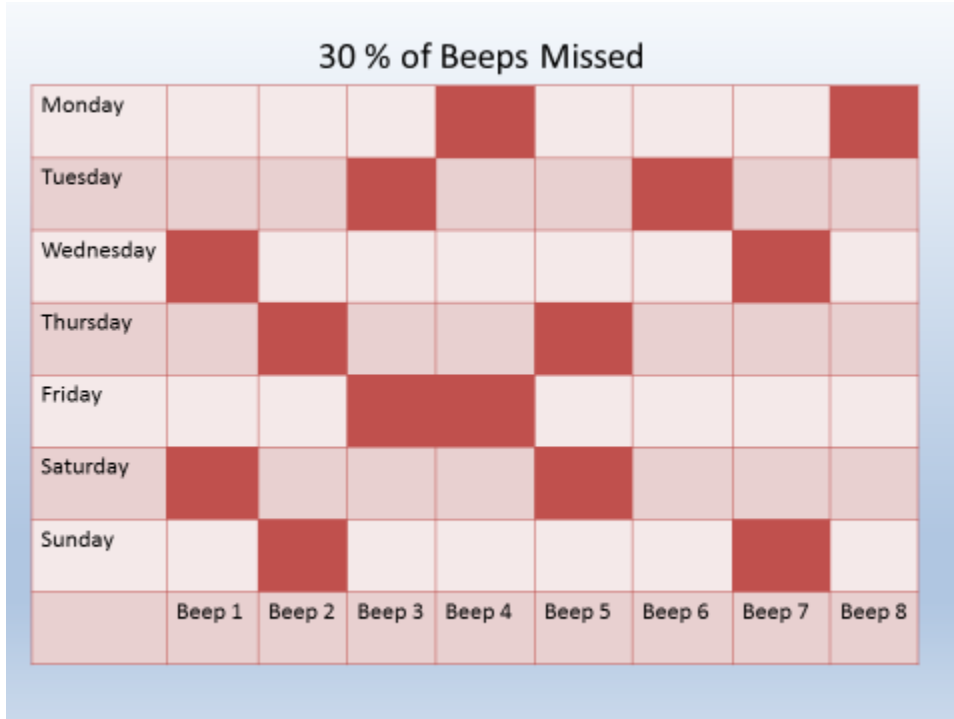
Appendix A: Figures

Figure 1.1 Graphic Representation of Missed Beeps, 10% Missed, Random Selection Process.



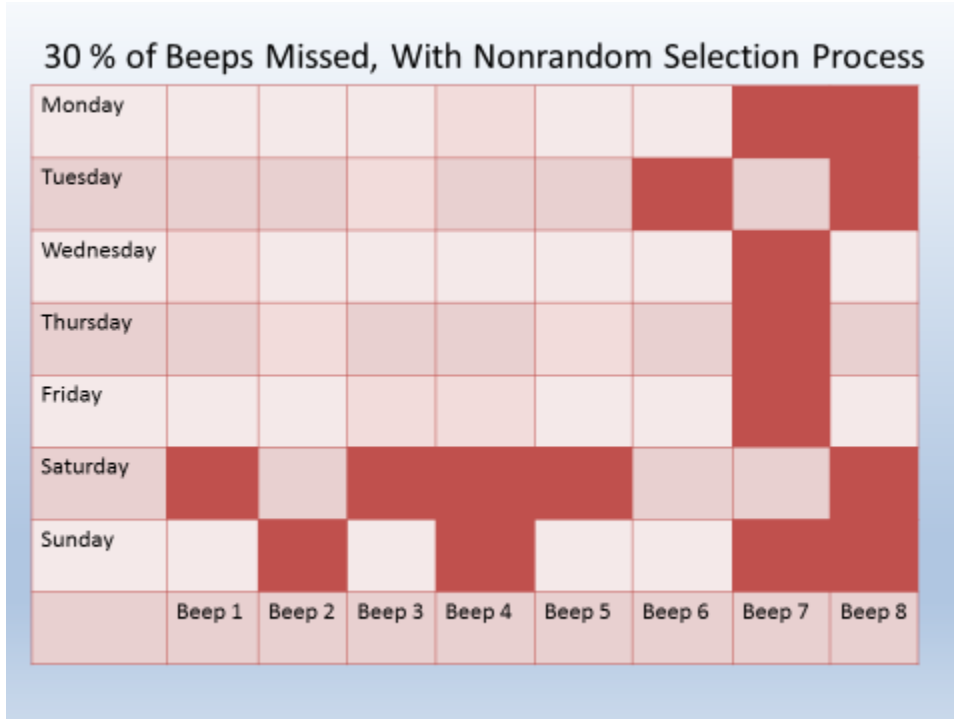
Note. Dark red squares indicate a missed beep.

Figure 1.2. Graphic Representation of Missed Beeps, 30% Missed, Random Selection Process.



Note. Dark red squares indicate a missed beep.

Figure 1.3. Graphic Representation of Missed Beeps, 30% Missed, Nonrandom Selection Process.



Appendix B: Tables

Table 1.1 Sample Characteristics for High Schools in the Study

	Northern H.S.	Lowrey H.S.	Meridian H.S.	Overall
School Classification	Urban	Rural	Suburban	
Total Participants	17	53	71	141
Total ESM Events (Beeps)	963	2,943	4,055	7,961
% Female	47%	47%	53%	51%
% Nonwhite	50%	8%	20%	19%
Collection Time Frame	December 2013	January 2014	February 2014	

Table 1.2 Descriptive Statistics for the Analytical Sample

Variable Name	N	Mean	SE	Minimum	Maximum
<i>Level 1 Predictors</i>					
MISSED_BEEP	7961	0.53	0.006	0.00	1.00
MORNING	7961	0.44	0.006	0.00	1.00
SCHOOL_OUT	7961	0.34	0.005	0.00	1.00
WEEKEND	7961	0.28	0.005	0.00	1.00
L.HAPPY	3584	2.65	0.018	1.00	4.00
L.ACTIVE	3584	2.09	0.018	1.00	4.00
L.ANXIOUS	3584	1.81	0.017	1.00	4.00
L.BORED	3584	2.16	0.019	1.00	4.00
L.LONELY	3584	1.55	0.015	1.00	4.00
<i>Level 2 Predictors</i>					
MALE	141	0.51	0.042	0.00	1.00
NONWHITE	140	0.19	0.033	0.00	1.00
NORTHERN	141	0.12	0.027	0.00	1.00
LOWREY	141	0.38	0.041	0.00	1.00
MEAN_HAPPY	141	2.68	0.052	1.26	5.00
MEAN_ACTIVE	141	2.12	0.060	1.00	4.83
MEAN_ANXIOUS	141	1.86	0.053	1.00	4.84
MEAN_BORED	141	2.21	0.049	1.00	4.81
MEAN_LONELY	141	1.54	0.049	1.00	4.50

Table 1.3 Results of 2-Level HGLM Predicting a Missed Beep, Conditional (Unit-Specific) Model

VARIABLES	(1) Time and Day Only		(2) Plus Demographics		(3) Plus Avg. States		(4) Plus Lag Effects	
	<u>OR</u>	<u>(95% CI)</u>	<u>OR</u>	<u>(95% CI)</u>	<u>OR</u>	<u>(95% CI)</u>	<u>OR</u>	<u>(95% CI)</u>
<u>Fixed Effects</u>								
morning	0.591***	(0.520, 0.672)	0.590***	(0.519, 0.672)	0.591***	(0.519, 0.672)	0.647***	(0.535, 0.783)
school_out	1.328***	(1.161, 1.520)	1.321***	(1.154, 1.513)	1.322***	(1.154, 1.513)	1.330***	(1.084, 1.632)
weekend	2.196***	(1.953, 2.470)	2.223***	(1.976, 2.500)	2.226***	(1.979, 2.503)	0.933	(0.768, 1.135)
male			0.821	(0.587, 1.148)	0.857	(0.607, 1.208)	0.954	(0.725, 1.257)
nonwhite			1.031	(0.653, 1.629)	1.028	(0.655, 1.613)	1.066	(0.744, 1.528)
northern			0.586*	(0.335, 1.025)	0.570**	(0.329, 0.987)	0.548***	(0.354, 0.848)
lowrey			0.965	(0.675, 1.380)	1.050	(0.725, 1.521)	1.250	(0.931, 1.679)
mean_happy					1.262	(0.880, 1.811)	1.088	(0.805, 1.471)
mean_active					0.958	(0.710, 1.293)	1.031	(0.799, 1.329)
mean_anxious					1.011	(0.750, 1.362)	1.047	(0.804, 1.362)
mean_bored					1.500**	(1.061, 2.122)	1.278	(0.945, 1.729)
mean_lonely					1.001	(0.716, 1.398)	1.014	(0.760, 1.355)
L.happy							1.039	(0.944, 1.143)
L.active							0.986	(0.898, 1.083)
L.anxious							0.983	(0.896, 1.078)
L.bored							0.930*	(0.855, 1.012)
L.lonely							0.976	(0.872, 1.092)
Constant	1.173	(0.967, 1.424)	1.371**	(1.004, 1.874)	0.305*	(0.0811, 1.151)	0.318**	(0.110, 0.920)
<u>Random Effects</u>								
ψ		0.923		0.874		0.830		0.355
σ		0.220		0.210		0.201		.097
Log likelihood		-4847.73		-4801.33		-4797.98		-2203.01
Observations		7,961		7,872		7,872		3,584
Number of ID		141		140		140		140

Notes. 95% confidence interval for odds ratio in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 1.4 Results of 2-Level HGLM Predicting a Missed Beep, Marginal (Population-Average) Model

VARIABLES	(1) Time and Day Only		(2) Plus Demographics		(3) Plus Avg. States		(4) Plus Lag Effects	
	<u>OR</u>	<u>(95% CI)</u>	<u>OR</u>	<u>(95% CI)</u>	<u>OR</u>	<u>(95% CI)</u>	<u>OR</u>	<u>(95% CI)</u>
morning	0.640***	(0.574, 0.715)	0.637***	(0.570, 0.713)	0.635***	(0.568, 0.711)	0.673***	(0.566, 0.799)
school_out	1.279***	(1.138, 1.438)	1.277***	(1.134, 1.438)	1.280***	(1.136, 1.442)	1.294***	(1.075, 1.557)
weekend	1.970***	(1.778, 2.183)	1.999***	(1.802, 2.218)	2.010***	(1.811, 2.230)	0.941	(0.790, 1.122)
male			0.902	(0.694, 1.172)	0.950	(0.724, 1.247)	0.945	(0.721, 1.237)
nonwhite			1.044	(0.730, 1.492)	1.067	(0.747, 1.525)	1.063	(0.747, 1.512)
northern			0.604**	(0.389, 0.937)	0.542***	(0.350, 0.841)	0.549***	(0.351, 0.857)
lowrey			0.930	(0.704, 1.228)	0.990	(0.739, 1.328)	1.210	(0.909, 1.611)
mean_happy					1.155	(0.869, 1.536)	1.095	(0.819, 1.465)
mean_active					0.970	(0.765, 1.229)	1.021	(0.799, 1.304)
mean_anxious					0.955	(0.754, 1.208)	1.049	(0.813, 1.353)
mean_bored					1.348**	(1.025, 1.773)	1.264	(0.942, 1.695)
mean_lonely					1.094	(0.840, 1.426)	1.010	(0.764, 1.335)
L.happy							1.034	(0.948, 1.127)
L.active							0.988	(0.908, 1.075)
L.anxious							0.983	(0.905, 1.068)
L.bored							0.937*	(0.868, 1.012)
L.lonely							0.978	(0.882, 1.083)
Constant	1.181**	(1.011, 1.379)	1.307**	(1.023, 1.671)	0.445	(0.155, 1.272)	0.340**	(0.120, 0.960)
Observations	7,961		7,872		7,872		3,584	
Number of ID	141		140		140		140	

Notes. 95% confidence interval for odds ratio in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 1.5 Predicted Probabilities Based on
Model 2 Log Likelihood Estimates

Variable	Predicted Probability of Nonresponse ^a
INTERCEPT	0.58
MORNING	0.45
SCHOOL_OUT	0.65
WEEKEND	0.75

a. Calculated using Formula 2 on page 14.

Appendix C: Instrument

Sample ESM Questionnaire

1. Take a picture of what you are doing.
2. Where are you?
3. Who are you with?
4. What are you doing?
5. What are you thinking?
6. Are you doing the main activity because...[You wanted to, You had to, You had nothing else to do]
7. Is the main thing you are doing... [more like school work, more like play, both, neither]

II. How do you feel about the main activity (4-point scale):

8. Challenge of the activity: [Low/High]
9. Your skills in the activity: [Low/High]
10. Is this activity interesting [Not at all/Very much]
11. Are you succeeding? [Not at all/Very much]
12. Do you feel capable? [Not at all/Very much]
13. Do you feel in control? [Not at all/Very much]
14. Do you enjoy what you are doing? [Not at all/Very much]

III. How do you feel about the main activity? (4-point scale: Not at all/Very much)

15. Is this activity important for you?
16. Do you feel competent in this activity?
17. How important is this activity in relation to your future goals/plans?
18. Are you living up to the expectations of others?
19. Are you living up to your expectations?

IV. How are you feeling? (4-point scale: Not at all/Very much)

20. Are you feeling...Happy
21. Are you feeling...Energetic
22. Are you feeling...Anxious
23. Are you feeling...Competitive
24. Are you feeling...Lonely
25. Are you feeling...Stressed
26. Are you feeling...Proud
27. Are you feeling...Cooperative
28. Are you feeling...Bored
29. Are you feeling...Self-confident
30. Are you feeling...Confused
31. Are you feeling...Active

REFERENCES

REFERENCES

- Allison, P. D. (2002). *Missing data* (Vol. 136). Thousand Oaks, CA: Sage Publications.
- Bethlehem, J., Cobben, F., & Schouten, B. (Eds.). (2011). *Handbook of nonresponse in household surveys* (Vol. 568). New York: John Wiley & Sons.
- Black, A.C., Harel, O., & Matthews, G. (2011). Techniques for analyzing intensive longitudinal data with missing values. In M. R. Mehl, & T. S. Conner (Eds.), *Handbook of research methods for studying daily life*. New York: Guilford Press.
- Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York: Guilford Press.
- Couper, M. P. (2005). Technology trends in survey data collection. *Social Science Computer Review*, 23(4), 486-501.
- Courvoisier, D. S., Eid, M., & Lischetzke, T. (2012). Compliance to a cell phone-based ecological momentary assessment study: The effect of time and personality characteristics. *Psychological assessment*, 24(3), 713.
- Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *The Journal of nervous and mental disease*, 175(9), 526-536.
- Csikszentmihalyi, M., & Schneider, B. (2000). *Becoming Adult*. New York: Basic Books.
- Groves, R.M. (2006). Nonresponse rates and nonresponse bias in household surveys. *The Public Opinion Quarterly* 70(5), 646-675.
- Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (Eds.). (2007). *Experience sampling method: Measuring the quality of everyday life*. Thousand Oaks, CA: Sage Publications.
- Hong, G., & Raudenbush, S. W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis*, 27(3), 205-224.
- Hox, J. (2010). *Multilevel analysis: Techniques and applications*. New York: Routledge.
- Jeong, J.G. (2005). Obtaining accurate measures of time use from the ESM. In B. Schneider & L.J. Waite (Eds.), *Being together, working apart: Dual-career*

- families and the work-life balance*. New York: Cambridge University Press, 461-482.
- Kreuter, F., Couper, M., & Lyberg, L. (2010). The use of paradata to monitor and manage survey data collection. In *Proceedings of the Joint Statistical Meetings, American Statistical Association* (pp. 282-296).
- Mehl, M.R., & Conner, T.S. (Eds.). (2011). *Handbook of research methods for studying daily life*. New York: Guilford Publications.
- Messiah, A., Grondin, O., & Encrenaz, G. (2011). Factors associated with missing data in an experience sampling investigation of substance use determinants. *Drug and alcohol dependence*, 114(2), 153-158.
- Mulligan, C. B., Schneider, B., & Wolfe, R. (2005). Non-response and population representation in studies of adolescent time use. *Electronic International Journal of Time Use Research*, 2(1), 33-53.
- Rabe-Hesketh, S., & Skrondal, A. (2008). *Multilevel and longitudinal modeling using STATA*. College Station, TX: STATA Press.
- Raudenbush, S.W., & Bryk, A. (2002). *Hierarchical linear models*. Thousand Oaks, CA: Sage Publications Inc.
- Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63(3), 581-592.
- Silvia, P. J., Kwapil, T. R., Eddington, K. M., & Brown, L. H. (2013). Missed Beeps and Missing Data: Dispositional and Situational Predictors of Nonresponse in Experience Sampling Research. *Social Science Computer Review* 31(4), 471-481.
- Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press.
- Wagner, J. (2010). The fraction of missing information as a tool for monitoring the quality of survey data. *Public Opinion Quarterly* 74(2), 223-243.
- Zirkel, S., Garcia, J. A., & Murphy, M. C. (2015). Experience-sampling research methods and their potential for education research. *Educational Researcher* 44(1), 7-16.

CHAPTER 2: A PRELIMINARY ANALYSIS OF THE EFFECTS OF RANDOMIZED MINDSET AND BELONGING INTERVENTIONS ON FRESHMEN STUDENTS' ACADEMIC SUCCESS

Introduction

Colleges and universities are more popular than ever before as the next step for recent US high school graduates. Since 1960, enrollment in 4-year degree programs has increased more than 500%, to more than 20 million students in 2012 (National Center for Education Statistics [NCES], 2013). Ballooning enrollment has driven similar dramatic increases in the number of bachelor's degrees granted (400%) and total expenditures by public higher education institutions, which (in 2011 dollars) has increased 1200% from \$25 billion in 1960 to over \$300 billion in 2013 (NCES, 2013). A key driving factor behind this growth is the growing public awareness of the monetary impacts of obtaining a four-year degree- in 2011, the median salary for fully employed 25 year olds with Bachelor's degrees was about \$53,000, about \$9,000 higher than those with an Associate's degree, and nearly \$20,000 higher than workers with a high school diploma. The "college for all" rhetoric has become a key element of federal policymaking, with consistent appeals for early college programs and increased access for low-income and minority students (Obama, 2015).

Despite the evidence of, and appeals for, increased four-year enrollment, in the past decade growing dissent has emerged about the value of a "college for all" philosophy (Rosenbaum, 2001). This criticism intensified following the 2008 financial crisis, which took a particular toll on personal savings and other assets that are frequently used to pay for college. Since then, earnings for recent graduates with four-year degrees have stagnated across almost all majors, and are consistently lower than comparable

earnings a decade earlier (NCES, 2012). Reduced savings and lower earnings have also coincided with increased college loan debt, which has significantly increased the debt burden for college graduates between 2000 and 2011 (NCES, 2013). This problem appears to be most acute for low-income and minority students, who have lower rates of persistence and completion in four-year colleges than their white counterparts.

In sum, given the growing concern for the rising costs of college, particularly for students from low-income backgrounds, and the parallel emphasis on the value of advanced postsecondary schooling for long-term benefit, many policymakers in higher education have begun to look for ways to increase the rates at which students are persisting in college and completing advanced degrees. The possible solutions are numerous and diffuse, however one relatively low-cost set of options that has gained popularity are social psychological interventions that target students' sense of belonging and growth mindsets in an effort to help them 1) reduce anxiety from stereotype threat, and 2) adopt a mindset on intelligence that views it as malleable and open to growth, as opposed to fixed and immutable (Wilson, 2006; Wilson, Damiani & Shelton, 2002; Yeager & Walton, 2011). This study examines preliminary impacts from a randomized control trial for incoming freshmen students at Michigan State University that incorporated belonging and mindset interventions in an attempt to increase preliminary measures of postsecondary persistence (in this case fall grade point average and course credits attempted and completed).

Relevant Literature

In the past ten years, there has been a resurgence of interest in the potential for social psychological interventions to impact students' educational outcomes (Yeager &

Walton, 2011). These interventions are usually very brief in nature (an hour or less), and structured to take advantage of students' motivation structures and social context. In their review of such interventions and their theoretical underpinnings, Yeager and Walton (2011) identify two main classes of social psychological interventions: those designed around 1) changing students' attributions and implicit theories of intelligence, and 2) mitigating stereotype threat in new or unfamiliar social and educational environments. While clearly related to one another, these two types of interventions operate in slightly different ways to impact students' educational outcomes.

Research on attribution theory and its extensions to student success in college is not new. In fact, early work by Wilson and Linville (1982, 1985) suggested that attribution-focused treatments could increase college students' grade point averages by as much as .30 points over a single year. Attribution interventions aim to shift the way in which students attribute academic success or failure from stable factors (typically one's fixed intelligence) to more unstable factors (such as effort). In other words, they aim to convince students that rather than fixed and finite, intelligence is malleable, and one can become smarter and more successful in school by working harder. The delivery mechanism for this message varies, and has included taking part in a brief neuroscience workshop (Blackwell, Trzesniewski, & Dweck, 2007); reading a brief scientific article on "building the brain" aimed at students (Nussbaum & Dweck, 2008); receiving different types of feedback and praise that rewards effort vs. fixed intelligence (Mueller & Dweck, 1998); and watching videos of upperclassmen who reflect on their experience as freshmen and promote the adoption of a growth mindset for academic success (Wilson & Linville, 1982, 1985). Effects in each study vary, but overall results seem to suggest that

attribution (or mindset) interventions could increase students' grade point average by as much as .20 or .30 points over the course of a year.

Along with attribution-based interventions, Yeager and Walton (2011) outline a second class of treatment based on the concept of stereotype threat. Here, the motivating logic is that (particularly racial and ethnic minority) students in educational settings often are worried that they are being “perceived through the lens of a negative intellectual stereotype” (Yeager & Walton, 2011, p. 279), and that this added psychological stress can impact academic performance. Interventions for college students that target stereotype threat usually attempt to assuage these anxieties by either 1) encouraging participants to reflect on their own values and what makes them unique (so-called “self-affirmation” interventions, see Cohen, Garcia, Apfel, & Master, 2006; Cohen, Garcia, Purdie-Vaughns, Apfel, & Brzustoski, 2009), or 2) exposing students to similar peers who share their own experiences with adjusting to a new school or community to encourage the idea that one is not alone in having anxieties about starting college (Walton & Cohen 2007, 2011). In both cases, this type of intervention typically includes written reflection intended to organize students' thinking and stimulate deeper interaction with the content.

The current study examines the preliminary impact of two interventions, one based on mindset/ attribution theory, and a second developed to target students' sense of belonging by mitigating the impact of stereotype threat. Both were developed to closely mimic the approaches advocated by Yeager and Walton (2011) in an effort to replicate the effects of similar work with the student population at Michigan State. As a preliminary analysis, this work focuses on the properties of the initial randomization,

potential attrition of study participants, and the initial treatment effects of the belonging and mindset interventions students' Fall semester GPA, course credits completed, and course credits attempted. The driving research questions are below:

Research Questions

- 1) How effective was the randomization process in balancing observed covariates between treatment and control groups?
- 2) How effective was the randomization process in balancing observed covariates between baseline and analytic samples?
- 3) What levels of attrition were experienced between baseline and analytic samples?
- 4) What were the preliminary effects of the belonging and mindset treatments?
- 5) How does the incorporation of additional statistical controls impact treatment estimates for the belonging and mindset intervention groups?

Below, I describe the methods of this study, including characteristics of the treatment and control conditions, the setting and participants, the randomization and formation of groups, and the overall analytic approach. Then, I present results pertaining to the quality of the randomization process, the initial analysis of treatment effects, and the addition of relevant statistical controls. Finally, I discuss the significance of these results for the extant research literature on belonging and mindset interventions, and offer additional suggestions for further research and implementation of similar experiments.

Methods

Study Characteristics

The summer prior to enrollment, incoming freshmen students at Michigan State are required to attend a two-day long Academic Orientation Program (AOP), which

includes information about course enrollment, academic programs, social and cultural resources, and other key features of campus life. The intervention being studied here was designed to function as a component of this orientation. Several weeks prior to their orientation session, students were sent a link from a university officer to an online survey on Qualtrics requesting their participation in the study. Students were permitted to complete the survey any time before their scheduled AOP session; those that did not complete the survey before AOP were given time to do so after arriving. After providing several pieces of demographic information, students were randomized (blocking on race/ethnicity) into one of three conditions: 1) a *mindset* condition, 2) a *belonging* condition, and 3) a *control* condition. Each condition is described in more detail below, along with additional information about the university setting and target participants.

Mindset Intervention Condition

Students in the mindset intervention group read a short scientific article on “Building the Brain” (adapted from Nussbaum & Dweck, 2008) that introduced the concept of brain plasticity, or the idea that the brain, similar to other muscles, can grow when given repeated practice. The purpose behind this article is to expose students to the idea that their intelligence is not fixed, and that extra effort and focus on their part can translate to significant growth in intelligence over time. After reading the brief article, students are asked several reflective questions, in which they are encouraged to identify moments in their own lives when they may have (or have not) adopted a growth mindset. Students are encouraged to write open-length responses to each reflective question. In total, students typically spent about twenty minutes on the mindset intervention activities.

Belonging Intervention Condition

Students in the belonging treatment group were given a series of quotes that were justified as responses from upper classmen on a recent survey investigating the challenges of starting out in college. These quotes dealt with a series of issues around leaving home, including finding friends, homesickness, fitting in socially, and trying to find one's identity as a member of a new community. Each quote the student reads is attributed to an upper classman at MSU given a pseudonym. Further, the first quote the student reads is matched with the reader's identified gender and race/ethnicity. In other words, if the student identifies as being male and Latino at the start of the survey, the first quote presented is from a Latino male. Later quotes in the series were from other race/ethnicity and gender pairings. After reading the quotes, students were then asked to reflect on their meaning for their own lives in a series of short reflective responses. In total, students typically spent between fifteen and twenty minutes completing the belonging treatment activities.

Comparison Condition

The comparison (or control) condition was a more benign version of the belonging intervention. In this case, instead of prompting specific reflection about people, places, or events that might trigger feelings of belonging, the comparison condition introduced students to much more general reflections of what it will be like to be on campus for the first time. For example, quotations were included that talked about the weather in East Lansing, adjusting to a new class schedule, and finding places to eat. As with the belonging intervention, students in the comparison condition were given a series of quotes to read and then asked to reflect on what the quotes may mean for them

as they begin their college experience. Typically, students completed the control condition activities in ten or fifteen minutes.

Setting

Michigan State University is a large flagship state university, offering more than 200 academic programs housed in 17 degree-granting units. In Fall 2014, more than 43,708 students were enrolled full-time, including 34,788 undergraduates. Of these undergraduates, about 66% are White, 6% are African-American, 4% are Hispanic, and 4% are Asian-American. About 15% of enrolled undergraduates were international students. In terms of gender, slightly more than half of undergraduates were females (50.3%, vs. 49.7% males).

Sample Formation

Eligible participants for this study were identified by their participation in MSU's summer Academic Orientation Program, which is required for all incoming students with "Freshman" academic status. In total, 8,331 students were scheduled to participate in AOP (and therefore eligible for participation in the study); of those 8,331 students 7,686 responded to the survey link and 6,879 completed surveys for a response rate of 92 percent, and a completion rate of 82 percent.

Outcome Measures

This study uses three primary outcome measures, all of which are strongly associated with a students' persistence to a second year of college and eventual completion of a B.A. All three outcomes were obtained via the MSU Office of the Registrar. The first outcome, *grade point average* (GPA), is calculated in the

conventional format, by multiplying the numerical course grade (ranging from 0-4, in increments of 0.50) by the number of credits for a given course, totaling the grade points, and then dividing by the number of credits taken for the semester. The second outcome, *course credits attempted*, is the number of total credits a student attempted in the Fall Semester. The final outcome, *course credits completed*, is the total number of credits a student attempted and received a passing grade (in this case 1.0, or a “D”).

Analytic Approach

This study uses several analytic approaches to investigate the impact of the belonging and mindset treatments. First, it employs mean comparison using independent samples *t*-tests according to Formulas 1 and 2 below:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_{x_1x_2} * \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (1)$$

where

$$s_{x_1x_2} = \sqrt{\frac{(n_1 - 1)s_{x_1}^2 + (n_2 - 1)s_{x_2}^2}{n_1 + n_2 - 2}} \quad (2)$$

In this case, $\bar{X}_1$ and $\bar{X}_2$ represent the sample means for the treatment and control groups, n_1 and n_2 represent the group sample sizes, and $s_{x_1}^2$ and $s_{x_2}^2$ are estimates of the group variances. For significance testing, a *t*-distribution with $(n_1 + n_2 - 2)$ degrees of freedom is used. This test was employed in two ways: 1) to test the baseline equivalence of treatment and control groups based on common covariates, and then 2) to obtain initial (pre-regression) estimates of the treatment effect for each of the three outcomes. In addition to *t*-tests, to better contextualize results with similar work, Cohen’s *d* effect sizes were calculated according to Formula 3 below:

(3)

$$d = \frac{\bar{X}_1 - \bar{X}_2}{SD} = \frac{2t}{\sqrt{N}}$$

Here, $\bar{X}_1$ and $\bar{X}_2$ again represent the sample means for the treatment and control groups, SD represents the pooled standard deviation, t represents the t -value corresponding to the mean difference, and N represents the combined sample size.

The above approach was used to calculate standardized means differences for interval measures, such as the Likert-scaled questions used to examine baseline equivalence. Some measures used in this paper reflect proportions of a group, such as percent female, percent African-American, percent Latino, etc. Here, the approach for calculating a d value was somewhat different. First pairs of proportions (treatment vs. control) were converted to odds ratios, and then the odds ratios were converted to d values using the following formula (Borenstein, Hedges, Higgins & Rothstein, 2009):

(4)

$$d = \text{LogOddsRatio} * \frac{\sqrt{3}}{\pi}$$

Where LogOddsRatio = the natural log of the odds ratio calculated for each set of proportions.

After applying preliminary t -tests, treatment effects are estimated using a series of multiple linear ordinary least squares (OLS) regressions (Murnane & Willett, 2010). These regression models, performed for separately for each treatment-control comparison (e.g. mindset vs. control, treatment vs. control) and each of the three outcomes, build in complexity, beginning with the most basic treatment effect estimator (Model 1). Models 2-5 add additional vectors of covariates that can improve power by absorbing additional residual variation in the outcomes. Models 1-5 are outlined below, and are written in

terms of student i . For all models below, the student-specific error term, ε_i , is assumed to be randomly and independently drawn from a normal distribution with mean 0 and homoscedastic variance, σ_ε^2 .

Model 1: Basic Treatment Model

$$GPA_i = \beta_0 + \beta_1 TREAT_i + \varepsilon_i \quad (5)$$

Where $TREAT_i = 1$ if student was randomized into a treatment group (either mindset or belonging).

Model 2: Treatment Plus Engagement

$$GPA_i = \beta_0 + \beta_1 TREAT_i + \beta_2 Response_pct_i + \beta_3 Word_pct_i + \varepsilon_i \quad (6)$$

Where $Response_pct_i$ = the percent of survey completed for student i , as calculated by dividing the number of survey questions responded to by the total number of survey questions; and $Word_pct_i$ = the percent of relevant content-specific words students use in their open responses, as calculated by dividing the number of content-related words student i uses by the total possible number of content-specific words.

Model 3: Treatment, Engagement, and Demographic Characteristics

$$GPA_i = \beta_0 + \beta_1 TREAT_i + \dots + \beta_4 Female_i + \beta_5 AfricanAm_i + \beta_6 Latino_i + \beta_7 NativeAm_i + \beta_8 AsianAm_i + \beta_9 Multiracial_i + \varepsilon_i \quad (7)$$

Where $Female_i$ = the percent of survey completed for student i , as calculated by dividing the number of survey questions responded to by the total number of survey questions; $AfricanAm_i$, $Latino_i$, $NativeAm_i$, and $AsianAm_i = 1$ if student i identifies as African-American, Latino, Native American, or Asian-American, respectively; and $Multiracial_i = 1$ if student i identifies as more than one of the racial/ethnic categories above.

Model 4: Treatment, Engagement, Demographic Characteristics, and Achievement Pre-Measures

$$GPA_i = \beta_0 + \beta_1 TREAT_i + \dots + \beta_{11} Mom_BA_i + \beta_{12} HS_GPA_i + \beta_{13} ACT_i + \varepsilon_i \quad (8)$$

Where $Mom_BA_i = 1$ if student i reported their mother as earning a Bachelor's degree or higher; HS_GPA_i = student i 's reported high school grade point average; and ACT_i = student i 's composite ACT score.

Model 5: Relevant Interaction Effects

$$GPA_i = \beta_0 + \beta_1 TREAT_i + \dots + \beta_{14} (TREAT * Response_{pct})_i + \beta_{15} (TREAT * Female)_i + \beta_{16} (TREAT * AfricanAm)_i + \beta_{17} (TREAT * Latino)_i + \beta_{18} (TREAT * Mom_BA)_i + \beta_{19} (TREAT * HS_GPA)_i + \beta_{20} (TREAT * ACT)_i + \varepsilon_i \quad (9)$$

Where $(TREAT * Response_{pct})_i$ is a treatment-by-response percent interaction effect; $(TREAT * Female)_i = 1$ if student i was in the treatment group and identified as female (a treatment-by-gender interaction effect); $(TREAT * AfricanAm)_i$ and $(TREAT * Latino)_i = 1$ if student i was in the treatment group and identified as African-American or Latino, respectively (a treatment-by-race/ethnicity interaction effect); $(TREAT * Mom_BA)_i = 1$ if student i was in the treatment group and reported their mother as earning a B.A. degree or higher (a treatment-by-mother's education interaction effect); $(TREAT * HS_GPA)_i$ is a treatment-by-high school GPA interaction effect; and $(TREAT * ACT)_i$ is a treatment-by-ACT score interaction effect.

Statistical Adjustments

Statistical adjustments were made using a series of treatment-related, demographic, and achievement predictors, as described in models one through five above. Table 2.1 presents means and standard deviations for all relevant predictors.

Missing Data

Overall, rates of missing data in this study were quite low- for the pre-randomization questions used to examine the balance of treatment and control groups, rates of missing were less than one percent. This was also the case for the three initial outcomes (GPA, credits attempted, and credits completed), which all had rates of missing below one percent.

Somewhat higher rates of missingness occurred for mother's education, high school GPA, and ACT composite score, which were obtained from a separate survey collection during the semester. Here, rates of missingness ranged between fifteen and seventeen percent, but did not differ significantly between the treatment and control conditions. Given that students were randomly assigned to these groups, and that rates of missingness do not differ, this suggests that the possible bias due to differential rates of missing responses is negligible. In all analyses, observations with missing responses were removed using listwise deletion.

Results

Equivalence of Treatment and Control Groups at Time 1

Table 2.2 presents means and standard deviations for the belonging intervention group (n=2,322) and the control group (n=2,353). For measures of gender and race/ethnicity, a *t*-test using a binomial distribution was used to determine if demographic proportions varied between the treatment and control group. Results suggest that the composition of the two groups does not differ significantly by gender or race. In addition to demographic measures, Table 2.2 also compares control group and experimental group estimates for a series of attitudinal measures related to one's predisposed inclination towards belonging and brain plasticity. Results suggest that across all eight pre-

measures, belonging treatment and control group values do not appear to vary significantly. Reported t-values are all below the $\alpha = .05$ threshold, and the largest reported effect size difference, $d=0.03$, is well below the $d = 0.05$ benchmark which would suggest additional statistical controls might be necessary to obtain unbiased estimates of the treatment effect.

Table 2.3 presents means and standard deviations for the mindset intervention group ($n=2,297$) and the control group ($n=2,353$). Here again, results suggest that the composition of the two groups does not differ significantly by gender or race. For the attitudinal variables, results suggest that for six of eight pre-measures, the mindset treatment and control group values do not appear to vary significantly. Two measures reported significant differences: “worry_not_belong” ($t = 1.85$, $d = 0.05$) was larger for the control group, and “can’t_change_smart” ($t = -2.15$, $d = -0.06$) was larger for the treatment group. In both cases, the reported effect sizes were very close to the $d = 0.05$ threshold, which suggests that some additional statistical controls may be necessary to clearly establish baseline equivalence between these two groups. However, this does not imply that the difference between the groups is enough to violate absolute assumptions of baseline equivalence. An effect size difference of 0.25 or larger is generally needed to reject baseline equivalence altogether.

Equivalence of Treatment and Control Groups at Time 2

Having established the baseline equivalence of the two groups at Time 1 (the initial randomization), it is now necessary to confirm that the treatment and control groups are still equivalent after obtaining initial outcomes. Tables 2.4 and 2.5 present frequencies, means, and standard deviations for the analytic samples for the belonging

treatment (n=2,302), mindset treatment (n=2,297), and control (n=2,330) groups. Here, the analytic sample is defined as all students randomized into the initial experiment that were able to be matched with relevant first semester outcomes.

Table 2.4 compares the belonging and control analytic samples. Here, we again see no significant differences between the two groups, both in terms of *t*-test results or effect size estimates. The largest effect size difference, $d = 0.03$, lies well below the $d = 0.05$ threshold mentioned above. This suggests strong equivalence between the two groups.

Table 2.5 compares the mindset and control analytic samples. Here, as above in the baseline sample, six of the eight baseline measures show no significant differences, while the “worry_not_belong” ($t = 2.04$, $d = 0.06$) and “can’t_change_smart” ($t = -2.08$, $d = -0.06$) items showed differences by group. In terms of effect size, both measures were associated with an effect size difference of 0.06, which suggests that some additional statistical controls may be necessary to reduce potential bias in treatment estimates.

Attrition Rates

Having established equivalence for pre-measures in both the baseline and analytic samples, next I examine rates of attrition between the two samples. Attrition is assessed in two forms- 1) overall attrition, which is the percentage of the overall sample (both treatment and control) that is lost from the initial randomization, and 2) differential attrition, which is the extent to which attrition occurs in an unbalanced fashion, with higher rates in one group than another. Table 2.6 presents sample sizes and attrition rates for both the mindset and belonging experiments. In both cases, overall attrition was very low, less than 1% in each case. The same is true for differential attrition- less than 1% in

both situations. Comparing these rates to established guidelines in Institute for Education Sciences (2014a; 2014b), these rates of attrition fall well within “tolerable” range for obtaining relatively unbiased treatment effect estimates.

Initial Results- Mean Differences

Table 2.7 Compares outcomes for the belonging treatment and control groups using *t*-tests. Students in the belonging treatment had slightly higher GPAs ($\bar{x}_T = 3.15$, $\bar{x}_C = 3.12$, $t = -1.23$), and attempted ($\bar{x}_T = 13.46$, $\bar{x}_C = 13.45$, $t = -0.25$) and completed ($\bar{x}_T = 13.00$, $\bar{x}_C = 12.97$, $t = -0.42$) slightly more credits than students in the control condition. However, these differences were quite small, and in all cases did not differ significantly from zero. Estimated effect sizes of the treatment effect range from 0.01 for credits attempted and completed, to 0.04 for Fall GPA.

Similar results appear for the mindset experiment in Table 2.8. Here, students in the treatment group again had slightly higher GPAs ($\bar{x}_T = 3.17$, $\bar{x}_C = 3.12$, $t = -2.04$), and attempted ($\bar{x}_T = 13.53$, $\bar{x}_C = 13.45$, $t = -1.38$) and completed ($\bar{x}_T = 13.08$, $\bar{x}_C = 12.97$, $t = -1.50$) more credits than students in the control condition. Again, these differences were quite small, although the *t*-test result for GPA suggests that this difference is in fact significant at the $\alpha = 0.05$ significance level. Estimated effect sizes were quite similar to the belonging experiment, ranging from 0.01 for the credit outcomes and 0.04 for Fall GPA.

Regression Models

Building off of these initial mean comparisons, treatment effects are next estimated using a series of regression models. These models begin with the most basic approach, OLS regression with a single predictor of treatment, and then progress by

adding vectors of relevant covariates. These include predictors related to treatment dosage and/or engagement, demographic characteristics, pre-measures of student achievement, and finally all relevant interactions between treatment and key predictors from previous models.

Belonging Intervention

Tables 2.9-2.11 present regression estimates for the impact of the belonging intervention on Fall semester GPA, credits completed, and credits attempted, respectively. Table 2.9 presents results for the impact on Fall semester GPA. Here, students in the treatment condition (“belong”) do not appear to earn significantly more (or fewer) credits than students in the control condition. This trend holds for Model 1 (see column 1), as well as for more complex models that account for engagement with treatment, demographics, and prior achievement. Model 5, which introduces treatment-specific interaction effects, suggests that the belonging treatment may have differential positive effects for Latino students. Compared to Model 4, in which Latino students are associated with a negative and significant difference in GPA points of about 0.15 ($p < .01$) relative to their white student peers, Model 5 suggests that for Latino students in the belonging intervention group are associated with a positive and significant impact of about 0.20 ($p < .10$) relative to Latino students in the control group, after controlling for all other covariates.

While Latino students may receive differential positive benefits to GPA from the belonging treatment, students who report their mother as earning a Bachelor’s degree seem to experience negative impacts. Model 4 suggests that overall, students in this sample who report their mother having earned a B.A. (or higher) are associated with a

positive and significant difference of about 0.12 GPA points ($p < .01$) compared to students with mothers who have not received a B.A. However, Model 5 suggests that for students in the belonging treatment group whose mother earned a B.A., this impact is reduced by about 0.08 GPA points ($p < .10$).

Table 2.10 presents results for the impact of the belonging treatment on completed credits in the Fall semester. Similar to GPA, there is no impact for the belonging treatment on completed credits in Model 1, as well as in Models 2-5 which introduce additional predictors. Model 5, which introduces treatment-specific interaction effects for survey engagement, gender, race, and achievement, shows no significant effects for any interactions. This suggests that the treatment did not have any differential impacts on completed credits.

Table 2.11 presents results for the impact of the belonging treatment on attempted credits in the Fall semester. Once again, the belonging treatment appears to have no significant impact in all models 1-5. However, Model 5 suggests several differential impacts. For example, the impact of survey engagement for students in the belonging treatment is a small and negative differential effect of about -0.01 ($p < .10$) compared to the overall impact of engagement, which is not different from zero. Further, for students in the belonging treatment, the impact of an increase of one unit in ACT score was increased by about 0.04 credits ($p < .05$) compared to the impact of ACT score overall, which was not significant. This suggests that the belonging treatment may have had a very small differential positive effect for students with higher ACT scores.

Mindset Intervention

Tables 2.12-2.14 present regression estimates for the impact of the mindset intervention on the three outcomes of interest. Table 2.12 presents results for the impact on Fall semester GPA. Here, students in the treatment condition (“mindset”) do appear to earn significantly more credits than students in the control condition. The coefficient for treatment, 0.0476 ($p < .05$), suggests that students in the mindset intervention group earned a GPA that was about 0.05 points higher than students in the control group. This estimate is similar in significance and magnitude for Model 1 (see column 1), as well as for more complex models that account for engagement with treatment, demographics, and prior achievement. Model 5, which introduces treatment-specific interaction effects, suggests that the mindset treatment may have differential positive effects for Latino students. Compared to Model 4, in which Latino students are associated with a negative and significant difference in GPA points of about 0.10 ($p < .10$) relative to their white student peers, Model 5 suggests that Latino students in the mindset intervention group are associated with a positive and significant differential impact of about 0.30 ($p < .01$) relative to Latino students in the control group, after controlling for all other covariates. The magnitude of this positive interaction mitigates, and somewhat reverses, the Latino student effect in the overall sample.

While Latino students may receive differential positive benefits to GPA from the mindset treatment, other treatment-specific interaction effects in Model 5 suggests negative and significant differential impacts from treatment for the high school GPA and ACT score covariates. For example, for students in the mindset treatment, the impact of the high school GPA coefficient on college GPA was reduced by about 0.15 GPA points. The main effect for GPA in Model 4, 0.704 ($p < .01$) suggests that an increase of one

point in high school GPA is associated with an increase of about 0.70 points in high school GPA, after controlling for all other covariates. For students in the mindset group, this increase was 0.15 GPA points smaller, or about 0.55 GPA points.

A similar differential effect appears to exist with ACT score and the mindset treatment. Model 4 reports a positive and significant main effect for ACT of 0.02 ($p < .01$). In other words, an increase of one point on the ACT is associated with an increase of about 0.02 GPA points, after controlling for all other covariates. However, for students in the mindset treatment, this effect is reduced by 0.015 ($p < .05$) GPA points. Thus, the mindset treatment appears to have small differential negative effects for student with higher ACT scores, even after controlling for all other covariates.

Table 2.13 presents results for the impact of the mindset treatment on completed credits in the Fall semester. Here, there appears to be no impact for the mindset treatment on completed credits in Model 1, as well as in Models 3-5 which introduce additional predictors. Model 5, which introduces treatment-specific interaction effects for survey engagement, gender, race, and achievement, shows a significant and positive differential effect for Latino students of 0.664 ($p < .10$). In Model 3, the main effect for Latino students was -0.418 ($p < .05$), which suggests that Latino students completed about half a credit less than their white counterparts. However, this difference is no longer significant in Model 4, after controlling for high school GPA and ACT score. Thus, the interpretation of the differential effect for Latino students in the mindset group is somewhat difficult. However, a conservative interpretation might hold that Latino students in the mindset group earned slightly more credits than students in the control group, after controlling for all other covariates.

Finally, Table 2.14 presents results for the impact of the mindset treatment on attempted credits in the Fall semester. Similar to completed credits, there is no impact for the belonging treatment on completed credits in Model 1, as well as in Models 3-5 which introduce additional predictors. Model 5, which introduces treatment-specific interaction effects for survey engagement, gender, race, and achievement, shows no significant effects for any interactions. This suggests that the treatment did not have any differential impacts on attempted credits.

Note on Survey Engagement Predictor

When examining the impact of the belonging intervention, the introduction of the survey engagement predictor (p_tot_response) in Model 2, which represents the proportion of treatment (or control) questions completed by the participant, did not appear to significantly impact the estimates of the treatment itself. In other words, the extent to which a student thoroughly completed the activities in the survey did not appear to impact the estimate of the treatment (although before introducing additional predictors, survey engagement did significantly predict GPA- see Table 2.9, Column 2). For the mindset intervention, survey engagement was consistently positive and significant in Model 2 for all three outcomes, however typically this significance disappeared in later models that introduce stronger covariates. In all cases except for one (see Table 2.11, Column 5), however, the treatment-by-engagement interaction was not significant, which suggests that there were not systematic differential effects resulting from students' engagement with the treatment activities versus the mindset activities.

Discussion

Previous research has suggested that brief, low-cost social psychological interventions that target sense of belonging and growth mindset may significantly impact key indicators of student persistence and completion of a postsecondary degree. This study extends that work by testing two such interventions separately with a large and diverse sample of incoming undergraduates at Michigan State. Initial results from this investigation indicate no significant impacts from the belonging intervention on any of the three outcomes of interest, and small but significant effects ($d = 0.04$) of the mindset intervention on Fall semester GPA, but no effects on credits attempted or completed.

Compared to previous research, this study differs in several important ways, which may help to contextualize the findings. First, this study intentionally examines the impact of belonging and mindset separately, using a three-group design (belonging treatment, mindset treatment, and comparison/ control). Some studies that report significant effects from similar interventions examine the so-called “double dose” of mindset and belonging treatments delivered to the same experimental group. In many ways, belonging and mindset interventions compliment one another in terms of their logic and approach; so combining the two seems a reasonable choice. In the case of this study, however, a key question was how the two interventions might function independently. Without a separate examination of these effects, it can be difficult to accurately discern which treatment (if any) is driving an effect. This study can provide new information about how the two treatments function in the context of the same student population, and in this case, it seems that the mindset intervention may be more impactful at least in terms of initial outcomes.

In addition to separating the belonging and mindset treatments, this study adopts a slightly different approach to measuring outcomes. Previous work in this field has often measured experimental impacts on the residual of outcomes such as grades. In our case and context, a more straightforward impact assessment was most useful, particularly as this work is used to inform institutional program improvement and needs to be widely accessible and interpretable. Future analysis might introduce additional outcomes such as residual grades in an effort to more closely compare results from this study with those currently underway at other institutions.

Limitations

This study has two key limitations: sample size and time. First, while the overall sample size of the study is quite robust, the demographics of the student body make it difficult to offer strong claims about the impact of the intervention on particular racial and ethnic subgroups. Overall, more than 75 percent of the sample is composed of white students, so one could reasonably expect that trends in this group would be the significant driver of overall treatment estimates. Several interaction effects suggest that differential impacts may exist for subgroups; this is certainly the case for Latino students, who were associated with strong differential effects in the mindset treatment. Starting this past summer, students were randomized by block according to their reported race/ethnicity, which will hopefully allow for building a larger, longitudinal experimental sample by block in later years. This will allow for much more robust estimation of treatment effects within each group.

Along with sample size, time is a key limitation in this work. This analysis represents an initial examination of key outcomes related to postsecondary persistence

and success. Additional time will provide more opportunities for meaningful measurement of other outcomes, including students' persistence into the second year of study at MSU, their choice of major, and their ongoing GPAs and rates of credit completion. Some research in this field has suggested the possibility of persistent but delayed effects resulting from social psychological interventions; following the current sample of students for the next few years will allow us to estimate this.

Conclusion

These findings represent an initial exploration at the university level of using randomized interventions to stimulate persistence and success. Early on, the mindset intervention appears to be more impactful, producing significant, albeit small positive impacts on Fall semester GPA. With the introduction of subsequent cohorts in the years to come, this study will hopefully provide even clearer insights into types of interventions that may support students as they progress through their postsecondary experience.

APPENDIX

Appendix

Table 2.1 Descriptive Statistics for and Outcomes and Controls used in Regression Analysis

<u>Belonging Intervention</u>					
Variable	N	Mean	S.E.	Min	Max
GPA	4633	3.135	0.012	0	4
Credits_Cmplt	4633	12.984	0.039	0	19
Credits_Att	4633	13.455	0.030	0	20
Belong	4633	0.497	0.007	0	1
Response_pct	4633	95.719	0.146	17.15	100
Female	4581	0.538	0.016	0	1
AfricanAm	4581	0.079	0.004	0	1
Latino	4581	0.044	0.003	0	1
NativeAm	4581	0.003	0.001	0	1
AsianAm	4581	0.052	0.003	0	1
Multiracial	4581	0.038	0.003	0	1
Mom_BA	3740	0.459	0.014	0	1
HS_GPA	3740	3.695	0.006	0	5.34
ACT	3740	26.018	0.059	15	36
<u>Mindset Intervention</u>					
Variable	N	Mean	S.E.	Min	Max
GPA	4607	3.144	0.012	0	4
Credits_Cmplt	4607	13.024	0.038	0	19
Credits_Att	4607	13.488	0.029	0	19
mindset	4607	0.494	0.015	0	1
Response_pct	4607	94.438	0.195	12	100
female	4560	0.536	0.016	0	1
AfricanAm	4560	0.077	0.004	0	1
Latino	4560	0.046	0.003	0	1
NativeAm	4560	0.004	0.001	0	1
AsianAm	4560	0.053	0.003	0	1
Multiracial	4560	0.037	0.003	0	1
Mom_BA	3715	0.461	0.015	0	1
HS_GPA	3715	3.693	0.006	0	5.34
ACT	3715	25.938	0.059	15	36

Table 2.2 Descriptive Statistics for Pre-Randomization Questions at Time 1, Belonging Experiment

Variables	Belonging Treatment (n=2,322)		Belonging Control (n=2,353)			
<u>Demographic Pre-Measures</u>	<u>Mean</u>	<u>S.E.</u>	<u>Mean</u>	<u>S.E.</u>	<u>z-score</u>	<u>Effect Size (Cohen's <i>d</i>)</u>
female	0.550	0.010	0.530	0.010	-1.06	0.015
Race/ ethnicity						
White	0.780	0.010	0.780	0.010	0.22	-0.004
African-American	0.080	0.010	0.080	0.010	-0.56	0.015
Latino	0.043	0.004	0.045	0.004	0.26	-0.009
Native American	0.003	0.001	0.004	0.001	0.75	-0.094
Asian-American	0.053	0.005	0.052	0.005	-0.03	0.001
Multiracial	0.039	0.004	0.038	0.004	-0.16	0.006
None specified	0.009	0.002	0.009	0.002	0.11	-0.008
<u>Belonging and Mindset Pre-Measures</u>	<u>Mean</u>	<u>S.E.</u>	<u>Mean</u>	<u>S.E.</u>	<u>t-value</u>	<u>Effect Size (Cohen's <i>d</i>)</u>
confident_belong	4.06	0.02	4.09	0.02	1.14	0.03
wonder_fit_in	2.37	0.02	2.34	0.02	-0.75	-0.02
worry_not_belong	1.89	0.02	1.91	0.02	0.94	0.03
certain_fit_in	3.67	0.02	3.66	0.02	-0.38	-0.01
smart or not	2.36	0.02	2.35	0.02	-0.06	-0.002
cant_change_intel	2.29	0.02	2.31	0.02	0.47	0.01
cant_change_smart	2.41	0.02	2.39	0.02	-0.77	-0.02
can_grow_intel	5.18	0.02	5.16	0.02	-0.84	-0.02

Note: Total N = 4675. + p-value ≤ 0.10 ; * p-value ≤ 0.05 ; ** p-value ≤ 0.01

Table 2.3 Descriptive Statistics for Pre-Randomization Questions at Time 1, Mindset Experiment

Variables	Mindset Treatment (n=2,297)		Belonging Control (n=2,353)		z-score	Effect Size (Cohen's <i>d</i>)
	Mean	S.E.	Mean	S.E.		
<u>Demographic Pre-Measures</u>						
female	0.54	0.01	0.53	0.01	-0.69	0.01
Race/ ethnicity						
White	0.779	0.009	0.784	0.009	0.44	-0.007
African-American	0.080	0.006	0.076	0.006	-0.48	0.012
Latino	0.047	0.004	0.045	0.004	-0.22	0.007
Native American	0.003	0.001	0.004	0.001	0.46	-0.056
Asian-American	0.054	0.005	0.052	0.005	-0.23	0.007
Multiracial	0.037	0.004	0.038	0.004	0.10	-0.004
None specified	0.006	0.002	0.009	0.002	1.27	-0.103
<u>Belonging and Mindset Pre-Measures</u>						
	Mean	S.E.	Mean	S.E.	t-value	Effect Size (Cohen's <i>d</i>)
confident_belong	4.08	0.02	4.09	0.02	0.16	0.005
wonder_fit_in	2.36	0.02	2.34	0.02	-0.53	-0.02
worry_not_belong	1.86	0.02	1.91	0.02	1.85*	0.05
certain_fit_in	3.67	0.02	3.66	0.02	-0.35	-0.01
smart or not	2.40	0.02	2.35	0.02	-1.27	-0.04
cant_change_intel	2.35	0.02	2.31	0.02	-1.40	-0.04
cant_change_smart	2.46	0.02	2.39	0.02	-2.15**	-0.06
can_grow_intel	5.12	0.02	5.16	0.02	1.50	0.04

Note: Total N = 4650. + p-value ≤ 0.10 ; * p-value ≤ 0.05 ; ** p-value ≤ 0.01

Table 2.4 Descriptive Statistics for Pre-Randomization Questions at Time 2, Belonging Experiment

Variables	Belonging Treatment (n=2,302)		Belonging Control (n=2,330)			
<u>Demographic Pre-Measures</u>	<u>Mean</u>	<u>S.E.</u>	<u>Mean</u>	<u>S.E.</u>	<u>z-score</u>	<u>Effect Size (Cohen's d)</u>
female						0.014
Race/ ethnicity						
White	0.784	0.009	0.786	0.009	0.13	-0.002
African-American	0.081	0.006	0.076	0.006	-0.55	0.014
Latino	0.042	0.004	0.045	0.004	0.64	-0.022
Native American	0.002	0.001	0.004	0.001	1.05	-0.138
Asian-American	0.053	0.005	0.052	0.005	-0.22	0.007
Multiracial	0.039	0.004	0.037	0.004	-0.23	0.009
None specified	0.009	0.002	0.009	0.002	0.11	-0.008
<u>Belonging and Mindset Pre-Measures</u>	<u>Mean</u>	<u>S.E.</u>	<u>Mean</u>	<u>S.E.</u>	<u>t-value</u>	<u>Effect Size (Cohen's d)</u>
confident_belong	4.06	0.02	4.09	0.02	0.98	0.03
wonder_fit_in	2.37	0.02	2.34	0.02	-0.72	-0.02
worry_not_belong	1.89	0.02	1.92	0.02	0.98	0.03
certain_fit_in	3.68	0.02	3.66	0.02	-0.42	-0.01
smart or not	2.36	0.02	2.36	0.02	-0.03	-0.001
cant_change_intel	2.29	0.02	2.31	0.02	0.41	0.01
cant_change_smart	2.41	0.02	2.39	0.02	-0.72	-0.02
can_grow_intel	5.18	0.02	5.16	0.02	-0.87	-0.03

Note. Total N = 4632. + p-value ≤ 0.10 ; * p-value ≤ 0.05 ; ** p-value ≤ 0.01

Table 2.5 Descriptive Statistics for Pre-Randomization Questions at Time 2, Mindset Experiment

Variables	Mindset Treatment (n=2,297)		Belonging Control (n=2,330)		z-score	<u>Effect Size</u> (Cohen's <i>d</i>)
	<u>Mean</u>	<u>S.E.</u>	<u>Mean</u>	<u>S.E.</u>		
<u>Demographic Pre-Measures</u>						
female	0.540	0.010	0.530	0.010	-0.69	0.010
Race/ ethnicity						
White	0.786	0.009	0.780	0.009	0.42	-0.007
African-American	0.078	0.006	0.076	0.006	-0.25	0.007
Latino	0.047	0.004	0.045	0.004	-0.29	0.010
Native American	0.003	0.001	0.004	0.001	0.46	-0.056
Asian-American	0.054	0.005	0.052	0.005	-0.36	0.011
Multiracial	0.037	0.004	0.037	0.004	0.03	-0.001
None specified	0.006	0.002	0.009	0.002	1.27	-0.104
<u>Belonging and Mindset Pre-Measures</u>	<u>Mean</u>	<u>S.E.</u>	<u>Mean</u>	<u>S.E.</u>	<u>t-value</u>	<u>Effect Size</u> (Cohen's <i>d</i>)
confident_belong	4.09	0.02	4.09	0.02	-0.03	-0.001
wonder_fit_in	2.36	0.02	2.34	0.02	-0.37	-0.01
worry_not_belong	1.86	0.02	1.92	0.02	2.04**	0.06
certain_fit_in	3.67	0.02	3.66	0.02	-0.44	-0.01
smart or not	2.40	0.02	2.36	0.02	-1.29	-0.04
cant_change_intel	2.35	0.02	2.31	0.02	-1.37	-0.04
cant_change_smart	2.46	0.02	2.39	0.02	-2.08**	-0.06
can_grow_intel	5.12	0.02	5.16	0.02	1.36	0.04

Note. Total *N* = 4606. + $p\text{-value} \leq 0.10$; * $p\text{-value} \leq 0.05$; ** $p\text{-value} \leq 0.01$

Table 2.6 Sample Sizes and Attrition Rates for Mindset and Belonging Experiments

<u>Mindset Experiment</u>						
Outcome	Randomized Sample		Analytic Sample		Attrition Rates	
	<u>N Treat</u>	<u>N Control</u>	<u>N Treat</u>	<u>N Control</u>	<u>Overall</u>	<u>Differential</u>
Fall GPA	2298	2353	2277	2330	0.91%	0.98%
Credits Completed	2298	2353	2277	2330	0.91%	0.98%
Credits Attempted	2298	2353	2277	2330	0.91%	0.98%
<u>Belonging Experiment</u>						
Outcome	Randomized Sample		Analytic Sample		Attrition Rates	
	<u>N Treat</u>	<u>N Control</u>	<u>N Treat</u>	<u>N Control</u>	<u>Overall</u>	<u>Differential</u>
Fall GPA	2323	2353	2303	2330	0.86%	0.98%
Credits Completed	2323	2353	2303	2330	0.86%	0.98%
Credits Attempted	2323	2353	2303	2330	0.86%	0.98%

Note: Control group (belonging control) is same for both experiments.

Table 2.7 Estimated Treatment Effects, Belonging Experiment

Outcome	Belonging Treatment (n=2,303)		Belonging Control (n=2,330)		<i>t</i> -value	Effect Size (<i>d</i>)
	<u>Mean</u>	<u>S.E.</u>	<u>Mean</u>	<u>S.E.</u>		
Fall GPA	3.15	0.02	3.12	0.02	-1.23	-0.04
Credits Completed	13.00	0.05	12.97	0.06	-0.42	-0.01
Credits Attempted	13.46	0.04	13.45	0.04	-0.25	-0.01

Note: Total *N* = 4633.

Table 2.8 Estimated Treatment Effects, Mindset Experiment

Outcome	Mindset Treatment (n=2,277)		Belonging Control (n=2,330)		<i>t</i> -value	Effect Size (<i>d</i>)
	<u>Mean</u>	<u>S.E.</u>	<u>Mean</u>	<u>S.E.</u>		
Fall GPA	3.17	0.02	3.12	0.02	-2.04**	-0.04
Credits Completed	13.08	0.05	12.97	0.06	-1.50	-0.01
Credits Attempted	13.53	0.04	13.45	0.04	-1.38	-0.01

Note: Total *N* = 4607.

Table 2.9 Regression Estimates for Impact of Belonging Treatment on Fall GPA

VARIABLES	(1) Treatment Only	(2) T plus engagement	(3) Demographics	(4) Achievement	(5) Interactions
belong	0.0291 (0.0236)	0.0298 (0.0236)	0.0227 (0.0229)	0.0156 (0.0218)	-7.44e-05 (0.341)
p_tot_response		0.00298** (0.00119)	0.00260** (0.00115)	0.00152 (0.00111)	0.00162 (0.00169)
female			0.152*** (0.0231)	0.112*** (0.0226)	0.116*** (0.0318)
african_am			-0.651*** (0.0430)	-0.247*** (0.0441)	-0.253*** (0.0625)
latino			-0.410*** (0.0564)	-0.150*** (0.0565)	-0.244*** (0.0773)
native_am			-0.132 (0.208)	-0.0877 (0.193)	-0.0866 (0.193)
asian_am			0.00973 (0.0518)	-0.00707 (0.0491)	-0.00708 (0.0491)
multiracial			-0.141** (0.0602)	-0.0550 (0.0578)	-0.0597 (0.0579)
mom_ba				0.117*** (0.0225)	0.154*** (0.0318)
hs_gpa				0.764*** (0.0373)	0.778*** (0.0518)
act				0.0190*** (0.00371)	0.0157*** (0.00521)
bresp_interaction					-0.000321 (0.00224)
bfemale_interaction					-0.00908 (0.0452)
bafam_interaction					0.0175

Table 2.9 (cont'd.)

					(0.0878)
blatino_interaction					0.199*
					(0.113)
bmomBA_interaction					-0.0762*
					(0.0451)
bgpa_interaction					-0.0252
					(0.0747)
bact_interaction					0.00654
					(0.00742)
Constant	3.120***	2.835***	2.867***	-0.377**	-0.365
	(0.0166)	(0.115)	(0.113)	(0.170)	(0.248)
Observations	4,633	4,633	4,581	3,740	3,740
R-squared	0.000	0.002	0.064	0.224	0.226

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.10 Regression Estimates for Impact of Belonging Treatment on Credits Completed

VARIABLES	(1) Treatment Only	(2) T plus engagement	(3) Demographics	(4) Achievement	(5) Interactions
belong	0.0331 (0.0778)	0.0343 (0.0778)	0.0116 (0.0769)	-0.0243 (0.0749)	0.609 (1.172)
p_tot_response		0.00594 (0.00392)	0.00509 (0.00387)	0.000157 (0.00381)	0.00274 (0.00580)
female			0.422*** (0.0774)	0.339*** (0.0776)	0.283*** (0.109)
african_am			-1.423*** (0.144)	-1.076*** (0.152)	-0.884*** (0.215)
latino			-0.896*** (0.189)	-0.496** (0.194)	-0.590** (0.266)
native_am			-0.126 (0.697)	-0.0568 (0.663)	-0.0380 (0.663)
asian_am			-0.241 (0.174)	-0.240 (0.169)	-0.236 (0.169)
multiracial			-0.171 (0.202)	-0.100 (0.199)	-0.109 (0.199)
mom_ba				0.235*** (0.0774)	0.346*** (0.109)
hs_gpa				1.145*** (0.128)	1.252*** (0.178)
act				-0.00617 (0.0127)	-0.0183 (0.0179)
bresp_interaction					-0.00470 (0.00769)
bfemale_interaction					0.122 (0.155)
bafam_interaction					-0.361

Table 2.10 (cont'd.)

					(0.302)
blatino_interaction					0.191
					(0.389)
bmomBA_interaction					-0.226
					(0.155)
bgpa_interaction					-0.203
					(0.257)
bact_interaction					0.0241
					(0.0255)
Constant	12.97***	12.40***	12.44***	8.925***	8.563***
	(0.0549)	(0.379)	(0.379)	(0.586)	(0.852)
Observations	4,633	4,633	4,581	3,740	3,740
R-squared	0.000	0.001	0.030	0.069	0.071

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.11 Regression Estimates for Impact of Belonging Treatment on Credits Attempted

VARIABLES	(1) Treatment Only	(2) T plus engagement	(3) Demographics	(4) Achievement	(5) Interactions
belong	0.0148 (0.0600)	0.0343 (0.0778)	0.00772 (0.0601)	-0.0352 (0.0523)	0.610 (0.819)
p_tot_response		0.00594 (0.00392)	0.000607 (0.00302)	-0.00375 (0.00266)	0.00157 (0.00405)
female			0.216*** (0.0604)	0.187*** (0.0542)	0.207*** (0.0763)
african_am			-0.605*** (0.113)	-0.402*** (0.106)	-0.362** (0.150)
latino			-0.257* (0.148)	-0.131 (0.136)	-0.108 (0.186)
native_am			0.0649 (0.545)	0.147 (0.463)	0.143 (0.463)
asian_am			-0.221 (0.136)	-0.233** (0.118)	-0.235** (0.118)
multiracial			0.0668 (0.158)	0.0908 (0.139)	0.0815 (0.139)
mom_ba				0.100* (0.0541)	0.100 (0.0763)
hs_gpa				0.424*** (0.0895)	0.508*** (0.124)
act				0.0108 (0.00890)	-0.00628 (0.0125)
bresp_interaction					-0.00952* (0.00537)
bfemale_interaction					-0.0357 (0.108)
bafam_interaction					-0.0487

Table 2.11 (cont'd.)

					(0.211)
blatino_interaction					-0.0666
					(0.271)
bmomBA_interaction					-0.00910
					(0.108)
bgpa_interaction					-0.166
					(0.179)
bact_interaction					0.0351**
					(0.0178)
Constant	13.45***	12.40***	13.35***	12.06***	11.67***
	(0.0423)	(0.379)	(0.296)	(0.409)	(0.595)
Observations	4,633	4,633	4,581	3,740	3,740
R-squared	0.000	0.001	0.010	0.028	0.030

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.12 Regression Estimates for Impact of Mindset Treatment on Fall GPA

VARIABLES	(1) Treatment Only	(2) T plus engagement	(3) Demographics	(4) Achievement	(5) Interactions
mindset	0.0476** (0.0233)	0.0534** (0.0234)	0.0492** (0.0227)	0.0440** (0.0216)	0.447 (0.318)
p_tot_response		0.00207** (0.000885)	0.00107 (0.000864)	-5.20e-05 (0.000824)	0.00160 (0.00166)
female			0.122*** (0.0227)	0.106*** (0.0221)	0.115*** (0.0312)
african_am			-0.682*** (0.0429)	-0.259*** (0.0439)	-0.254*** (0.0613)
latino			-0.282*** (0.0540)	-0.101* (0.0539)	-0.245*** (0.0758)
native_am			-0.0676 (0.192)	0.0480 (0.198)	0.0301 (0.198)
asian_am			-0.0230 (0.0509)	-0.0213 (0.0480)	-0.0228 (0.0480)
multiracial			-0.142** (0.0602)	-0.0768 (0.0574)	-0.0786 (0.0574)
mom_ba				0.133*** (0.0221)	0.154*** (0.0312)
hs_gpa				0.704*** (0.0365)	0.778*** (0.0508)
act				0.0231*** (0.00365)	0.0157*** (0.00511)
resp_interaction					-0.00215 (0.00191)
female_interaction					-0.0184 (0.0442)
afam_interaction					-0.00187

Table 2.12 (cont'd.)

latino_interaction					(0.0871) 0.286***
					(0.107)
momBA_interaction					-0.0454 (0.0441)
gpa_interaction					-0.154** (0.0730)
act_interaction					0.0151** (0.00728)
Constant	3.120*** (0.0164)	2.921*** (0.0864)	3.029*** (0.0856)	-0.112 (0.154)	-0.361 (0.243)
Observations	4,607	4,607	4,560	3,715	3,715
R-squared	0.001	0.002	0.062	0.226	0.229

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.13 Regression Estimates for Impact of Mindset Treatment on Credits Completed

VARIABLES	(1) Treatment Only	(2) T plus engagement	(3) Demographics	(4) Achievement	(5) Interactions
mindset	0.114 (0.0758)	0.135* (0.0762)	0.112 (0.0752)	0.0745 (0.0717)	0.869 (1.060)
p_tot_response		0.00754*** (0.00287)	0.00569** (0.00286)	0.00284 (0.00274)	0.00271 (0.00551)
female			0.302*** (0.0752)	0.206*** (0.0736)	0.275*** (0.104)
african_am			-1.419*** (0.142)	-1.050*** (0.146)	-0.863*** (0.204)
latino			-0.418** (0.179)	-0.240 (0.179)	-0.575** (0.253)
native_am			0.185 (0.633)	0.555 (0.658)	0.535 (0.658)
asian_am			-0.168 (0.168)	-0.247 (0.160)	-0.248 (0.160)
multiracial			0.105 (0.199)	0.104 (0.191)	0.103 (0.191)
mom_ba				0.409*** (0.0736)	0.345*** (0.104)
hs_gpa				1.098*** (0.122)	1.261*** (0.169)
act				-0.00875 (0.0122)	-0.0171 (0.0170)
resp_interaction					0.000158 (0.00635)
female_interaction					-0.134 (0.147)
afam_interaction					-0.363

Table 2.13 (cont'd.)

latino_interaction					(0.290) 0.664*
					(0.357)
momBA_interaction					0.122 (0.147)
gpa_interaction					-0.323 (0.243)
act_interaction					0.0155 (0.0242)
Constant	12.97*** (0.0533)	12.25*** (0.281)	12.41*** (0.283)	8.877*** (0.511)	8.498*** (0.810)
Observations	4,607	4,607	4,560	3,715	3,715
R-squared	0.000	0.002	0.027	0.076	0.079

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2.14 Regression Estimates for Impact of Mindset Treatment on Credits Attempted

VARIABLES	(1) Treatment Only	(2) T plus engagement	(3) Demographics	(4) Achievement	(5) Interactions
mindset	0.0798 (0.0577)	0.135* (0.0762)	0.0799 (0.0580)	0.0497 (0.0477)	0.378 (0.705)
p_tot_response		0.00754*** (0.00287)	0.00386* (0.00220)	0.00176 (0.00182)	0.00160 (0.00367)
female			0.166*** (0.0580)	0.131*** (0.0489)	0.204*** (0.0690)
african_am			-0.479*** (0.109)	-0.341*** (0.0972)	-0.349** (0.136)
latino			0.0754 (0.138)	0.0192 (0.119)	-0.0972 (0.168)
native_am			0.0989 (0.488)	0.383 (0.438)	0.369 (0.438)
asian_am			-0.0539 (0.130)	-0.182* (0.106)	-0.183* (0.106)
multiracial			0.228 (0.153)	0.196 (0.127)	0.194 (0.127)
mom_ba				0.174*** (0.0489)	0.101 (0.0691)
hs_gpa				0.424*** (0.0808)	0.513*** (0.113)
act				0.000307 (0.00808)	-0.00586 (0.0113)
resp_interaction					0.000281 (0.00423)
female_interaction					-0.148 (0.0978)
afam_interaction					0.0305

Table 2.14 (cont'd.)

					(0.193)
latino_interaction					0.226
					(0.238)
momBA_interaction					0.143
					(0.0978)
gpa_interaction					-0.181
					(0.162)
act_interaction					0.0122
					(0.0161)
Constant	13.45***	12.25***	13.02***	11.79***	11.63***
	(0.0406)	(0.281)	(0.218)	(0.340)	(0.539)
Observations	4,607	4,607	4,560	3,715	3,715
R-squared	0.000	0.002	0.008	0.030	0.032

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

REFERENCES

REFERENCES

- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78(1), 246-263.
- Borenstein, M., Hedges, L.V., Higgins, J. P. T., & Rothstein, H.R. (2009). *Introduction to Meta-Analysis*. New York: John Wiley & Sons.
- Cohen, G. L., Garcia, J., Apfel, N., & Master, A. (2006). Reducing the racial achievement gap: A social-psychological intervention. *Science*, 313(5791), 1307–1310.
- Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009). Recursive processes in self-affirmation: Intervening to close the minority achievement gap. *Science*, 324(5925), 400–403.
- Institute of Education Sciences. (2014a). *What Works Clearinghouse (WWC) procedures and standards handbook*. Washington, D.C.: Author. Retrieved February 1, 2015 from <http://ies.ed.gov/ncee/wwc/DocumentSum.aspx?sid=19>.
- Institute of Education Sciences. (2014b). *Assessing attrition bias*. Washington, DC: Author. Retrieved February 1, 2015 from <http://ies.ed.gov/ncee/wwc/documentsum.aspx?sid=243>.
- Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, 75(1), 33.
- Murnane, R. J., & Willett, J. B. (2010). *Methods matter: Improving causal inference in educational and social science research*. New York: Oxford University Press.
- National Center for Education Statistics. (2012). *Digest of Education Statistics 2011* (NCES 2012-001). Washington, D.C.: U.S. Department of Education.
- _____. (2013). *Digest of Education Statistics 2012* (NCES 2014-015). Washington, D.C.: U.S. Department of Education.
- Nussbaum, A. D., & Dweck, C. S. (2008). Defensiveness versus remediation: Self-theories and modes of self-esteem maintenance. *Personality and Social Psychology Bulletin*, 34(5), 599-612.
- Obama, B. (2015). *State of the union address 2015*. Washington, D.C.: The White House, Office of the Press Secretary.

- Rosenbaum, J. E. (2001). *Beyond college for all: Career paths for the forgotten half*. New York: Russell Sage Foundation.
- Walton, G. M., & Cohen, G. L. (2007). A question of belonging: Race, social fit, and achievement. *Journal of Personality and Social Psychology*, 92(1), 82–96.
- Walton, G. M., & Cohen, G. L. (2011). A brief social-belonging intervention improves academic and health outcomes among minority students. *Science*, 331(6023), 1447–1451.
- Wilson, T. D. (2006). The power of social psychological interventions. *Science*, 313(5791), 1251–1252.
- Wilson, T. D., & Linville, P. W. (1982). Improving the academic performance of college freshmen: Attribution therapy revisited. *Journal of Personality and Social Psychology*, 42(2), 367–376.
- Wilson, T. D., & Linville, P. W. (1985). Improving the performance of college freshmen with attributional techniques. *Journal of Personality and Social Psychology*, 49(1), 287–293.
- Wilson, T. D., Damiani, M., & Shelton, N. (2002). Improving the academic performance of college students with brief attributional interventions. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 88–108). San Diego, CA: Academic Press.
- Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, 81(2), 267–301.

CHAPTER 3: TEACHERS' SOCIAL NETWORKS, COLLEGE ATTAINMENT, AND THE DIFFUSION OF A SCHOOL-BASED REFORM INITIATIVE

Introduction

Given schools' diverse social interactions, clustered groups of students and teachers, and embedded networks, using social network theory to understand how students and teachers share resources in regard to college-going behaviors can provide insight into the complex transition period for students from high school to college. Previous network research in schools has often focused on student peer groups (Frank, 1998; Moody & White, 2003), based either on one-mode data such as friendship networks or two-mode data such as course taking patterns. Frank, Zhao, and Borman (2004) extended school network analysis to include teachers. Their work focused on the impact of teachers' social networks on knowledge diffusion and innovation in teaching practices. This study builds on previous research that examines the relationship between teacher networks and the implementation of school reform. However, unlike previous research, which has focused primarily on teachers' classroom practices (i.e. curriculum or technology reforms), this study examines how teacher networks impact levels of knowledge, school norms, and practices related to college-going and subsequently influence the school-wide college-going culture.

The transition from high school to postsecondary school is complex, and while students from middle- and high- socioeconomic families often have access to resources (such as private counselors and preparation programs for entrance exams) to help support them in attaining their educational and occupational goals, many students from low-income families, who have similarly high ambitions, often lack access to resources that help to align their educational plans with future career goals (Schneider & Stevenson,

2000). Given the limited exposure to college-educated individuals that students from low-SES families often have, teachers, particularly those in lower-SES communities, may play an important role in providing resources and support in the college-going process (Schneider, 2000; Farmer-Hinton, 2008). This study will look at the complex relationship between teachers' own knowledge and practices and those of their peers in the context of a school-wide reform intervention, the College Ambition Program (CAP) (Schneider, Broda, Judy, & Burkander, 2013). Specifically, this study analyzes teachers' collegial networks, their levels of interaction with CAP program staff, and the impact of both on the flow of knowledge and practices that promote college-going.

The College Ambition Program (CAP) is a whole-school intervention that provides resources and support to students, staff, parents, and community stakeholders. The CAP model was developed based on empirical research on transitions to adulthood (Schneider, 2007; Csikszentmihalyi & Schneider, 2000; Schneider & Stevenson, 1999) which found that an adolescent's decision to enroll in a postsecondary institution is associated with being able to *visualize* oneself as a college student, transform interests into *realistic actions*, and create *strategic plans*. This intervention places a CAP Center in each treatment school, located within the building and accessible to all students. The center is staffed by graduate students in education who work with students and school personnel to increase awareness about the college-going process. The program is comprised of four components and designed to supplement existing school resources in order to enhance the college-going culture of the school. Briefly, the four major components of CAP are: (1) course counseling, (2) financial aid advising, (3) mentoring and tutoring, and (4) college visits. The section below reviews additional relevant

literature on developing college-going culture in high schools, and then advances a theory and hypothesis about how teachers' collegial networks might impact this development.

Relevant Literature

College-going interventions and school-wide college culture

In recent years, national awareness has grown that a college degree is a prerequisite for securing stable long-term employment (U.S. Census Bureau, 2002). However, not all schools are able to ensure that students have access to the support and preparation they need for college while in high school. While white college-going rates have steadily increased since the 1970s, members of historically underrepresented groups have not seen a similar increase. Despite the presence of college preparation programs throughout the country for several decades, these trends have not changed significantly (Gandara & Bial, 2001; Perna & Titus, 2005). These programs typically offer some combination of mentoring, tutoring, course counseling, and financial advising, but they vary in implementation from programs that target specific minority groups to school-wide interventions that are available to the entire student body. Evaluations of many of these programs are mixed. While most college outreach programs show modest effects, school-wide interventions seem to be more effective than those that target individual students (Domina, 2009). In a review of research on four widely used college outreach programs (Gear Up, Talent Search, Upward Bound, and Quantum Opportunities Program), Domina (2009) suggests that larger, school-wide programs have more potential for “spillover effect” (p. 127), which could mean that unmotivated students (the most difficult group to impact) are more likely to experience changes in belief or actions related to college. In contrast, most targeted interventions identify students who are

motivated to attend college but may lack sufficient knowledge or resources, effectively leaving a portion of students out of the treatment.

School-wide interventions have the potential to be more effective than targeted programs as they can leverage school-wide college-going culture in supporting students to manage the college search, application, and matriculation process. Particularly for students whose parents did not attend college, the school resources dedicated to college guidance are critical supports. Often these students rely heavily on school networks, where administrators, counselors and teachers with college experience can supplement the students' own family or local networks with limited college experience (Farmer-Hinton, 2008). School size and location, course offerings and tracks, and racial and socioeconomic status of students together impact the college-going culture of the school, which is reinforced by counselors and teachers (Holland & Farmer-Hinton, 2009). Other factors include district expectations and priorities regarding college-going, relationships that counselors have with postsecondary institutions, and the prevalence of counselors dedicated solely to college preparation rather than scheduling, discipline, or other administrative tasks (McDonough, 1997, 2005). Several recent studies suggest that the presence of a strong college-going culture at the high school is an important factor for first-generation students, as well as minority students and students from low-income backgrounds (O'Connor, 2000; Farmer-Hinton, 2008; Plank & Jordan, 2001).

While most research on promoting college access for low-income youth is based on work in urban settings, rural communities can also face similar issues. A trend in research on rural students is that their educational and occupational ambitions are often lower compared with students in non-rural settings (Cobb, McIntire & Pratt, 1989;

Chenoweth & Galliher, 2004). Students with lower career ambitions often demonstrate lower expectations of educational attainment (Cobb, McIntire, & Pratt, 1989). Students in rural schools also face several other factors that may also contribute to the development and attainment of lower educational goals including poverty; lack of family and professional role models; lack of self-confidence; valuation of work over education; and lack of support/encouragement from influential individuals (Cobb, McIntire, & Pratt, 1989; Chenoweth & Galliher, 2004). Further, the occurrence of lower ambitions may be magnified during difficult economic conditions, especially in rural communities when top students leave for college, resulting in a perceived brain drain (Sherman & Sage, 2011).

The literature is also consistent that counselors are not the only ones providing college guidance to students and encouraging (or discouraging) their ambitions. This may be due to the notably high ratio of students to guidance counselors in schools today. While the American School Counselor Association's suggested ratio is 250 students to one counselor, the national average in 2008-2009 was 457:1 and Michigan's average was considerably higher at 638:1 (American School Counselor Association, 2005). Ideally, counselors can devote time to meet with all students individually and provide advice on course selection, college applications, scholarships, and financial aid. However, as a counselor's caseload increases, the feasibility of significant one-on-one contact decreases. McDonough (1997) suggests that this case-overload, coupled with extra administrative responsibilities such as discipline and scheduling, mean that few counselors get to spend the kind of one-on-one time with students that they really deserve. Consequently, students often turn to teachers for this support. Students who successfully transition to college often credit one or two helpful teachers or school

personnel who guided them through the process (Freeman, 1997; Levine & Nidiffer, 1996; Farmer-Hinton, 2008). However, little is known about how teachers get the knowledge needed to support these students and how resources to support students are shared among colleagues within a school.

The role of teacher networks in school-wide reforms

Because teachers play a critical role in implementing school-level reforms, particularly those that target school culture, one must consider how to measure the influence of teachers on a school-wide reform such as CAP and understand the different mechanisms that might shape it. Previous research has identified multiple factors that contribute to how teachers respond to reforms, including organizational structure and dissemination of information (Bryk & Schneider, 2002; Daly, Moolenaar, Bolivar, & Burke, 2010), professional development (Kearns et al., 2010; Wilson, 1990), and ideological alignment to reform initiatives (Kennedy, 2005; Olsen & Sexton, 2009). While acknowledging influence of all these factors, this study focuses on knowledge, norms, and behaviors (specially college-going knowledge, behaviors, and the alignment of teachers to school norms), given that recent research suggests that teachers adapt reforms through experimentation and by drawing on what is learned from colleagues (Frank, Zhao, Penuel, Ellefson, & Porter, 2011).

Research Questions

Since a primary goal of the CAP model is to support schools in developing a strong college-going culture, it follows that teachers and their social context would play a critical role in culture change. This assumption drives the following research questions below:

- 1) What are the underlying collegial social structure(s) within CAP high schools?
- 2) Given these social structures, how do teachers' colleagues influence their own beliefs and practices related to college-going?
- 3) How do teachers' interactions with CAP staff influence their own beliefs and practices related to college-going?

Theory and Hypothesis

CAP is designed to be a collective intervention, based on the diffusion of knowledge and practices. Using CAP as a hub for resources about college, the program aims to distribute knowledge and resources that impact the college-going mindsets of both students and teachers. It has been observed at the CAP treatment schools that within a faculty, teachers possess varying attitudes and levels of information related to the college search, application and selection processes. It is hypothesized that teachers with similar mindsets will cluster together (homophily), creating pockets in the building which are more or less supportive of the aims of our project and the college-going culture at each school. In addition, teachers' interactions with CAP staff should drive growth in college-going mindsets and practices. Therefore, teachers who interact more often with CAP staff should see growth in these areas, even after controlling for their time-1 beliefs and practices.

Theoretical Framework

Drawing on the work of Cobb et al. (1989) and Frank, Zhao, & Borman (2004) this study argues that knowledge in schools is transmitted in networks of teachers, often in the form of knowledge flow from teachers with high expertise in an area to those with less expertise. This suggests that an influence model (Frank, 1998) may be the best way

to understand collegial networks in CAP schools. Given that teachers will likely have varying levels of knowledge about college-going strategies and differing mindsets about their students' ability to succeed in college, how will their interaction with colleagues and CAP program staff influence these beliefs and mindsets over time? Also, do different types of interactions (e.g. informal conversations about college with colleagues, vs. co-planning a lesson that incorporates college-going themes, vs. wearing your alma mater's t-shirt on Friday) seem to generate more growth in teacher knowledge and change in mindsets?

According to this theory, teacher behavior in a school with a strong college-going culture is different than behavior in a traditional urban or rural high school. Teachers hold the expectation that all students have the ability to attain a postsecondary credential and they clearly communicate this expectation in their behaviors with the students, administrators, and parents. Teachers discuss their own collegiate experiences and help to stimulate students' thinking, as early as ninth grade, about their college aspirations and interests. Teachers are familiar with entrance exam and scholarship deadlines, the process for applying for Federal student aid, and college entrance requirements. Teachers in these schools do not reserve college counseling to the guidance counselor, but see themselves as a critical component of the students' support system during this critical transition.

While teachers are not the direct outcome of interest of the CAP intervention, teachers have a role in how the intervention evolves in each of the treatment sites. How teachers interact with their closest colleagues, their administrators, and with the CAP intervention may influence how the intended enhancement of the college-going culture for each treatment school evolves. A shared focus and vision among teachers created

through a strong foundation of trust also contributes to the effectiveness of school-wide reforms (Bryk & Schnieder, 2002). The design of the CAP intervention builds on these principles that link the complex nature of reform implementation in a school-wide setting and the critical role that teachers play in proposed changes to their school culture.

Methods

Sample

CAP was designed to be a whole-school intervention that influences not only students but also teachers, staff and families. For the 2011-2012 school year (the year this study was conducted) the CAP intervention was implemented in four public secondary schools (two urban and two rural) in mid-Michigan. Table 3.1 provides additional descriptive statistics on the four schools included in the study. The two urban schools can be classified as “midsize” serving between 800 and 1200 students, are located in the same district, and serve a racially diverse population of students (around 40 percent white, 35 percent black, 20 percent Hispanic, 5 percent other races). One of the urban schools has a slightly higher percentage of black students and fewer white and Hispanic students (50 percent black, 25 percent white, 10 percent Hispanic). The rural treatment schools serve between 400 and 600 students, nearly all of whom are white. The urban schools serve a large percentage of economically disadvantaged students, with 80 percent of their students eligible for free and reduced lunch. At the rural schools, around 20 percent of the students are eligible for free and reduced lunch. Additionally, all treatment schools have a substantial number of students who would be the first in their family to go to college.

Development of Survey Instruments

Data used in this study come from teachers in the four schools participating in the CAP intervention in 2011-2012 (henceforth “CAP schools”). The survey instruments were developed using survey items from the Survey of Chicago Public Schools—High School Teacher Edition and the High School Longitudinal Study (2009) teacher questionnaire with a few items specifically developed to measure each teacher’s interaction with the CAP program (e.g. student referrals, using CAP mentors, participating in professional development). Items used in the teacher questionnaire measure three areas related to the school’s college-going culture: (1) teacher’s knowledge about the process, (2) teacher behaviors (e.g. writing letters of recommendation), and (3) perceptions of schoolwide norms. In addition, teachers were asked to list their five closest colleagues in the building, and indicate the frequency with which they interact. A sample of the “wave one” survey is available in Appendix 3.C.

Data Collection

Data was collected from teachers and instructional staff in two waves. Wave one, a one-page paper survey, collected prior knowledge, behaviors, and norms from all of the teachers in CAP schools in the fall of 2011 and winter of 2012. Wave two, an online instrument using SurveyMonkey, collected a second measurement of teacher knowledge, behaviors, and norms in addition to background information and teacher network data. Teacher names were coded with a teacher-specific identifier and then removed from the survey responses to protect the identity of each participant. The response rate for the first wave was 75% of the total teaching faculty across all schools. For the second wave, the total response rate was 70% of all teachers.

Measures

The outcomes in this study fall into two general categories, teacher *practices* and *beliefs* related to college-going. Within each group, three separate outcomes were used, for a total of six different outcomes tested. Teacher practices were measured using three separate approaches: the first outcome, PSUM_TIME2, is a composite measure representing the sum of teachers' reports on six survey items asking how often they engage in a variety of teaching practices related to college-going (see Appendix 3.C for a copy of the teacher survey). All practice items were developed on a 1 to 5 scale: 1 "Never", 2 "A few times a year", 3 "Monthly", 4 "Weekly", and 5 "Daily." A second outcome, PMEAN_TIME2, represents the mean of teachers' responses to the same six practice items. The third practice outcome, T2_RECLETTERS, represents teacher reports of how many letters of recommendation they wrote over the course of the current school year. Responses here were also placed on a 1-5 scale: 1 "0 letters", 2 "1-3 letters", 3 "4-6 letters", 4 "7-9 letters", and 5 "10 or more letters". Each of these three outcomes was collected at both time-1 and time-2. Time-1 responses were included as prior measures in final regression models.

In addition to the three teacher-practice outcomes above, this study also examined three outcomes related to teacher beliefs about college-going. The first belief outcome, T2_CAPACITY, represents teachers' response to the item "My students have the capacity to do college preparatory work." Responses were on a scale from 1-4: 1 "Strongly Disagree", 2 "Disagree", 3 "Agree", and 4 "Strongly Agree". The second outcome, T2_MYJOB, is constructed using teachers' responses to the following question: "I feel that it is part of my job to help prepare students to succeed in college." Responses were on the same 1-4 scale above. The final outcome, FAMILMEAN_TIME2,

represents teachers' mean levels of reported familiarity with five key components of the college-going process: 1) the Free Application for Federal Student Aid (FAFSA); 2) Deadlines for college applications; 3) Dates for college admissions exams (e.g. ACT and SAT); 4) Online resources for the college search process; and 5) Resources for scholarship and grant opportunities. For each area above, teachers ranked their familiarity on a scale of 1-5, with 1 representing "Not Familiar" and 5 representing "Very Familiar." As with the practice outcomes above, each belief outcome was collected at both time-1 and time-2. Time-1 responses were included as prior covariates in the final model.

Focal Covariates

This study is particularly interested in how teachers' colleagues influence their college-going beliefs and practices. As such, a key covariate is each teacher's exposure to the beliefs and practices of their colleagues at time-1. Exposure terms were calculated for each of the six outcomes listed above by calculating the mean values of all teachers that are listed as important colleagues by teacher i . For example, if teacher i listed teachers j , k , l , m , and o as colleagues, teacher i 's estimated network exposure is calculated as the mean of responses for teachers j through o at time-1. In addition to network exposure, this study as tests the impact of teachers' interactions with CAP program staff on their beliefs and practices at time-2. Here, interactions with CAP staff are operationalized by the survey item "How often do you interact with CAP center staff?" Responses were on a scale from 1-4: 1 "Never", 2 "Rarely", 3 "Sometimes", and 4 "Frequently".

Other covariates

Along with the focal covariates described above, a series of additional covariates were included as statistical controls. These include dichotomous predictors of gender (*FEMALE*), race/ethnicity (*NONWHITE*), and school assignment (*MIDDLEBROOK*, *SCHERZER* and *KELLY*). Several additional dichotomous predictors were included to account for teachers' experience (*YEAR10LESS*, *YEAR20MORE*), subject area assignment (*SCIMATH*, *NONINSTRUCT*). A final covariate, *KNOWT1*, represents the teachers' ability to accurately rank a series of student credentials as criteria for the college admissions process. These criteria include extracurricular involvement, senior year grades, number of college preparatory classes, application essay, and ACT or SAT score. Teachers who ranked these in the conventional order (the one recommended by CAP curriculum) received a "1", other teachers received a "0". This measure serves as a broad control for teachers' alignment with CAP-centric college-going knowledge at time-1. Table 3.2 presents means and standard errors for all variables used in the analysis.

Within-School Cluster Analysis

To better understand the dynamics of teachers' social structures in the four high schools, the wave two survey asked teachers to identify their five closest colleagues in the building, and the frequency with which they interacted with them (a Likert scale from 1 "monthly" to 4 "several times per day"). This data was then used to construct a sociogram for each school, which illustrates each teacher (represented by dots) and their respective collegial ties (represented by lines from nominator to nominee). Along with a visual representation, a cluster analysis was performed using *Kliquefinder* (Frank 1995, 1996) that tested potential teacher subgroups against a null hypothesis that the observed

groups could have formed at random. Using a Monte Carlo simulation, a distribution and test statistic (θ) are generated to evaluate this null hypothesis.

Modeling teacher network effects

Cluster analysis provides one lens on the social structure of the schools in this sample by providing a picture of the overall social structure, including subgroups of teachers, and also by indicating the likelihood that a school has a robust social structure with distinct subgroups of teachers. Next, a series of influence models (Frank, 1998) were estimated to examine the impact of teachers' collegial networks and their interaction with CAP program staff on their college-going-related beliefs and practices. The influence model takes the form of an ordinary least squares (OLS) regression, with the outcome of teacher beliefs and practices at time-2 predicted by a series of covariates, including network exposure and time-1 prior measures. The formal model (Equation 1) and explanation of components are below, specified in terms of teacher i :

$$\begin{aligned}
 Y_i = & \beta_0 + \beta_1(TIME1PRIOR_i) + \beta_2(YEAR10LESS_i) + \beta_3(YEAR20MORE_i) \\
 & + \beta_4(NONINSTRUCT_i) + \beta_5(FEMALE_i) + \beta_6(NONWHITE_i) \\
 & + \beta_7(SCIMATH_i) + \beta_8(KNOWT1_i) + \beta_9(MIDDLEBROOK_i) \\
 & + \beta_{10}(SCHERZER_i) + \beta_{11}(KELLY_i) + \beta_{12}(MISS_T1PRIOR_i) \\
 & + \beta_{13}(MISS_EXPOSURE_i) + \beta_{14}(EXPOSURE_i) + \beta_{15}(T2_CAPSTAFF_i) \\
 & + \varepsilon_i
 \end{aligned}
 \tag{1}$$

Where:

Y_i = teacher i 's time-2 measure of outcome (practices or beliefs)

$TIME1PRIOR_i$ = teacher i 's time-1 measure of outcome

$YEAR10LESS_i = 1$ if teacher i reports ten or fewer years of teaching experience
 $YEAR20MORE_i = 1$ if teacher i reports twenty or more years of teaching experience
 $NONINSTRUCT_i = 1$ if teacher i works in a noninstructional teaching role (guidance counselor, content coach)
 $FEMALE_i = 1$ if teacher i identifies as female
 $NONWHITE_i = 1$ if teacher i identifies as nonwhite
 $SCIMATH_i = 1$ if teacher i reports teaching math or science
 $KNOWT1_i = 1$ if teacher i correctly identified priority of application materials
 $MIDDLEBROOK_i = 1$ if teacher i works in Middlebrook HS
 $SCHERZER_i = 1$ if teacher i works in Scherzer HS
 $KELLY_i = 1$ if teacher i works in Kelly HS
 $MISS_T1PRIOR_i = 1$ if teacher i was missing time-1 measure of outcome
 $MISS_EXPOSURE_i = 1$ if teacher i was missing time-1 measure of network exposure
 $EXPOSURE_i =$ mean of time-1 outcomes for teacher i 's reported colleagues
 $T2_CAPSTAFF_i =$ teacher i 's reported level of interaction with CAP program staff
 $\varepsilon_i =$ person-specific error term for teacher i , independent and identically normally distributed with common variance, σ_ε^2 .

For each of the six outcomes described above in measures, regressions were run in two stages, first with time-1 prior measures and other covariates, and second with the addition of the focal covariates, network exposure and interaction with CAP staff. This two-step approach allows one to examine the relative impact of the focal covariates on the overall model by examining changes in adjusted R^2 values, which represent the overall proportion of model variance explained by the included predictors. Relevant

tables for each model are included in Appendix 3.B (see Tables 3.3 to 3.8). The results section below refers to the models collectively, as summarized by outcome in Table 3.9.

Missing Data

Following previous precedent in missing data techniques relevant to network analysis (J. Cohen, Cohen, West, & Aiken, 2013; Penuel, Frank, Sun, Kim, & Singleton, 2013), and given the already limited sample size in this study, teachers with missing data at time-1 were accounted for with a series of dummy variables for missing data, one for missing time-1 outcomes, and one for missing time-1 exposure. Percent missing ranged from 27% to 31%, depending on the outcome being measured. Table 3.2 on the previous page presents proportions of missingness for relevant time-1 and time-2 measures under the section “Missing Data Predictors.”

Results

Results are presented in three stages. First, the results of cluster analyses at each school are described, including a sociogram to represent the collegial structure of the teaching staff in each building. Next, the results of the six final regression models are presented and summarized, with three outcomes related to teacher practices and three outcomes related to teacher beliefs. Finally, I use Frank, Maroulis, Duong, & Kelcey’s (2013) framework for measuring the robustness of regression estimates to evaluate the strength of the association I find between teachers’ interactions with CAP staff and their college-related instructional practices.

Within-School Cluster Analysis and Sociograms

Figures 3.1-3.4 are sociograms for each of the four high schools in the study sample. As the diagrams illustrate, each school is characterized by a unique social

clustering structure. Monte Carlo simulations produced strong evidence that the observed subgroups in all schools were not formed by random chance (Frank, 1995). In other words, all four schools showed evidence of clustering among teachers.

Figure 1, Middlebrook High School, shows evidence of four distinct teacher clusters, with high levels of within-group ties and somewhat lower levels of between-group ties. Figure 2, Drew High School, has five teacher clusters, with one cluster isolated from the other four. This means that teachers in this cluster only nominate each other, and not other faculty. Scherzer High School also appears to have five distinct clusters, but a different arrangement, with one cluster (the red dots) of teachers both giving and receiving high levels of nominations from the other subgroups. Finally, Kelly High School has nine distinct teacher clusters, the most by far of any school in the sample. Two subgroups, denoted by green and gray dots in the sociogram, appear highly central to the social structure, and both give and receive the majority of nominations.

Influence Model for Teacher Practices and Beliefs

Moving from cluster analysis to influence modeling, I estimate the impact of teachers' colleagues and CAP staff interactions on their beliefs and practices related to college-going. Table 3.9 presents results from influence models that estimated the impact of teachers' colleagues on their own practices and beliefs. Models 1, 2, and 3 (see columns 1-3 in Table 3.9) examined three teacher practice outcomes. Teachers' exposure to their colleagues' practices at time-1 was positively related to their practices (sum) and number of letters written at time-2, however these estimates were not statistically significant. However, for the same outcomes, interactions with CAP staff were positive and statistically significant ($p < 0.10$). In this case, an increase of one unit in level of

interaction with CAP staff was associated with a 0.863 increase in time-2 practices, and a 0.263 increase in letters written at time-2, after controlling for all other covariates, including time-1 prior measures.

Models 4, 5, and 6 (see columns 4-6 in Table 3.9) examined teacher belief outcomes. Here, while the coefficients for network exposure and CAP staff interaction are both positive in all three cases, none of the estimates are significantly different from zero. However, two other variables appear to be significant. Teachers with ten or fewer years of teaching experience appeared to answer about 0.20 units ($p < 0.10$) more positively on the question “I feel that it is part of my job to help prepare students to succeed in college” at time-2 after controlling for all other covariates. Also, science and math teachers report feeling about 0.40 units ($p < 0.05$) less familiar with college-going support and resources, again after controlling for covariates.

Robustness Check- What Would it take to change the inference?

Building off of the finding that teachers’ interactions with CAP staff appear to have a positive and significant effect on their college-related instructional practices, several methods from Frank (2014) and Frank, Maroulis, Duong, and Kelcey (2013) were used to examine the robustness of this association. The primary threat to significance is the possibility that a confounding variable left out of the model may reduce the effect and render the finding insignificant. As Frank (2000) demonstrates, in order for a confounding variable to significantly change an outcome, it must have a strong association with both the outcome and the predictor of interest. If these conditions are satisfied, a confounder may in fact undo a previously significant association.

While every possible confounder can never be truly known, it is possible to estimate how strong an association a hypothetical confounder must have with the outcome and the predictor of interest in order to invalidate the finding (Frank, 2000). There are several ways to conceptualize this association. One approach is sample replacement- in other words, how much of the overall sample of teachers would need to be replaced with teachers for whom there is no CAP staff effect to invalidate the inference? In this case (assuming a confidence threshold of 0.10, the level of significance of the finding) about 7% of the cases would have to be replaced with cases in which there is an effect of zero (Frank, Maroulis, Duong, & Kelcey, 2013). Another option is to frame robustness in terms of correlation. Here, the question is how strong the correlation must be between a confounder and the outcome and the confounder and the predictor of interest to invalidate the inference. Framed in this way, and again assuming a 0.10 confidence threshold, to invalidate the inference, an omitted variable would have to be correlated at 0.125 with $t2_capstaff$ and 0.125 with the outcome, conditional on covariates (Frank, 2000). These robustness checks suggest that even at the $\alpha = .10$ threshold these findings may be vulnerable to confounders with small to moderate correlations. However, it is important to note that the CAP staff findings come from models that already employed many covariates, including a particularly strong prior measure of the outcome that serves as a powerful control. This suggests that the findings, while small, have proven quite durable.

Discussion and Conclusion

This study offers limited, but supportive evidence for previous research findings (Frank, Zhao, & Borman 2004; Frank et al. 2008) which held that teachers' peers can

significantly influence their instructional practices and expertise in a particular area, such as the use instructional technology. However, it also extends previous studies by analyzing the influence of teacher networks in a new arena, developing school-wide beliefs and practices that encourage college-going. Here, it seems that the impact of CAP program staff appears to be slightly more beneficial than the influence of one's collegial network. While this study finds a significant relationship between two measures of teacher practices and their level of interaction with CAP staff, it finds no such connection between CAP staff or collegial interaction and a change in teachers' beliefs. It would seem that in this context, practices and behaviors are easier to influence than beliefs, which makes sense, given that beliefs seem to be more deeply held ideas that would take much longer to change.

To estimate the robustness of these findings, Frank's (2000) correlation-based approach was used to examine how strong a potential confounding variable would have to be to invalidate the inference about teacher practices and CAP staff interactions. The goal here was to offer additional perspective on what it might take to change the inference (Frank, Maroulis, Duong, & Kelcey, 2013) by suggesting just how strong a confounding variable would have to be to alter the overall significance. This shifts the conversation about validity of the findings from identifying any single possible confounder (after all, any study might have many) to thinking about whether or not a confounder truly has the impact to alter the inference. Like all inferences, this one is subject to interpretation and debate, but as Frank et al. (2013) suggest, this type of dialogue is necessary to maintaining a pragmatic focus in education policy research.

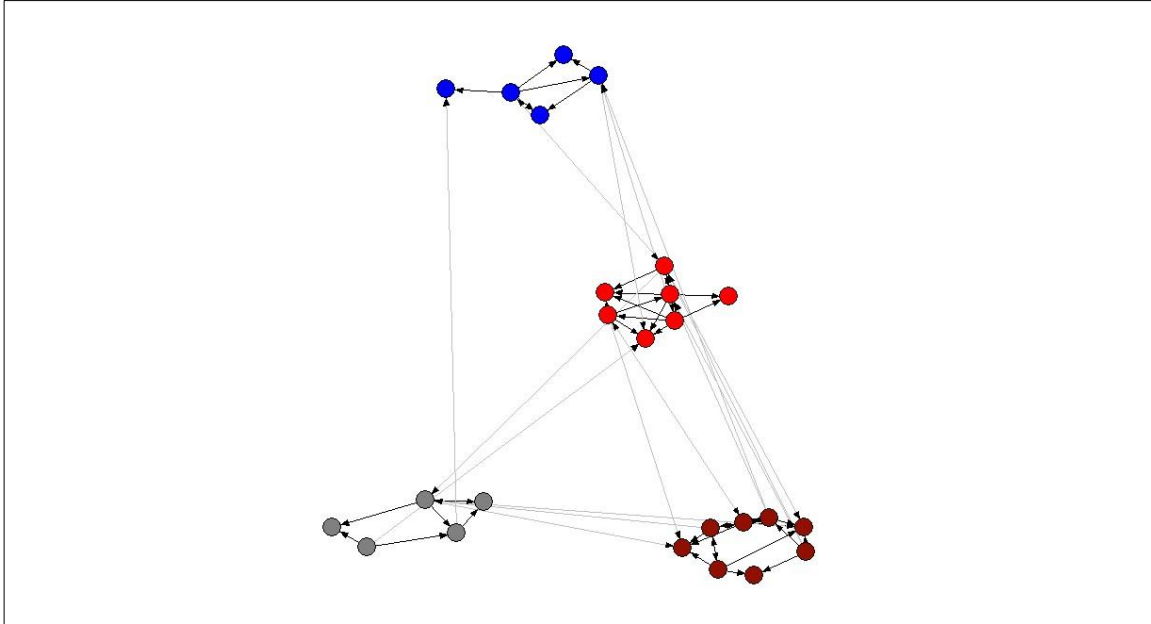
Limitations

Finally, this study has several limitations that should be addressed. First, because the subjects of interest (teachers) are a part of already-formed social cliques and subgroups, the sample can not be seen as random; in fact, it is intentionally not so. To help alleviate the risk of selection bias and/or other latent variables influencing our outcome, two waves of data were collected, which provided a prior measure to control for when estimating the teacher network effect on beliefs and practices. Second, this study makes use of teacher self-reports of both beliefs and practices. Beliefs, at least in this case, seem on one hand to be fairly straightforward (teachers know about financial aid resources, or they do not), but also encompass broader views and perspectives on building-wide culture, which could much harder to understand systematically. Finally, even after controlling for pre-test measures at time one it is important to note that the findings should not be interpreted as causal, but rather associational. Teachers' interactions with CAP staff may not be causally linked to their own practices, but they certainly appear to be a modest, yet significant positive factor.

APPENDICES

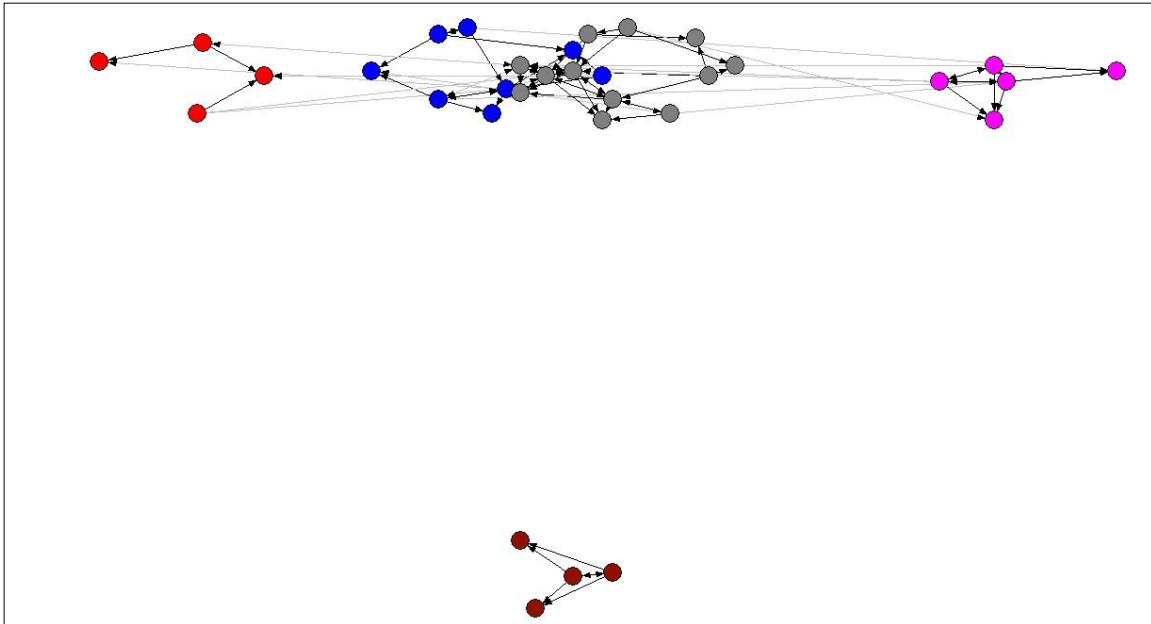
Appendix A: Figures

Figure 3.1 Sociogram of Collegial Ties at Middlebrook HS



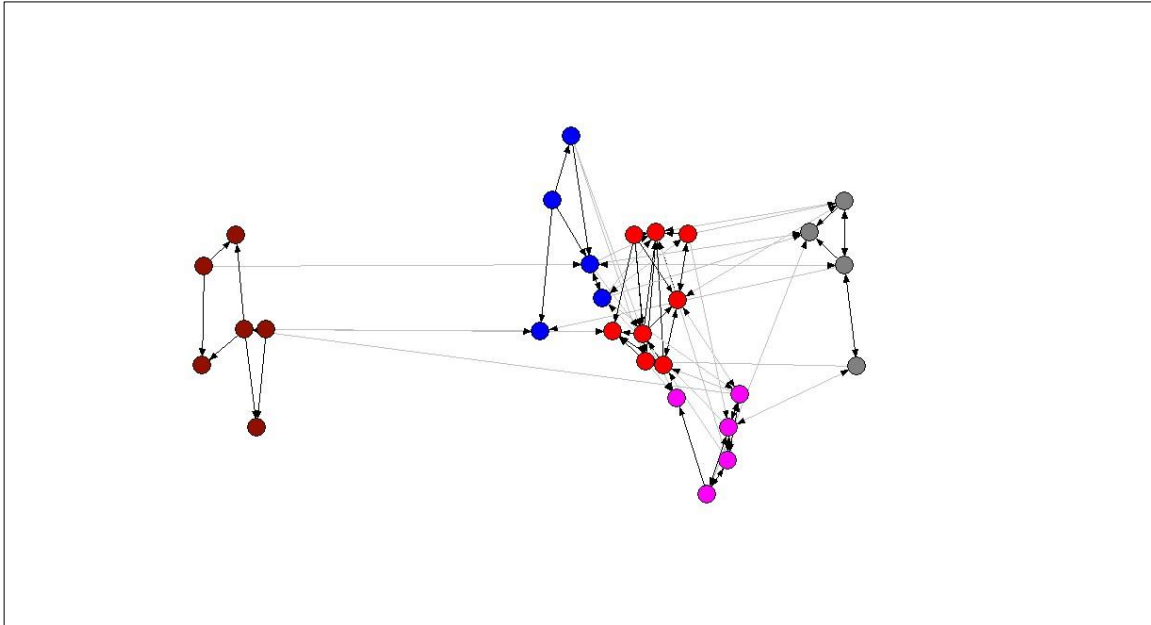
Note: Ties are indicated by an arrow and solid line from nominator to nominee. Different colors represent statistically significant within-school clusters, as determined by Frank's (1995, 1996) clustering algorithm in Kliqefinder.

Figure 3.2 Sociogram of Collegial Ties at Drew HS



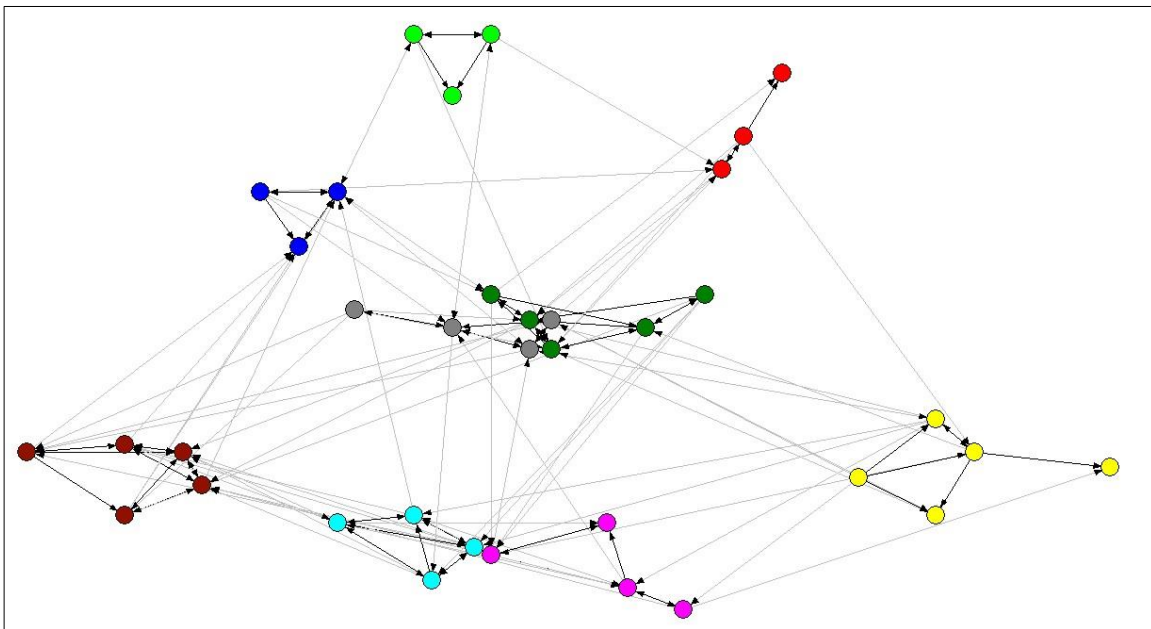
Note: Ties are indicated by an arrow and solid line from nominator to nominee. Different colors represent statistically significant within-school clusters, as determined by Frank's (1995, 1996) clustering algorithm in Kliqefinder.

Figure 3.3 Sociogram of Collegial Ties at Scherzer HS



Note: Ties are indicated by an arrow and solid line from nominator to nominee. Different colors represent statistically significant within-school clusters, as determined by Frank's (1995, 1996) clustering algorithm in Kliqfinder.

Figure 3.4 Sociogram of Collegial Ties at Kelly HS.



Note: Ties are indicated by an arrow and solid line from nominator to nominee. Different colors represent statistically significant within-school clusters, as determined by Frank's (1995, 1996) clustering algorithm in Kliqfinder.

Appendix B: Tables

Table 3.1 Relevant School Characteristics for Study Sample

	Urban schools		Rural schools	
	<u>Middlebrook</u>	<u>Drew</u>	<u>Scherzer</u>	<u>Kelly</u>
Total Enrollment	800	1,200	400	550
FRPL	56.0%	70.0%	20.0%	20.0%
ELL	18.0%	25.0%	<1%	<1%
Minority	84.0%	75.0%	<1%	3.0%

Note. FRPL = free or reduced price lunch;
ELL = English language learner.

Table 3.2 Descriptive Statistics for the Analytical Sample

Variable Name	N	Mean	SE	Minimum	Maximum
<u>Outcomes</u>					
psum_time2	104	14.54	0.533	0	28
pmean_time2	99	2.555	0.075	1	4.667
t2_reclatters	99	2.919	0.137	1	5
t2_capacity	96	2.833	0.074	1	4
t2_myjob	99	3.404	0.052	2	4
familmean_time2	99	3.021	0.115	1	5
<u>Focal Covariates</u>					
exposure_sum	104	5.47	0.433	0	16
exposure_pmean	104	1.065	0.076	0	2.667
exposure_letters	104	1.371	0.115	0	4
exposure_capacity	104	1.674	0.126	0	3.5
exposure_job	104	2.025	0.147	0	4
exposure_famil	104	1.895	0.146	0	5
t2_capstaff	99	2.465	0.092	1	4
<u>Time-1 Prior Measures</u>					
psum_time1	104	5.923	0.541	0	24
pmean_time1	101	1.017	0.091	0	4
t1_2letters	104	1.25	0.136	0	4
t1capacity	101	1.931	0.138	0	4
t1job	104	2.308	0.162	0	4
familmean_time1	101	2.2	0.171	0	5
<u>Other Covariates</u>					
year10less	101	0.426	0.049	0	1
year20more	101	0.198	0.040	0	1
noninstruct	101	0.277	0.045	0	1
female	98	0.653	0.048	0	1
nonwhite	101	0.0891	0.028	0	1
scimath	101	0.257	0.043	0	1
knowt1	101	0.208	0.040	0	1
middlebrook	104	0.154	0.035	0	1
scherzer	104	0.212	0.040	0	1
kelly	104	0.327	0.046	0	1
<u>Missing Data Predictors</u>					
misst1sum	104	0.296	.045	0	1
miss_exposure_sum	104	0.264	.043	0	1
misst1mean	104	0.288	0.044	0	1
miss_exposure_pmean	104	0.288	0.044	0	1
miss_letterst1	104	0.327	0.046	0	1
miss_exposure_letters	104	0.317	0.046	0	1
misst1capacity	104	0.298	0.045	0	1

Table 3.2 (cont'd.)

miss_exposure_capacity	104	0.288	0.044	0	1
misst1job	104	0.317	0.046	0	1
miss_exposure_job	104	0.288	0.044	0	1
missfamilt1	104	0.298	0.045	0	1
miss_exposure_famil	104	0.288	0.044	0	1

Table 3.3 Regression Estimates for the Influence of Colleagues and CAP Staff on Teachers' Practices (Sum)

VARIABLES	(1) Pretest plus covariates	(2) Colleague and CAP exposure
psum_time1	0.662*** (0.0973)	0.625*** (0.0997)
year10less	0.689 (0.821)	0.185 (0.861)
year20more	0.109 (1.097)	0.0470 (1.096)
noninstruct	1.835** (0.897)	1.258 (0.970)
female	0.521 (0.803)	0.821 (0.825)
nonwhite	-0.149 (1.597)	0.153 (1.645)
scimath	-0.243 (0.906)	-0.394 (0.911)
knowt1	1.331 (1.024)	1.290 (1.021)
middlebrook	-2.164* (1.130)	-1.851 (1.149)
scherzer	-0.110 (1.055)	-0.409 (1.108)
kelly	0.210 (0.962)	0.173 (0.967)
misst1sum	7.065*** (1.218)	6.917*** (1.255)
miss_exposure_sum		0.0841 (1.409)
exposure_sum		0.0489 (0.131)
t2_capstaff		0.863* (0.483)
Constant	8.053*** (1.550)	6.167*** (2.070)
N	98	98
R ²	0.446	0.469
Adjusted R ²	0.368	0.372

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.4 Regression Estimates for the Influence of Colleagues and CAP Staff on Teachers' Practices (Mean)

VARIABLES	(1) Pretest plus covariates	(2) Colleague and CAP exposure
pmean_time1	0.676*** (0.0972)	0.644*** (0.101)
year10less	0.125 (0.137)	0.0517 (0.144)
year20more	0.0314 (0.183)	0.0221 (0.184)
noninstruct	0.286* (0.149)	0.196 (0.162)
female	0.113 (0.134)	0.150 (0.137)
nonwhite	-0.0679 (0.266)	-0.00100 (0.272)
scimath	-0.0275 (0.151)	-0.0473 (0.152)
knowt1	0.266 (0.171)	0.262 (0.174)
middlebrook	-0.298 (0.188)	-0.258 (0.193)
scherzer	-0.0216 (0.176)	-0.0851 (0.187)
kelly	0.0459 (0.160)	0.0303 (0.168)
misst1mean	1.195*** (0.203)	1.176*** (0.218)
miss_exposure_pmean		-0.0578 (0.337)
exposure_pmean		-0.0190 (0.195)
t2_capstaff		0.133 (0.0805)
Constant	1.291*** (0.258)	1.088** (0.424)
N	98	98
R ²	0.449	0.468
Adjusted R ²	0.372	0.370

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.5 Regression Estimates for the Influence of Colleagues and CAP Staff on Number of Recommendation Letters Written

VARIABLES	(1) Pretest plus covariates	(2) Colleague and CAP exposure
t1_2letters	0.796*** (0.104)	0.736*** (0.106)
year10less	0.135 (0.241)	-0.0189 (0.251)
year20more	-0.0654 (0.326)	-0.105 (0.321)
noninstruct	0.131 (0.266)	-0.120 (0.285)
female	-0.0458 (0.236)	0.00299 (0.237)
nonwhite	-0.197 (0.473)	-0.0750 (0.470)
scimath	0.118 (0.267)	0.110 (0.264)
knowt1	0.378 (0.307)	0.331 (0.305)
middlebrook	-1.409*** (0.336)	-1.407*** (0.350)
scherzer	-0.217 (0.310)	-0.247 (0.314)
kelly	-0.378 (0.283)	-0.454 (0.281)
miss_letterst1	1.571*** (0.319)	1.367*** (0.336)
miss_exposure_letters		0.696 (0.451)
exposure_letters		0.234 (0.176)
t2_capstaff		0.266* (0.139)
Constant	1.660*** (0.407)	0.757 (0.552)
N	98	98
R ²	0.486	0.520
Adjusted R ²	0.413	0.432

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.6 Regression Estimates for the Influence of Colleagues and CAP Staff on Teachers' Perception of Students' Capacity for College-Level Work

VARIABLES	(1) Pretest plus covariates	(2) Colleague and CAP exposure
t1capacity	0.365** (0.155)	0.375** (0.156)
year10less	0.134 (0.165)	0.130 (0.176)
year20more	-0.273 (0.222)	-0.298 (0.224)
noninstruct	0.169 (0.187)	0.167 (0.201)
female	-0.224 (0.163)	-0.206 (0.168)
nonwhite	-0.237 (0.348)	-0.253 (0.353)
scimath	-0.0521 (0.190)	-0.108 (0.194)
knowt1	0.0574 (0.214)	0.0887 (0.218)
middlebrook	0.347 (0.234)	0.328 (0.238)
scherzer	0.211 (0.216)	0.159 (0.222)
kelly	0.123 (0.199)	0.0653 (0.213)
misst1capacity	0.853* (0.445)	0.976** (0.458)
miss_exposure_capacity		0.0431 (0.335)
exposure_capacity		0.111 (0.119)
t2_capstaff		0.0141 (0.0955)
Constant	1.832*** (0.444)	1.578*** (0.558)
N	95	95
R ²	0.171	0.195
Adjusted R ²	0.0501	0.0420

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.7 Regression Estimates for the Influence of Colleagues and CAP Staff on Teachers' Alignment with College-Going Priorities

VARIABLES	(1) Pretest plus covariates	(2) Colleague and CAP exposure
t1job	0.199* (0.113)	0.185 (0.118)
year10less	0.241** (0.118)	0.210* (0.125)
year20more	0.115 (0.158)	0.0811 (0.159)
female	0.0147 (0.118)	0.0521 (0.121)
nonwhite	-0.196 (0.230)	-0.172 (0.232)
noninstruct	0.147 (0.130)	0.0666 (0.141)
scimath	-0.0285 (0.131)	-0.0495 (0.132)
knowt1	0.0458 (0.149)	0.0857 (0.152)
middlebrook	-0.0762 (0.162)	-0.0458 (0.165)
scherzer	-0.0452 (0.152)	-0.0876 (0.154)
kelly	-0.105 (0.139)	-0.154 (0.144)
misst1job	0.656 (0.395)	0.626 (0.402)
miss_exposure_job		0.287 (0.262)
exposure_job		0.0994 (0.0763)
t2_capstaff		0.0789 (0.0698)
Constant	2.630*** (0.387)	2.229*** (0.465)
N	98	98
R ²	0.141	0.175
Adjusted R ²	0.0199	0.0237

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.8 Regression Estimates for the Influence of Colleagues and CAP Staff on Teachers' Familiarity with College-Going Resources

VARIABLES	(1) Pretest plus covariates	(2) Colleague and CAP exposure
familmean_time1	0.782*** (0.107)	0.764*** (0.110)
year10less	0.0533 (0.195)	-0.00320 (0.205)
year20more	0.398 (0.251)	0.386 (0.254)
female	-0.143 (0.183)	-0.109 (0.188)
nonwhite	-0.311 (0.365)	-0.274 (0.380)
noninstruct	-0.0430 (0.213)	-0.105 (0.229)
scimath	-0.418** (0.208)	-0.434** (0.213)
knowt1	0.304 (0.234)	0.302 (0.239)
middlebrook	-0.442* (0.258)	-0.400 (0.266)
scherzer	0.129 (0.239)	0.0876 (0.245)
kelly	-0.257 (0.221)	-0.259 (0.228)
missfamilt1	2.365*** (0.388)	2.328*** (0.411)
miss_exposure_famil		0.0383 (0.339)
exposure_famil		0.0299 (0.0979)
t2_capstaff		0.109 (0.111)
Constant	0.780* (0.428)	0.524 (0.543)
N	98	98
R ²	0.571	0.577
Adjusted R ²	0.510	0.499

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.9 Summary of All Regression Estimates, by Outcome

VARIABLES	(1) Sum of Practices	(2) Mean of Practices	(3) Letters Written	(4) Student Capacity	(5) Job Role	(6) CG Familiarity
Time1 pretest	0.625*** (0.0997)	0.644*** (0.101)	0.736*** (0.106)	0.375** (0.156)	0.185 (0.118)	0.764*** (0.110)
year10less	0.185 (0.861)	0.0517 (0.144)	-0.0189 (0.251)	0.130 (0.176)	0.210* (0.125)	-0.00320 (0.205)
year20more	0.0470 (1.096)	0.0221 (0.184)	-0.105 (0.321)	-0.298 (0.224)	0.0811 (0.159)	0.386 (0.254)
noninstruct	1.258 (0.970)	0.196 (0.162)	-0.120 (0.285)	0.167 (0.201)	0.0521 (0.121)	-0.109 (0.188)
female	0.821 (0.825)	0.150 (0.137)	0.00299 (0.237)	-0.206 (0.168)	-0.172 (0.232)	-0.274 (0.380)
nonwhite	0.153 (1.645)	-0.00100 (0.272)	-0.0750 (0.470)	-0.253 (0.353)	0.0666 (0.141)	-0.105 (0.229)
scimath	-0.394 (0.911)	-0.0473 (0.152)	0.110 (0.264)	-0.108 (0.194)	-0.0495 (0.132)	-0.434** (0.213)
knowt1	1.290 (1.021)	0.262 (0.174)	0.331 (0.305)	0.0887 (0.218)	0.0857 (0.152)	0.302 (0.239)
middlebrook	-1.851 (1.149)	-0.258 (0.193)	-1.407*** (0.350)	0.328 (0.238)	-0.0458 (0.165)	-0.400 (0.266)
scherzer	-0.409 (1.108)	-0.0851 (0.187)	-0.247 (0.314)	0.159 (0.222)	-0.0876 (0.154)	0.0876 (0.245)

Table 3.9 (cont'd.)

kelly	0.173 (0.967)	0.0303 (0.168)	-0.454 (0.281)	0.0653 (0.213)	-0.154 (0.144)	-0.259 (0.228)
misst1sum	6.917*** (1.255)	1.176*** (0.218)	1.367*** (0.336)	0.976** (0.458)	0.626 (0.402)	2.328*** (0.411)
miss_exposure	0.0841 (1.409)	-0.0578 (0.337)	0.696 (0.451)	0.0431 (0.335)	0.287 (0.262)	0.0383 (0.339)
exposure	0.0489 (0.131)	-0.0190 (0.195)	0.234 (0.176)	0.111 (0.119)	0.0994 (0.0763)	0.0299 (0.0979)
t2_capstaff	0.863* (0.483)	0.133 (0.0805)	0.266* (0.139)	0.0141 (0.0955)	0.0789 (0.0698)	0.109 (0.111)
Constant	6.167*** (2.070)	1.088** (0.424)	0.757 (0.552)	1.578*** (0.558)	2.229*** (0.465)	0.524 (0.543)
N	98	98	98	95	98	98
R^2	0.469	0.468	0.520	0.195	0.175	0.577
Adjusted R^2	0.372	0.370	0.432	0.0420	0.0237	0.499

Notes. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix C: Instrument

The College Ambition Program Teacher Survey—Fall 2011

1. In this school, how often do you?

<i>Check one box on each line</i>	Never	A few times a year	Monthly	Weekly	Daily
Talk about your experiences in college with students					
Talk about college admissions requirements and application deadlines with students					
Talk with students about scholarship and financial aid opportunities					
Help students with their college application essays or personal statements					
Talk to students about what classes they should take to prepare for certain colleges					
Integrate information about college into regular classroom lessons					

2. How many letters of recommendation did you write last year (2010-2011) for students applying to college? (please circle one)

0	1-3	4-6	7-9	10 or more
---	-----	-----	-----	------------

3. To what extent do you agree with the following statements?

<i>Check one box on each line</i>	Strongly Disagree	Disagree	Agree	Strongly Agree
Teachers expect most students in this school to go to college				
Teachers at this school help students plan for college outside of class time				
The curriculum at this school is focused on helping students get ready for college				
All students at this school have access to courses that prepare them for college				
My students have the capacity to do college preparatory work				
I feel that it is part of my job to help prepare students to succeed in college				

4. What do you consider to be the most important criteria for the college admissions process?

Please rank by placing numbers 1-5, where 1 is the most important and 5 is the least important
Involvement in extracurricular activities
Senior year grades
Number of college preparatory classes
Application essay
SAT or ACT test score

5. How familiar are you with the following parts of the college-going process?

<i>Circle one number on each line.</i>	Not Familiar Very Familiar				
Free Application for Federal Student Aid (FAFSA)	1	2	3	4	5
Deadlines for college applications (including early decision)	1	2	3	4	5
Dates for college admissions exams (e.g. ACT and SAT)	1	2	3	4	5
Online resources for the college search process	5	1	2	3	4
Resources for scholarship and grant opportunities	1	2	3	4	5

REFERENCES

REFERENCES

- Bryk, A. & Schneider, B. (2002). *Trust In Schools*. New York: Russell Sage Foundation.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. New York: Routledge.
- Csikszentmihalyi, M. & Schneider, B. (2000). *Becoming adult: How teenagers prepare for the world of work*. New York: Basic Books.
- Daly, A., Moolenaar, N., Bolivar, J., & Burke, P. (2010). Relationships in reform: the role of teachers' social networks. *Journal of Educational Administration*, 48(3), 359-391.
- Domina, T. (2009). What works in college outreach: Assessing targeted and schoolwide interventions for disadvantaged students. *Educational Evaluation and Policy Analysis*, 31(2), 127-152.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motivation* (pp. 75–146). San Francisco, CA: W. H. Freeman.
- Holland, N.E. & Farmer-Hinton, R.L. (2009). Leave no schools behind: The importance of a college culture in urban public high schools. *The High School Journal*, 92(3), 24-43.
- Farmer-Hinton, R.L. (2008). Social capital and college planning: Students of color using school networks for support and guidance. *Education and Urban Society*, 41(1), 127-157.
- Frank, K.A. (1995). Identifying cohesive subgroups. *Social Networks*, 17(1), 27-56.
- Frank, K.A. (1996). Mapping interactions within and between cohesive subgroups. *Social Networks*, 18(2), 93-119.
- Frank, K. A. (2000). Impact of a confounding variable on a regression coefficient. *Sociological Methods & Research*, 29(2), 147-194.
- Frank, K.A. (2014). *KonFound-it!*. Available at <http://www.msu.edu/~kenfrank/research.htm#causal>
- Frank, K. A., & Min, K. (2007). Indices of robustness for sample representation. *Sociological Methodology*, 37, 349-392.

- Frank, K. A., Zhao, Y., & Borman, K. (2004). Social capital and the diffusion of innovations within organizations: Application to the implementation of computer technology in schools. *Sociology of Education*, 77(2), 148–171.
- Frank, K. A., Maroulis, S. J., Duong, M. Q., & Kelcey, B. M. (2013). What would it take to change an inference? Using Rubin’s Causal Model to interpret the robustness of causal inferences. *Educational Evaluation and Policy Analysis*, 35(4), 437-460.
- Frank, K. A., Zhao, Y., Penuel, W. R., Ellefson, N., Porter, S. (2011). Focus, fiddle and friends: Knowledge to perform the complex task of teaching. *Sociology of Education*, 84(2), 137-156.
- Frank, K. A., Muller, C., Schiller, K.S., Riegle- Crumb, C., Mueller, A.S., Crosnoe, R., & Pearson, J. (2008). The social dynamics of mathematics coursetaking in high school. *American Journal of Sociology*, 113(6), 1645–1696.
- Frank, K.A., Sykes, G., Anagnostopoulos, D., Cannata, M., Chard, L., Krause, A. & McCrory, R. (2008). Extended influence: National Board Certified Teachers as help providers. *Education, Evaluation, and Policy Analysis*, 30(1), 3-30.
- Freeman, K. (1997). Increasing African-Americans’ participation in higher education: African-American high-school students’ perspectives. *The Journal of Higher Education*, 68(5), 523-550.
- Gandara, P. & Bial, D. (2001). *Paving the way to postsecondary education: K-12 intervention programs for underrepresented youth* (NCES 2001-205). Washington, DC: U.S. Department of Education, National Center for Education Statistics.
- Kearns, D., Fuchs, D., McMaster, K., Saenz, L., Fuchs, L., Yen, L., Meyers, C. Stein, M., Compton, D., Berends, M., & Smith, T. (2010). Factors contributing to teachers’ sustained use of kindergarten peer-assisted learning strategies. *Journal of Educational Effectiveness*, 3(4), 315-342.
- Kennedy, M. (2005). *Inside Teaching*. Cambridge: Harvard University Press.
- Levine, A., & Nidiffer, J. (1996). *Beating the odds: How the poor get to college*. San Francisco: Jossey-Bass.
- McDonough, P.M. (1997). *Choosing colleges: How social class and schools structure opportunity*. Albany, NY: SUNY Press.
- McDonough, P.M. (2005). *Counseling and college counseling in America’s high schools*. Alexandria, VA: National Association for College Admission Counseling.

- McGrath, D.J., Swisher, R.R., Elder, G.H., & Conger, R.D. (2001). Breaking new ground: Diverse routes to college in rural America. *Rural Sociology*, 66(2), 244-267.
- Moody, J., & White, D.R. (2003). Structural cohesion and embeddedness: A hierarchical conception of social groups. *American Sociological Review*, 68(1), 103-127.
- Olsen, B., & Sexton, D. (2009). Threat rigidity, school reform, and how teachers view their work inside current education policy contexts. *American Educational Research Journal*, 46(1), 9-44.
- Penuel, W. R., Frank, K. A., Sun, M., Kim, C., & Singleton, C. (2013). The organization as a filter of institutional diffusion. *Teachers College Record*, 115(1), 306-339.
- Perna, L. W., & Titus, M. A. (2005). The relationship between parental involvement as social capital and college enrollment: An examination of racial/ethnic group differences. *The Journal of Higher Education*, 76(5), 485-518.
- Plank, S.B., & Jordan, W.J. (2001). Effects of information, guidance, and actions on postsecondary destinations: A study of talent loss. *American Educational Research Journal*, 38(4), 947-980.
- Schneider, B. (2007). *Forming a college-going community in U.S. public high schools*. Seattle, WA: Bill & Melinda Gates Foundation.
- Schneider, B., & Stevenson, D. (1999). *The ambitious generation: America's teenagers motivated but directionless*. New Haven, CT: Yale University Press.
- Schneider, B., Broda, M., Judy, J., & Burkander, K. (2013). Pathways to college and STEM careers: Enhancing the high school experience. *New Directions for Youth Development*, 2013(140): 9-29.
- Shadish, W.R., Clark, M.H., & Steiner, P.M. (2008). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association*, 103(484), 1334-1344.
- U.S. Census Bureau. (2000). The big payoff: Educational attainment and synthetic estimates of work-life earnings. *Current Population Reports, Special Studies*.
- Wilson, S. (1990). A conflict of interests: The case of mark black. *Educational Evaluation and Policy Analysis*, 12(3), 293-310.