REPRESENTATIVENESS TWO WAYS: AN ASSESSMENT OF REPRESENTATIVENESS AND MISSING DATA MECHANISMS IN A STUDY OF AN AT-RISK POPULATION

by

ANNELIESE CHRISTINE BOLLAND

JAMES E. MCLEAN, CHAIR

JOHN DANTZLER
JUDY GIESEN
STEVE THOMA
SARA TOMEK
JOE WEBER

A DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in the
Department of Educational Studies in
Psychology, Research Methodology,
and Counseling in the Graduate
School of The University
of Alabama

TUSCALOOSA, ALABAMA

2012

ABSTRACT

Vulnerable populations like at-risk adolescents are often difficult to study, yet the data they provide are invaluable to researchers in a number of fields, including, but not limited to health and education. Because these populations are difficult to study, some argue that any studies of these populations (or samples from these populations) have an inherent selection bias, suggesting that the results may not be generalizable to the populations studied. These arguments are made stronger when random sampling techniques are not used to identify a sample of at-risk adolescents, for example.

This study examines the representativeness of a sample of at-risk adolescents in a community-based longitudinal study (the Mobile Youth Survey, or MYS) of poverty and adolescent risk in Mobile, Alabama. Further, this study examines the missing data patterns that exist in 10 waves of data collected in this study to determine which missing data mechanisms exist in this dataset, in order to conclude whether these missing data are ignorable. With over 20,000 data points, and items measuring developmental, behavioral, and psychosocial constructs, the MYS can be used by researchers in several fields, but only to the extent that the data are of high quality, that is, representative of the population in terms of demographic (grade level, gender, race, free lunch eligibility status, neighborhood type) and functional (cognitive and behavioral) characteristics.

Results show that while there are concerns about the demographic representativeness of the MYS sample to the population, overall, these results are not alarming, and in fact, are somewhat expected. Further, these differences suggest that perhaps the population should be re-

defined. Overall, these results demonstrate that (a) there is not an inherent sampling bias in studies of vulnerable populations; (b) in the MYS, while demographic characteristics may not always be representative of the defined population, there are no consistent differences between the sample and population with respect to functional characteristics once demographic factors have been statistically controlled; and (c) missing data can be studied as it relates to representativeness.

LIST OF ABBREVIATIONS AND SYMBOLS

GEE ..... Generalized Estimating Equations

GIS ..... Geographic Information Systems

MAR ..... Missing At Random

MCAR ..... Missing Completely At Random

MCPSS ..... Mobile County Public School System

MI ..... Multiple Imputation

MLE ..... Maximum Likelihood Estimation

MNAR ..... Missing Not At Random

MYS ..... Mobile Youth Survey

SI ..... Single Imputation

ACKNOWLEDGMENTS

Several people and organizations were instrumental to the completion of this dissertation. I express gratitude to the chair of my dissertation, Dean Jim McClean. I also express gratitude to Sara Tomek, who not only served on my committee, but guided me through the complex nature of the analyses used in this study. Other members of my committee, John Dantzler, Judy Giesen, Steve Thoma, and Joe Weber were supportive and available and I cannot thank them enough for their service on my committee. The data used in this study would not have been available without the contribution of numerous individuals, including interns, project investigators, participants, lab employees and volunteers, and others associated with the Mobile Youth Survey. Similarly, without the funding from NICHD (HD 30060, HD 058857), CSAT/SAMHSA (TI 13340), NIDA (DA 017428, DA 018920), and the CDC (CE 000191), data would not have been collected, and I would not have had funding to work on this study.

Finally, I could not have completed this dissertation without the support and encouragement from my friends and family. Specifically, Kim Bissell has not only been encouraging, but inspiring to me professionally and personally. Rebecca Ballard has not only been encouraging and inspiring, but also very kind to look through countless pages and tables to edit my formatting errors. All of my co-workers and friends at the University of Alabama Rec Center are acknowledged as well for their encouragement, support, and the occasional much-needed distraction from SAS. Last, but certainly not least, my parents are my inspiration, and their patience, support, encouragement, edits, guidance, and love are not unnoticed and are directly related to the completion of this study.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Statement of the Problem and Significance of the Project

Nonrepresentativeness is a potential threat to virtually all research and especially that involving humans as it may undermine the validity of research conclusions. Research conclusions often are based on the idea that results from studying a sample also apply to the population from which the sample was drawn. Inferences from the sample to the population assume that the sample is representative of the population. Nonrepresentativeness, or a sample being different from a population on one or more important characteristics, may occur for a number of reasons. First, a study may produce a nonrepresentative sample if the sample is not randomly selected and therefore, certain types of respondents (e.g., those with more extreme behavior) are undersampled. Second, a study may produce a nonrepresentative sample, at least for certain variables, if respondents selectively choose which questions to answer for nonrandom reasons. Third, a study may produce a nonrepresentative sample even with a randomly selected sample if the response rate varies significantly from unity for nonrandom reasons (e.g., people with more extreme behaviors, opinions, or attitudes are less likely to participate). In this case, nonrepresentativeness equates to data missingness, which is likely to occur in survey research, especially in longitudinal studies, where not only might data be missing at an item level, but also at a case level, when participants miss one or more wave of data collection. Fourth, a study may produce a nonrepresentative sample if some respondents choose not to participate (or are unable

to participate) in longitudinal follow-up for nonrandom reasons (e.g., people with more extreme behaviors, opinions, or attitudes participate and then drop out). Finally, a follow-up sample in a longitudinal study may not be representative of the previous sample(s). While not an exhaustive list, these five examples illustrate how nonrepresentativeness may occur when the sample deviates from randomness, either due to sampling strategy or due to data missingness. But, it is critical to realize that deviations from randomness do not necessarily create nonrepresentativeness. Thus, the challenge for researchers is to determine whether deviations from randomness in their studies (and there are almost always deviations from randomness in survey or field studies) results in nonrepresentativeness. In this study, missing data mechanisms are assessed through an examination of representativeness for three reasons: to allow researchers who use this data to use it more confidently, to empirically examine the relationship between nonrandomness and nonrepresentativeness, and to show methodologically how representativeness can be assessed in similar studies. Specifically, this study is important because (a) it tests whether non-randomness necessarily leads to bias; (b) it develops a protocol for using auxiliary data to test the representativeness of a sample; and (c) it identifies the level of non-representativeness that exists in the Mobile Youth Survey (MYS).

It is important to assess deviations from randomness, because to the extent that deviations from randomness bias results, the generalizabilty of the findings is limited and remedial action should be taken. To the extent that deviations from randomness do not result in nonrepresentativeness, they can be effectively ignored. In discussing these issues, the terms "data missingness" and "missing data" are used to indicate all of the threats to representativeness previously discussed. While this is a departure from the traditional use of these terms (i.e., discussing missing data in terms of three categories: missing at random, missing completely at

random, and missing not at random), it helps to clarify the relationship between the concepts of data missingness and sample representativeness. That is, that when there are any deviations from randomness in sampling strategies or deviations from 100% participation, the sample may not be representative of the population studied and any missing data may be meaningful.

While researchers may often be aware of potential limitations caused by missing data, they may be unsure of the best ways to handle the missing data issues, including whether it is acceptable to ignore the issues. One key issue researchers need to understand about their missing data, so they can decide how to handle the issues, is the extent to which their available data are representative of their populations. When small amounts of data are missing from a random sample, the missing data issues are fairly simple to handle. When samples are not random but still thought to be representative, however, and when there are multiple types of missing data, the problem is compounded. Researchers would be well served to try to determine whether data are missing randomly or non-randomly (McKnight, McKnight, Sidani, & Figueredo, 2007) because the mechanisms that generate missing data affect how data should be analyzed and interpreted.

When the situations surrounding missing data are not considered, resulting parameter estimates can be biased. That is, samples not representative of the population lead to biased parameter estimates; consequently, generalizations are not well supported. In longitudinal studies, representativeness should be approached in two ways: cross-sectionally (i.e., is the sample in the wave being considered representative of the population) and longitudinally (i.e., are all the waves representative of the entire sample), because one does not necessarily guarantee the other.

Numerous texts and articles discuss missing data and techniques for handling missing data (e.g. Allison, 2000; Buhi, Goodson, & Neilands, 2008; Enders, 2001; Graham, 2009; Little,

1988; Little & Rubin, 2002; Molenberghs, & Kenward, 2007; O'Rourke, 2003; Roth & Switzer, 1995; Schafer & Graham, 2002). Some of these provide examples of real data and others use simulated data to show the effects of different solutions. Rarely, however, does a researcher have both a complicated missing data situation and available data from other sources to use to investigate the degree of bias caused by the missing data. Such an analysis provides what much of the literature does not: a readily accessible and potentially more valid diagnostic procedure to assess missing data mechanisms and to guide analysis in the presence of missing data.

In this study, missingness and its surrounding situations are examined in a longitudinal dataset, providing researchers who find themselves with missing data with an example of how to confidently approach a diagnosis of that missing data. Here, missingness is examined in the first 10 years of data from the MYS, a multiple-cohort longitudinal study of poverty and adolescent risk in Mobile, Alabama (Bolland, 2008).

This study examines data from the MYS and data from the Mobile County Public School System (MCPSS) to identify the degree to which the MYS sample is biased. Specifically, two research questions are addressed: (a) To what extent are enrollment and yearly participation in the MYS (1998-2007) representative of the population of adolescents living in MYS neighborhoods, in terms of demographic (grade level, gender, race, free lunch eligibility status, neighborhood type) and functional (cognitive and behavioral) characteristics? and (b) Longitudinally, can the between-wave missing data mechanisms in the MYS data be considered to be missing at random? Without answers to these questions, conclusions reached in a number of published studies using the MYS data (e.g., Bolland, 2003; Bolland, Lian, & Formichella, 2005; Bolland, Bryant, Lian, McCallum, Vazsonyi, & Barth, 2007; Drummund, Bolland, & Harris, 2011; Park, Lee, Bolland, & Vazsonyi, 2008; Park, Lee, Sun, Vazsonyi, & Bolland, 2010;

Spano, Freilich, & Bolland, 2008; Spano, Rivera, & Bolland, 2006, 2010; Spano, Rivera, Vazsonyi, & Bolland, 2008; Stewart & Bolland, 2003; Stoddard, Henley, Sieving, & Bolland, 2010; Vazsonyi, Pickering, & Bolland, 2006) and future studies are less convincing than they might otherwise be.

Researchers began paying closer attention to the effects of missing data in the 1970s (e.g., Dempster, Laird, & Rubin, 1977; Heckman, 1979; Rubin, 1976), but the subject of missingness and how to address missing data solidified as a subfield of statistics in the 1980s with the publication of Little and Rubin's (1987) *Statistical Analysis with Missing Data*, updated with a second edition in 2002. Many studies of missingness in data have been concerned with clinical studies (e.g., Molenberghs & Kenward, 2007), while not as many studies have focused on longitudinal survey data. Further, not many studies have considered both cross sectional and longitudinal representativeness.

Longitudinal datasets consist of data collected at more than one time point, often with the goals of describing the trajectories of behavioral or attitudinal constructs, treatment effects, and how covariates affect trajectories or treatment effects (Diggle & Kenward, 1994). Longitudinal data can be viewed as having the advantages of both cross-sectional and time-series data, that is, not only can researchers examine the relationships within and between variables, but they can also examine patterns over time. Longitudinal designs can be advantageous because they can provide evidence of causation rather than merely a description of relationships (Barry, 2005).

Longitudinal data, while advantageous to researchers, often cannot be collected completely. Researchers may choose to analyze only the available data or to impute values for the missing data. However, imputation becomes complex when reasons for missingness are unknown. Thus, the existence of missing data forces researchers decide how to manage any

missing data to result in the least amount of bias. Often, with large longitudinal datasets, no one method can be used to (a) determine how missing data are missing, or (b) handle all missing data or instances of dropout efficiently (Hogan, Roy, & Korkontzelou, 2004).

The analyses in this study are useful for multiple reasons. Not only do they provide a general strategy for assessing representativeness in other nonrandom longitudinal datasets with missing data (with independent sources of data for validation), but they also provide useful information for the many researchers who have used and will use the MYS for a variety of purposes. Like the National Longitudinal Dataset of Adolescent Health (Add Health) ("Add Health," 2012), the MYS is an important dataset that can inform researchers about several issues regarding youth at risk. The dataset is particularly useful for those who can use the data from the MCPSS along with the MYS data. Arguably the representativeness of the sample in any large study intended for use by multiple individuals should be examined, allowing it to be used with more confidence by multiple researchers. Unfortunately, the complete independent dataset necessary for a thorough analysis of representativeness often is not available. Add Health is a national dataset that is presumably representative, but it is virtually impossible to determine the representativeness of this dataset because there is no complete independent dataset that can be used in conjunction with Add Health to independently assess the representativeness of the sample. Such a dataset is available, however, for the MYS. Results of the analyses allow researchers who use the MYS to have confidence in its representativeness. The analyses also provide a model for similar analyses of other datasets when an independent dataset is also available. Further, this study explores the advisability of using such steps to determine representativeness of data in similar studies. That is, this study examines whether using these steps is an effective way of establishing representativeness in a longitudinal dataset and

consequently, whether others with similar datasets can explore representativeness using the same steps.

The MYS, like Add Health, is a longitudinal study; however, what makes the MYS unique is that it focuses on a vulnerable (often described as "hard to reach") population in a single geographic area. While this dissertation study may seem to be quite limited in its applicability, studying the MYS data and assessing its representativeness has implications for how the MYS data is studied in the future, the conclusions that are drawn from these data, and the future studies that are conducted. Further, while many researchers who have conducted studies have not gained access to an auxiliary dataset with which to assess representativeness, this dissertation study highlights the importance of obtaining such a dataset and suggests how an auxiliary dataset can be used to assess representativeness.

Additionally, given that MYS researchers are studying a vulnerable population, missing data may be particularly prevalent, leading to concerns about generalizability of conclusions. Some may argue that when studying vulnerable populations there is an inherent selection bias, resulting in inherently non-representative data. That is, sample selection, sample enrollment, and longitudinal dropout mechanisms indicate or are consistent with informative (and therefore non-ignorable) missingness[1]. Moreover, aspects of what makes the people in the population vulnerable may be related to why some of them cannot be identified, fail to enroll in or drop out of the study, or fail to respond to some of the items in a survey. That missing data from a vulnerable population must be non-ignorable is only an assumption, however. It is possible that even with a vulnerable population, the missing data are missing at random. These analyses

[1] The term *ignorable* does not suggest that the missingness can be ignored, rather the term is used to indicate that the reasons for the missingness can be ignored and that analyses will produce unbiased results. When missingness is termed *non-ignorable*, it is an indication that only certain types of strategies are appropriate for dealing with the missingness and even then, limitations of using the strategies must be acknowledged.

examined a decade of data from the MYS, thus exploring the argument that vulnerable populations inherently yield nonrepresentative samples. Because it is impossible, in most cases, to determine the true mechanisms underlying data missingness, this study provides an alternative means for examining the extent to which the missing data in the MYS are ignorable or non-ignorable. If they are ignorable, support is provided for the validity of community-based studies of vulnerable populations.

The MYS is one of the largest longitudinal community-based surveys of an at-risk population living in poverty, and an examination of this dataset provides others with (a) a reasonable approach to diagnosing the representativeness of a population that is often difficult to recruit (e.g., Hatchett, Holmes, Duran, & Davis, 2000; Kerkorian, Traub, & McKay, 2007; Pottick & Lerman, 1991) and (b) the effects of non-enrollment and dropout for this population. The diagnosis of the representativeness is possible because access to measures such as the MCPSS records allows for the comparison of survey participants with non-participants and non-dropouts with dropouts based on demographic and functional characteristics. If the techniques used to address representativeness and missing data in this sample can be demonstrated to be effective, the implications are great for how researchers can diagnose their missing data in a longitudinal dataset and demonstrate representativeness cross-sectionally and longitudinally. Further, if the sampling strategies used in the MYS resulted in a representative sample, future studies can be designed to study an at-risk population with greater confidence in similar sampling strategies.

Overview of the Mobile Youth Survey

Background

Health disparities have been the subject of many descriptive and clinical research studies over the past several decades. Those who live in poverty continue to experience health problems to a greater extent that those living above the poverty line (Aber & Bennett, 1997). The statistics that help quantify this information and the conclusions reached from the data provide researchers with a foundation for defining and studying at-risk populations and designing and assessing interventions. While the census ("U.S. Census," 2012) reports that the poverty rate has declined through the past decades (22% in 1959 to 12.5% in 2007), this drop in poverty rate is misleading. Although there was a notable drop in the poverty rate from 1959 (22%) to 1968 (12.8%), this rate has remained relatively constant over the past 40 years and issues that affect this population continue to exist at non-ignorable levels. Further, concentrated poverty has increased in some census tracts, particularly in inner cities among minorities (Jargowsky, 1997; 2003). While the national poverty rate has declined, the census ("U.S. Census," 2012) reports that there has been an increase in the percentage of people who live in extreme poverty, that is, the percentage of people who live below 50% of the poverty level has increased from 30.2% in 1977 to 41.8% in 2007. The number of the Nation's poorest and the health disparities that occur in the Nation's poorest have not been ignored and steps have been taken to positively affect this group of individuals. For example, *Healthy People 2020* is a set of public health goals set forth by the U.S. Department of Health and Human Services (2011) to reduce the nation's most significant and preventable threats to health. It is built on past initiatives including the 1979 Surgeon General's Report and *Healthy People*. The two overarching goals for *Healthy People 2020* are to (a) increase quality and years of life, and (b) eliminate health disparities. Health disparities

have been shown to have negative consequences for educational outcomes as well (Braveman, Cubbin, Egerter, Williams, & Pamuk, 2010).

Evidence-based interventions are required to meet these goals, which have focused attention on research and evaluation. An approach to this evaluation is identifying high-risk areas and then determining to what extent living in a particular area can affect a child's behavior trends (e.g., engagement in risk behaviors such as substance use). Minority youth living in poverty has proven to be a difficult population to study, resulting in studies with nonrandom samples and high attrition rates, which is problematic when those who participate each year and across years are not representative of the population and when those who do not participate are not participating for reasons that cannot be ignored (Bor, Najman, O'Callaghan, Williams, & Anstey, 2001; Buhi & Goodson, 2006; Teitler, Reichman, & Sprachman, 2003; Way & Robinson, 2003). Even when response rates are high, the possibility of response bias, whether or not overtly detectable is not wiped out (Groves, 2006; Groves & Peytcheva, 2008).

The data collected from the MYS addresses some of the goals of *Healthy People 2020* in that the sample studied is minority youth living in poverty. The goals of the MYS are twofold: (a) to describe the characteristics, circumstances, and behaviors of its participants (disadvantaged urban adolescents, aged 10 - 18); and (b) to describe the etiology of these characteristics, circumstances, and behaviors. With over 8,700 youths enrolled in the MYS and 23,500 surveys conducted through the summer of 2007, the MYS has reached a large segment of the disadvantaged population in Mobile's poorest neighborhoods (Bolland, 2008, estimated this figure at more than 70% in Mobile's largest public housing neighborhood). The large sample size "allows analysis of low-prevalence events (e.g., marijuana use among 10 and 11 year olds) and how those events affect trajectories" (Bolland, 2008, p. 4). This survey is not school-based,

allowing data to be collected from some members of the population who would not be reached in school-based surveys (e.g., high school dropouts, chronically absent students). Finally, the longitudinal nature of this survey has allowed researchers to maintain a presence in the communities where adolescents live, which has likely been a factor in the high cooperation rates in this survey.

Thus, the data collected from this project may "shed valuable light on the developmental, social, and health conditions of these youths, and how those conditions may be improved" (Bolland, 2008, p. 2). The benefits of these data, however, are only as good as the quality of the data. And, one primary threat to the quality of the data consists of their possible non-representativeness and non-random loss to follow-up.

Justification for Analysis

Generally, for data to be useful, the dataset has to be representative of the population so that analyses do not produce biased estimates. Representativeness can be assessed by examining who has participated in a study compared to the population, leading a researcher to identify who is missing from the study. Especially when random sampling techniques are not used to obtain a sample, representativeness should be assessed (although representativeness should also be assessed even when random sampling is used but the response rate deviates significantly from unity).

The sampling strategy used in the MYS was not random. With both active and passive recruitment, no one who was eligible to participate in the survey (i.e., lived in or was connected to the target neighborhoods, was age eligible, and had parental consent) was turned down. This type of sampling strategy, though aggressive and resulting in a large number of participants, may not have yielded a representative sample in each of the years the study was conducted or overall

throughout the study. Further, there are missing data in this dataset (i.e., some eligible adolescents did not participate in the survey at all, some did not participate every year they were eligible to participate, and not all items on the survey were completed by each participant). Neither of these factors means that the MYS sample is not representative. They do, however, suggest that the representativeness of this dataset should be diagnosed.

There may be a situation where researchers have evidence of longitudinal representativeness absent a distinct analysis. If each enrollment year's participation is representative, for example, and the missing data are missing at random (i.e., if the reasons for missing data are ignorable), then by definition, the longitudinal data must be also be representative of the sample. This, however, is not always the case. When assessing longitudinal data, researchers should not be interested only in participants with more than one data point, but they should also be concerned with participants with only one data point, to the extent that they drop out. That is, even when researchers have access to a longitudinal dataset, they should also be concerned with cross-sectional representativeness and should not dismiss those participants with only one data point without an examination of why there is only one data point. The individuals with only one data point may not be crucial in a longitudinal analysis, because no trajectories can be constructed; however, these participants should still be considered in assessing the sample cross-sectionally and to the extent possible, the reasons why these individuals did not continue participation should be explored.

Consider the hypothetical example (see Table 1) of a population consisting of two groups: adolescents at high risk and those classified as at low risk. It may be the case that in Year 1, the sample is completely representative of the population. But, say for example, those participants at high risk only participate in Year 1 and then drop out of the study. They are

replaced by a new group of high risk participants in Year 2. But, the same low risk participants continue to participate in Year 2. When assessing cross-sectional representativeness in Year 2, the sample would still be representative of the population. But, when the dataset is examined longitudinally, the researcher will notice that those in the highest-risk group have only one data point, making longitudinal analysis impossible for those in the high risk category. Therefore, the longitudinal sample is not longitudinally representative of the population because of this dropout. The next year, Year 3, the same pattern occurs. That is, the same adolescents in the low risk category participate, giving them three data points. Those in the high risk category again drop out and a new group of adolescents in the high risk category who have never participated in the survey take the survey. Again, the cross-sectional sample is representative of the population; but the longitudinal sample is not representative of the population: high risk respondents have only one year of data and cannot be included in the longitudinal analysis, while low risk participants have three years of data and can be included in all longitudinal analyses. It is not likely that an entire subpopulation will drop out of a longitudinal study each year, and in this hypothetical example, it would be obvious that the dropout process was not random. However, selected dropout could occur on a less obvious, but no less important level, especially in a dataset where there are multiple subpopulations, each defined by several variables. This example highlights the complexity of representativeness and missing data and provides cause to investigate the representativeness of the MYS dataset on both a cross-sectional and longitudinal level.

Table 1

Hypothetical Example of Survey Participation by Year

Participant	Risk Status	Year 1	Year 2	Year 3
1	[a]Low	[b]O	O	O
2	Low	O	O	O
3	Low	O	O	O
4	[c]High	O	[d]X	X
5	High	O	X	X
6	High	O	X	X
7	High	X	O	X
8	High	X	O	X
9	High	X	O	X
10	High	X	X	O
11	High	X	X	O
12	High	X	X	O

[a]Indicates participant falls into a low-risk category
[b]Indicates observation
[c]Indicates participant falls into a high-risk category
[d]Indicates non-participation

One might question how representativeness can be determined when there are missing values in a dataset. When researchers define a population, they have an idea of the characteristics of that population that are meaningful for their studies. Therefore, on such characteristics, the sample can be compared when resources are available. Assessing representativeness from a cross-sectional approach can most easily be done through the use of proxy measures, records that provide researchers with multiple characteristics of the population. If proxy measures are complete, such as school system records, researchers can compare those who participated in a study and those who did not participate in a study based on relevant characteristics.

When assessing representativeness, researchers often use demographic characteristics because this information is more easily obtained for a population than other variables. In other words, representativeness is often based solely on variables like gender, age, and race. These factors, however, may not account for the reasons people decide not to participate in or to drop

out of research studies. That is, just because participants are similar on race, age, or gender, does not mean they do not differ on non-ignorable factors (e.g., exposure to or perpetration of school violence). Therefore, researchers should examine their research questions and purpose for conducting a study, being mindful of the factors that should be compared for participants and non-participants in order to achieve a representative sample. If those who did not participate in the study are similar to those who did participate in the study on these relevant characteristics (e.g., age, gender, race, cognition, and behavior), then there is increased evidence of a representative sample. If those who did not participate in the study are not similar to those who did participate in terms of these relevant characteristics, then the researcher must examine those deviations more closely to determine how to whether those deviations can be ignored or not. Often, researchers using longitudinal data are only concerned with assessing the longitudinal representativeness of their samples and do not attempt to assess or verify cross-sectional representativeness.

When a sample is not representative, results may be biased and a study's validity may be compromised. If the sample is proven to not be representative on a cross-sectional level, counter measures (e.g., weighting, deletion of graphical areas where representativeness is lacking) can be taken to ensure unbiased analyses. On the other hand, when considering longitudinal representativeness, part of the sample may be representative (i.e., a certain group—age or race, for example—may be representative in the sample), thus, unbiased conclusions may be drawn about a specific group, if the whole sample is determined to be unrepresentative of the population. When longitudinal data are not representative of any group, it is difficult to draw conclusions about change over time. For example, if dropout rates are different for participants with certain characteristics, change can still be measured, but results from analyses may be

biased. To continue this example, if adolescents engaged in the riskiest behaviors drop out of a study, then the researcher cannot draw conclusions about that risky behavior.

Only after patterns of representativeness and dropout have been assessed can the data be analyzed to identify the trends that exist overtime in this at-risk population without bias. The MYS was not conducted with a random sampling technique which may suggest that the resulting sample is not representative of the population; further, researchers might identify more than one type of missingness occurring within this dataset. Missingness and how it affects representativeness on both a cross-sectional and longitudinal level using the MYS dataset with the addition of an auxiliary dataset (i.e., MCPSS records[2]) were examined in this project.

In the past, many studies surrounding missing data have been conducted in clinical settings (e.g., mastisis in dairy cattle, Fluvoxomine trials, and depression trials[3]) where researchers may have better access to their participants and therefore, more ability to control random factors (lack of transportation to assessment) that might result in dropout. Further, in clinical studies, researchers often have more knowledge about participant characteristics and their responses to stimuli (e.g., treatment effects) than they do in community studies. This results in the ability to better model dropout in clinical analyses. In a community field study (e.g., the MYS), the possibility of higher levels of missing data exists, particularly in the case of longitudinal studies where participants are surveyed only once a year.

Many researchers address representativeness only in their sampling strategies (i.e., initial representativeness at enrollment) with no follow-up diagnoses to ensure samples are indeed representative of the population. When sampling procedures are random and there is complete

---

[2] It is important to mention that the recruitment of participants for the MYS was not targeted by enrollment in the MCPSS. Because adolescents are legally required to attend school until the age of 16, most of those who fall in the 10 through 15 year-old age bracket do attend a MCPSS school, but that was not a requirement of participation in the MYS.

[3] Molenberghs & Kenward, 2007

participation, as sample size increases, the probability of non-representativeness approaches zero. However, as the response rate decreases from unity, representativeness can be undermined, regardless of sampling strategy.

Several techniques are used to address missing data problems and in some cases, researchers might have to employ multiple techniques to respond to different missing data mechanisms. Researchers have conducted Monte Carlo studies (e.g., Roth & Switzer, 1995) to determine how missing data techniques respond to randomly and non-randomly removing data; however, it is rare that researchers have been able to diagnostically evaluate a real-world longitudinal community-based dataset to determine what types of missingness are functioning and then follow that diagnostic assessment with the implementation of different techniques with which to handle the missing data.

In fact, missingness may be completely overlooked in a research study, particularly when different mechanisms of missingness exist; such is the case in many longitudinal datasets. Missingness may reflect a dataset that is not representative of its population. This study seeks to address missing data mechanisms and representativeness in a community-based longitudinal sample, including defining these two concepts, diagnosing a longitudinal dataset's representativeness from a cross-sectional and longitudinal perspective, and describing techniques to address non-representativeness and/or missing data that is not missing at random. In Chapter 2, current research on topics surrounding representativeness and missing data are described and discussed to provide a basis and defensible strategy for the MYS data, or missing data, to be analyzed. Then, in Chapter 3, the research questions of the study are defined and methods used to address these questions are described. Results of these analyses are reported in Chapter 4, and finally Chapter 5 is a discussion of those results and the implications for future research.

CHAPTER 2

REVIEW OF THE LITERATURE

An exploration of representativeness using a community-based longitudinal dataset requires a conversation about missing data, including: definitions of missingness and how missingness affects representativeness. This can be most effectively done by discussing how the issues of representativeness (or nonrepresentativeness) and missingness have been addressed in the literature, including discussions of why the previously suggested methods are not the most efficient or effective to use with a longitudinal dataset when a proxy dataset is available. Further, the discussion would be incomplete without addressing those more efficient and effective ways to address the issues of representativeness and missingness in datasets similar to the Mobile Youth Survey (MYS).

The problem addressed in this study is how to determine cross-sectional and longitudinal representativeness in a longitudinal dataset so that appropriate analyses can be conducted with minimal bias. As suggested previously, this is most often discussed in the literature in terms of data missingness, and that is where this review initially focuses. Dependent on the type of missingness (e.g., within-wave missingness, between-wave missingness) and the pattern(s) of missingness, different techniques can be used to adequately manage that missingness. In this chapter, some of the techniques more frequently used to assess and manage missing data are discussed.

Considering which patterns or mechanisms of missingness exist in a dataset is important because of the limitations some of these mechanisms present to researchers. In order to rule out

a mechanism or provide evidence supporting a certain mechanism of missingness, three missing data mechanisms are described. Because so much of the analysis depends on which missing data mechanisms are functioning in a dataset, it is important to define these mechanisms and how they affect the data, but also to describe how researchers can determine which mechanisms of missingness are functioning in their datasets. That is, it is important to describe how to diagnose the mechanism or mechanisms of missing data functioning within a dataset.

Once missing data mechanisms have been identified, researchers can then make decisions about how to manage those mechanisms. Several techniques can be used to manage missing data and when considering which techniques are appropriate to use in a longitudinal community-based survey, it is practical to first rule out techniques that are simplest to apply but often not appropriate: (a) available-data methods and (b) single imputation methods. When these two techniques are dismissed, use of an augmentation method (e.g., multiple imputation, techniques using a maximum likelihood framework, and modeling) is a more complex but viable option. The focus of this study, however, is the diagnosis of cross-sectional representativeness and exploration of missing data mechanisms in a longitudinal dataset, which may allow researchers to avoid the statistically complex and time consuming techniques for replacing missing data that are often described in missing data literature. Still, a discussion of missing data mechanisms and the techniques used to manage missing data is necessary to provide context for this study.

Background: Missingness

Defining Missingness

Missingness is a term that has been defined and classified in different ways, and these classifications are not always mutually exclusive. One classification system separates missing data into what de Leeuw (2001) describes as first-level and second-level missingness. First-level

missingness is described as unit non-response and second-level missingness is described as item-non-response. Wave non-response can be described as a type of unit non-response, and may occur for several reasons (Schafer & Graham, 2002), which are described in the following sections. Of particular interest in this exercise is between-wave missingness, which is unit-missingness over the span of one participant's longitudinal data, and though other classifications are identified, unit-missingness or wave missingness is the primary focus of this study.

Item or second-level missingness may often occur because of the sensitive nature of some questions on surveys (e.g., substance use). In other cases, demographic information could be skipped or be illegible. Often, survey participants do not wish to fill in their income, for example, therefore, if an income variable or variables is present in a survey, there may be many missing measurements. In the case of item missingness, and in the example just described, however, some data most likely still exists for a participant (e.g., work hours and job type variables); therefore, often researchers are able to estimate a value of demographic missingness from other variables.

Another classification system separates between-wave missing data into monotone and non-monotone missingness (see Table 2). Longitudinally, monotone missingness occurs when one wave of missingness is never followed by an observed measurement (Kenward, 1998). Several situations may occur where missingness can be described as monotone, and it is important to note that a description of monotone missingness does not indicate a missing data mechanism of missing at random (MAR) or missing not at random (MNAR). One example of monotone MAR missingness in a study involving risky behavior in adolescents might be the relocation of an individual to a place of his or her own. The counter to that, the MNAR monotone missingness would be the relocation of an individual to jail. Monotone missingness

could also be reflected in an increase in the participants' commitments that prevent them from participation in their final year of eligibility. For example, the participant may have entered college. Another example of monotone missingness is the aging out of an individual from a study, or the last year of data collection. These two cases, while fitting the characteristics of monotone missingness, cannot be considered truly missing data or dropout. Finally, there may be ineligible participants who participate in a study, for example, adolescents who are outside the sampling frame of a survey who are visiting relatives during survey collection times. Data from these "participants" should not be included in further analyses and their missing data should not be considered truly missing.

Non-monotone missingness can be described, in longitudinal datasets, as interior missingness, where a missing wave of measurement is followed by a measurement, and a participant with missing interior observations is somewhat less problematic in terms of dropout analysis than one who has trailing missing data. In some situations, non-monotone missingness can also be described as intermittent missingness (Diggle & Kenward, 1994). Just as is the case with monotone missingness, a description of non-monotone missingness does not indicate a missing data mechanism of MAR or MNAR. Non-monotone missingness can occur in several different patterns. The dropout may be for one or more waves before dropping back in. Further, dropout, if for multiple waves, may not be consecutive. Similarly to monotone missingness, there are several different situations that may account for non-monotone missing data.

First, a missing interior observation may indicate that a participant was simply unavailable for one wave of data, supporting the case that the participant did not drop out permanently from the study. A missing interior observation due to unavailability can occur for several reasons: (a) a relatively small amount of time was spent in each of the target

neighborhoods and the participant was simply unavailable at that time; (b) the participant was punished and not allowed to participate; (c) the participant was sick; (d) the participant was otherwise engaged. Each of these reasons may point to an MAR mechanism. However, this unavailability may be due to reasons suggesting an MNAR mechanism too. For example, if the participant was incarcerated during the time of data collection, an MNAR might be presumed, although this type of unavailability would probably be associated with "abnormal level of behavior (e.g., a high level of illegal behavior that caused the incarceration; a low level of illegal behavior during the incarceration)" (Bolland, 2008, p. 14). Another pattern of non-monotone missingness includes two or more observations with one or more interior missing observations.

Table 2

Monotone and Non-Monotone Missingness Patterns

Participant	Year 1	Year 2	Year 3	Year 4
Complete Case	[a]O	O	O	O
Complete Case	[b]X	X	O	O
Non-Participation	X	X	X	X
Monotone 1	O	X	X	X
Monotone 2	O	O	X	X
Monotone 3	O	O	O	X
Non-monotone 1	O	O	X	O
Non-monotone 2	O	X	X	O
Non-monotone 3	O	X	O	X
Non-monotone 4	O	X	O	O

[a]Indicates observation
[b]Indicates non-participation

Finally, dropout has been classified in terms of how it could impact analyses: completely random, random, and informative (Little & Rubin, 2002; Rubin, 1976). This classification system is further explored in terms of missing data mechanisms. Still other definitions of missingness exist, and while researchers should not ignore missing data in their studies, it is up

to them to define missingness as it applies to their studies. In this exercise, between-wave missingness is primarily considered, both monotonically and non-monotonically.

Reasons for Between-Wave Missingness

Causes for missing data often depend on the type of study (Hogan et al., 2004). Researchers have identified several reasons for attrition in a longitudinal study. First, on a data collection level, surveys, for example, may be lost or responses may be illegible. Buhi, Goodson, and Neilands (2008) mention that a participant may drop out from a survey study because he/she was not able to connect with the researcher giving the survey or he/she may drop out of the study for reasons that are directly related to the instrument. These identified reasons for attrition help researchers decide how to define dropout, which, in turn, helps them manage that dropout and determine how to appropriately analyze a dataset with missing data. Further, in Chapter 3, these reasons and several others are identified as direct reasons for dropout.

Effects of Between-Wave Missingness

Missing data can affect longitudinal analyses leading to poor parameter estimates, standard errors, and test statistics (Lin & Schaeffer, 1995; Patrician, 2002). Holes in datasets can lead to inaccurate or biased results surrounding observations, causal inferences, and any generalizations made beyond the survey (McKnight et al., 2007). Missing data can also result in a reduction of precision (Jackson, White, & Leese, 2010) and the loss of statistical power (Baron, Ravaud, Samson, & Giradeau, 2008; Streiner, 2002). Different techniques can minimize the effects of missing data, however, the usefulness of these techniques often depends on correct classification of the missing data and correct identification of the missing data pattern. As previously mentioned, Rubin (1976) and Little and Rubin (2002) described dropout as being completely random, random, or informative. These classifications can guide researchers'

management of missing data in a way that minimizes the previously mentioned effects of missing data. Accordingly, dropout classifications are now referred to as missing data mechanisms.

Missing Data Mechanisms

Between-wave missingness often occurs in longitudinal datasets and may be in a monotonic or non-monotonic pattern. Further, an informative or non-informative mechanism or reason accounts for that missingness. It is these mechanisms or reasons for missingness that researchers need to determine. Further, diagnosing a missing data mechanism or mechanisms is important for researchers, because these classifications influence what types of techniques can appropriately be used to address any missing data. Buhi et al. (2008) describe these missing data mechanisms in terms of randomness: complete randomness, conditional randomness, and non-randomness or bias. McKnight et al. (2007) explain that missing data should be classified according to the variables with missing data, any associated variables (e.g., covariates), and the reason(s) for which data are missing.

Little and Rubin's (2002) descriptions of completely random dropout, random dropout, and informative dropout have evolved into the terms of data that are missing completely at random (MCAR), MAR, or MNAR. Data that are missing either completely at random or at random are considered to be ignorable; data that are missing not at random are considered non-ignorable[4]. The extent to which researchers can determine whether those who dropped out of a study and those who did not are different and whether those differences are informative or not is the biggest issue when determining a missing data mechanism. In this discussion, following Rubin's (1976) framework, missingness can be specified as a probability function. Further, in

[4]Collins, Shafer, and Kam (2001) write that the term ignorable does not mean, in missing data terms, that the missing data can be ignored. Further, missing at random, does not imply that the missingness can be thought of as purely and truly unrelated to any variables within a dataset or totally unrelated to the measurement process.

this discussion, R is referred to as a binary matrix where R_{ij}=1 if Y_{ij} is observed and R_{ij}=0 if Y_{ij} is missing.

Missing At Random

Data are missing at random (MAR) when the probability that Y is missing does not depend on the value of Y. Thus, the missing data may be associated with the observed data, but not to the respondent's true score of that variable (Schafer & Olsen, 1998; Weisberg, 2005). In other words, these missing data are not related to the unobserved outcomes[5] (Buhi et al., 2008; Enders, 2001; Molenberghs & Kenward, 2007). Mathematically, Rubin's (1976) framework defines MAR data as

$$P(R \mid Y_{\text{com}}) = P(R \mid Y_{\text{obs}})$$

where complete data is denoted as Y_{com}, which is made up of observed (Y_{obs}) and missing (Y_{mis}) data (also see Schafer & Graham, 2002). Graham (2009) cautions that the term random in MAR can be a little troublesome to some researchers, and prefers to refer to MAR data as conditionally MAR, conditional on an independent variable. One key descriptor of data that are MAR is that they can be considered ignorable and a missing data technique can be employed to manage the missingness, without too much concern about bias (Buhi et al.).

Assumptions and Analysis

According to Honaker and King (2007), MAR data assumes ignorability and non-confounding. Either of these assumptions may be violated, and by definition, it is impossible to definitively conclude a MAR mechanism when looking at the data alone. Thus, testing for the MAR data mechanism would require follow-up data from the missing participants or require the inclusion of an unverifiable model into any analysis.

[5] Weisberg shares an example where missing data could be defined as missing at random. He writes that an example of MAR data would occur "if the probability of having missing data on income were related to the person's education and gender but not to the person's actual income" (p. 141).

When researchers assume that missing data is missing at random, it allows them to have full use of the data and analysis is generally straightforward (Jackson et al., 2010). However, there should be an examination of whether outcomes or predictors (or both) are MAR (Piggot, 2001). If a MAR mechanism is operating within a dataset, researchers can use the estimated relationships among observed variables to confidently predict the values of any missing variables without the use of a modeling process (Molenberghs & Kenward, 2007; Schafer & Olsen, 1998).

Researchers have options when faced with a MAR mechanism. For example, in some situations, proxy scores can be used in place of missing variables. For example, Patrician (2002) writes that when an income variable is missing for a particular respondent, if the respondent has given his/her zip code and/or job information, those variables may be able to serve as proxy scores for the income variable. The use of proxy scores with datasets with MAR missing data is accepted because the proxy scores should be reasonable given other observed data, therefore, the missing at random mechanism stands (Jackson et al., 2010).

Missing Completely At Random

A special case of MAR is called missing completely at random (MCAR). An MCAR mechanism occurs when the missingness with respect to one variable cannot be predicted from any of the observed variables for a respondent (Molenberghs & Kenward, 2007; Weisberg, 2005). Mathematically, using Rubin's (1976; see also Schafer & Graham, 2002) framework

$$P(R \mid Y_{\text{com}}) = P(R).$$

In other words, in this mechanism, there are no systematic differences, on one variable, between those who did and did not provide data (McKnight et al., 2007). Here, those who did provide data on that variable can be seen as a random sample of the original sample (Weisberg, 2005). For example, bad weather on a day that a survey was being given might lead to MCAR

missing data. Unlike the other two mechanisms, the mechanism of MCAR missingness can be reliably detected (see Little, 1988). When data are MCAR, analyses do not result in biased estimates (Streiner, 2002); however, whenever data are missing, regardless of the missing data mechanism, there could be a decrease in statistical power (Graham, 2009).

Ignorability

Missing data that satisfy either the MAR or MCAR mechanisms can be described as ignorable. Again, the term ignorable does not indicate that missing data can be ignored. Rather, in the case of ignorable missing data, the reasons for missingness can be ignored and the model can be estimated without ramifications of biased estimates (Buhi et al., 2008; Hogan et al., 2004; Piggot, 2001). When missing data are termed ignorable, the missing values can often be filled in using other measured values from the dataset (Streiner, 2002). Though it is difficult to meet the assumptions of MCAR or MAR, when evidence points to either of these mechanisms, Piggot explains that both maximum likelihood and multiple imputation methods can be used in analysis. Further, these conditions allow certain types of inferences to be made without having to resort to complex modeling (Heitjan & Basu, 1996).

On the other hand, missing data are said to be non-ignorable when, Weisberg (2005) writes, the probability of data being missing is related to any estimated parameters. This situation may lead to a sample that is not representative of the population. If there is no follow-up with survey non-respondents, researchers cannot justifiably classify the missing data as ignorable. Only logic (e.g., diagnostic procedures, sometimes statistically guided), in this situation, can guide the researcher to make a decision whether the missing data might be classified as ignorable (Schafer, 1997).

Missing Not at Random

As previously mentioned, not all missing data mechanisms are ignorable. Missing data that are missing not at random (MNAR) are also referred to as non-ignorable. This type of missing data occurs when there is a non-ignorable reason for the data to be missing, that is, the missingness is not due to chance. In other words, data are considered MNAR when the probability of the data being missing is dependent on the estimated parameter (Nakagawa & Freckleton, 2008; Weisberg, 2005). Often the information that would enable the researcher to determine whether the data are MNAR is unavailable (McKnight et al., 2007).

Buhi et al. (2008) provide a good example of an MNAR mechanism, occurring when the values on a weight variable are missing. In their example, people who weigh more may be reluctant to disclose their weight on a survey, and thus, their failure to report weight is related to their weight and not related to any other measure in the survey, the probability of the missing value would depend on that value only, which is missing.

In any dataset, researchers strive for representativeness and few options are available to guide the researcher to make a decision whether the missing data might be classified as ignorable (see, McKnight et al., 2007; Molenberghs & Kenward, 2007; Schafer, 1997). When there is missing data in a dataset, it is especially difficult to prove unbiased parameters (Graham, 2009, Jamshidian & Mata, 2008). Techniques that may best be used to deal with the MNAR mechanism is described in more detail in following sections, however, these techniques are often employed with caution because of the untestable assumptions that come with the MNAR mechanism (Buhi et al., 2008; Molenberghs & Kenward, 2007). In a case of MNAR missingness, distributions for missingness should be established when thinking about how to analyze the data; and, these distributions can be established using either selection or pattern-

mixture models (Schafer & Graham, 2002). If researchers want to avoid complex modeling, they are well advised to design studies so that as much data as possible can be collected, reducing the probability of missing data (Buhi, et al.; Molenberghs & Kenward, 2007).

Distinguishing Missing Data Mechanisms

MAR vs. MNAR. Though MAR is considered an ignorable mechanism and MNAR is considered a non-ignorable mechanism, thus appearing almost completely opposite, it is very important to first note that these two mechanisms cannot be distinguished in a dataset (Molenberghs & Kenward, 2007). If MCAR mechanisms have been ruled out, researchers can classify the data as being either MAR or MNAR, but cannot definitively assign either of these labels[6] because the missingness mechanism is a function of the unobserved missing value for MNAR. In MAR, the missingness mechanism is not a function of the unobserved missing values. But, because those missing values have not been observed, researchers cannot judge whether the missingness mechanism is a function of the unobserved missing values (McKnight et al., 2007).

There are diagnostic tools that researchers may use to aid in classifying missing data as either ignorable or non-ignorable (see McKnight et al., 2007). Even with these diagnostic tools, when missing data is beyond researchers' control, classification as MAR (rather than MNAR) cannot be certain (Schafer, 1997). The main issue that arises with these two missing data mechanisms, because they are indistinguishable, is that the representativeness of the sample may be compromised with either mechanism.

[6]Molenberghs (2008) describes that when given an MNAR dataset, the researcher can manipulate the conditional distribution of those unobserved outcomes based on the observed ones, conforming to an MAR mechanism, thus resulting in a model that has the same fit as the original MNAR model, and thus providing evidence that distinguishing between MNAR and MAR mechanisms is statistically impossible.

MCAR vs. MAR/MNAR. The MNAR and MAR mechanisms cannot be distinguished from each other, but MCAR can be distinguished from these. That is, if the missing data mechanism is found to not be MCAR, then it must be either MAR or MNAR. In the MNAR mechanism, the value of *R* (0 or 1) depends on the true value of *Y* for that variable, which has not been observed. If the true value of *Y* is correlated with whether *Y* is missing or not, then it is missing not at random. If the value of *R* is unrelated to the true value of *Y*, then it is missing at random. McKnight and colleagues (2007) and Little (1988) proposed statistical methods for helping distinguish an MNAR mechanism from a MAR mechanism. Other methods of determining missing data mechanisms include studying the means and covariances and looking for homogeneity between groups that have the same missing data patterns (see Kim & Bentler, 2002; Little, 1988). Homogeneity provides evidence for an MCAR mechanism, although this example does not stand for every case (Jamishidian & Mata, 2008). This logic suggests that the more similar the groups are with respect to some characteristics, the more similar they are with respect to other characteristics. Consequently, if the groups with missing data are similar to the groups with complete data, there is evidence that the missingness occurs randomly.

However, another suggested method to distinguish MCAR data from MAR/MNAR is to use graphical procedures (see McKnight et al., 2007). Allowed in some statistical packages, Nakagawa and Freckleton (2008), recommend

> replacing missing observations with "outlier" values in a multivariate data set and scatterplotting each bivariate combination in the data set will easily detect whether or not missing data are MCAR (i.e. clusters of the outlier values will be found if not MCAR). (p. 595)

Approach to Analysis: Parameter Estimation and Representativeness

Again, when missingness depends on the missing values themselves, the underlying mechanism of missingness can be described as MNAR, or non-ignorable non-response (Jamshidian & Mata, 2008). Because it is possible to distinguish between missing data that are MCAR and MAR or MNAR, it is critical for researchers to assess their missing data when they begin to think about managing any missing data within a dataset. Researchers want to avoid biased parameters in their conclusions, and according to May and colleagues. (2009), "MCAR results in a correct sampling distribution for Y_{obs} and poses no problem for parameter estimation (except of resulting in reduced sample size)" (p. 79). When MCAR cannot be accepted, it is still possible that the resulting distributions produce unbiased estimates (May et al.), when procedures, for example maximum likelihood estimation (MLE), are employed, though use of MLE has been contested (see Enders, 2001; Jamshidian & Mata).

As discussed, missing data classified as MAR or MNAR are conceptually different because the probability that data are missing is different in these two mechanisms. MAR data are akin to data that are considered MCAR, that is, ignorable. Without information about why the data is missing, however, researchers cannot be confident that missing data are ignorable. Researchers should be cautious when approaching data that is considered MNAR when diagnosing, analyzing, and drawing conclusions about their data.

Missing Data and Validity

When discussing parameter estimation, a discussion of how missing data can affect the validity of a study is warranted. Validity can be defined in many ways; however, one informal definition of validity is the extent to which an instrument measures what it purports to measure. When data are missing non-randomly, it becomes difficult for researchers to conclude whether

their instruments are measuring what they were designed to measure and therefore whether their conclusions are warranted.

Internal Validity

Internal validity, informally, whether the independent variable, as opposed to some other variable(s) caused or influenced the dependent variable, can also be threatened by missing data. One way missing data influences internal validity is that the inferences that are drawn suffer from selection bias. This may happen, for example, because only a select group of individuals participate in a survey, and there may be a difference between the group of individuals who participated and the group of individuals who did not participate. Inferences, then, would be based on only a truncated set of data, and one might conclude that a relationship does not exist when in reality, one does. Attrition or mortality can also threaten internal validity because there may be differences in those who drop out of a longitudinal survey and those who remain in the study (McKnight et al., 2007).

Graham (2009) shows, though, that even though biases in the data may result from attrition and dropout mechanism, these biases may be low enough to be endured. Two types of bias described by Graham are inflation and suppression bias. Inflation bias makes an ineffective program, for example, appear to be effective. This type of bias calls into question the internal validity of a study, that is, the ability of the researcher to interpret any group differences in a dependent variable as being a function of an independent variable , rather than any exogenous variables. Dropout diminishes internal validity when the patterns of dropout are linked to the variables being studied (Barry, 2005). One way to reduce bias resulting from attrition is to measure any intent of attrition, because any information resulting from that measure may lead to evidence to support imputation for some measures (Graham).

External Validity

Missing data also threatens external validity, or the generalizability of findings. When there is non-ignorable missing data, the results of a study would apply only to those who participated, not to the population from which the sample was drawn. As McKnight et al. (2007) explain, when there is a lack of representativeness in a sample, external validity is limited, and parameter estimates can be biased.

Missing Data Mechanisms Conclusions

Without question, survey researchers are most confident when analyzing a complete dataset, one with no missing data. Because this situation rarely occurs, the next best option for researchers is for missing data to be MCAR. When completely random dropout is ruled out, and any form of dependence between the dropout process and measurement process is probable, labeling the dropout mechanism as random rather than informative is suspect (Diggle & Kenward, 1994). Missing data classified as MAR are ignorable and likely produces more valid results than missing data classified as MNAR because there are no systematic reasons for the missingness. Ignorable missing data produces more valid results on both an internal and external level. Concerning internal validity, the relationships between variables are likely to be biased if the groups who participated in a survey or dropped out of a longitudinal survey differ from those who did not participate in that survey or who remained in a longitudinal survey study. In terms of external validity, or generalizability, if missing data is non-ignorable, conclusions from the sample of participants are not likely generalizable to the population.

Because the way that missingness is classified affects the validity of conclusions drawn from a research study, it is important not only to understand that there are differences in missingness, but also to know how to diagnose these different mechanisms within a dataset.

Missingness, even when ignorable, is not meaningless. Streiner (2002) writes that for the most part, people do not drop out of studies for trivial reasons, therefore, those who have dropped out are different, and it is up to the researcher to expose those differences. However, this is approach opens up a discussion, because some would argue that it is those trivial differences (e.g., a cold or bad weather) that are representative of the most ideal MCAR mechanism of missingness. Often researchers use techniques to manage missing data that are statistically simple and available without assessing the appropriateness of using these techniques. Further, often only cross-sectional missingness is addressed without attention to longitudinal missingness. In the following sections, a few of the limitations of using some missing techniques absent diagnosis are identified, because these techniques are so widely implemented sans diagnosis, which can result in considerable bias.

Common Methods Suggested for Handling Missing Data

An Introduction

Again, the phenomenon of missing data is common to virtually all research studies, especially longitudinal studies, resulting in inflated Type I error, decreased statistical power, underestimated standard deviations, and other biased parameters (Streiner, 2002). Missing data are often managed using one of three techniques and within each of these general techniques, there are multiple methods. The first is an available-data technique, which has also been identified as a deletion technique, inference restricted to complete data, or direct analysis of the incomplete data. Using this technique, missing values are identified by a software program and these values are left as gaps in the data (Little, 1988).

Imputation is the second missing data technique and several methods fall under the categories of single and multiple imputation. Imputation methods estimate the values of the

missing data and those values are then used in analyses. Finally, modeling, using augmentation or likelihood-based approaches, is the third technique for handling missing data. Using modeling approaches, sometimes the missing data mechanism itself can lead the researcher to make some decisions about some of the missing values (de Morais & Aussem, 2009).

Researchers should select a missing data technique based on the type and level of missing data, how data are missing, and why data are missing. The decision about which missing data technique to use grows more important as the amount of missing data in a dataset increases. Raymond and Roberts (1987) report that when less than 5% of the data are missing, one's choice of missing data technique seems to make little difference when computing correlations and regression weights. According to Roth and Switzer (1995), when the amount of missing data approaches 10%, the choice of missing data technique can have dramatic effects on the conclusions that are reached. Many analyses, according to de Morais and Aussem (2009), are run with the assumption that any missing data are missing at random and the mechanism surrounding the missing data are ignorable, and therefore, any missing value can be reasonably inferred. This assumption, however, as previously discussed, is difficult to test and violation of this assumption can have serious implications for the conclusions drawn from analyses.

Some missing data techniques require or assume a specific missing data mechanism. Buhi et al. (2008) suggest that available-data methods require any missing data to be missing completely at random, which is rarely the case. When missing data are believed to be MAR, Shafer and Olsen (1998) report that principled methods (e.g., multiple imputation and maximum-likelihood) lead to better analysis than ad hoc techniques (e.g., listwise deletion and mean imputation) suggesting that even within the broad categories of available-case, imputation, and modeling, some methods are suggested over others.

Available data techniques. Researchers using an available-data technique use, as the name implies, only the available data for any analysis (i.e., they do not try to fill in a dataset's missing values). Two missing data techniques have traditionally been used for multivariate analyses with survey data when the analyses begin with the creation of a correlation or variance/covariance matrix: listwise and pairwise deletion. Listwise and pairwise deletion methods are often used because of their simplicity and availability in software packages (e.g., SPSS). While these methods are typically acceptable when missing data is considered MCAR, researchers are cautioned to examine the prevalence of missing data even if they believe their missing data are MCAR to avoid issues like loss of information, small sample size, biased estimates, matrix instability, and threatened validity (de Leeuw, 2001; Nakagawa & Freckleton, 2008; Weisberg, 2005).

Using listwise deletion, only cases with complete data are used in analyses. In a survey dataset, it is entirely possible that several respondents have at least one missing value; therefore, using a listwise deletion approach can result in a large loss of data and less precise estimates than originally planned (Piggot, 2001; Roth & Switzer, 1995).Roth and Switzer (1995) reported that "18.3% of cases from a dataset may be lost to analyses when 2% of the data are missing randomly and entire cases with missing data are deleted" (p. 1004). While MAR is seen as the most common missing data mechanism in social science research (McKnight et al., 2007), listwise deletion methods are most appropriate when missing data are classified as MCAR (Schafer & Graham, 2002) otherwise, eliminating cases may introduce bias (Graham, Cumsille, & Elek-Fisk, 2003, Weisberg, 2005). When cases are eliminated from the dataset, the remaining sample is likely to be representative of the population only if the missing data were missing completely at random. Even when missing data is determined to be MCAR, using listwise

deletion results in decreased power, i.e., inflated Type II error (Buhi et al., 2008; McKnight, et al., Weisberg, 2005).

Like listwise deletion, pairwise deletion, or available-case analysis, uses only available data for statistical analyses, where missingness is defined by missing values in each variable (McKnight et al., 2007). When using a pairwise deletion, only variable-level missing data are discarded, meaning that when compared with listwise deletion methods, a larger proportion of the sample remains available for statistical analyses. Statistically, when comparing a data set with deleted missing data on a variable level to a full data set, McKnight et al. (2007) found differences in item-total correlations and Cronbach's alpha. Because only available data are used, sample size for each variable is likely inconsistent and unless the missing data are MCAR, it is likely that the samples for each variable are different (Buhi et al., 2008). When sample sizes are not equal across variables, analyzing covariance or correlation matrices is problematic at best (see McKnight et al., 2007; Tanguma, 2000; Weisberg, 2005). That is, correlations and R^2 may lie outside an acceptable range. Further, matrices may be found to be not positive definite, which has serious implications in some statistical packages (Tanguma). As with listwise deletion, pairwise deletion is only advantageous if missing data are MCAR so estimates are consistent and are comparable as long as sample size is large (Allison, 2002).

To ensure unbiased statistics, available data techniques should be used very cautiously, when there is minimal missing data and the missing data are MCAR. In this study, using available data techniques would not be advisable, regardless of missing data mechanism. In the case of the MYS data, first, using available data techniques would lead to a loss of considerable data; second, these techniques are often used for item-level missingness, rather than unit-level missingness, even though using listwise deletion does remove the entire case even if only one

value is missing. Finally, the use of available data techniques often overlooks some important characteristics and implications of longitudinal datasets.

Imputation techniques. Another technique used to manage missing data is imputation and although this method is typically thought of for item-level missingness rather than unit-level missingness, it is still worth briefly describing these methods. Imputation is a simulation technique that seeks to fill in missing data, rather than delete instances of it, theoretically allowing researchers to continue with any analyses as if the dataset were complete (Schafer & Olsen, 1998). However, statistically, using imputation techniques without thought or grounds can result in distorted test statistics, leading to biased results (Schafer, 1999). Any imputation used should be, according to Little (1988), within the distribution of the variable of interest.

Single imputation. There are two major categories of imputation methods: single and multiple imputation. Researchers use single imputation to fill in a missing value with only one replacement (Gheyas & Smith, 2009). Under the umbrella of single imputation (SI) are two major categories: deterministic and stochastic methods. Then, within those categories, there are specific techniques to manage missing data. One of the main concerns about single imputation, before any nuances are described, is that when using single imputation, the completed dataset does not reflect the fact that a value was imputed, thus the results do not reflect that there is any uncertainty about the imputed value (Gheyas & Smith).

Deterministic imputation describes the simplest imputation methods and within this classification exist several different approaches (e.g., alternative source, optimization, regression, mean). Because deterministic imputation assigns values in a manner where those values are uniquely determined, the same value is imputed regardless of how many times an imputation were done, therefore, only a single imputation is necessary (Weisberg, 2005).

One deterministic imputation method is an alternative source procedure, where one person's missing data in a variable is determined by another source, a highly accurate method when the information to use the method is available (see O'Rourke, 2003; Roth & Switzer, 1995; Weisberg, 2005).Mean imputation is one of the most common SI procedures because of its simplicity and availability. However, studies using this approach do not provide conclusive evidence that using this procedure either reduces or increases variance (Roth & Switzer, 1995). Further, because any extreme value is underestimated, variance is reduced, leading to further incorrect calculations (e.g., correlations with other variables), even under both MCAR and MAR missing data mechanisms (Enders & Bandalos, 2001; McKnight et al., 2007; Piggot, 2001; Streiner, 2002; Tanguma, 2000; Weisberg, 2005).

Deterministic imputation often underestimates variable variance, leading to "overly narrow confidence intervals and falsely rejected null hypotheses" (Weisberg, 2005, p. 144). There is an element of uncertainty that comes with the deterministic methods; and, even when missing data are determined to have an MCAR mechanism, deterministic approaches still may lead to biased estimates because where one value is imputed for a missing value, there might be more than one reasonable value to impute. Standard errors are generally underestimated which leads to an overestimation of test statistics, which then leads to achieving statistical significance when there should not be (Weisberg).

Stochastic procedures add probability to the previously described SI procedures, which helps to remove bias leading to consistent estimates under the missing data mechanisms of MCAR and MAR (Little and Rubin 2002). Here, instead of imputing the mean value on a variable, the researcher can add a random error term to (or subtract it from) the mean, and there is less distortion in the distribution. Stochastic imputation increases variance, leading to better

estimates of standard error, and resulting in more accurate significance tests. Still, the standard error estimates are too small, leading to large test statistics, possibly resulting in a false rejection of the null hypothesis (Weisberg, 2005).

If there is very little missing data in a dataset, SI methods may be reasonable. The decision surrounding the use of SI methods also depends on what type of data is missing and if there are reasonable sources with which to impute data. However, when SI is used, variability is underestimated (Schafer, 1999). Further, biased estimates are likely (Rubin & Schenker, 1986). To expand SI methods, if you will, multiple imputation (MI) methods use variability to "adjust the standard error serves as a corrective to the attenuated standard error estimates under single imputation" (Weisberg, 2005, p. 150).

Multiple imputation. MI methods are nondeterministic methods and they are thought of as the most statistically complex imputation procedures, involving imputing multiple values, which has been argued to be a good method for estimating nonlinear models (Allison, 2002). MI methods used to be quite complicated for researchers because the probability distributions used to produce multiple imputations were complicated[7]. MI methods use regression-like statistics to estimate a missing value, multiple times, using different starting values each time. That is, using MI, the original dataset is not the only one used; new datasets are created and new parameter estimates and standard errors are computed. It is the mean of these different values that is the

[7] Often in multiple imputation, auxiliary variables, or additional variables (often from other sources) measured along with the desired outcomes variables, are used and that once introduced, these auxiliary variables affect all subsets of imputed data (Collins et al., 2001). These auxiliary variables, which can reduce bias, may be related to the missingness mechanism or correlated with a variable that has missing values, but in any case, researchers should have reasons for introducing them into a dataset (e.g., they are measures of the dependent variable that may not have been included in the analysis model) (Graham, 2009). It is often difficult, however, to assume non-ignorability in a model and in some cases, one may not be able to estimate certain parameters, thus auxiliary variables can be introduced which reduces bias without having to specify a MCAR mechanism, (Ibrahim et al., 2001). Auxiliary variables can be introduced into datasets using MI methods and maximum likelihood analyses. In MI analyses, auxiliary variables are simply added to the imputation model. These auxiliary variables remain in the imputation model and can benefit further analyses, whether or not they directly appear in any further analysis models (Graham, 2009).

final imputed value of the missing observation (Buhi et al., 2008, Nakagawa & Freckleton, 2008). Because these methods use different starting values, resulting in different final values, and then take the mean of those, variability can be built into the model, leading to less or unbiased parameter estimates (Streiner, 2002). Schafer and Graham (2002) explain that a key feature of MI is that the missing values for a respondent are based on his or her other observed values, with the added variability.

MI methods do assume MAR data, though Carpenter (2009) argues that that assumption is not required. MI methods also assume multivariate normality, however when this assumption is violated, findings may still be robust so long as the models used to analyze any imputed data account for non-normality (Buhi et al., 2008; Honaker & King).

Conclusions. As with SI methods, MI methods allow researchers to use a complete dataset for analysis, allowing almost any statistical analysis to be run. One of the advantages of MI is that it has fewer statistical disadvantages than other missing data techniques, because it allows the researcher to assess the uncertainty regarding the imputation (Patrician, 2002; Weisberg, 2005). MI methods also may give the researcher insight about how missing data affect parameter estimation (Nakagawa & Freckleton, 2008; Piggot, 2001). Finally, MI methods can reduce the likelihood of Type I error (McKnight et al., 2007), because a more accurate view of true variance is obtained. On the other hand, MI is not always easy to employ. If imputation estimates vary widely, MI may not be considered an appropriate mechanism to deal with missing data because replication of results would not be likely (Patrician, 2002). Further, if the estimates vary widely, the missing values are possibly non-ignorable, which may result in misleading conclusions.

When considering MI techniques, even if the results are robust to violated assumptions, researchers should still consider and report how their data are distributed (Buhi et al., 2008; Schafer & Graham, 2002). In order to ensure a variety of statistical analyses, "a rich imputation model that preserves a large number of association is desirable" (Schafer & Olsen, 1998, p. 10).

Maximum Likelihood

Maximum likelihood (ML) techniques are often used when assumptions (e.g., normality, independence) are violated. For example, if data are clustered, an ML algorithm can be used because least squares analyses lead to biased estimates which may result in incorrect conclusions. One commonly used algorithm is the expectation maximization (EM) algorithm, which is a deterministic, model-fitting technique typically used with missing data (see Allison, 2000; Frees, 2004; Graham, 2009; Honaker & King, 2007; Schafer & Olsen, 1998).

The goal of ML techniques is to maximize a result (or increase the likelihood of the result to 100%) given other factors. In the case of missing data, ML estimates the population parameters based on available data and is therefore not made dysfunctional by missing data unless the missing values are MNAR. When a missing data mechanism can be labeled as MAR, ML can be used to infer or estimate probable values for those that are missing. Those algorithms that ML employs assume that information contained within the observed values can be used to infer the missing values (Enders, 2001). As the name implies, ML methods seek to maximize the likelihood of the estimates that are inferred based on the observed values. ML is comparable to MI when an imputation model is formulated in terms of the parameters of interest and in this case, Carpenter (2009) explains that ML and MI agree.

ML estimation differs from imputation methods, however, in that the missing values are not filled in; rather, the parameters are estimated (Piggot, 2001) and the missing values are

treated as if they were random variables that need to be integrated from "the likelihood function as if they were never sampled" (Schafer & Graham, 2002, p. 148). If missing values are few in a dataset and if there are good covariates, then even if a missing data mechanism is ignored, resulting biases may be small (Little, 1995).

The techniques used to manage missing data that have been discussed are progressively more statistically complex and progressively more acceptable for use in longitudinal datasets with wave missingness. ML estimation with an EM algorithm is appropriate when missing data are MAR or MCAR and can be used when entire cases are missing and when imputation methods can be argued to be inappropriate. Also, modeling, discussed below, can be used in conjunction with ML estimation techniques if researchers find that the assumption of MAR or MCAR data has been violated.

Modeling

A final approach to dealing with missing data is modeling, which can be used if researchers know or can hypothesize why data are missing. Using model-based techniques, a model is developed as if data were complete and a component for unit non-response is added (Weisberg, 2005). Like other methods for managing missing data, there are multiple procedures that fall under the category of modeling (see Heckman, Ichimura, Smith, & Todd, 1998; Little, 1993). Modeling procedures are advantageous because they can be used even when missing data are non-ignorable. On the other hand, some modeling procedures require several good predictor variables, which may or may not exist in the data. If no predictor variables exist, modeling procedures cannot be used. One of the advantages of modeling procedures is that some of them consider how variables are interrelated (Gheyas & Smith, 2009). It has been suggested that a model that remains stable under relatively small modifications is preferred (Molenberghs et al.,

2002). Like many other techniques, one of the assumptions for using modeling techniques is multivariate normality (Piggot, 2001).

Little (1995) classified these models broadly as random-coefficient selection models (see Heckman, 1979) and random-coefficient pattern-mixture models, writing that these classifications depend on the factoring of the joint-distribution of the data and the missingness mechanism of the data. Little further writes that inference in these models is likelihood-based either using maximum likelihood or Bayesian methods. In these models, a hypothetical complete dataset is modeled and then joined to a model for the missing data mechanism, which is conditional on that hypothetical complete dataset (Little; Jamshidian & Mata, 2008; Molenberghs et al., 2002).

Selection models produce results different from imputation methods because rather than imputing individual missing values, selection models show "how the coefficients for predictors change when the missing data process is taken into account" (Weisberg, 2005, p. 152). This procedure works only to the extent that the researcher can specify a missing data mechanism (Weisberg). Further, this procedure is often statistically complex, discouraging some researchers from using these models to manage missing data (Hogan et al., 2004).

Pattern-mixture models equate missingness mechanisms with missingness patterns. Thus, they stratify the population conditional on the missingness mechanism, resulting in a model for the entire population that takes into account a mixture of missing data patterns (Jamshidian & Mata, 2008; Little, 1995). Because there are often different missing data patterns within a dataset, pattern mixture models allow difference response models to be introduced, and the observed data reflects a mixture of those (Kenward, 1998). These procedures do not solve the questions surrounding missing data mechanisms, per se, but they do allow the researcher to

"describe the observed responses in each missingness group and extrapolate aspects of this behavior to unseen portions of the data" (Schafer & Graham, 2002, p. 172). Pattern-mixture models are not as sensitive to the assumption of a normal distribution as selection models are; and, they can be used with longitudinal data (see Hogan et al., 2004; Schafer & Graham, 2002). It should be noted, however, that there may be several different patterns of missing data that occur in just one longitudinal dataset. When there are just four years of longitudinal data, there are eight different patterns of missingness possible for each participant. The number of missingness patterns increases with each added year of data collection, making the results difficult to interpret.

Sensitivity Analysis

Ideally, any missing data in a dataset is MCAR, allowing researchers to carry on with analyses without the fear of biased parameter estimates. It is irresponsible, however, for researchers to assume that missing data is missing completely at random without employing any diagnostic procedures. Because researchers do not typically have the resources to ask people why they did not participate in a study or did not respond to specific items, it is important for researchers to use diagnostic tools that can provide support for a certain missing data mechanism. However, as previously described, there are some techniques available to researchers that allow missing data to be MNAR, but still produce unbiased (or at least less biased) parameter estimates (e.g., selection and pattern mixture models). Still, problems may arise using these techniques; these problems have been referred to as "sensitivity to unverifiable modeling assumptions" (Molenberghs & Kenward, 2007). Research surrounding these problems has been coined "sensitivity analysis" and allows researchers to describe how missing data affect their analyses, consequently providing evidence of the robustness of any results (Baron et al.,

2008). Sensitivity analyses can also be used to gather more information even when a missing data mechanism has already been determined (see Ibrahim, Lipsitz, & Horton, 2001).

There are several suggestions about how to approach a sensitivity analysis. For example, with longitudinal datasets like the MYS dataset, Molenberghs and Kenward (2007) suggest estimating a joint measurement with dropout model. In another approach to sensitivity analysis, researchers examine how the model responds to the deletion of outliers. If outliers are the only values that cause a model to appear to function under an MNAR mechanism, the researcher should consider deleting those outliers (Molenberghs & Kenward, 2007). Yet another approach researchers may use is a postmodeling sensitivity analysis which allows for distinction between data that are MCAR and data that are, contrastingly, informatively missing (IM), or a dataset that has non-ignorable missingness just by examining the dataset in its current model, not through modeling a mechanism. Here, the researcher may assume that if missing data can be described as MCAR, the asymptotic distributions of maximum likelihood estimates are the same size for randomly selected subsets or subsamples of the sample (de Morais & Aussem, 2009; Jamshidian & Mata, 2008).

When conducting a sensitivity analysis, researchers may choose to assume a general missing data model, assigning particular values to the coefficients of the outcome of interest, or, they may assign prior distributions to these parameters (Jackson et al., 2010). While implementing priors can be approached in several different ways, some authors look at this method in a negative light and regard it as just a mathematical convenience (Schafer & Olsen, 1998). Honaker and King (2007) approach priors implementation using data augmentation, which works within an EM algorithm.

Again, it is important to remember that once an MCAR mechanism has been ruled out, the researcher cannot statistically demonstrate that either an MAR or an MNAR mechanism is functioning. For example, a sensitivity analysis may give evidence that an MAR mechanism is not operating while not necessarily providing evidence that an MNAR mechanism is operating (Molenberghs & Kenward, 2007). Therefore, it is appropriate to use a sensitivity analysis while also using logic when making decisions about missing data mechanisms. Once researchers have some information about what patterns of missingness exist within their datasets, they can make more informed decisions about how to manage that missing data. There are several different techniques that can be used to "solve" missing data issues, but not all of these are advisable for use with between-wave missingness or with all missing data mechanisms. Further, not all of these are possible for the typical researcher or evaluator given their complex statistical nature and time constraints.

In order to analyze a longitudinal dataset appropriately, missing data must be diagnosed and managed. Addressing missing data is very important because the generalizability to a population depends on the representativeness of its sample. Therefore, it is important to understand missing data mechanisms and the techniques that have been offered as a way to manage missing data. Researchers who do not diagnose the missing data patterns that exist in their data may not approach managing it in an appropriate way, which can lead to biased parameters and incorrect conclusions. Just because researchers know the difference between missing data mechanisms does not provide them with the means to diagnose the mechanism absent a combination of logic and sensitivity analysis techniques. Further, even when researchers know which technique should be used to appropriately handle missing data, they may not be able to employ the proper technique to handle the missing data. This does not mean,

however, that a diagnosis of missing data should be overlooked. There has been much discussion over the past years about addressing missing data on a cross-sectional level and there has been discussion of how to address missing data in a longitudinal clinical study, but there has been less examination of wave missingness and between-wave missingness in a community-based longitudinal field study, like the MYS.

Representativeness: An Alternative Way of Viewing Data Missingness

Missing data affects different aspects of a research study, namely the analyses that can be appropriately performed and the conclusions that are drawn from the analyses. One way missing data affects conclusions that are drawn from a study is how missing data can lead to a sample that is not representative of the population being studied. Missing data can also lead to causal inferences that are incorrect if the factor that causes a sample to be biased or nonrepresentative interferes with appropriate conclusions about the outcome(s) (Shadish, Cook, & Campbell, 2002). While missingness has been described to this point in terms of mechanisms and techniques to handle missing data, it can also be approached through assessing representativeness. Often, missing data occur when eligible people drop out of a study after already enrolling. In a clinical study, researchers usually know who those dropouts are. In a community study, like the MYS though, there may be no definitive sampling frame, making it difficult to know who should be enrolled. Therefore, in this study, non-participants can be defined as those who are eligible to participate in the survey in a given year, but do not. Thus non-participants have missing data. Those who are not eligible include those (a) who live outside the geographical boundaries of the neighborhoods being studied, and (b) who fall outside the defined age range (10 through 18). In a longitudinal study, one way to examine missingness is to look at patterns of missingness per case. For example, in five years or waves of data

collection, a participant may have five data points; or, he/she may have three consecutive data points, followed by one year of no data collection, followed by a final year of data collection. In the MYS, if missing data occurs whenever there is a lack of data for any eligible participant, missingness may occur before an adolescent is enrolled in the survey; there may also be missingness after the adolescent is enrolled in the survey. In the case of an eligible adolescent who never enrolled in the survey, there is absolute missingness.

The questions of why the data are missing and whether those reasons are relevant to the study still exist; and, these reasons affect the degree to which researchers can conclude their samples to be representative of the population. If a sample has been found to be representative, it is more likely that any missing data within the dataset can be considered missing at random, allowing researchers more freedom in choice of analyses. In a longitudinal study, there are two types of representativeness: cross-sectional and longitudinal. When any data are missing and the missing data mechanism is unknown, longitudinal representativeness does not guarantee cross-sectional representativeness, nor does cross-sectional representativeness guarantee longitudinal representativeness, and both are quite important to establish before moving on to other analyses.

Representativeness is how a sample compares to a population on various parameters and is one of the characteristics of a study by which it is judged. Representativeness is also discussed in terms of selection bias and population validity (Bracht & Glass, 1986). When a sample is representative of its population, generalizations can be made from the dataset to the population; in fact, in order for researchers to make any statements about a project beyond any observed measurements, evidence of a representative sample should be provided. Generalizations can be made on a cross-sectional level and across time. Thus, in a longitudinal study, researchers should strive to achieve both within-and between-wave representativeness.

Often researchers approach representativeness with respect to a sampling strategy without really assessing the representativeness of the final sample. When assessing representativeness, if 95% of those recruited to participate consistently participated in a study, diagnosing the representativeness of the dataset is not as important as if the response rate was only 70%, which may not guarantee representativeness on all variables.

Simple random or stratified sampling techniques are probability techniques where each member of the population has a known chance of being selected or recruited to participate in the study. Theoretically, if each member of the population has a known chance of being sampled, and that sample is determined randomly, and the response rate is unity, the sample is representative of the population (within known sampling error intervals). Practically, even in simple random sampling, though each individual in the population may have the same probability of being chosen to participate in the study, not all of those chosen may participate, and there is no guarantee that those who actually do participate are random. That is, there may be reasons related to the outcomes measured accounting that keep eligible participants from enrolling or that lead to their dropping out if they do enroll. In a longitudinal sample, the originally recruited sample might respond in large numbers the initial year of the survey, but may not continue participating. Frees (2004) mentions that when the same subjects are studied over time, it is likely for their response rates to decline over time, noting an example where in the first year (1968) of a panel study, non-response was at a rate of 24% and by 1985, that non-response rate had grown to 50% (p. 12). In other words, by the laws of probability (p(a and b) = p(a) × (pb)), the probability of complete data over two years equals the probability of complete data during any given year squared, and the probability of complete data during three consecutive years equals the probability of complete data during any given year cubed, etc.

Further, in some community-based samples where a population is difficult to recruit, the proposed sample size must often be increased to the point where inadequate resources exist to ensure a high response rate. As a result, randomness may suffer. Even when an attempt is made to recruit an entire population to participate in a study, there are most likely still some members of the population who do not participate. Therefore, researchers should try to examine the reasons for nonparticipation, in the event that those reasons may be informative, or non-ignorable, and related to the outcomes measures.

In the case of a representative sample, still there may be missing data in the final dataset. In a longitudinal dataset, there might be unit- and wave-level missingness, with drop-out occurring in some waves. Depending on whether those with missing data dropped out for random or non-random reasons (i.e., with respect to the responses they would have given had they been surveyed), the resulting dataset may not be truly representative of the population, even if recruitment and sampling was random. Several techniques can be used to address the lack of representativeness in a dataset.

One technique used to address cross-sectional representativeness is weighting, which is used more often with unit non-response than with item non-response. Weighting can be used in datasets where non-ignorable missingness cannot be ruled out and can result in less bias in observed variables, not in unmeasured variables[8]. Representativeness is also important on a longitudinal level, that is, over time with cases with multiple data points. One way of examining whether single-observations cases are legitimately part of the sample is to remove all those cases with only one data point and determining whether cross-sectional representativeness increases when those cases are removed.

[8]For more discussion on weighting, see Lin & Schaeffer, 1995; Little, 1988; Patrician, 2002; Schafer, 1999; Schafer & Graham, 2002.

Again, if a representative sample were achieved in a survey study *and* if there were no incomplete cases, analysis of the dataset could take place and conclusions resulting from any analyses would be unbiased. This is, however, rarely the case, and even when a logical analysis of the sampling strategy and participant numbers suggest that a sample is representative of the population (e.g., a large sample, aggressive recruitment strategy, high response rate), the sample should be assessed for representativeness. Proxy measures such as school system records can be used to compare participants with non-participants and can provide support for either a representative or non-representative sample. If researchers find that one wave of data is non-representative, methods such as weighting can be employed to make a sample representative of the population on key characteristics. Once a dataset has been determined to have within-wave representativeness, the researcher is still interested in whether the dataset has between-wave representativeness. When sampling is geographically determined (e.g., by neighborhood, as in the MYS), the researcher must determine which cases in an auxiliary dataset (i.e., school records in the case of the MYS) are associated with each neighborhood. This can be done most effectively based on the address associated with each school record, and assignment of those addresses to specific neighborhoods using a geographical information system (GIS) framework. Once neighborhood boundaries are delimited and spatial coordinates (i.e., latitude and longitude) for each address are calculated, GIS software automatically assigns each address to one of the defined neighborhoods. However, the process of calculating spatial coordinates, called geocoding, is neither trivial or error free.

Geographic Information Systems (GIS)

Because addresses do not automatically indicate the poverty designation of a person's neighborhood, GIS can be used to identify high-poverty neighborhoods and assign addresses to

those neighborhoods. While a more complete description of the geocoding processes used in this study can be found in Chapter 3, a general description of geocoding and creating a GIS is discussed in this section. In a study with a geographically stratified sampling strategy, like one focused on respondents living in impoverished neighborhoods, GIS can be used to (a) define neighborhoods, (b) assign participants to neighborhoods, (c) assign school district addresses to neighborhoods, which then allows (d) the analysis of participant and non-participant school records jointly by neighborhood. GIS is used to accomplish (a), (b), and (c) in this study, and of particular importance is geocoding, which is used to accomplish (b) and (c); and because of its importance, it is important to discuss how geocoding is done and what possible problems are associated with it. In one approach to geocoding, researchers first match addresses using software that produces a geographic coordinate for each referenced address. Then GIS software is used to plot each coordinate (Hay, Kypri, Whigham, & Langley, 2008). Matches are made because geocoding software introduces an address locator into a shape file, which includes geographic information (e.g., latitude and longitude coordinates). This address locator has roads and road segments, each containing two address ranges (odd and even numbered addresses). A match is made generating a geocode through the interpolation between two identified addresses at the end of the part of the road (Strickland, Siffel, Garner, Berzen & Correa, 2007). Geocoding is not always error-free, however.

Data-Level Errors

Errors in geocoding can occur on a data-level for several reasons, including the misspelling of names, colloquial street names, or incompletely reported addresses. Errors may also exist internally in the geocoding software (e.g., unrecognized streets or intersections) (Hay et al., 2008). Further, errors might occur when researchers match addresses to the wrong

geocode (errors in latitude, longitude, census block group, or census tract) (Krieger, Waterman, Lemieux, Zierler, & Hogan, 2001). Also, an address may be incorrectly recorded, but that address may still match to a viable address (Strickland et al., 2007). Finally, errors may occur when matching criteria are relaxed in order to increase the match rate in a geocoded dataset. While the matching success rates increase, accuracy may be sacrificed (Cayo & Talbot, 2003; Strickland et al.).

Positional Error

Next, some addresses that do not match in geocoding software for any of the previously listed reasons still may be hand matched in geocoding software. This method of matching may prove to be inaccurate when considering the precise coordinates of an address compared to the coordinates that were hand-matched. Even when addresses are geocoded correctly, there still may be errors in the actual point coordinates. Cayo and Talbot (2003) explain that the information used to match addresses comes from state and local governments and is compiled into a file (e.g., TIGER) and that these files have a positional accuracy of ±167 feet. Because these files are being compiled from different data sources, there may be different amounts of error in different address matches. Therefore, researchers are often unable to report exactly how much error exists in the geocoded file.

Geographic or Cartographic Confounding

Geographic confounding is another type of geocoding error related to location, where incomplete addresses (e.g., missing street numbers) are matched using a coarser match geocode. That is, rather than using a point-system for matching, that is, an exact coordinate to match an address to, a polygon or street segment centroid is used as the matching factor. While this type

of matching may not be inaccurate, Strickland et al. (2007) found that this type of geocoding resulted in approximately 5% error, that is, addresses being placed into the wrong census tract.

Selection Bias

Errors in geocoding may also occur on a software level "due to interpolation mechanisms used by most geocoding software to calculate coordinates using street segment and street number range data...[which] assume a homogeneous distribution of addresses along a street segment" (Hay et al., 2008, p. 562). These positional errors tend to occur more often in rural areas than in urban areas (Cayo & Talbot, 2003; Hay, et al.). This discrepancy is due to several reasons as described by Hay et al.: (a) rural addresses can be less precise and often people who live on these routes use post office boxes; (b) there may be more use of unofficial street names in rural areas; (c) longer street segments may increase interpolation errors; and (d) "roadway reference data for rural areas are less accurate than they are for urban areas" (p. 563).

Errors in Geocoding Analyses

When thinking about analyses using areas rather than people as the unit of analysis, researchers should be cautious about committing Type I and Type II errors. When researchers use a large spatial unit as the unit for analysis, for example a polygon rather than a point, they can overlook important differences that may occur within that polygon (Hay et al., 2008). On the other hand, when an exact match is not possible and matches are less precise, there may be misclassification and addresses may be matched incorrectly. In terms of Types I and II errors, "a Type II error will occur if the misclassification is large and unsystematic and, Type I error may result if misclassification is not random" (Hay et al., p. 565). Further, data may be lost because of geocoding error or failure to match addresses, causing sample bias as well as reduced statistical power to detect important associations (Cayo & Talbot, 2003). This sample bias, or

inefficiency of geocoding in certain areas can result in non-random missingness of geocoded addresses (Oliver, Matthews, Siadaty, Hauck, & Pickle, 2005). Again, because geocoding is used in this study, it is important to acknowledge the possible errors that are associated with geocoding, so that special attention can be made to minimize these errors when completing the geocoding process with the addresses given in the datasets.

Conclusions

In missing data research, several methods or techniques have been proposed to handle missing data. Some of these procedures are as simple as a couple of point-and-clicks in software (available data techniques); however, these methods are not widely advised for handling missing data in longitudinal survey research, especially using a hard-to-reach population. As techniques used to handle missing data become more acceptable, they also become more statistically complex and require greater sophistication and more resources. These techniques, however, involve diagnosing the missing data and/or missing data mechanisms that might exist in a dataset. An alternative framework for considering missingness involves assessing the representativeness of a sample each year of a longitudinal sample and determining the missing data patterns in that longitudinal sample. This approach, however, requires access to a relatively complete dataset to supplement the dataset that is being assessed. When such a dataset is available, an alternative or novel way of approaching missingness and representativeness can be used without having to have the complex statistical knowledge of modeling and larger amounts of resources, including time and money. In the next chapter, the method used for assessing representativeness and missing data patterns using this newer approach is described.

CHAPTER 3

METHODS

Mobile Youth Survey

Before discussing the methods used to address the research questions in this study, it is important to first describe in more detail the Mobile Youth Survey (MYS). Specifically, it is important to discuss the geographical context of the MYS, the MYS survey instrument, and MYS recruitment and administration procedures; these shed light on why there may be missing data in this study and how the missing data may have come about.

Geographical Context

The greater Mobile area (the Metropolitan Statistical Area (MSA), has a population of 540,258. The largest city within this MSA is Mobile, which has a population of approximately 200,000. In 2000, 46.1% of the residents of Mobile were African American and 22.4% of this population lived in poverty ("U.S. Census," 2012). The median household income in Mobile in 2000 was $31,445. Prichard, the second largest city in the Mobile MSA, borders Mobile to the north. In 2000, 83.3% of its residents were African American and 44.1% of this population lived in poverty. The median income for the city of Prichard was $19,544 ("U.S. Census," 2012).

In 1998, the MYS was conducted in the 13 most impoverished neighborhoods in the Mobile MSA and during that summer, the first wave of participants (98% African American) in the MYS was surveyed. Characteristics of the individual neighborhoods are presented in Table

3. Overall, at this time in these neighborhoods[9], the population was over 95% African American, where the 1990 median household income was $5,190 and the 1990 poverty rate was 73% ("U.S. Census" 2012).

Sample Recruitment

When it began in 1998, the sampling goal of the MYS was to recruit the entire eligible population of adolescents from the target neighborhoods. Eligibility is defined as adolescents between the ages of 10 and 18[10] who either lived in or were otherwise connected to one or more of those targeted neighborhoods.

Recruitment of MYS participants was approached in two ways in 1998. Specifically, the actively recruited sample consisted of about half of the residences (apartments) in public housing neighborhoods that housing authority records identified as housing youths between the ages of 10 and 18. Further, in non-public housing neighborhoods, about half of the residences (houses and apartments) were actively recruited (Bolland, Bolland, Tomek, Devereaux, Mrug, & Wimberly, n.d.). Passive recruitment was also a strategy where signs were posted inviting adolescents to contact the researchers to come to survey sites (word of mouth from actively recruited participants played into this second approach as well). All those who were eligible had to have informed consent from an adult caregiver in order to participate in the MYS.

In 1998, 1,771 adolescents completed the MYS. The response rate is difficult to calculate because researchers did not have a definitive sampling frame. Bolland (2008) reports a refusal rate of 9.1% of the 1,526 actively recruited respondents; further, 11.5% of those were never

[9] The 13 neighborhoods were defined by boundaries based on three primary criteria: (a) natural boundaries (e.g., large roads; public housing boundaries) (b) correspondence with census boundaries (which often proved to be difficult); (c) time spent in neighborhoods (after spending considerable time in the neighborhoods, the researchers attempted to delineate housing styles, land use patterns, etc., see Bolland, 2008).

[10] Nine year-olds were allowed to participate as long as they turned 10 before August 16 of the year in which they took the survey; 19-year-olds were allowed to participate as long as they turned 19 after May 16 of the year in which they took the survey.

contacted and another 20.4% agreed to participate (and obtained informed consent) but never completed the survey. The researchers conservatively report a 60% wave response rate amongst those actively recruited in 1998.

In subsequent years, each of those participants who were still age-eligible was actively recruited to participate until he/she turned 19, even if residential mobility took the respondent beyond one of the 13 target neighborhoods. New participants were added each year, resulting in approximately 8,700 adolescents completing more than 23,500 surveys between 1998 and 2007. Thus, the MYS is one of the largest and longest running longitudinal studies conducted with at-risk impoverished adolescents. A summary of participants through 2007 is presented in Table 4.

As the participants moved from the place of residence at the point of enrollment, they continued to be actively recruited to participate in the survey in subsequent waves of the survey. When found, these individuals were asked to participate in the survey again, as were any siblings or relatives in the household who met the age criteria for participation. By 2005, 38.8% of the MYS participants lived outside of the 13 MYS target neighborhoods. Further, moves tended to be clustered. Therefore expansion neighborhoods were identified and former participants were actively recruited to come to survey sites in these neighborhoods beginning in 2006. In addition, in 2006, all the eligible adolescents on selected streets were recruited[11]. Eligible walk-ins continued to be accepted as survey participants[12].

[11] Names were identified by the Mobile County Public School System (MCPSS) records.
[12] In order to verify residence, sources including the MCPSS records and Mobile Housing Board (MHB) records were consulted after surveys were completed.

Table 3

Description of MYS Target Neighborhoods: 2000 Census

	Census tracts (Block groups)	Population	African-American population	Poverty rate (individuals)	Extreme poverty rate (individuals)	Median household income
Non-Public Housing						
Plateau [a](M)	12	2,511	88%	56.7%	28.3%	$13,810
Harlem [a](P)	39.02 (1)	1,203	85.6%	47.1%	11.2%	$18,426
Martin Luther King (M)	4.01 (2, 3, 4)	2,827	97.2%	49.5%	30.6%	$12,157
	5 (1)					
Snug Harbor (P)	43 (1)	535	100.0%	65.2%	24.2%	$11,597
Alabama Village (P)	47 (1)	2,565	84.5%	70.7%	39.0%	$10,793
	48 (1, 2)					
Trinity Gardens (M)	39.01 (1, 2, 3)	2,479	97.9%	31.5%	12.2%	$18,374
Public Housing						
Orange Grove (M)	4.01 (1, 2)	3,517	98.7%	76.3%	59.2%	$6,696
	4.02 (1, 2)					
Josephine Allen Homes (M)	12	2,511	88.8%	56.7%	28.3%	$13,810
Roger Williams Homes (M)	6 (2, 3)	2,326	97.2%	56.7%	30.3%	$11,236
Oaklawn Homes (M)	13.02 (2)	1,816	98.2%	44.2%	22.9%	$14,648
R.V. Taylor Plaza (M)	15.01 (2, 4)	3,139	95.6%	64.6%	36.9%	$9,963
	15.02 (1)					
Gulf Village (P)	48 (1)	943	94.7%	81.4%	44.1%	$8,783
Bessemer Apartments (P)	40 (4)	1,487	98.0%	57.7%	30.3%	$11,950

Note. From "Testing methodological assumptions about the Mobile Youth Survey," by J. M. Bolland, 2008, unpublished grant proposal submitted to NICHD. Reprinted with permission.
[a]M = Mobile; P = Prichard

Table 4

Mobile Youth Survey Multiple Cohort Design

			Data Points									
			10	9	8	7	6	5	4	3	2	1
Year	*N*	New Cohort										
1998	1,771	1,771	9	31	75	88	170	173	237	255	349	384
1999	2,454	1,213		46	45	74	81	119	128	151	195	374
2000	2,185	615			40	49	53	55	72	63	92	191
2001	2,456	871				89	81	77	107	105	138	274
2002	2,256	686					90	83	66	89	122	236
2003	2,267	648						94	100	114	124	216
2004	2,306	535							166	118	107	144
2005	2,632	733								269	226	238
2006	2,351	587									336	251
2007	3,088	1,066										1,066
Total	23,766	8,725	9	77	160	300	475	601	876	1,1164	1,689	3,374

Note. From "Testing methodological assumptions about the Mobile Youth Survey," by J. M. Bolland, 2008, unpublished grant proposal submitted to NICHD. Reprinted with permission.

Assessing Representativeness of the MYS

Given that there were two recruitment strategies for participation in the MYS: one active (random) and one passive (non-random), the MYS does not necessarily contain a representative sample of the population. Even when sampling strategies are random, and especially when they are not, the representativeness of the sample should be evaluated. In this study, sample representativeness is linked to missing data assessment. Given the complex statistical nature of the suggested methods used to assess missing data, often assumptions are made that missing data is missing at random without any assessment of those assumptions (e.g., Horne, Sugai, Smolkowski, Eber, Nakasato, Todd, & Esperanza, 2009). Theoretically, these assumptions allow researchers to analyze their data "without bias;" but, only to the extent that the assumption is valid.

The current study presents an alternative approach to assessing missing data by looking at the representativeness of a sample. Because there was no attempt to limit participation in the MYS (except for eligibility requirements), the final sample should be studied in terms of who participated in the study, and maybe more importantly, who is missing from the study and why.

Assessing the representativeness of a community-based longitudinal study of an at-risk population can be quite daunting, and some might question whether it can be done at all. In the current study, representativeness is assessed using an auxiliary dataset (MCPSS records) and comparing participants to non-participants living in the recruited and assessing longitudinal dropout.

Measures

Researchers have been able to link adolescent development and behavior to different factors, one being geographic area (Ellen & Turner, 1997). Like most datasets, there is missing

data in the MYS dataset. Therefore, especially absent random sampling techniques and a high response rate, the representativeness can and should be assessed in this survey. In addition to the MYS dataset, the MCPSS records are available, providing more or less complete records that can be used to compare MYS participants to non-participants, ages 10 through 15, within the sampling frame of the MYS neighborhoods. These records provide information about each of the neighborhoods annually, whereas census records only provide information for each decade. As in any research study, before describing how each research question is addressed, it is important to first describe how the raw data was integrated to form the dataset that was used in this study. Then, it is also important to describe and define the variables used in the analyses, particularly in the cases when multiple data sources are being used.

In order to select the students who met the requirements for this study, the process of geocoding was used. That is, the initial sample was limited to adolescents who live in MYS neighborhoods, initially defined as the 13 poorest neighborhoods in Mobile, with neighborhood boundaries established by major roads and by housing type (public housing vs. non-public housing) (see Bolland, 2008). High rates of residential mobility made it necessary for researchers to expand their recruitment of participants for the MYS beyond the 13 initially targeted neighborhoods and beginning in 2006, new neighborhoods were added for active recruitment. The exact boundaries of these new neighborhoods were identified post hoc by aggregating census blocks where actual MYS participants lived. Major roads, major barriers and landmarks, and housing types were still used to establish the boundaries of these expansion neighborhoods.

Using GIS software, the addresses connected to each student who attended a MCPSS school attended by five or more MYS participants were geocoded, providing an exact latitude

and longitude for each given address. Then, those points were joined to polygons established by neighborhood boundaries, resulting in an identification of the population of MCPSS students aged 10 through 15 living in each neighborhood.

This geocoding showed that the majority of these addresses (over 95%) could be matched to a composite address locator, and showed that the majority of unmatched addresses are rural addresses in parts of the county where very few of the MYS participants live. This geocoding is analyzed with caution due to possible sources of error previously described.

The first aim of this study is to determine the representativeness of the sample of students (a) who enrolled in the MYS and (b) who participated each year in the MYS. Any student who at any time participated in the MYS was counted as enrolled in the MYS (enrollees: $E = 1$; non-enrollees: $E = 0$). Thus, enrollment is time invariant. In contrast, yearly participation ($P_{t)}$ depends on time (e.g., $P_1 = 1$ when a student participates in the MYS in Year 1) and each year, participants are compared to non-participants (participants: $P_t = 1$; non-participants: $P_t = 0$).

Second, between-wave dropout is analyzed for participants in the MYS. Students were also identified as participating longitudinally ($P_{t-1, t}$) in the MYS. That is, if a student participated in the MYS for two consecutive summers, he/she was identified as longitudinally participating in the survey ($P_{t-1, t} = 1$). If a student participated in one summer, but not the next summer and was still eligible to participate (e.g., participated in the summer of Year 1, but not in the summer of Year 2), he/she was identified as not participating longitudinally ($P_{t-1, t} = 0$). For the purposes of this study, there is only an interest in whether someone dropped out of the survey (but was still age eligible), rather than dropping into the survey in the second of a consecutive pair of years.

To examine representativeness, two categories of characteristics are considered: demographic and functional. In this study, most of demographic[13] and functional characteristics are reported and/or derived from the MCPSS records. To further address the research questions, dropout was studied as it relates to one more set of functional characteristics, reported by the MYS.

Demographic Variables

Demographic variables examined in this study include gender (G), race (R), grade level (Gr), free lunch eligibility status (LE), and neighborhood type (NT). Gender is identified by the MCPSS records (Male: G = 0; Females: G = 1). Race is also identified by the MCPSS records[14]. In this study, race is further collapsed into two categories: African American (R = 1) and non-African American (R = 0). Race was treated dichotomously because of the low prevalence of Caucasians, Hispanics, Asians, and other races in the study neighborhoods. Grade level is also identified by MCPSS records and treated as a continuous variable in all analyses. Finally, free lunch eligibility status is identified in MCPSS records. The MCPSS takes part in the National School Lunch Program (NSLP), which provides reduced-cost or free meals to students who are eligible according to federally established standards ("National School Lunch Program," 2012). In this study, there are three categories of free lunch eligibility status: those who do not qualify for any federal assistance for school lunch (LE = 0), those who qualify for a reduced cost lunch (LE = 1), and those who receive a free lunch (LE = 2)[15].

Neighborhood differences are important to consider; however, given that there are 48 neighborhoods included in the MYS sampling frame, in these analyses, it is cumbersome to

[13] Neighborhood type is included in the demographic characteristics and this characteristic is based on GIS analysis.
[14] The MCPSS does not treat mixed race as a category; therefore, most mixed race African American-Caucasian students are treated as African American in these records.
[15] For the longitudinal analyses, the categories of free lunch eligibility status were collapsed into two categories, paid and free/reduced.

report neighborhood differences. Therefore, as a convenience to the reader, neighborhoods were divided into neighborhood types (NT), consisting of the originally targeted neighborhoods (NT = 1) and the expansion neighborhoods (NT = 0). Individual neighborhood identification is included in the analyses nested within neighborhood type, and is reported in significance tables, but is not discussed further.

Functional Characteristics: MCPSS

In addition to examining demographic characteristics of those enrolled and participating in the MYS compared to those not enrolled and not participating in the MYS, respectively, functional characteristics are also examined. These functional characteristics include school achievement, as indicated by Stanford Achievement Test (SAT) percentile ranks and by weighted school violation and weighted disciplinary action scores. All of these functional characteristics are reported by the MCPSS.

First, students in 3rd through 8th grades are assessed annually using the SAT[16]. Two major subscales of the SAT are reading and math. Total Percentile Rank[17] is used in this study to indicate student achievement in reading and math[18].

Next, the MCPSS records school violations and disciplinary actions for each student annually. Violations were classified as A, B, C, D, or E[19] violations, in order of severity. As an

[16] The SAT 9th edition was given to students in Alabama during the 1998-1999, 1999-2000, 2000-2001, and 2001-2002 school years ("Reports," n.d.). Available for the SAT 9th edition are 1995 norms and updated 2000 norms ("Stanford Achievement Test," 2012). The SAT 9th edition was given, in Alabama, to students in 3rd through 11th grades. In the Spring of 2003, students, only in grades 3 through 8, in the state of Alabama were given the Stanford Achievement Test, 10th edition ("Reports," n.d.). The SAT 10th edition is based on 2002 norms, and updated 2007 norms are available ("Stanford Achievement Test," 2012). Because the SAT 10th edition was given only to students in grades 3 through 8 in those years, for this study, the sample for years when the SAT 9th edition is also limited to grades 3 through 8. Further, while it is cautioned that results for the SAT 9th edition should not be directly compared to results from the SAT 10th edition ("Reports," n.d.), in this study, students are being compared within a single year, so it is appropriate to conduct these analyses.

[17] SAT 9 only reported the percentile ranks while SAT 10 reports percentile ranks, stanines, normal curve equivalencies, and raw scores, etc. Even though more detailed information is available for years when the SAT 10 was used, in order to be use consistent models, percentile ranks are used in all analyses.

[18] Some students had a 0 for one of their Total Percentile Rank scores in either reading or math. Based on consultation with MCPSS staff, scores of 0 were treated as missing data for these analyses.

overall measure, violations were weighted (A = 1, B = 2, C = 3, D = 4, E = 5) for severity and then summed for each year. School violations can result in a number of disciplinary actions. Rather than summing the number of office referrals as a measure of discipline, in this study, only more severe disciplinary actions are considered: (a) detention, (b) referral for conference, (c) in-school suspension, (d) out of school suspension, (e) alternative school placement, and (f) expulsion. As an overall measure, disciplinary actions were weighted (a = 1, b = 2, c = 3, d = 4, e = 5, **f** = 6) for severity and then summed for each year.

Functional Characteristics: MYS

To supplement the analyses conducted to assess longitudinal representativeness, five items from the MYS items were used in additional analyses. Three of these are indicators of risky behaviors. One item is related to neighborhood affiliation; and one final item is related to paid employment (See Appendix A).

Age Restrictions

Age is a limiting factor in this study. First, age is limited by the MYS because adolescents are eligible to participate between the ages of 10 and 18. However, in the current study, data are limited to students ages 10 through 15. This limiting of the data was done because the goal of this study is to determine the representativeness of the MYS to the population. Because students are legally required to attend school through the age of 15, for adolescents in that age range, neighborhood population and MCPSS population are coincident. Unlike other populations, there is little home schooling in these neighborhoods, which would otherwise lead to deviations between the sample and population. Further, there is little private school attendance by adolescents living in these types of neighborhoods ("Statistical Abstract,"

[19] A violations indicate minor infractions, e.g., tardies or dress-code violation. B violations indicate similar, but more serious violations, e.g., willful disobedience. C violations indicate violence and aggression (without weapons). D violations indicate drugs and/or alcohol. E violations indicate bringing weapons to school.

2012). Beyond age 15, however, because students can drop out of school and because the MCPSS records provide no indication of dropout, the neighborhood population and the MCPSS population begin to deviate, and to include students 16 - 18 in the study would introduce bias. While there may be some error in reported MCPSS addresses (e.g., students moving and not notifying the school), Bolland (2008) estimates this error to be less that 13%. Age eligibility was calculated by date of birth given by MCPSS records. Students were eligible if they turned 10 by August 15 of the school year and were not eligible if turned 16 after May 16th of the school year.

Data Analysis

In this section, the two research questions are addressed separately. For reference, the two research questions addressed in this study are:

> Research Question 1: To what extent is enrollment and yearly participation in the MYS (1998-2007) representative of the population of adolescents living in the MYS neighborhoods, in terms of demographic (grade level, gender, race, free lunch eligibility status, and neighborhood type) and functional (cognitive and behavioral) characteristics?
>
> Research Question 2: Longitudinally, can the between-wave missing data mechanisms in the MYS data be considered to be missing at random (that is, reasons for missing data are not related to outcomes)?

Both of these research questions are concerned with missing data, three components of which have been identified:

> (a) unit non-response (i.e., failure of recruited individuals to participate in the study); (b) item non-response (i.e., failure of individual respondents to answer specific questions); and (c) wave non-response (i.e., failure of participants to complete any given wave of data collection. (Bolland, 2008, p. 12)

Research Question 1 focuses on unit non-response, although it also addresses elements of wave non-response[20]. Research Question 2 exclusively addresses wave non-response. Item non-response is not expected to highly affect any analyses because there is so little of it (Bolland, 2008). Further, diagnosing the missing data mechanisms in terms of item non-response does not influence the diagnosis of between-wave missingness. Procedures used to diagnose the missing data mechanisms in item-level non-response are described by Little (1988).

Because the missing data mechanism (i.e., missing completely at random, missing at random, missing not at random) influences the analyses that appropriately can be performed, it is necessary to examine the data, to determine whether missingness is ignorable or non-ignorable for each of these types of missingness. Rather than reviewing these missing data mechanisms using traditional methods (including sensitivity analysis), because there is access to an auxiliary dataset (MCPSS records), these mechanisms are studied in terms of representativeness.

Research Question 1: Enrollment and Yearly Participation

Research question 1 addresses the extent to which the MYS is a representative sample of the adolescents in MYS neighborhoods on a number of characteristics. This question is approached in two ways. First, enrollees are compared to non-enrollees on demographic and functional characteristics. Second, each year, participants are compared to non-participants on demographic and functional characteristics.

Because of the substantial residential mobility of this population, there are cases where one individual is reported in multiple neighborhoods, resulting in non-independence of the observations. Therefore, a Generalized Estimating Equations (GEE) approach is used, as

[20] Wave non-response can first be classified into two broad categories: complete data and incomplete (i.e., dropout or temporary dropout) data. The complete data category can be further broken down into two categories: those with one observation who aged out or who started participation the final year of data collection and those with two or more observations until they aged out or until final year of data collection.

implemented in SAS PROC GENMOD and a goodness of fit is calculated for each analysis with a χ^2 distribution, which is used to test the statistical significance of the effects specified in the model (Liang & Zeger, 1986; Diggle, Heagerty, Liang, Zeger, 2002)[21].

As noted in the literature review, steps can be taken to alleviate some of the biases caused by missing data when the missingness is not ignorable. One of the least statistically complex and most accessible ways of doing this is by weighting the sample, which can only be done to the extent that population characteristics are known. In general, these characteristics are limited to demographic variables. Therefore, in addressing Research Question 1, first, the demographic representativeness of the MYS sample is assessed by regressing enrollment/participation on demographic characteristics. Specifically, main effects to be included are:

1. grade level (Gr);
2. race (R);
3. gender (G);
4. neighborhood (N);
5. neighborhood type (NT);
6. free lunch eligibility status (LE);

interaction effects to be included are:

7. LE × NT;
8. G × NT.

If the sample is found not to be representative in terms of demographics, remediation steps can be taken. A more complex assessment of representativeness is in terms of functional

[21] The dependent variables in the demographic analyses are dichotomous, therefore a logit link function and a binomial error distribution are used. For the cognitive analyses, no link function or error distribution are specified. Finally, for the behavioral analyses, a logit link function and a Poisson error distribution are used.

characteristics, because these characteristics are often unknown in the population. Second, the representativeness of the MYS sample is assessed using proxy functional characteristics, controlling for demographic characteristics; the effects of which could be removed through remediation techniques. Here, functional characteristics (SAT percentile ranks, weighted school violation scores, and weighted disciplinary action scores) were regressed on enrollment/participation, controlling for demographic characteristics. Specifically, main effects included are:

1. enroll (E) or participate (P);
2. grade level (Gr);
3. race (R);
4. gender (G);
5. neighborhood (N);
6. neighborhood type (NT);
7. free lunch eligibility status (LE);

interaction effects to be included are:

8. LE × NT;
9. G × NT;
10. E/P × NT;
11. E/P × G;
12. E/P × Gr;
13. E/P × LE;
14. E/P × R;
15. E/P × NT × Gr;
16. E/P × NT × G.

Demographic main effects were selected on the basis of their availability in the MCPSS dataset. Racial, gender, and/or socioeconomic (i.e., free/reduced lunch status) differences between sampled and unsampled students living in the MYS neighborhoods would likely indicate a sampling failure, with some demographic groups enrolled or participating at higher levels than others. On the other hand, neighborhood differences (and differences in neighborhood type) were largely planned by the MYS researchers, and, by themselves, do not suggest a failure to proportionally recruit different demographic groups. Type of neighborhood may, however, prove to be important in conjunction with race, gender, grade level, and socioeconomic status; hence selected interactions on these are included in the model.

These are not saturated models, and statistically significant results for some main effects and/or interactions may be more important than others when considering the representativeness of the MYS sample and whether missingness can be ignored. First, significant main effects and/or interactions indicate different rates of enrollment or participation for the different

categories of each of these variables. Further, significant effects may suggest that the MYS sample is not representative of the population. However, statistical significance should be considered in conjunction with effect size, with the consistency of patterns of estimates, and with the number of years of significance before conclusions of non-ignorable missingness are drawn. It is expected that for the main effect of neighborhood type, there will be differences between those enrolled and not enrolled and those participating and not participating. These results are expected because of the sampling strategies used in the MYS. As such, for statistically significant differences in neighborhood type to be particularly meaningful, they must be part of an interaction term.

Notably, neighborhood was not included in any interaction; nor were any four- or more-way interactions. Given the structure of the data (i.e., small cell sizes), model convergence would not be possible if all these interactions were included. Even so, some models did not converge and these are discussed on a case-by-case basis. For example, in Year 2, there were no enrollees, and therefore participants, living in neighborhood 48. Therefore, neighborhood 48 was eliminated from all Year 2 analyses. Similarly, in Years 8 and 9, there were a very small number of enrollees, and therefore participants, living in neighborhood 2. Therefore, neighborhood 2 was eliminated from all Year 8 and Year 9 analyses.

The number of participants each year, and therefore the number of longitudinal participants each year, was far smaller than the number of enrollees in the MYS. Therefore, the cell size restrictions were more pronounced for these analyses than for analyses involving enrollment and, further restrictions were implemented. First, in Years 1 through 7, analyses were limited to those who lived in the targeted neighborhoods because recruitment for participation in the MYS was focused on those neighborhoods in those years. Over time, however, increasing

numbers of participants moved to expansion neighborhoods, therefore, in Year 8, enough participants lived in the expansion neighborhoods for the models to converge and for neighborhood type to be included in the models. Next, in the yearly participation analyses, the race by neighborhood type interaction was removed from all of the Year 9 analyses due to lack of model convergence.

Ten years of data were used in this study. For simplicity, years of analyses are referred to consecutively. Table 5 shows how the MYS and MCPSS were paired to each $Time_t$ for analyses.

Table 5

MYS/MCPSS Correspondence to t

t	MYS Year	MCPSS Year
1	1998	1998-1999
2	1999	1999-2000
3	2000	2000-2001
4	2001	2001-2002
5	2002	2002-2003
6	2003	2003-2004
7	2004	2004-2005
8	2005	2005-2006
9	2006	2006-2007
10	2007	2007-2008

There are two different approaches to examining the research questions, in terms of statistical significance. First, a year-by-year approach can be taken, where each year's results are considered separately in order to identify yearly deviations from representativeness. This approach is useful in determining weighting schemes, should that remediation technique be warranted. Second, a study-wise approach can be taken to determine whether the MYS in its entirety is representative; in this approach, sporadic and inconsistent deviations from representativeness are considered anomalous. That is, sporadic and inconsistent deviations from representativeness, while they may be interesting and warrant further exploration, do not, by

themselves, undermine the overall representativeness of the sample on a specific characteristic. When determining whether statistically significant results are anomalous and can effectively be ignored, there are three other considerations. First, the functional forms in statistically significant effects should be examined for consistency across years. Second, similar measures (e.g., weighted school violation and weighted disciplinary action scores) should be examined to determine whether statistically significant results occur in both. Third, very small to small effect size fails to support a conclusion of non-ignorable missingness.

To expand, as many as 10 separate analyses were conducted (one for each year) to assess representativeness in demographic and functional characteristics. Therefore, when setting $\alpha = .05$ for yearly analyses, the probability of making a family-wise Type I error equals .401. To control for α-inflation, individual α-levels are set to .005, yielding a family-wise (across 10 years) Type 1 error equal to .049 ($= 1 - (1 - .005)^{10}$). That is, controlling for α-inflation allows variables for a specific year to be identified as non-representative (Shaffer, 1995).

To examine the sample's representativeness in its entirety (i.e., across years), consistent deviations from representativeness are explored. That is, it is important to look at the consistency of annual deviations, rather than individual deviations. To do this, α is adjusted in a different manner. Specifically, given $\alpha = .05$ for yearly analyses, the probability of committing a Type I error in one of the 10 years of data equals .401 ($= 1 - .95^{10}$); the probability of an additional Type I error in the remaining nine years of data equals .370. And, the probability of committing a third Type I error in the remaining eight years of data equals .337. Therefore, the probability of committing a Type I error for at least three out of 10 years equals .0500 ($= .401 \cdot .370 \cdot .337$)[22]. Thus, if significance is found at an unadjusted α-level = .05 in three or more out

[22] Even when there are fewer than 10 years in a set of analyses, three significant results are still required for further examination. The smallest number of years used in any set of analyses in this study was seven and the probability of

of 10 years, the overall sample may be biased with respect to that variable and results are examined more carefully to determine the nature of the bias.

To examine practical significance, it is useful to calculate effect size, however there is a lack of consensus about the appropriate measure of effect size for likelihood based statistics (Nakagawa & Cuthill, 2007). Cohen (1988) seems to suggest that the appropriate statistic for effect size stemming from a χ^2 statistic used as a measure of goodness of fit in complex models (perhaps including likelihood ratio tests) is **w** and provides a basis for determining whether **w** is small, medium, or large (pp.224-226)[23]. Cohen shows that Cramér's φ (φ_C) is equal to **w** when degrees of freedom is equal to 1. However, when degrees of freedom is greater than 1, **w** equals $(\varphi_C) \bullet (k - 1)^{½}$ where k = degrees of freedom (p. 223). While Johnston, Berry, and Mielke (2006) argue that **w** is not defined for likelihood ratio goodness of fit tests, Howell (2002) writes that for large sample sizes, the likelihood ratio test and Pearson χ^2converge. Thus, with a large sample, **w** may be an appropriate statistic for effect size (i.e., deriving **w** from a likelihood ratio χ^2 provides a reasonable estimate of effect size) (e.g., Aarts, Close, & Wallis, 2009).

Research Question 2: Longitudinal Missingness

The second research question concerns longitudinal missing data patterns (i.e., loss to follow-up) and whether the missing data can be considered to be missing at random. The logic used to address this research question follows the same logic that was used to address Research Question 1, even though meaning of representativeness is different[24].

---

committing a Type 1 error in three out of seven analyses is .018, which is more conservative than using $p < .05$; however, the probability of finding two significant results out of seven is .08.

[23] Cohen (1988) shows examples of small, medium, and large effect sizes for departures from an equal distribution in a two-category classification, such that a small effect size (**w** = .10) corresponds to .45 versus .55; a medium effect size (**w** = .30) corresponds to .35 versus .65; and, a large effect size (**w** = .50) corresponds to .25 versus .75.

[24] In Research Question 1, representativeness is defined as whether the sample of MYS enrollees and participants is representative of the population. In Research Question 2, representativeness is defined as whether the sample of MYS participants at $Time_t$ is representative of the sample of MYS participants at $Time_{t-1}$.

In these analyses, the demographic and functional characteristics of dropouts are compared to the demographic and functional characteristics of non-dropouts, again using a GEE approach. Unlike the analyses used in Research Question 1, the analyses for Research Question 2, for each year ($Time_t$) only considered the cases where $P_{t-1} = 1$. Otherwise, all analyses are conducted as in Research Question 1, with additional model restrictions. Because the number of participants and longitudinal participants each year was far smaller than the number of enrollees in the MYS, the cell size restrictions were more pronounced for these analyses than for analyses involving enrollment and, further restrictions were implemented. As in the yearly participation analyses, for the longitudinal analyses, Years 1 through 7 analyses were limited to those who lived in the target neighborhoods. Also, in all of the longitudinal participation analyses, race as a main effect or component of an interaction was omitted from all models due to the low prevalence of non-African Americans who longitudinally participated in the MYS. Similarly, neighborhood as a main effect was omitted from all models due to the low prevalence of individuals in specific neighborhoods[25]. Finally, few repeat participants in the MYS had a reduced cost lunch status. Therefore, this variable was dichotomized by combining free and reduced cost lunch categories.

In addition to these primary analyses, two sub-analyses were conducted. First, monotonic and non-monotonic patterns of missingness (see Table 2) were explored, with an examination only of those cases who are dropouts. This sub-analysis involves examining whether the $Time_t$ demographic and functional characteristics predict the $Time_{t-1,\, t}$ pattern of missingness using a GEE approach. If these analyses show differences in the characteristics

[25] Neighborhood as a main effect was included in these analyses originally, but only converged in four out of nine years. Therefore, it was removed from all analyses. When neighborhood was included in these four analyses, as a main effect, it was not consistently statistically significant. However, in the four cases when the model did converge, other substantive results were the same with and without neighborhood as a main effect.

between those participants with monotonic and non-monotonic missingness patterns, the monotonicity can be modeled in substantive analyses to control for bias.

A second sub-analysis uses alternative proxy measures to extend the ability to interpret dropout mechanisms. The primary analyses use auxiliary datasets to assess representativeness, but these analyses are limited to 10 through 15 year olds. Also, because these analyses do not use MYS data, they can only be used to model dropout mechanisms to the extent that there is a direct correspondence between the two datasets. It is reasonable to assume this correspondence with respect to unchanging demographic characteristics. However, the functional characteristics as measured by MCPSS data are only proxies of beliefs, attitudes, and behaviors assessed by the MYS. Because the data are missing, there is no definitive way to determine whether missingness is ignorable or non-ignorable. However, in any longitudinal study, one of the best predictors of behavior at $Time_t$ is behavior at $Time_{t-1}$. Thus, still assessing dropout, it is possible to use MYS data at $Time_{t-1}$ as a proxy for MYS data at $Time_t$. In this analysis, using a GEE approach, dropout at $Time_t$ was regressed on the following MYS variables measured at $Time_{t-1}$:

1. Time spent in residence in neighborhood (residential tenure);
2. Arrests in the last year;
3. Carrying a gun in the last three months;
4. Using marijuana in the last year (for complete items and response options, see Appendix A).

These analyses have the additional benefit of using all eligible participants, rather than just those ages 10 through 15. Because paid employment may have a clearer meaning for 16 through 18 year-olds, a second analysis limits the sample to only this age group, including the previously

listed MYS items and adding one to measure amount of hours in paid employment each week. Lack of significant effects can be interpreted as support for an ignorable missingness mechanism.

Interpretation of Results

To discuss the results reported in Chapter 4, criteria for meaningful significance must be established. That is, before concluding that deviations from representativeness exist or result in non-ignorable missingness, several criteria should be met. First, for any given main effect or interaction, statistical significance ($\alpha = .05$) should exist in any given analysis in three or more years. Second, for analyses of functional characteristics, any given main effect or interaction should be statistically significance for both measures of cognitive ability or for both measures of behavior[26]. Third, for any given main effect or interaction, statistical significance should occur in both enrollment and yearly participation for significant results to be deemed meaningful. When statistical significance occurs in an analysis of longitudinal representativeness, it should also occur in an analysis of yearly participation with similar results[27]. Fourth, for any given main effect or interaction to be meaningful, the functional form of estimates should be consistent across years. Fifth, effect size should be small or small-to-moderate. As suggested by Cohen (1988), a small effect (**w** = .10) corresponds to a .45 versus .55 distribution in a two-category classification. A medium effect (**w** = .30) corresponds to a .35 versus .65 distribution. For the purposes of this study, a .45 versus .55 distribution in a two-category classification is not sufficient to conclude that the sample is not representative of the population in a meaningful way. In order for a deviation from representativeness to support non-ignorable missingness, a

[26] This criterion does not apply to the analyses of demographic characteristics.
[27] Similar results should results in the same years and in the same functional form. If there is statistical significance in an analysis of longitudinal participation but not in yearly participation or if results for both analyses were significant but not similar in functional form, it would suggest that the reasons for dropout and for non-participation are different and should be explored further.

small-to-moderate effect size of .15, corresponding to a .425 versus .575 distribution should be consistently found in years of significance.

CHAPTER 4

RESULTS

Before presenting the results for the two research questions in this study, it is important to fully describe the samples that were used to answer these questions. Next, the results for the two identified research questions are presented. In considering Research Question 1, results for enrollment and yearly participation in the Mobile Youth Survey (MYS) are presented separately; within each of these presentations, demographic and functional characteristics are also considered separately. In considering Research Question 2, demographic and functional characteristics as predictors of dropout are presented separately. Following this, the results of the sub-analyses are presented.

Description of Samples

All descriptive summaries are limited to Mobile County Public School System (MCPSS) students, ages 10 through 15, living in MYS neighborhoods. Table 6 shows the annual unadjusted sample size, and for these analyses, sample and population converge[28]. Shown in Table 7, most of these students live in the expansion neighborhoods rather than the target neighborhoods, with the annual rate varying between 75.3% and 81.9% living in expansion neighborhoods. Not surprising, most of the MYS participants live in the target neighborhoods, rather than the expansion neighborhoods Thus, differences in enrollment and participation are expected in neighborhood and neighborhood type (Research Question 1). Table 8 shows the

[28] Subsequent sample size tables are adjusted for missing data in race and grade level.

distribution of these students by race, the majority being African American (*Mdn* = 95.0%), with the annual rate varying between 93.5% and 99.0%. Overall, the proportion of African American students in these neighborhoods is slightly less than the proportion of African American participants in the MYS (*Mdn* = 98.75%, Range = 98.2% - 99.4%). Table 9 shows the gender distribution of these students each year to be relatively equal (*Mdn* = 50.9% male), with the annual rate varying between 50.1% and 51.6% male. The MYS sample reflects a relatively equal gender distribution as well (*Mdn* = 51.4%, Range = 49.5% - 53.0% male). Table 10 shows the free lunch eligibility status distribution of these students. For most years, the vast majority of these students qualify for free or reduced-cost lunch (*Mdn* = 83.9%), with the annual rate varying between 80.2% and 88.4%[29]. Overall the proportion of students receiving free or reduced-cost lunch is slightly less than the proportion of students receiving free or reduced cost lunch in the MYS (*Mdn* = 92.9%, Range = 91.1% - 94.8%). Table 11 shows the mean grade level of these students (*Mdn* = 7.2, Range = 7.2 - 7.3). The MYS sample reflects a relatively equal grade level mean as well (*Mdn* = 7.1, Range = 7.0 - 7.4). Tables 12 and 13 show the mean percentile ranks for reading (SAT_{TR}) and math (SAT_{TM}), respectively. Tables 14 and 15 show the mean weighted scores for school violation scores (WSV) and disciplinary action scores (WDA), respectively.

When the MYS began, the sample consisted of individuals from the 13 poorest neighborhoods in Mobile County (target neighborhoods). As participants moved, the MYS sampling frame expanded to include 35 additional neighborhoods (expansion neighborhoods). As expected, there are small racial and socio-economic differences between target and expansion neighborhoods. The target neighborhoods have higher proportions of African American students (*Mdn* = 98.95%, Range = 98.2% - 99.9%) than do the expansion neighborhoods (*Mdn* = 93.85%,

[29] In Year 4, 67.2% of the MCPSS students ages 10 through 15 living in MYS neighborhoods qualified for free or reduced-cost lunch.

Range = 91.8% - 98.8%). In addition, the target neighborhoods have higher proportions of students qualifying for a free or reduced-cost lunch (*Mdn* = 90.5%, Range = 87.5% - 93.1%) than the students in the expansion neighborhoods (*Mdn* = 82.5%, Range = 76.7% - 87.3%)[30].

In the analyses of enrollment, the proportions of African Americans and those receiving free or reduced-cost lunch reflect those reported in Tables 8 and 10. In the analyses of yearly participation, because analyses were limited to students living in target neighborhoods in Years 1 through 7, the proportions of African Americans and those qualifying for a free or reduced-cost lunch are higher than what is reflected in these tables (see previous paragraph). In the analyses of longitudinal participation and the supplemental analyses (Research Question 2), proportions of demographic characteristics reflect those of the MYS sample. For all analyses, proportions of gender remain stable. As will be discussed, mean grade level may vary from the means presented in Table 11 for Research Question 2 analyses. Further, because only students in grades 3 through 8 are considered in analyses of SAT data, mean grade level for these analyses will differ as well.

Table 6

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods

	Year									
	1	2	3	4	5	6	7	8	9	10
n	11,231	11,039	9,292	12,971	11,084	10,506	10,667	10,633	11,276	9,009
[a]*n**	10,890	10,730	9,087	11,641	10,209	10,213	10,286	10,262	10,264	8,717

[a]Non-repeated cases are indicated here as a point of description, but in this table and all subsequent tables, *n* denotes all cases to be used for analyses, which includes replicates.

[30] In Year 4, 73.6% of the students in the target neighborhoods qualified for free or reduced cost lunch, while 65.2% of the students in the expansion neighborhoods qualified for free or reduced cost lunch.

Table 7

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods by Neighborhood Type

	Year									
	1	2	3	4	5	6	7	8	9	10
Expansion Neighborhood	8,460	8,317	7,014	9,830	8,477	8,145	8,422	8,588	9,180	7,382
Target Neighborhood	2,771	2,722	2,278	3,141	2,607	2,361	2,245	2,045	2,096	1,627

Table 8

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods by Race

	Year									
	1	2	3	4	5	6	7	8	9	10
Non-African Americans	729	695	600	776	627	451	441	465	107	383
African Americans	10,495	10,344	8,683	12,174	10,440	10,046	10,212	10,154	10,733	8,611
Missing Race	7	0	9	21	17	9	14	14	436	15

Table 9

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods by Gender

	Year									
	1	2	3	4	5	6	7	8	9	10
Males	5,652	5,611	4,726	6,684	5,693	5,416	5,470	5,392	5,732	4,509
Females	5,579	5,428	4,566	6,287	5,391	5,090	5,197	5,241	5,544	4,500

Table 10

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods by Free Lunch Eligibility Status

	Year									
	[a]1	2	3	[b]4	5	6	7	8	9	10
Not Eligible	2,298	2,060	1,495	4,251	2,177	1,368	1,406	1,614	1,872	1,049
Reduced	670	262	638	655	635	618	561	564	527	483
Free	8,261	8,717	7,159	8,064	8,272	8,520	8,700	8,455	8,877	7,477

[s]Two students' free lunch eligibility status was unknown in Year 1.
[b]One student's free lunch eligibility status was unknown in Year 4.

Table 11

Mean Grade Level of MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods

	Year									
	1	2	3	4	5	[a]6	7	[b]8	9	10
n	11,231	11,039	9,292	12,971	11,084	10,506	10,667	10,623	11,276	9,009
M	7.27	7.27	7.23	7.16	7.18	7.19	7.19	7.24	7.27	7.33
SD	1.81	1.82	1.87	1.86	1.81	2.03	1.82	1.81	1.80	1.78

[a]One student's grade level was inconclusive and therefore considered missing.
[b]In Year 8, 10 students' grade levels were missing.

Table 12

Mean Percentile Ranks (Total Reading) of MCPSS Students Grades 3 through 8 (Ages 10 through 15) Living in MYS Neighborhoods

	Year									
	1	2	3	4	5	6	7	8	9	10
n	8,357	8,154	7,266	7,410	7,025	6,736	7,041	6,622	7,101	5,937
M	34.14	35.72	34.09	34.68	33.22	34.60	35.74	35.72	38.20	38.84
SD	23.25	23.42	23.46	24.64	24.11	25.04	25.04	25.26	24.70	24.69

Note. It is important to remember that students started taking the SAT 10th edition in the spring of 2003 (Academic Year 5) and caution is given against comparing scores from the 9th and 10th editions of the SAT.

Table 13

Mean Percentile Rank (Total Math) of MCPSS Students Grades 3 through 8 (Ages 10 through 15) Living in MYS Neighborhoods

	Year									
	1	2	3	4	5	6	7	8	9	10
n	8,345	8,161	7,222	7,646	7,103	6,746	7,014	6,576	7,054	5,916
M	39.92	41.96	40.72	38.88	33.94	36.21	39.40	40.87	41.13	41.84
SD	24.32	24.43	24.60	25.60	23.69	24.82	25.63	24.86	25.40	25.26

Note. It is important to remember that students started taking the SAT 10th edition in the spring of 2003 (Academic Year 5) and caution is given against comparing scores from the 9th and 10th editions of the SAT.

Table 14

Mean Weighted School Violation Scores of MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods

	Year									
	1	2	3	4	5	6	7	8	9	10
n	11,231	11,039	9,292	12,971	11,084	10,506	10,667	10,633	11,276	9,009
M	1.68	2.31	2.65	3.02	3.47	4.17	4.52	4.18	5.54	6.45
SD	3.74	4.27	4.68	5.09	5.58	6.30	6.52	6.03	7.26	8.12

Note. Students' violations were weighted and summed.

Table 15

Mean Weighted School Disciplinary Action Scores of MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods

	Year									
	1	2	3	4	5	6	7	8	9	10
n	11,231	11,039	9,292	12,971	11,084	10,506	10,667	10,633	11,276	9,009
M	0.46	2.74	3.02	3.47	4.22	5.20	5.17	4.83	5.62	5.94
SD	1.79	5.63	6.08	6.71	7.37	8.67	8.47	8.15	8.80	9.26

Note. Students' disciplinary actions were weighted and summed.

Research Question 1

Enrollment (E)

The first part of Research Question 1 is concerned with the extent to which those enrolled in the MYS (1998 - 2007) are representative of adolescents living in MYS neighborhoods. Table 16 shows the number of MYS-enrolled and non-enrolled students from the MYS neighborhoods each year. Notable, between 21.1% and 31.7% of all residents aged 10 through 15 were enrolled in the MYS.

Table 16

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods by MYS Enrollment Status

	Year									
	1	2	3	4	5	6	7	8	9	10
E = 1	2,657	2,904	2,447	2,889	3,504	2,936	2,772	2,993	2,282	2,352
E = 0	8,565	8,063	6,836	10,060	7,563	7,561	7,881	7,607	8,548	6,642

Note. E = enrolled. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Demographics. Table 17 shows the goodness of fit statistics (χ^2), significance levels, and effect sizes (**w**) for the GEE results regressing enrollment on demographic variables[31]. A power analysis shows power to detect each of the main effects to be greater than or equal to .995 (**f** = .15, α = .05)[32].

Results demonstrate that there was a significant departure from representativeness in neighborhood type, race, and for free lunch eligibility status in each year and for grade level in nine of the 10 years. Table 18 shows the estimates for those variables that produced statistically significant results. For neighborhood type, race, and free lunch eligibility status, these estimates represent proportions for each category that are enrolled in the MYS. For grade level, the estimate is an odds ratio. As expected, the rates of enrollment in the target neighborhoods are higher than the rates of enrollment in the expansion neighborhoods. Further, the rate of enrollment for African Americans is much higher than the rate of enrollment for non-African

[31] Ideally, the correlation structure in this model is specified as unstructured; however, given the small number of repeated subjects, the SAS procedure would not estimate the model using any correlation structure other than independent. Misspecification of correlation matrix, however, does not bias the estimates because the estimates for the regression coefficients and the estimates for the correlation coefficients are orthogonal (Hardin & Hilbe, 2003). The resulting estimates, while not biased, are less efficient (Diggle, Heagerty, Liang, & Zeger, 2002), however Liang and Zeger (1986) demonstrate that the gains in efficiency are slight. Further, as sample sizes increase, however, the efficiency of the estimators converges (Yang & Tsiatis, 2001).

[32] In this and subsequent power analyses, a GLM framework is used rather than a GEE framework for the purposes of simplification. Cohen (1988) suggests **f** as a measure of effect size with small, medium, and large effect sizes corresponding to **f** = .10, **f** = .25, **f** = .4, respectively. Extrapolating from corresponding values for **w**, it is estimated that an effect of **f** = .15 is approximately equivalent to a small-to-moderate effect size. Cohen further shows that **f** = .15 corresponds to η^2 = .022.

Americans. Also, the rate of enrollment for those eligible for free lunch are higher than the rate of enrollment for both those who are eligible for reduced lunch and those who are eligible for neither. While these differences in rates of enrollment are expected, they do suggest that the sample is not representative of the population. However, these differences may not be indicative of non-ignorable missingness. Finally, the odds ratios for grade level in Year 1 through Year 5 are less than 1.0, indicating a negative relationship between enrollment and grade level. In contrast, the odds ratios for grade level in Year 6 through Year 10 are greater than 1.0, indicating a positive relationship between enrollment and grade level. That is, during the earlier waves of MYS data collection, younger students were enrolled at disproportionately high rates, while during the later waves, older students were enrolled at disproportionately high rates. The inconsistency in direction in the results raises questions about the meaning of these findings as they relates to ignorable missingness.

In order for statistically significant results to be meaningful, effect sizes should consistently be at least small-to-moderate ($\mathbf{w} \geq .15$). However, in these analyses, effect sizes are small: (a) for neighborhood type: $.03 \leq \mathbf{w} \leq .05$; (b) for race: $.08 \leq \mathbf{w} \leq .12$; (c) for lunch eligibility: $.07 \leq \mathbf{w} \leq .11$; (d) grade: $.01 \leq \mathbf{w} \leq .12$; (e) and for gender: $\mathbf{w} \leq .03$.

Results also demonstrate that there was a significant departure from representativeness in gender for Years 1 and 3, with males having higher rates of enrollment than females, with very small effect sizes ($\mathbf{w} = .03$). Finally, in Year 4, there was a significant free lunch status × neighborhood type[33]. Figure 1 depicts this interaction. This figure shows that the biggest discrepancy in enrollment status occurs for those with free lunch. Specifically, rates of enrollment for those living in expansion neighborhoods are relatively stable, no matter what the status of free lunch eligibility status. For those living in the target neighborhoods, rates of

[33] For this result, because it only occurred in one year, estimates are not reported in any table.

enrollment were higher for those qualifying for free lunch, while those who did not qualify for free lunch were relatively similar. Because this interaction is significant in only one year and because the effect size is small, it is treated as an anomaly. That is, threat of non-representativeness across years falls within the range of statistical chance (based on the criteria specified in the previous chapter). In this case, this statistically significant interaction does not undermine overall representativeness (or missingness being non-random) because of the one occurrence in 10 years and the low effect size.

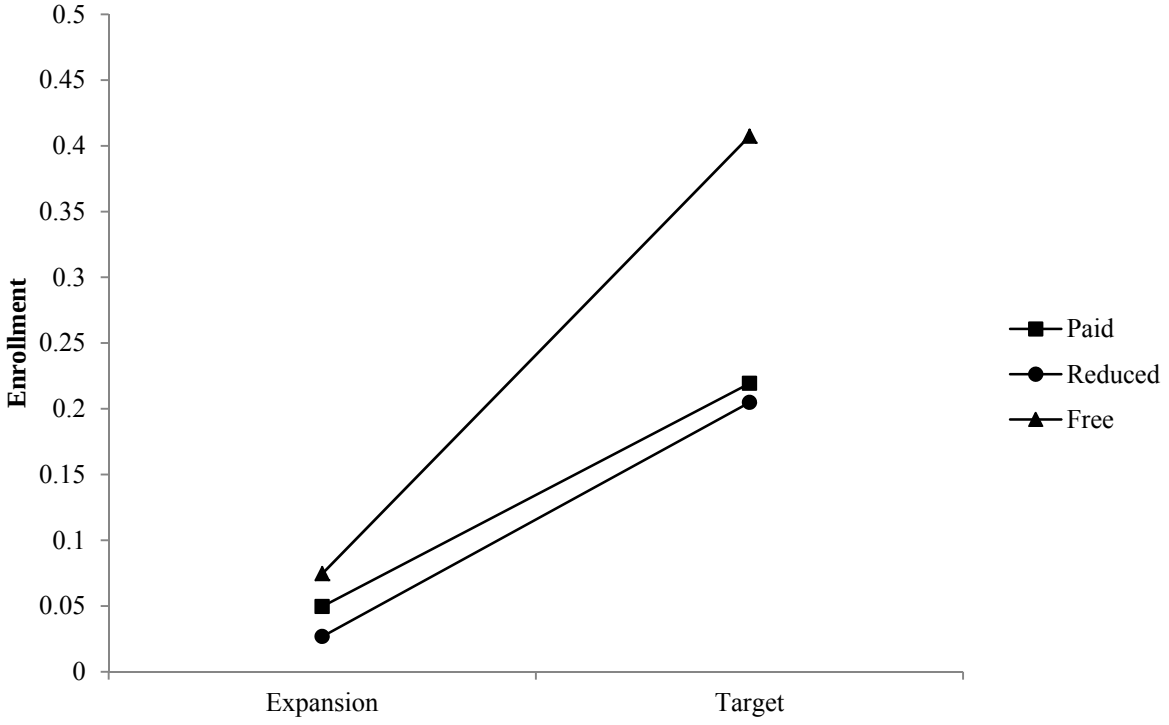

Figure 1. Estimates for Demographic Analysis for Year 4 Enrollment: Neighborhood Type × Lunch Eligibility

It is important to also consider the main effects and interactions that are not statistically significant. There were no consistent departures from representativeness (in three or more years) for gender or either interaction effect (free lunch status × neighborhood type and gender ×

neighborhood type) included in this analysis of demographic representativeness. That is, there are no consistent differences in the rates of enrollment for males and females. Next, while there were significant departures from representativeness for free lunch eligibility status, when included as a component of interaction with neighborhood type, there are no statistically significant results. This indicates that across neighborhood types, rates of enrollment for students qualifying (or not) for free lunch are similar; and this interaction does not contribute anything beyond the main effects. Finally, while there were statistically significant differences in neighborhood type, the rates of enrollment for males were similar in both neighborhood types, and this interaction does not contribute anything beyond the main effects. This lack of significance does suggest that the MYS sample is representative of the population on these characteristics and fails to suggest that missingness is non-ignorable.

Table 17

Significance and Effect Size of Demographic Characteristics on MYS Enrollment

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
G	1	8.65	.003	.03	2.51	.113	.02	8.97	.003	.03	1.80	.179	.01	0.20	.656	.00
R	1	79.99	<.001	.08	82.42	<.001	.09	73.55	<.001	.09	113.08	<.001	.09	140.78	<.001	.11
LE	2	63.98	<.001	.08	50.59	<.001	.07	90.60	<.001	.10	161.94	<.001	.11	95.55	<.001	.09
NT	1	11.01	.001	.03	9.01	.003	.03	12.04	.001	.04	10.99	<.001	.03	13.16	<.001	.03
LE × NT	2	0.22	.895	.00	1.91	.384	.01	2.85	.241	.02	14.91	<.001	.03	5.04	.081	.02
G × NT	1	0.23	.632	.00	0.43	.512	.01	0.12	.734	.00	1.07	.301	.01	0.88	.349	.01
Gr	1	104.21	<.001	.10	110.03	<.001	.10	35.69	<.001	.06	0.86	.354	.01	10.58	.001	.03
N	[a]46	507.07	<.001	.21	527.51	<.001	.22	405.72	<.001	.21	367.12	<.001	.17	467.20	<.001	.21

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
G	1	2.74	.098	.02	3.42	.065	.02	0.12	.731	.00	0.87	.350	.01	0.67	.413	.01
R	1	97.57	<.001	.10	127.39	<.001	.11	163.62	<.001	.12	40.78	<.001	.06	126.00	<.001	.12
LE	2	52.57	<.001	.07	46.87	<.001	.07	124.99	<.001	.11	53.99	<.001	.07	91.19	<.001	.10
NT	1	15.16	<.001	.04	8.03	.005	.03	14.58	<.001	.04	14.50	<.001	.04	6.01	.014	.03
LE × NT	2	0.27	.874	.01	2.37	.306	.01	2.65	.266	.02	1.34	.511	.01	4.21	.122	.02
G × NT	1	1.19	.275	.01	0.88	.348	.01	0.24	.626	.00	0.42	.519	.01	0.56	.453	.01
Gr	1	25.42	<.001	.05	55.07	<.001	.07	21.54	<.001	.05	165.89	<.001	.12	65.33	<.001	.09
N	46	448.87	<.001	.21	514.76	<.001	.22	430.98	<.001	.20	388.69	<.001	.19	463.43	<.001	.23

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. G = gender; R = race; LE = free lunch eligibility status; NT = neighborhood type; Gr = grade level, N = neighborhood.

[a]Because not all neighborhoods existed at all times for all analyses, *df* for N is inconsistent across years. In Years 2, 5, 8, 9, *df* was 45.

Table 18

Adjusted Estimates and Standard Errors for Selected Demographic Characteristics on MYS Enrollment

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Expansion Neighborhood	***0.03***	0.14	***0.04***	0.15	***0.04***	0.12	***0.05***	0.11	***0.07***	0.11
Target Neighborhood	***0.36***	0.19	***0.43***	0.20	***0.40***	0.17	***0.27***	0.16	***0.36***	0.18
Non-African American	***0.06***	0.25	***0.08***	0.24	***0.08***	0.22	***0.06***	0.21	***0.08***	0.22
African American	***0.24***	0.08	***0.29***	0.11	***0.27***	0.08	***0.23***	0.07	***0.34***	0.08
Not Eligible	***0.12***	0.14	***0.13***	0.14	***0.13***	0.14	***0.11***	0.11	***0.16***	0.13
Reduced Lunch	***0.08***	0.24	***0.12***	0.30	***0.10***	0.22	***0.08***	0.22	***0.11***	0.22
Free Lunch	***0.20***	0.13	***0.22***	0.12	***0.26***	0.11	***0.19***	0.10	***0.27***	0.12
Grade Level	***0.85***	0.02	***0.85***	0.02	***0.90***	0.02	0.99	0.01	***0.95***	0.01

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Expansion Neighborhood	***0.06***	0.12	***0.05***	0.13	***0.06***	0.14	***0.04***	0.36	***0.04***	0.19
Target Neighborhood	***0.36***	0.17	***0.26***	0.21	***0.22***	0.18	***0.19***	0.38	***0.17***	0.25
Non-African American	***0.08***	0.23	***0.05***	0.25	***0.05***	0.27	***0.03***	0.71	***0.03***	0.38
African American	***0.29***	0.08	***0.26***	0.09	***0.28***	0.07	***0.22***	0.08	***0.23***	0.10
Not Eligible	***0.15***	0.15	***0.13***	0.15	***0.10***	0.16	***0.07***	0.36	***0.08***	0.22
Reduced Lunch	***0.11***	0.21	***0.07***	0.26	***0.08***	0.21	***0.07***	0.40	***0.04***	0.30
Free Lunch	***0.24***	0.11	***0.19***	0.13	***0.21***	0.14	***0.13***	0.35	***0.17***	0.19
Grade Level	***1.06***	0.02	***1.11***	0.01	***1.06***	0.01	***1.22***	0.02	***1.14***	0.02

Note. Boldfaced italicized estimates are statistically significant, $p < .05$.

SAT percentile ranks. Table 19 shows the number of students included in these analyses: those aged 10 through 15 who are in grades 3 through 8 and have either a total reading percentile rank (SAT_{TR}) or a total math percentile rank (SAT_{TM}).

Table 19

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods with SAT Data

	Year									
	1	2	3	4	5	6	7	8	9	10
With SAT_{TR} Percentile Ranks	6,066	5,768	5,227	7,380	6,996	6,730	7,036	6,606	6,853	5,927
With SAT_{TM} Percentile Ranks	6,074	5,786	5,221	7,615	7,074	6,740	7,009	6,560	6,812	5,906

Note. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Tables 20 and 22 show goodness of fit statistics (χ^2), significance levels, and effect sizes (**w**) for the GEE results regressing SAT_{TR} and SAT_{TM}, respectively, on enrollment (main effects and interactions) controlling for demographic characteristics. A power analysis shows power to detect each of the main effects to be greater than or equal to .995 (**f** = .15, α = .05) for both cognitive ability analyses.

Results demonstrate a significant difference in Year 6 in SAT_{TR} and SAT_{TM} for those enrolled and not enrolled, with those not enrolled scoring higher on each subtest than those who were enrolled[34]. While these results are consistent across measures, they occur only in one year and effect sizes are very small (**w** ≤ .05); therefore, based on criteria described in the previous chapter, these results do not undermine the representativeness of the sample on cognitive ability. Results also demonstrate a significant enrollment × neighborhood type interaction for SAT_{TM} percentile ranks in Year 2[35]. For this interaction, because the statistical significance occurred only in one year, for one measure, and because effect size was small (**w** = .04), further

[34] For these results, because they only occurred in one year, estimates are not reported in any table.
[35] For this result, because it only occurred in one year, estimates are not reported in any table.

exploration of the school context in Year 2 may be warranted, but overall representativeness is not undermined and missingness may be treated as ignorable. More consistent departures from representativeness are evident in the enrollment × free lunch eligibility status interaction for SAT_{TR} and the enrollment × gender × neighborhood type interaction for SAT_{TM}.

Tables 21 and 23 show the estimates for those variables that produced statistically significant results for the SAT_{TR} and SAT_{TM} analyses, respectively. First, regarding the enrollment × free lunch eligibility status interaction, Table 21 shows very different patterns of interaction with Years 1 and 3 producing similar results and Years 2 and 6 producing similar results. Figures 2a through d show the interactions between enrollment × free lunch eligibility status for SAT_{TR} for Years 1, 2, 3 and 6. The divergence is clear from these figures. For example, in Years 1 and 2, for those receiving free lunch, non-enrollees have higher percentile ranks for SAT_{TR} than enrollees, but, the divergence is more substantial in Year 2. However, the relationship between lunch status and enrollment is reversed for the reduced-cost lunch and not eligible for free lunch categories in the Years 1 and 2. Years 1 and 3 follow a similar pattern and Years 2 and 6 follow a similar pattern. These inconsistencies in functional form, coupled with the lack of significance across measures and the small effect sizes, fail to suggest non-ignorable missing data mechanisms. Further support for the lack of study-wise practical significance is gained through an examination of Figure 2e, the average of weighted yearly estimates for enrollment × free lunch eligibility status for SAT_{TR}. Differences between those enrolled and not enrolled who receive free lunch or reduced lunch are less than eight points. Only in the case of those receiving no federal assistance for lunch do the differences for those enrolled and not enrolled exceed 10 points.

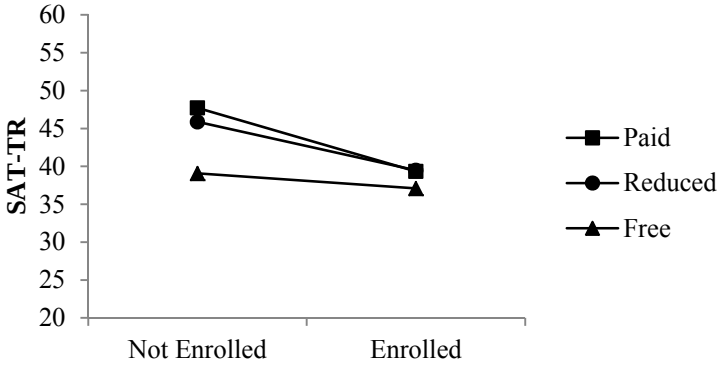

Figure 2a. Estimates for SAT_{TR} Analysis for Year 1: Enrollment × Lunch Eligibility

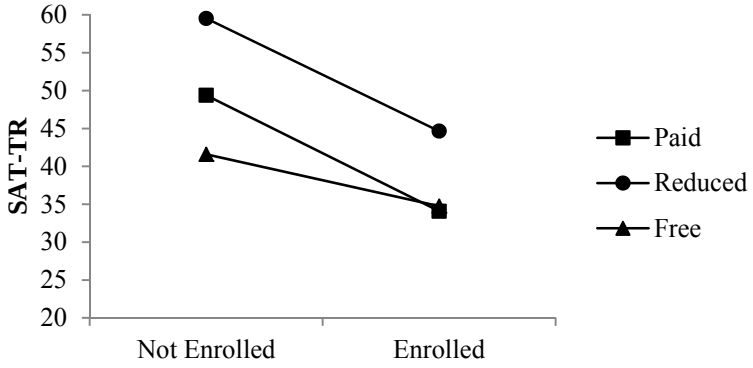

Figure 2b. Estimates for SAT_{TR} Analysis for Year 2: Enrollment × Lunch Eligibility

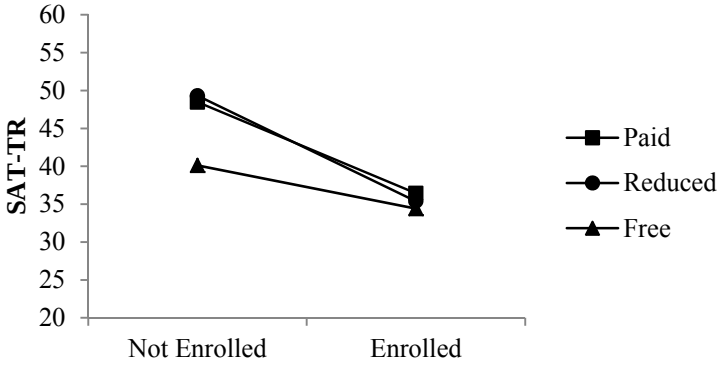

Figure 2c. Estimates for SAT_{TR} Analysis for Year 3: Enrollment × Lunch Eligibility

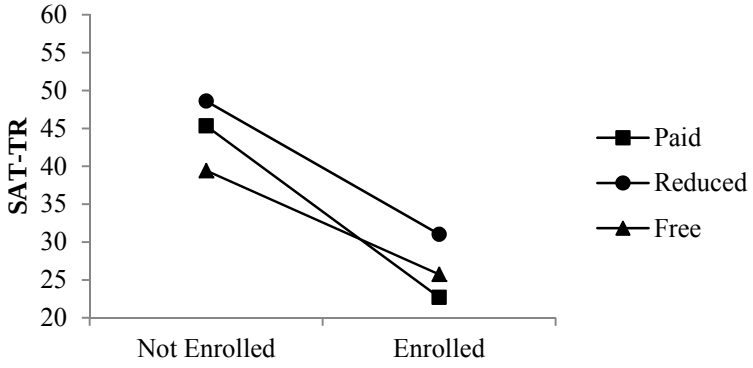

Figure 2d. Estimates for SAT_{TR} Analysis for Year 6: Enrollment × Lunch Eligibility

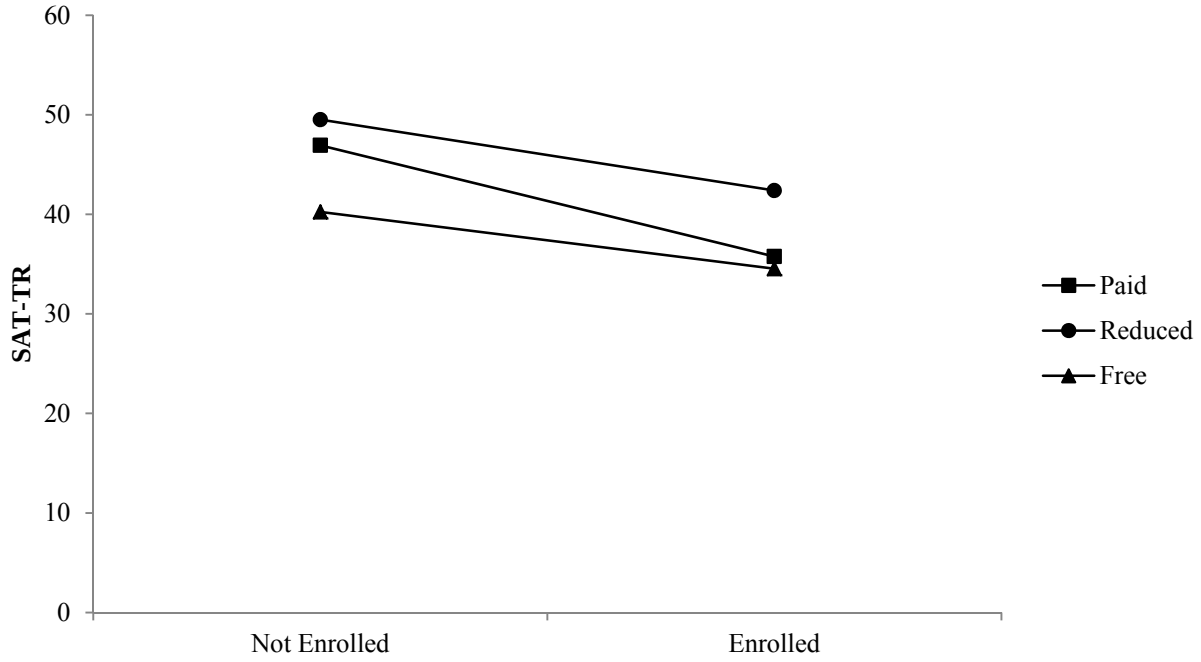

Figure 2e. Weighted Estimates for SAT_{TR} Analysis for All Years: Enrollment × Lunch Eligibility

For SAT_{TM} percentile ranks, the three significant results (Years 2, 3, and 4) suggest little difference between enrolled males and un-enrolled males in the target neighborhoods, but in the expansion neighborhoods, the enrolled males have worse SAT_{TM} percentile ranks than the un-enrolled (Figures 3a and b). The trend for each of the significant years is consistent. While enrollees have lower percentile ranks than non-enrollees, the results were only significant in three years. Differences seemed to localize in three years for males in expansion neighborhoods compared with males in target neighborhoods. These results, while statistically significant in three or more years and consistent in functional form, were not consistent across measures of cognitive ability. Further, effect sizes in these years were $\leq .05$, suggesting that these deviations from representativeness are not meaningful in terms of non-ignorable missingness.

While it is important to describe the statistically significant results, it is also important to note the lack of consistent significance in enrollment or many of its interactions. There is no consistent statistical significance in any interaction across measures (SAT_{TR} and SAT_{TM}) and across years. This lack of statistical significance suggests that those enrolled in the MYS are representative of the population on cognitive characteristics and that missingness may be ignorable.

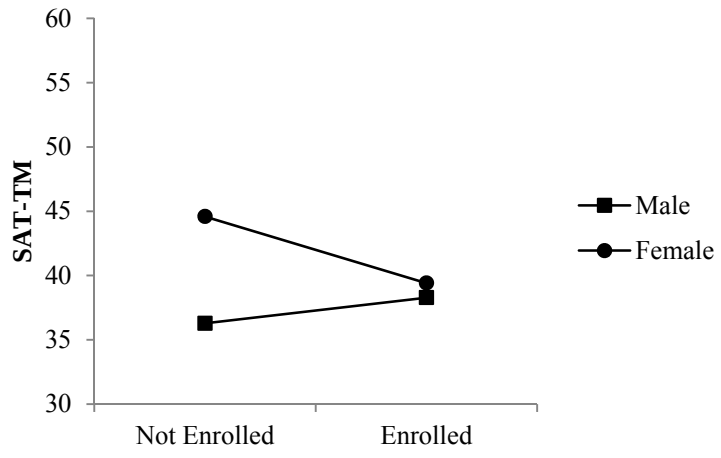

Figure 3a. Estimates for SAT_{TM} Analysis for Year 3: Enrollment × Gender in Expansion Neighborhoods

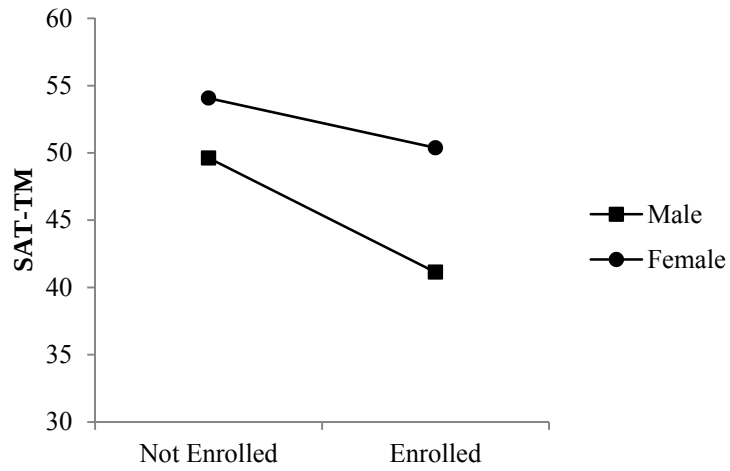

Figure 3b. Estimates for SAT_{TM} Analysis for Year 3: Enrollment × Gender in Target Neighborhoods

Table 20

Significance and Effect Size of MYS Enrollment and Demographic Characteristics on Total Reading Percentile Ranks

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
E	1	0.26	.607	.01	3.70	.054	.03	2.98	.085	.02	0.02	.896	.00	0.13	.721	.00
NT	1	9.37	.002	.04	0.60	.439	.01	0.43	.512	.01	0.05	.820	.00	0.28	.597	.01
Gr	1	3.43	.064	.02	0.02	.882	.00	4.55	.033	.03	1.22	.269	.01	5.89	.015	.03
G	1	26.46	<.001	.07	35.36	<.001	.08	64.58	<.001	.11	77.37	<.001	.10	80.68	<.001	.11
R	1	4.88	.027	.03	6.47	.011	.03	8.09	.005	.04	5.05	.025	.03	3.74	.053	.02
LE	2	18.02	.000	.05	8.81	.012	.04	10.48	.005	.04	11.77	.003	.04	30.86	.001	.07
N	[a] 46	96.23	<.001	.13	115.57	<.001	.14	155.08	<.001	.17	172.51	<.001	.15	98.60	<.001	.12
LE × NT	2	2.56	.278	.02	0.89	.642	.01	2.89	.235	.02	3.01	.222	.02	2.96	.228	.02
G × NT	1	0.41	.522	.01	1.21	.271	.01	3.10	.078	.02	4.23	.040	.02	4.87	.027	.03
E × G	1	1.93	.164	.02	2.07	.150	.02	0.30	.584	.01	0.02	.902	.00	0.06	.802	.00
E × Gr	1	0.55	.458	.01	0.03	.872	.00	0.17	.678	.01	6.04	.014	.03	4.15	.042	.02
E × LE	2	6.61	.037	.03	10.07	.007	.04	6.35	.042	.03	1.67	.435	.02	1.38	.501	.01
E × NT	1	1.40	.236	.02	3.45	.063	.02	0.27	.602	.01	0.01	.930	.00	0.03	.874	.00
E × R	1	0.31	.581	.01	1.27	.260	.01	0.08	.776	.00	0.30	.583	.01	0.93	.334	.01
E × NT × Gr	2	4.41	.110	.03	5.74	.057	.03	0.18	.912	.01	3.26	.196	.02	0.33	.850	.01
E × G × NT	1	1.32	.251	.01	0.05	.823	.00	7.05	.008	.04	4.38	.036	.02	2.18	.140	.02

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
E	1	17.29	<.001	.05	2.27	.132	.02	0.17	.679	.01	1.55	.213	.02	1.77	.183	.02
NT	1	0.00	.971	.00	0.23	.633	.01	0.63	.427	.01	0.75	.388	.01	0.75	.386	.01
Gr	1	0.43	.513	.01	0.48	.488	.01	5.19	.023	.03	5.78	.016	.03	17.88	<.001	.05
G	1	58.05	<.001	.09	69.16	<.001	.10	72.17	<.001	.10	69.49	<.001	.10	74.54	<.001	.11
R	1	0.53	.465	.01	4.41	.036	.03	4.55	.033	.03	1.62	.204	.02	3.46	.063	.02
LE	2	8.30	.016	.04	24.20	<.001	.06	13.31	.001	.04	6.92	.032	.03	12.32	.002	.05
N	46	188.66	<.001	.17	102.66	<.001	.12	153.93	<.001	.15	83.56	<.001	.11	93.43	<.001	.13
LE × NT	2	6.85	.033	.03	2.84	.241	.02	3.49	.175	.02	6.20	.045	.03	0.41	.816	.01
G × NT	1	0.53	.467	.01	1.72	.189	.02	0.26	.610	.01	0.12	.734	.00	2.61	.106	.02
E × G	1	2.54	.111	.02	2.27	.132	.02	0.06	.802	.00	0.73	.392	.01	0.03	.868	.00
E × Gr	1	2.29	.130	.02	0.25	.614	.01	0.14	.708	.00	3.82	.051	.02	0.02	.882	.00
E × LE	2	8.31	.016	.04	5.46	.065	.03	4.77	.092	.03	0.46	.795	.01	4.07	.131	.03
E × NT	1	0.27	.606	.01	0.15	.697	.00	3.49	.062	.02	0.53	.468	.01	0.03	.872	.00
E × R	1	6.79	.009	.03	0.91	.339	.01	0.01	.917	.00	0.90	.342	.01	0.22	.640	.01
E × NT × Gr	2	1.52	.469	.02	1.47	.481	.01	5.39	.067	.03	0.70	.705	.01	0.23	.893	.01
E × G × NT	1	1.47	.226	.01	3.30	.069	.02	0.74	.389	.01	0.34	.558	.01	0.13	.719	.00

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. E = enroll; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status; N = neighborhood.

[a]Because not all neighborhoods existed at all times for all analyses, *df* for N is inconsistent across years. In Years 2, 5, 8, and 9, *df* was 45.

Table 21

Adjusted Estimates and Standard Errors for MYS Enrollment × Free Lunch Eligibility Status on Total Reading Percentile Ranks

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Not Enrolled, Not Eligible	***47.72***	1.56	***49.39***	1.70	***48.47***	1.89	42.30	1.09	43.66	1.54
Not Enrolled, Reduced Lunch	***45.87***	2.41	***59.54***	6.67	***49.32***	2.60	45.15	2.72	49.15	2.94
Not Enrolled, Free Lunch	***39.07***	0.98	***41.60***	1.01	***40.12***	1.02	38.82	0.99	37.70	1.09
Enrolled, Not Eligible	***39.34***	4.03	***34.09***	3.29	***36.43***	3.89	34.46	4.02	33.16	3.33
Enrolled, Reduced Lunch	***39.48***	5.75	***44.66***	9.36	***35.39***	5.64	37.27	5.95	41.96	4.87
Enrolled, Free Lunch	***37.09***	3.45	***34.75***	2.42	***34.41***	2.79	33.11	3.93	29.42	3.05

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Not Enrolled, Not Eligible	***45.33***	1.92	52.44	2.00	44.28	2.04	43.94	2.43	53.82	2.46
Not Enrolled, Reduced Lunch	***48.61***	2.31	54.35	2.82	46.97	2.48	47.44	3.69	50.11	2.79
Not Enrolled, Free Lunch	***39.43***	1.10	42.72	1.15	39.46	1.19	41.94	2.13	42.02	1.21
Enrolled, Not Eligible	***22.70***	3.13	36.55	4.56	35.61	4.22	47.11	6.33	38.84	5.11
Enrolled, Reduced Lunch	***31.02***	5.02	46.92	5.81	47.47	5.04	53.94	8.51	44.85	5.81
Enrolled, Free Lunch	***25.73***	2.35	34.22	3.71	36.53	3.51	44.31	5.91	36.55	3.00

Note. Boldfaced italicized estimates are statistically significant, $p < .05$.

Table 22

Significance and Effect Size for MYS Enrollment and Demographic Characteristics on Total Math Percentile Ranks

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
E	1	1.87	.172	.02	1.78	.182	.02	0.27	.604	.01	0.10	.753	.00	0.01	.936	.00
NT	1	8.48	.004	.04	3.95	.047	.03	0.02	.893	.00	0.59	.443	.01	0.04	.843	.00
Gr	1	19.10	<.001	.06	27.97	<.001	.07	54.56	<.001	.10	37.43	<.001	.07	1.69	.194	.02
G	1	26.90	<.001	.07	23.11	<.001	.06	46.81	<.001	.09	53.20	<.001	.08	28.06	<.001	.06
R	1	1.87	.171	.02	6.00	.014	.03	6.74	.009	.04	2.76	.097	.02	2.65	.104	.02
LE	2	7.90	.019	.04	1.78	.412	.02	9.60	.008	.04	8.39	.015	.03	28.99	<.001	.06
N	[a]46	96.69	<.001	.13	115.38	<.001	.14	142.43	<.001	.17	163.14	<.001	.15	138.59	<.001	.14
LE × NT	2	5.66	.059	.03	9.25	.010	.04	3.02	.221	.02	7.48	.024	.03	0.02	.992	.00
G × NT	1	0.16	.691	.01	0.46	.498	.01	1.72	.190	.02	3.30	.069	.02	1.29	.255	.01
E × G	1	0.90	.342	.01	1.65	.199	.02	0.51	.476	.01	0.16	.689	.00	0.12	.727	.00
E × Gr	1	0.03	.858	.00	3.29	.070	.02	0.04	.843	.00	6.28	.012	.03	10.28	.001	.04
E × LE	2	7.77	.021	.04	4.14	.126	.03	1.07	.586	.01	0.73	.693	.01	1.35	.508	.01
E × NT	1	0.00	.974	.00	7.45	.006	.04	0.55	.457	.01	0.06	.804	.00	0.82	.364	.01
E × R	1	0.01	.904	.00	0.03	.857	.00	0.25	.615	.01	0.28	.594	.01	4.86	.028	.03
E × NT × Gr	2	2.19	.334	.02	10.44	.005	.04	4.46	.108	.03	7.99	.018	.03	2.66	.265	.02
E × G × NT	1	.23	.628	.01	4.32	.038	.03	13.37	<.001	.05	6.85	.009	.03	0.19	.661	.01

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
E	1	13.25	<.001	.04	0.60	.437	.01	0.88	.349	.01	0.26	.609	.01	0.04	.834	.00
NT	1	0.38	.536	.01	0.78	.379	.00	0.06	.806	.00	1.81	.179	.02	0.00	.998	.00
Gr	1	11.28	.001	.04	23.15	<.001	.09	52.27	<.001	.09	21.35	<.001	.06	26.23	<.001	.07
G	1	41.17	<.001	.08	34.74	<.001	.08	40.82	<.001	.08	46.48	<.001	.08	46.21	<.001	.09
R	1	1.02	.313	.01	2.22	.137	.02	2.11	.147	.02	0.04	.840	.00	0.83	.363	.01
LE	2	4.96	.084	.03	8.84	.012	.04	11.88	.003	.04	8.68	.013	.04	10.94	.004	.04
N	46	142.33	<.001	.15	118.33	<.001	.13	144.05	<.001	.15	113.96	<.001	.13	170.33	<.001	.17
LE × NT	2	5.85	.054	.03	3.25	.197	.04	10.83	.005	.04	3.21	.201	.02	0.23	.894	.01
G × NT	1	1.43	.232	.01	0.27	.607	.00	0.08	.779	.00	0.00	.951	.00	4.70	.030	.03
E × G	1	0.01	.929	.00	0.12	.732	.00	0.86	.353	.01	3.56	.059	.02	0.00	.973	.00
E × Gr	1	1.34	.247	.01	1.15	.283	.02	3.40	.065	.02	3.71	.054	.02	1.98	.160	.02
E × LE	2	8.69	.013	.04	7.08	.029	.02	1.99	.370	.02	0.07	.968	.00	1.43	.490	.02
E × NT	1	3.45	.063	.02	3.90	.048	.02	1.66	.198	.02	0.24	.622	.01	0.82	.365	.01
E × R	1	4.08	.043	.02	0.58	.448	.01	0.00	.950	.00	1.35	.246	.01	0.25	.619	.01
E × NT × Gr	2	4.67	.097	.03	3.02	.221	.03	5.87	.053	.03	2.61	.272	.02	1.22	.544	.01
E × G × NT	1	0.04	.841	.00	1.46	.227	.01	0.75	.388	.01	1.91	.167	.02	0.33	.565	.01

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. E = enroll; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status; N = neighborhood.

[a]Because not all neighborhoods existed at all times for all analyses, *df* for N is inconsistent across years. In Years 2, 5, 8, 9, *df* was 45.

Table 23

Adjusted Estimates and Standard Errors for MYS Enrollment × Gender × Neighborhood Type on Total Math Percentile Ranks

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Not Enrolled, Male, Expansion Neighborhood	48.64	1.04	***51.25***	1.17	***49.62***	1.09	***46.21***	1.01	42.91	1.03
Not Enrolled, Male, Target Neighborhood	39.19	2.46	***36.03***	3.64	***36.28***	2.62	***38.64***	2.08	39.69	2.54
Not Enrolled, Female, Expansion Neighborhood	52.88	1.02	***54.02***	1.16	***54.07***	1.08	***50.92***	1.00	47.48	1.04
Not Enrolled, Female, Target Neighborhood	44.78	2.52	***43.16***	3.50	***44.59***	2.48	***44.58***	2.17	42.14	2.57
Enrolled, Male, Expansion Neighborhood	40.01	4.59	***48.20***	4.50	***41.13***	3.98	***36.28***	4.57	31.36	2.79
Enrolled, Male, Target Neighborhood	33.98	4.72	***39.76***	4.28	***38.28***	4.27	***34.94***	4.69	29.00	3.22
Enrolled, Female, Expansion Neighborhood	43.48	4.60	***52.19***	4.51	***50.38***	3.98	***45.30***	4.56	35.83	2.81
Enrolled, Female, Target Neighborhood	37.32	4.74	***41.54***	4.24	***39.41***	4.20	***37.68***	4.61	32.53	3.29

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Not Enrolled, Male, Expansion Neighborhood	45.12	1.17	47.25	1.23	47.18	1.19	48.13	2.32	48.72	1.28
Not Enrolled, Male, Target Neighborhood	40.86	2.31	46.64	2.71	39.34	2.28	47.77	3.23	42.03	2.84
Not Enrolled, Female, Expansion Neighborhood	49.24	1.15	52.90	1.23	52.49	1.18	53.80	2.31	53.37	1.28
Not Enrolled, Female, Target Neighborhood	46.42	2.27	49.69	2.75	42.89	2.42	51.06	3.24	49.37	2.82
Enrolled, Male, Expansion Neighborhood	28.70	3.05	36.65	4.23	43.07	4.48	32.51	5.14	43.38	6.09
Enrolled, Male, Target Neighborhood	25.98	3.27	38.64	4.52	39.43	4.74	33.15	5.42	38.11	6.32
Enrolled, Female, Expansion Neighborhood	32.40	3.07	41.01	4.19	48.46	4.56	39.29	5.23	47.11	6.12
Enrolled, Female, Target Neighborhood	31.69	3.36	44.04	4.56	45.72	4.75	42.10	5.36	46.50	6.45

Note. Boldfaced italicized estimates are statistically significant, $p < .05$.

School violations and disciplinary actions. Table 24 shows the number of students included in analyses comparing the weighted school violation scores (WSV) of and weighted disciplinary action scores (WDA) for those enrolled and those not enrolled in the MYS. Disciplinary actions analyses for Year 1 were omitted because there were not sufficient data available for the model to converge[36].

Table 24

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods with School Violation and Disciplinary Action Data

	Year									
	1	2	3	4	5	6	7	8	9	10
n	11,222	10,967	9,283	12,949	11,067	10,497	10,653	10,600	10,830	8,994

Note. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Tables 25 and 27 show the goodness of fit statistics (χ^2), significance levels, and effect sizes (**w**) for the GEE results regressing WSV and WDA scores, respectively, on enrollment, controlling for demographic characteristics. A power analysis shows power to detect each of the main effects to be greater than or equal to .995 (**f** = .15, α = .05) for both analyses of behavior.

Results demonstrate that there was a significant difference in WSV and in WDA for the enrollment × race interaction in Year 2 and a significant enrollment × grade interaction for WDA in Year 9[37]. Based on the criteria discussed in the previous chapter, while these results are statistically significant, they do not suggest a conclusion of non-ignorable missingness and do not undermine the overall representativeness of the sample on behavioral characteristics. More consistent departures from representativeness are evident in WSV scores for those who were enrolled and those who were not enrolled. Further, more consistent departures from

[36] Variables could have been systematically removed from the model to force convergence, but then results for Year 1 would not have been comparable to other years.
[37] For these results, because they only occurred in one year, estimates are not reported in any table.

representativeness are evident in the enrollment × free lunch eligibility status interaction for both WSV and WDA and for grade × enrollment × neighborhood type for WSV.

Tables 26 and 28 show the estimates for those variables that produced statistically significant results for the WSV and WDA analyses, respectively. First, considering WSV significance, for enrollment and the enrollment × free lunch eligibility status interaction, these estimates represent mean scores for each category. For the grade × enrollment × neighborhood type interaction, the estimate is an odds ratio. For the enrollment main effect, those who are enrolled have a consistently higher WSV score than those who are not enrolled. While this main effect was significant in three years (Years 5, 7, and 10), statistical significance was not found across measures and effect sizes are small ($\mathbf{w} \leq .03$); thus, despite statistical significance, there appears to be little practical meaning and missingness may well be ignorable.

In Years 2, 4, and 5, there was a significant interaction for grade level × enrollment × neighborhood type for weighted WSV scores. Odds ratios were positive across all conditions in these years, indicating that the WSV scores increase as grade level increases. The effect of neighborhood type and enrollment, however, on this relationship, was not consistent across significant years. In Years 2 and 5, the relationship is strongest for those enrolled, regardless of neighborhood type. In Year 4, however, in the expansion neighborhoods, the relationship is strongest for those who are enrolled, while in the target neighborhoods, the relationship is strongest for those who are not enrolled. Given the number of years that demonstrated significant effects, this result should be ignored and may have implications for the representativeness of the sample on behavioral characteristics. However, the small effect sizes ($.03 \leq \mathbf{w} \leq .04$), lack of consistency in functional form, and lack of statistical significance across measures of behavior suggest that the threat to representativeness is small.

Considering the interaction of enrollment × free lunch eligibility status for WSV, there were significant results in Years 2, 5, and 7. Figures 4a, b, and c depict this interaction. In Years 5 and 7, there is a similar pattern where those not qualifying for free or reduced-cost lunch and who were enrolled had higher WSV scores. Those enrolled who qualified for reduced-cost lunch had the lowest WSV scores. When considering those not enrolled, however, those qualifying for free lunch had higher WSV scores than those who did not qualify for a reduction in cost of lunch. In each category of free lunch eligibility status, those who were enrolled had higher WSV scores than those not enrolled. In Year 2, however, although statistically significant, the WSV scores were relatively similar for those enrolled and those not enrolled in the MYS. Even though this interaction was statistically significant in both measures of behavior, there is little consistency across years and all effect sizes are very small ($.03 \leq \mathbf{w} \leq .04$); thus, there is little concern for the presence of a non-ignorable missing data mechanism. Further support for the lack of study-wise practical significance is shown by Figure 4d, the average of weighted yearly estimates for enrollment × free lunch eligibility status for WSV. When examining the average of yearly estimates for this interaction, those not enrolled in the MYS always had lower WSV scores than those enrolled, by about two points. The differences between categories of free lunch eligibility status are less than 1.0 points for those not enrolled and less than 1.2 points for those enrolled in the MYS.

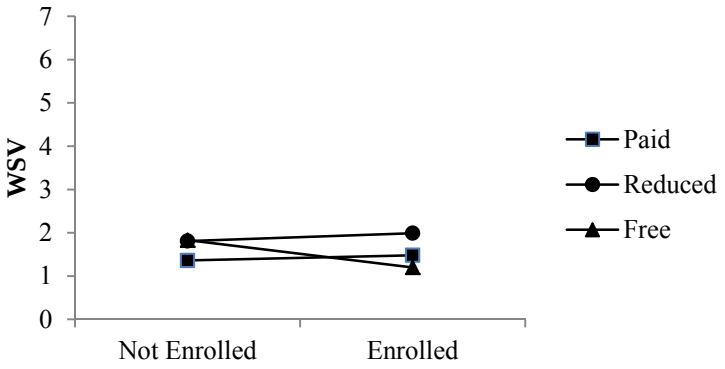

Figure 4a. Estimates for WSV Analysis for Year 2: Enrollment × Lunch Eligibility

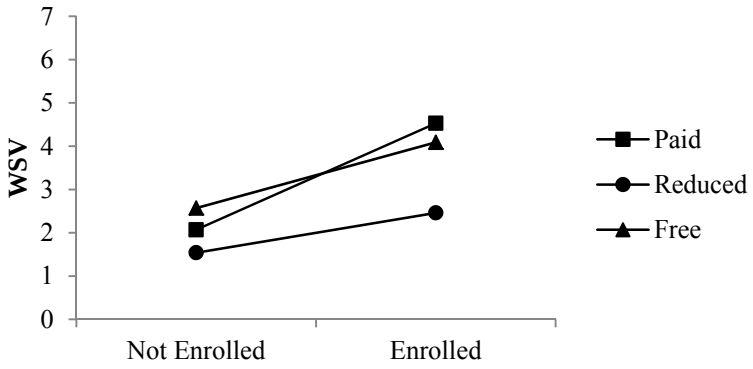

Figure 4b. Estimates for WSV Analysis for Year 5: Enrollment × Lunch Eligibility

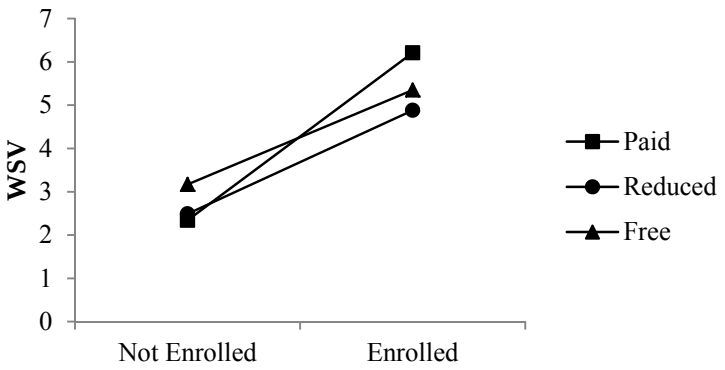

Figure 4c. Estimates for WSV Analysis for Year 7: Enrollment × Lunch Eligibility

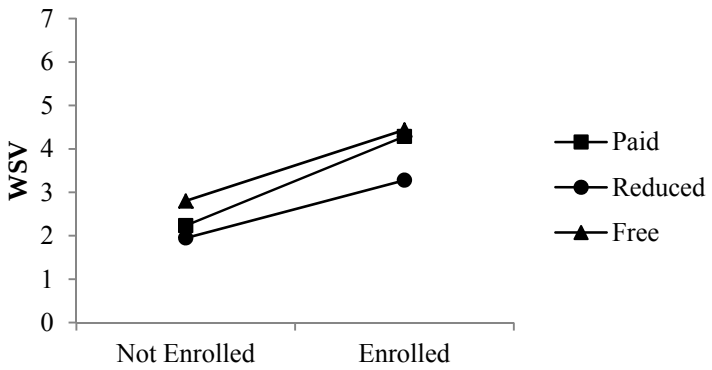

Figure 4d. Weighted Estimates for WSV Analysis for All Years: Enrollment × Lunch Eligibility

Finally, considering the interaction of enrollment × free lunch eligibility status for WDA, there were significant results in Years 2, 5, 7, and 8. Figures 5a, b, c, and d depict this interaction. Like the results for WSV scores, there is a similar pattern in Years 5 and 7 for WDA scores, with those not qualifying for free or reduced-cost lunch and who were enrolled having higher WDA scores. Those enrolled who qualified for reduced-cost lunch had the lowest WSV scores. When considering those not enrolled, however, those qualifying for free lunch had higher WDA scores than those who did not qualify for a reduction in cost of lunch. In each category of free lunch eligibility status, those who were enrolled had higher WDA scores than those not enrolled. In Year 2, however, those not enrolled had higher WDA scores for each category of free lunch eligibility status than those enrolled. Further, those qualifying for a reduced-cost lunch had the highest WDA scores regardless of enrollment status, compared to the other free lunch eligibility status categories. Finally, in Year 8, in Year 8, those qualifying for a free lunch or not qualifying for any federal assistance in lunch had higher WDA scores, regardless of enrollment, than the other two categories of free lunch eligibility status, with those qualifying for a free lunch having consistently higher scores than those not qualifying. Those qualifying for a reduced-cost lunch, however, had very similar WDA scores, whether enrolled or not enrolled in the MYS. There is little consistency across years and effect sizes are also very small; thus, these results can also be treated as anomalous. That is, despite statistical significance in individual years and across measures, for these results to suggest that missingness in the MYS sample is non-ignorable or that the MYS sample to be non-representative of the population, one should expect that these estimates would follow a similar pattern across years. One should also expect that effect sizes be larger than those observed ($.02 \leq \mathbf{w} \leq .04$). Further support for the lack of study-wise practical significance is shown by Figure 5e, the average of

weighted yearly estimates for enrollment × free lunch eligibility status for WDA. When examining the average of yearly estimates for this interaction, those enrolled in the MYS always had lower WDA scores than those enrolled by approximately 2.0 points or less. The differences between categories of free lunch eligibility status are approximately 1.0 to about 0.2 points for those not enrolled and approximately 0.2 to 1.5 points for those enrolled in the MYS.

While it is important to describe the statistically significant results, it is also important to describe the lack of consistent significance in enrollment or in many of its interactions. There is no consistent statistical significance in most interaction across measures (WSV and WDA). With the exception of the enrollment × free lunch interaction, there were no consistent interactions across years and measures. This lack of statistical significance suggests that with respect to these behavioral characteristics, those enrolled in the MYS are representative of the population. Further, this lack of statistical significance fails to support the conclusion of non-ignorable missingness.

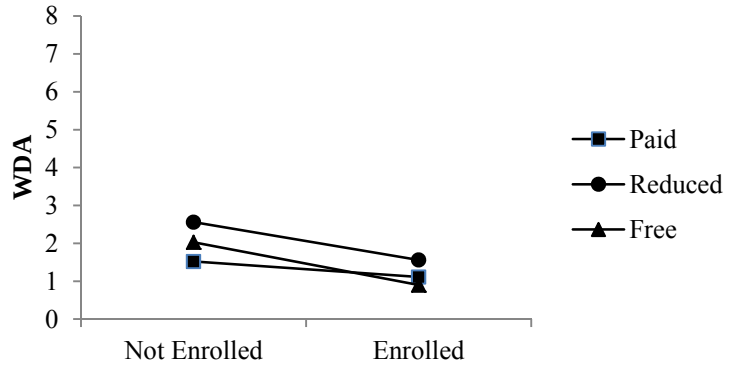

Figure 5a. Estimates for WDA for Year 2: Enrollment × Lunch Eligibility

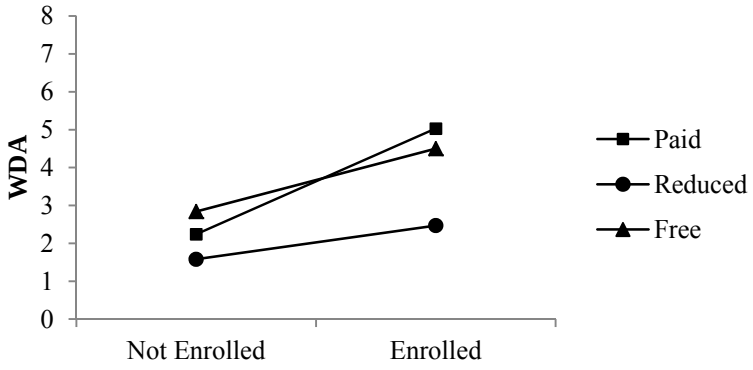

Figure 5b. Estimates for WDA for Year 5: Enrollment × Lunch Eligibility

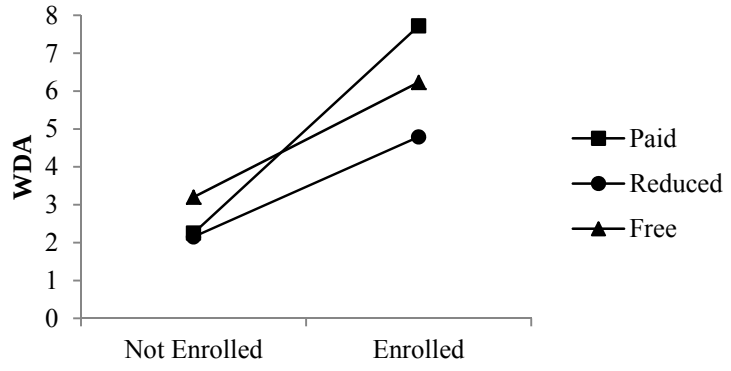

Figure 5a. Estimates for WDA for Year 7: Enrollment × Lunch Eligibility

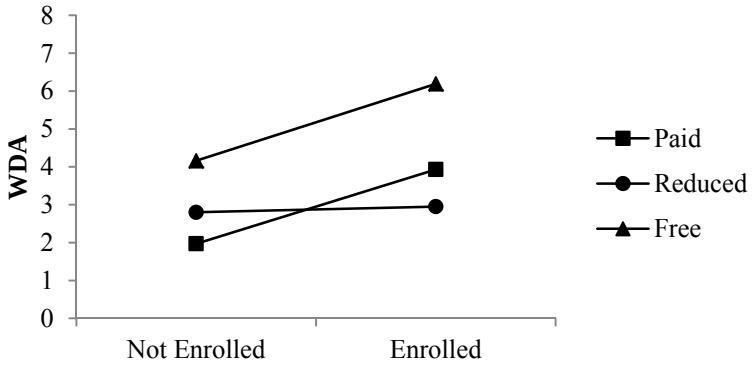

Figure 5b. Estimates for WDA for Year 8: Enrollment × Lunch Eligibility

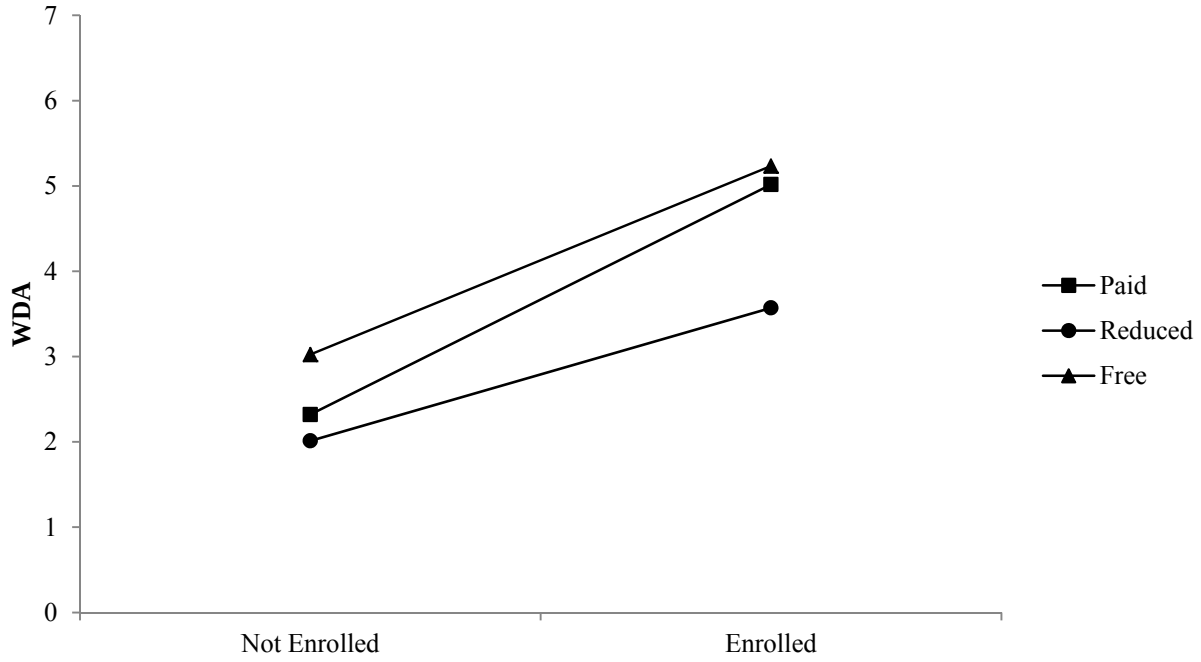

Figure 5e. Weighted Estimates for WDA for All Years: Enrollment × Lunch Eligibility

Table 25

Significance and Effect Size for MYS Enrollment and Demographic Characteristics on Weighted School Violation Scores

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
E	1	0.77	.379	.01	0.50	.481	.01	0.42	.518	.01	2.88	.090	.01	6.77	.009	.02
NT	1	1.37	.242	.01	4.76	.029	.02	1.22	.269	.01	2.74	.098	.01	11.35	.001	.03
Gr	1	144.19	<.001	.11	191.86	<.001	.13	141.00	<.001	.12	97.96	<.001	.09	137.02	<.001	.11
G	1	68.89	<.001	.08	123.69	<.001	.11	99.31	<.001	.10	170.22	<.001	.11	93.30	<.001	.09
R	1	3.82	.051	.02	12.12	.001	.03	6.56	.010	.03	1.76	.185	.01	1.62	.203	.01
LE	2	11.77	.003	.03	1.59	.451	.01	2.29	.318	.02	17.50	<.001	.04	14.19	.001	.04
N	[a]46	421.23	<.001	.19	254.64	<.001	.15	179.89	<.001	.14	237.60	<.001	.14	263.75	<.001	.15
LE × NT	2	2.77	.251	.02	2.84	.242	.02	8.74	.013	.03	7.09	.029	.02	0.14	.933	.00
G × NT	1	0.10	.749	.00	0.05	.818	.00	0.36	.550	.01	0.09	.769	.00	1.61	.205	.01
E × G	1	0.69	.405	.01	2.13	.145	.01	0.01	.935	.00	0.91	.339	.01	0.75	.388	.01
E × Gr	1	3.51	.061	.02	0.83	.361	.01	6.21	.013	.03	0.99	.321	.01	1.45	.229	.01
E × LE	2	3.57	.168	.02	11.23	.004	.03	0.74	.689	.01	2.00	.367	.01	9.82	.007	.03
E × NT	1	1.51	.219	.01	0.15	.697	.00	1.11	.292	.01	1.08	.300	.01	4.38	.036	.02
E × R	1	0.49	.486	.01	8.46	.004	.03	0.55	.457	.01	0.99	.320	.01	0.22	.643	.00
E × NT × Gr	2	0.67	.716	.01	11.96	.003	.03	0.90	.639	.01	8.46	.015	.03	16.91	<.001	.04
E × G × NT	1	0.00	.991	.00	5.35	.021	.02	0.56	.456	.01	0.48	.487	.01	0.88	.348	.01

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
E	1	3.30	.069	.02	6.73	.010	.03	1.16	.281	.01	1.31	.252	.01	5.56	.018	.02
NT	1	0.11	.741	.00	0.46	.498	.01	0.01	.912	.00	0.34	.558	.01	1.53	.217	.01
Gr	1	4.94	.026	.02	20.81	<.001	.04	45.53	<.001	.07	45.70	<.001	.06	129.11	<.001	.11
G	1	107.85	<.001	.10	132.41	<.001	.11	137.82	<.001	.11	131.74	<.001	.11	71.30	<.001	.08
R	1	5.81	.016	.02	3.15	.076	.02	1.68	.195	.01	0.23	.632	.00	1.17	.279	.01
LE	2	16.43	<.001	.04	2.94	.231	.02	103.59	<.001	.10	23.07	<.001	.05	15.11	.001	.04
N	46	293.86	<.001	.17	264.48	<.001	.16	301.37	<.001	.17	228.31	<.001	.15	228.83	<.001	.16
LE × NT	2	5.59	.061	.02	2.20	.333	.01	0.48	.786	.01	0.67	.716	.01	1.74	.418	.01
G × NT	1	0.32	.572	.01	0.04	.851	.00	0.94	.331	.01	0.09	.759	.00	0.89	.345	.01
E × G	1	0.24	.625	.00	0.17	.682	.00	2.74	.098	.02	2.27	.132	.01	0.01	.937	.00
E × Gr	1	0.75	.385	.01	0.28	.600	.01	0.00	.951	.00	11.59	.001	.03	2.11	.146	.01
E × LE	2	1.95	.377	.01	14.97	.001	.04	3.98	.137	.02	0.38	.827	.01	0.93	.627	.01
E × NT	1	0.25	.618	.00	2.37	.124	.01	2.59	.108	.02	0.51	.475	.01	0.11	.736	.00
E × R	1	0.59	.444	.01	0.58	.447	.01	0.14	.710	.00	0.86	.353	.01	2.56	.109	.02
E × NT × Gr	2	2.00	.368	.01	1.91	.384	.01	0.63	.732	.01	2.26	.323	.01	2.02	.365	.01
E × G × NT	1	0.57	.448	.01	1.04	.308	.01	0.44	.508	.01	0.00	.979	.00	0.05	.827	.00

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. E = enroll; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status; N = neighborhood.

[a]Because not all neighborhoods existed at all times for all analyses, *df* for N is inconsistent across years. In Years 2, 5, 8, 9, *df* was 45.

Table 26

Adjusted Estimates and Standard Errors for Selected MYS Enrollment Effects on Weighted School Violation Scores

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Grade, Not Enrolled, Target Neighborhood	1.21	0.08	***1.36***	0.12	1.50	0.13	***1.80***	0.11	***1.87***	0.12
Grade, Not Enrolled, Expansion Neighborhood	1.30	0.03	***1.58***	0.05	1.63	0.05	***1.57***	0.05	***1.88***	0.06
Grade, Enrolled, Target Neighborhood	1.43	0.08	***1.68***	0.07	1.70	0.10	***1.57***	0.09	***3.13***	0.11
Grade, Enrolled, Expansion Neighborhood	1.30	0.09	***1.88***	0.12	2.50	0.11	***1.99***	0.13	***2.77***	0.13
Not Enrolled	0.98	0.11	1.65	0.09	1.84	0.07	1.92	0.07	***2.01***	0.07
Enrolled	1.15	0.21	1.52	0.21	2.39	0.21	2.88	0.21	***3.57***	0.17
Not Enrolled, Not Eligible	1.17	0.11	***1.36***	0.09	1.97	0.08	2.24	0.06	***2.07***	0.07
Not Enrolled, Reduced Lunch	0.65	0.25	***1.81***	0.23	1.67	0.15	1.33	0.17	***1.54***	0.16
Not Enrolled, Free Lunch	1.22	0.07	***1.83***	0.05	1.91	0.06	2.37	0.05	***2.57***	0.06
Enrolled, Not Eligible	1.53	0.20	***1.48***	0.18	2.72	0.22	3.65	0.21	***4.53***	0.17
Enrolled, Reduced Lunch	0.80	0.36	***1.99***	0.46	2.13	0.31	1.91	0.28	***2.46***	0.25
Enrolled, Free Lunch	1.24	0.18	***1.20***	0.15	2.34	0.20	3.45	0.20	***4.09***	0.15

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Grade, Not Enrolled, Target Neighborhood	1.68	0.18	1.58	0.15	1.33	0.15	2.57	0.18	1.93	0.24
Grade, Not Enrolled, Expansion Neighborhood	2.28	0.06	1.82	0.07	1.86	0.06	3.52	0.08	2.51	0.08
Grade, Enrolled, Target Neighborhood	1.50	0.17	2.00	0.17	1.78	0.17	3.80	0.25	3.24	0.23
Grade, Enrolled, Expansion Neighborhood	3.01	0.17	2/81	0.18	3.20	0.13	4.19	0.25	3.22	0.19
Not Enrolled	2.56	0.08	***2.64***	0.08	2.68	0.06	3.15	0.10	***3.67***	0.06
Enrolled	3.58	0.16	***5.46***	0.17	3.52	0.19	7.85	0.28	***8.24***	0.15
Not Enrolled, Not Eligible	2.91	0.09	***2.34***	0.09	1.86	0.09	3.27	0.10	3.35	0.09
Not Enrolled, Reduced Lunch	1.76	0.19	***2.49***	0.16	2.75	0.10	2.49	0.14	3.38	0.10
Not Enrolled, Free Lunch	3.28	0.06	***3.17***	0.06	3.76	0.05	3.84	0.09	4.38	0.05
Enrolled, Not Eligible	4.44	0.17	***6.21***	0.16	3.03	0.20	8.06	0.28	7.90	0.17
Enrolled, Reduced Lunch	2.44	0.27	***4.88***	0.27	2.86	0.28	6.53	0.32	7.74	0.22
Enrolled, Free Lunch	4.23	0.14	***5.35***	0.15	5.03	0.17	9.18	0.27	9.15	0.13

Note. Boldfaced italicized estimates are statistically significant, $p < .05$. Odds ratios and standard errors for the grade interaction effects were estimated separately for each combination of enrollment and neighborhood type.

Table 27

Significance and Effect Size for MYS Enrollment and Demographic Characteristics on Weighted Disciplinary Action Scores

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
E	1				2.58	.108	.02	0.03	.872	.00	3.39	.065	.02	6.01	.014	.02
NT	1				0.41	.521	.01	0.67	.412	.01	4.16	.042	.02	9.08	.003	.03
Gr	1				253.70	<.001	.15	230.03	<.001	.16	207.12	<.001	.13	180.65	<.001	.13
G	1				96.83	<.001	.09	87.71	<.001	.10	139.33	<.001	.10	81.82	<.001	.09
R	1				11.15	.001	.03	3.73	.053	.02	1.40	.237	.01	3.97	.046	.02
LE	2				2.58	.276	.02	2.12	.346	.02	20.94	<.001	.04	15.64	<.001	.04
N	46				222.20	<.001	.14	159.81	<.001	.13	224.41	<.001	.13	221.33	<.001	.14
LE × NT	2				3.68	.159	.02	6.47	.039	.03	10.06	.007	.03	0.24	.885	.00
G × NT	1				0.01	.937	.00	1.35	.246	.01	0.00	.948	.00	0.29	.588	.01
E × G	1				2.19	.139	.01	0.82	.365	.01	0.85	.355	.01	2.35	.125	.01
E × Gr	1				1.33	.249	.01	5.04	.025	.02	0.88	.349	.01	4.96	.026	.02
E × LE	2				10.24	.006	.03	0.99	.610	.01	2.57	.277	.01	10.44	.005	.03
E × NT	1				0.52	.471	.01	0.09	.760	.00	0.00	.963	.00	5.35	.021	.02
E × R	1				8.42	.004	.03	0.19	.661	.00	1.18	.277	.01	0.00	.986	.00
E × NT × Gr	2				3.49	.175	.02	0.16	.921	.00	2.62	.270	.01	13.54	.001	.03
E × G × NT	1				1.76	.185	.01	0.67	.414	.01	0.63	.426	.01	0.33	.567	.01

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
E	1	2.27	.132	.01	7.01	.008	.03	2.20	.138	.01	1.61	.205	.01	3.60	.058	.02
NT	1	7.44	.006	.03	0.18	.670	.00	0.13	.714	.00	0.05	.816	.00	1.29	.256	.01
Gr	1	34.68	<.001	.06	46.07	<.001	.07	74.69	<.001	.08	8.26	.004	.03	78.47	<.001	.09
G	1	99.85	<.001	.10	128.37	<.001	.11	132.29	<.001	.11	143.89	<.001	.12	76.96	<.001	.08
R	1	5.67	.017	.02	3.25	.071	.02	0.94	.331	.01	0.20	.657	.00	3.69	.055	.02
LE	2	14.69	.001	.04	5.35	.069	.02	87.56	<.001	.09	20.23	<.001	.04	16.97	<.001	.04
N	46	288.04	<.001	.17	225.75	<.001	.15	304.68	<.001	.17	144.75	<.001	.12	172.49	<.001	.14
LE × NT	2	5.96	.051	.02	2.84	.242	.02	0.48	.785	.01	0.05	.976	.00	1.36	.508	.01
G × NT	1	0.31	.577	.01	0.47	.493	.01	0.64	.425	.01	0.44	.508	.01	0.33	.566	.01
E × G	1	0.57	.450	.01	0.08	.772	.00	4.29	.038	.02	0.82	.366	.01	0.11	.741	.00
E × Gr	1	0.02	.885	.00	2.26	.133	.01	0.30	.585	.01	17.57	<.001	.04	0.14	.707	.00
E × LE	2	2.14	.343	.01	17.66	<.001	.04	6.19	.045	.02	0.03	.984	.00	1.59	.451	.01
E × NT	1	0.01	.916	.00	2.63	.105	.02	3.54	.060	.02	0.39	.530	.01	0.00	.945	.00
E × R	1	0.30	.584	.01	0.85	.356	.01	0.48	.488	.01	1.28	.259	.01	1.18	.277	.01
E × NT × Gr	2	6.91	.032	.03	2.09	.352	.01	2.29	.318	.01	0.03	.987	.00	3.15	.207	.02
E × G × NT	1	0.29	.593	.01	0.23	.635	.00	0.31	.580	.01	0.02	.901	.00	1.21	.272	.01

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. E = enroll; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status; N = neighborhood.

[a]Because not all neighborhoods existed at all times for all analyses, *df* for N is inconsistent across years. In Years 2, 5, 8, 9, *df* was 45.

Table 28

Adjusted Estimates and Standard Errors for MYS Enrollment × Free Lunch Eligibility Status on Weighted Disciplinary Action Scores

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Not Enrolled, Not Eligible			***1.52***	0.10	2.03	0.10	2.24	0.07	***2.24***	0.08
Not Enrolled, Reduced Lunch			***2.56***	0.22	1.66	0.16	0.99	0.21	***1.58***	0.18
Not Enrolled, Free Lunch			***2.03***	0.06	1.97	0.07	2.44	0.06	***2.84***	0.07
Enrolled, Not Eligible			***1.11***	0.34	3.01	0.30	4.22	0.25	***5.03***	0.20
Enrolled, Reduced Lunch			***1.56***	0.54	2.51	0.38	1.93	0.33	***2.47***	0.29
Enrolled, Free Lunch			***0.90***	0.33	2.55	0.29	3.98	0.23	***4.50***	0.18

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Not Enrolled, Not Eligible	3.36	0.09	***2.25***	0.10	***1.97***	0.10	2.96	0.12	2.33	0.12
Not Enrolled, Reduced Lunch	2.01	0.19	***2.15***	0.19	***2.80***	0.13	2.28	0.19	2.30	0.15
Not Enrolled, Free Lunch	3.81	0.07	***3.20***	0.07	***4.16***	0.07	3.57	0.11	3.27	0.08
Enrolled, Not Eligible	5.47	0.19	***7.72***	0.19	***3.93***	0.22	8.67	0.26	6.13	0.25
Enrolled, Reduced Lunch	3.15	0.30	***4.79***	0.31	***2.95***	0.31	6.57	0.33	7.00	0.30
Enrolled, Free Lunch	5.16	0.17	***6.23***	0.18	***6.19***	0.18	10.29	0.25	7.68	0.22

Note. Boldfaced italicized estimates are statistically significant, $p < .05$.

Research Question 1: Yearly Participation (P_t)

The second part of Research Question 1 is concerned with the extent to which those who participate yearly in the MYS (1998-2007) are representative of adolescents living in MYS neighborhoods. Table 29 shows the number of students from the MYS neighborhoods that are yearly participants and that are not yearly participants in the MYS each year. As indicated in Chapter 3, analyses in Years 1 through 7 were limited to target neighborhoods; beginning in Year 8, expansion neighborhoods were included as well. This accounts for the substantial increase in the number of total cases in Years 8 through 10. In Years 1 through 7, between 33.0% and 43.6% of all residents aged 10 through 15 living in the target neighborhoods participated yearly in the MYS. In Years 8 through 10, between 13.8% and 18.3% of all residents aged 10 through 15 living in any of the MYS neighborhoods participated yearly in the MYS with the percentage increasing each year.

Table 29

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods by MYS Yearly Participation Status

	Year									
	1	2	3	4	5	6	7	8	9	10
$P_t = 1$	915	1,188	824	1,065	1,024	986	904	1,462	1,541	1,645
$P_t = 0$	1,855	1,534	1,452	2,073	1,582	1,375	1,340	9,138	9,289	7,349

Note. For Years 1 through 7, only students living in target neighborhoods were included in the analyses. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Demographics. Table 30 shows the goodness of fit statistics (χ^2), significance levels, and effect sizes (**w**) for the GEE results regressing yearly participation on demographic variables. A power analysis shows power to detect each of the main effects to be greater than or equal to .995 (**f** = .15, α = .05).

Results demonstrate that there was a significant departure from representativeness in race for several years (Years 2, 3, 8, 9, 10), for free lunch eligibility status each year except Year 6, for neighborhood type in Years 8 through 10, and for grade level in each year except Year 1.

Table 31 shows the estimates for those variables that produced statistically significant results. For race, free lunch eligibility status, and neighborhood type, estimates represent proportions for each category that are yearly participants in the MYS. For grade level, the estimate is an odds ratio. As expected, the rate of yearly participation for African Americans is higher than the rate of yearly participation for non-African Americans. Further, the rate of yearly participation for those eligible for free lunch is higher than the rate of yearly participation for both those who are eligible for reduced lunch and those are eligible for neither. Also, as expected, in Years 8 through 10, the rates of yearly participation in the target neighborhoods are higher than the rates of yearly participation in the expansion neighborhoods. While these differences in rates of yearly participation are expected, they do suggest that the sample is not representative of the population. However, these differences may not be indicative of ignorable missingness, because of small effect sizes (**w**): (a) for race: $.01 \leq \mathbf{w} \leq .09$; (b) for free lunch eligibility status: $.05 \leq \mathbf{w} \leq .20$; (c) and for neighborhood type: $.02 \leq \mathbf{w} \leq .03$. Even though $\mathbf{w} > .15$ in Year 4 for free lunch eligibility status, this small-to-moderate effect size was not consistently found across years.

Finally, the odds ratio for grade level in all years is greater than 1.0 indicating a positive relationship between yearly participation and grade level. That is, during MYS data collection each year, older students participated at disproportionately high rates. While results were consistent across years and in functional form, the effect sizes (**w**) tend to be small: $.05 \leq \mathbf{w} \leq .14$, limiting any conclusions about non-ignorable missingness.

It is important to also consider the main effects and interactions that are not statistically significant. There were no consistent departures from representativeness (in three or more years) for gender or either interaction effect (free lunch eligibility status × neighborhood type and gender × neighborhood type) included in this analysis of demographic representativeness. That is, there are no differences in the rates of yearly participation for males and females. Next, while there were significant departures in representativeness for free lunch eligibility status, when included as a component of interaction with neighborhood type, there are no statistically significant results. This indicates that across neighborhood types, rates of yearly participation for students qualifying (or not) for free lunch are similar; and this interaction does not contribute anything beyond the main effects. Finally, while there were statistically significant differences in neighborhood type, the rates of yearly participation for males were similar in both neighborhood types, and this interaction does not contribute anything beyond the main effects. This lack of significance suggests that the MYS sample is representative of the population on these characteristics and fails to suggest that missingness is non-ignorable.

Table 30

Significance and Effect Size for Demographic Characteristics on MYS Yearly Participation

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
G	1	0.29	.589	.01	0.46	.497	.01	0.34	.560	.01	0.05	.828	.00	1.89	.169	.03
R	1	0.92	.337	.02	14.55	<.001	.07	12.46	<.001	.07	5.36	.021	.04	1.99	.158	.03
LE	2	13.33	.001	.07	30.57	<.001	.11	31.71	<.001	.12	122.51	<.001	.20	35.77	<.001	.12
NT	1															
LE × NT	2															
G × NT	1															
Gr	1	6.00	.014	.05	13.82	<.001	.07	33.55	<.001	.12	59.85	<.001	.14	19.63	<.001	.09
N	[a]12	369.19	<.001	.18	453.03	<.001	.20	446.68	<.001	.22	200.83	<.001	.12	290.38	<.001	.16

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
G	1	1.92	.165	.03	0.00	.959	.00	3.42	.065	.02	0.48	.488	.01	0.68	.411	.01
R	1	1.25	.263	.02	0.38	.538	.01	63.54	<.001	.08	25.90	<.001	.05	68.03	<.001	.09
LE	2	6.93	.031	.05	12.73	.002	.08	127.29	<.001	.11	45.58	<.001	.06	78.23	<.001	.09
NT	1							7.04	.008	.03	10.17	.001	.03	5.02	.025	.02
LE × NT	2							0.64	.727	.01	1.00	.607	.01	2.35	.309	.02
G × NT	1							5.61	.018	.02	0.02	.900	.00	0.00	.960	.00
Gr	1	43.47	<.001	.14	26.97	<.001	.11	38.85	<.001	.06	31.32	<.001	.05	41.85	<.001	.07
N	12	354.47	<.001	.18	321.58	<.001	.17	357.56	<.001	.18	344.77	<.001	.18	447.72	<.001	.22

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. G = gender; R = race; LE = free lunch eligibility status; NT = neighborhood type; Gr = grade level; N = neighborhood.

[a]Because not all neighborhoods existed at all times for all analyses, *df* for N is inconsistent across years. Further, because only neighborhoods in the target neighborhoods were used in the Year 1 – 7 analyses, there are further inconsistencies in degrees of freedom. In Years 8 and 9, *df* = 45; in Year 10, *df* = 46.

Table 31

Adjusted Estimates and Standard Errors for Selected Demographic Characteristics on MYS Yearly Participation

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Non-African American	0.20	0.56	***0.08***	0.56	***0.06***	0.49	***0.12***	0.42	0.18	0.45
African American	0.26	0.17	***0.36***	0.17	***0.20***	0.24	***0.24***	0.15	0.29	0.16
Not Eligible	***0.23***	0.32	***0.12***	0.32	***0.14***	0.28	***0.13***	0.21	***0.20***	0.26
Reduced Lunch	***0.16***	0.53	***0.19***	0.53	***0.04***	0.71	***0.12***	0.48	***0.17***	0.48
Free Lunch	***0.32***	0.28	***0.25***	0.28	***0.24***	0.22	***0.30***	0.20	***0.35***	0.22
Expansion Neighborhood										
Target Neighborhood										
Grade Level	***1.06***	0.02	***1.09***	0.02	***1.17***	0.03	***1.19***	0.02	***1.12***	0.03

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Non-African American	0.25	0.46	0.22	0.49	***0.02***	0.47	***0.01***	1.02	***0.02***	0.48
African American	0.36	0.14	0.27	0.19	***0.12***	0.09	***0.13***	0.09	***0.14***	0.11
Not Eligible	0.28	0.29	***0.24***	0.29	***0.03***	0.25	***0.04***	0.52	***0.05***	0.27
Reduced Lunch	0.26	0.43	***0.17***	0.54	***0.03***	0.32	***0.04***	0.55	***0.03***	0.35
Free Lunch	0.37	0.23	***0.34***	0.24	***0.09***	0.23	***0.07***	0.51	***0.11***	0.24
Expansion Neighborhood					***0.02***	0.24	***0.02***	0.52	***0.02***	0.25
Target Neighborhood					***0.11***	0.27	***0.11***	0.53	***0.12***	0.30
Grade Level	***1.19***	0.03	***1.15***	0.03	***1.12***	0.02	***1.11***	0.02	***1.12***	0.02

Note. Boldfaced italicized estimates are statistically significant, $p < .05$.

SAT percentile ranks. Table 32 shows the number of students included in these analyses: those aged 10 through 15 who are in grades 3 through 8 and have either a total reading percentile rank (SAT_{TR}) or a total math percentile rank (SAT_{TM}). Tables 33 and 34 show goodness of fit statistics (χ^2), the significance levels, and the effect sizes (**w**) for the GEE results regressing SAT_{TR} and SAT_{TM}, respectively, on yearly participation (main effects and interactions) controlling for demographic characteristics. A power analysis shows power to detect each of the main effects to be greater than or equal to .995 (**f** = .15, α = .05) for both analyses of cognitive ability.

Results demonstrate no significant main effects for yearly participation in SAT_{TR} or SAT_{TM}. There are also no significant interactions for SAT_{TR}. However, results demonstrate a more consistent significant yearly participation × grade interaction for SAT_{TM} in Years 8 through 10. Table 35 shows the regression coefficients for the SAT_{TM} analyses for participants and non-participants. These regression coefficients show that for the significant years, there was a stronger negative relationship between SAT_{TM} and grade level for those participating yearly than for those not participating yearly. Even though statistically significant, effect sizes for all significant effects are very small (**w** = .03) and the statistical significance was not found across measures, thus, there is support for ignorable missingness.

While it is important to describe the statistically significant results, it is also important to note the lack of significance in all but one of the interactions that include yearly participation as a component. There is no consistent statistical significance in any interaction across measures (SAT_{TR} and SAT_{TM}) and across years. This lack of statistical significance suggests that those who are yearly participants in the MYS are representative of the population on cognitive characteristics and that missingness may be ignorable.

Table 32

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods Participating Yearly in the MYS with SAT Data

	Year									
	1	2	3	4	5	6	7	8	9	10
With SAT_{TR} Percentile Ranks	1,484	1,392	1,317	1,842	1,757	1,594	1,567	6,606	6,853	5,927
With SAT_{TM} Percentile Ranks	1,496	1,408	1,317	1,919	1,786	1,597	1,560	6,560	6,812	5,906

Note. For Years 1 through 7, only students living in target neighborhoods were included in the analyses Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Table 33

Significance and Effect Size for MYS Yearly Participation and Demographic Characteristics on Total Reading Percentile Ranks

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
P_t	1	0.03	.864	.00	0.99	.320	.03	0.05	.822	.01	1.46	.228	.03	1.08	.299	.02
NT	1															
Gr	1	2.73	.098	.04	2.41	.121	.04	4.09	.043	.06	4.43	.035	.05	3.41	.065	.04
G	1	13.17	<.001	.09	19.33	<.001	.12	12.59	<.001	.10	16.98	<.001	.10	19.79	<.001	.11
R	1	1.93	.165	.04	0.99	.321	.03	4.90	.027	.06	1.76	.185	.03	3.44	.064	.04
LE	2	4.41	.111	.05	1.39	.500	.03	1.55	.462	.03	0.16	.925	.01	5.10	.078	.05
N	[a]12	33.44	<.001	.15	32.42	.001	.15	83.47	<.001	.25	71.50	<.001	.20	33.74	<.001	.14
LE × NT	2															
G × NT	1															
P_t × G	1	0.19	.665	.01	1.33	.249	.03	1.65	.199	.04	1.82	.177	.03	0.16	.688	.01
P_t × Gr	1	3.17	.075	.05	0.05	.826	.01	1.11	.292	.03	1.51	.220	.03	0.03	.857	.00
P_t × LE	2	2.19	.335	.04	1.31	.520	.03	1.09	.579	.03	1.20	.549	.03	0.60	.740	.02
P_t × NT	1															
P_t × R	1	0.02	.890	.00	0.87	.351	.03	1.22	.269	.03	2.21	.137	.03	0.05	.823	.01
P_t × NT × Gr	2															
P_t × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
P_t	1	4.71	.030	.05	0.57	.452	.02	1.40	.236	.01	1.41	.234	.01	0.68	.408	.01
NT	1							0.51	.474	.01	0.30	.584	.01	1.43	.232	.02
Gr	1	0.26	.611	.01	0.17	.681	.01	2.04	.154	.02	8.08	.005	.03	17.96	<.001	.06
G	1	25.62	<.001	.13	25.15	<.001	.13	53.07	<.001	.09	68.95	<.001	.10	63.98	<.001	.10
R	1	5.17	.023	.06	0.10	.754	.01	0.69	.406	.01	1.06	.303	.01	2.49	.115	.02
LE	2	5.24	.073	.06	4.15	.126	.05	4.61	.100	.03	2.78	.249	.02	14.29	.001	.05
N	12	42.38	<.001	.16	21.09	.049	.12	148.48	<.001	.15	85.54	<.001	.11	94.66	<.001	.13
LE × NT	2							3.74	.154	.02	4.48	.107	.03	0.55	.760	.01
G × NT	1							0.50	.481	.01	0.77	.380	.01	1.80	.180	.02
P_t × G	1	0.37	.541	.02	1.24	.265	.03	0.21	.647	.01	2.05	.152	.02	0.24	.622	.01
P_t × Gr	1	0.66	.416	.02	1.17	.280	.03	0.01	.943	.00	5.01	.025	.03	0.05	.828	.00
P_t × LE	2	4.91	.086	.06	1.80	.406	.03	0.67	.716	.01	1.43	.489	.01	2.73	.256	.02
P_t × NT	1							0.40	.525	.01	1.18	.278	.01	0.43	.514	.01
P_t × R	1	1.92	.166	.03	1.53	.216	.03	0.46	.498	.01	1.05	.306	.01	0.70	.404	.01
P_t × NT × Gr	2							1.29	.524	.01	1.43	.490	.01	0.96	.618	.01
P_t × G × NT	1							5.47	.019	.03	0.05	.815	.00	0.06	.806	.00

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. P_t = yearly participation; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status; N = neighborhood.

[a]Because not all neighborhoods existed at all times for all analyses, *df* for N is inconsistent across years. Further, because only neighborhoods in the target neighborhoods were used in the Year 1 – 7 analyses, there are further inconsistencies in degrees of freedom. In Years 8 and 9, *df* = 45; in Year 10, *df* = 46.

Table 34

Significance and Effect Size for MYS Yearly Participation and Demographic Characteristics on Total Math Percentile Ranks

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
P_t	1	1.18	.277	.03	0.44	.506	.02	0.97	.324	.03	0.50	.479	.02	1.18	.278	.03
NT	1															
Gr	1	2.95	.086	.04	2.76	.097	.04	35.84	<.001	.16	30.52	<.001	.13	7.11	.008	.06
G	1	8.37	.004	.07	7.34	.007	.07	5.39	.020	.06	11.01	.001	.08	8.32	.004	.07
R	1	0.24	.625	.01	1.16	.281	.03	4.04	.044	.06	1.25	.264	.03	1.74	.187	.03
LE	2	1.85	.396	.04	0.36	.833	.02	1.81	.405	.04	0.26	.877	.01	6.26	.044	.06
N	12	39.48	<.001	.16	30.00	.003	.15	30.19	.003	.15	44.38	<.001	.15	41.26	<.001	.15
LE × NT	2															
G × NT	1															
P_t × G	1	2.47	.116	.04	2.72	.099	.04	4.67	.031	.06	0.70	.404	.02	0.87	.350	.02
P_t × Gr	1	0.40	.526	.02	2.82	.093	.04	0.03	.860	.00	0.24	.626	.01	0.04	.838	.00
P_t × LE	2	1.79	.408	.03	0.31	.856	.01	1.12	.571	.03	1.09	.580	.02	2.83	.243	.04
P_t × NT	1															
P_t × R	1	0.07	.786	.01	0.61	.433	.02	0.50	.480	.02	1.45	.228	.03	3.30	.070	.04
P_t × NT × Gr	2															
P_t × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
P_t	1	0.83	.362	.02	0.08	.771	.01	2.08	.149	.02	0.61	.436	.01	0.43	.514	.01
NT	1							0.08	.772	.00	1.02	.313	.01	0.01	.939	.00
Gr	1	11.70	.001	.09	16.94	<.001	.10	47.23	<.001	.08	24.64	<.001	.06	31.64	<.001	.07
G	1	23.35	<.001	.12	11.69	.001	.09	27.09	<.001	.06	46.84	<.001	.08	40.09	<.001	.08
R	1	0.48	.489	.02	0.31	.580	.01	1.87	.171	.02	0.97	.326	.01	0.00	.948	.00
LE	2	3.33	.189	.05	0.21	.902	.01	3.92	.141	.02	6.95	.031	.03	13.47	.001	.05
N	12	62.89	<.001	.20	24.95	.015	.13	141.39	<.001	.15	114.10	<.001	.13	171.45	<.001	.17
LE × NT	2							11.28	.004	.04	3.07	.215	.02	0.26	.877	.01
G × NT	1							0.55	.458	.01	0.07	.793	.00	5.18	.023	.03
P_t × G	1	0.33	.567	.01	0.03	.863	.00	0.00	.965	.00	4.61	.032	.03	0.01	.942	.00
P_t × Gr	1	0.08	.778	.01	0.10	.754	.01	5.29	.021	.03	4.58	.032	.03	4.54	.033	.03
P_t × LE	2	1.08	.584	.03	2.47	.291	.04	0.30	.861	.01	0.22	.894	.01	3.01	.222	.02
P_t × NT	1							0.08	.782	.00	0.01	.916	.00	0.55	.459	.01
P_t × R	1	1.49	.223	.03	0.44	.507	.02	0.47	.493	.01	0.39	.532	.01	1.60	.206	.02
P_t × NT × Gr	2							3.17	.205	.02	2.45	.294	.02	0.70	.704	.01
P_t × G × NT	1							2.71	.100	.02	0.65	.419	.01	0.58	.445	.01

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. P_t = yearly participation; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status.

[a]Because not all neighborhoods existed at all times for all analyses, *df* for N is inconsistent across years. Further, because only neighborhoods in the target neighborhoods were used in the Year 1 – 7 analyses, there are further inconsistencies in degrees of freedom. In Years 8 and 9, *df* = 45; in Year 10, *df* = 46.

Table 35

Regression Coefficients and Standard Errors for MYS Yearly Participation × Grade Level on Total Math Percentile Ranks

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
$P_t = 1$	-0.53	0.73	0.07	0.74	-2.89	0.81	-2.92	0.75	-0.98	0.57
$P_t = 0$	-1.13	0.56	-1.63	0.61	-3.11	0.58	-2.42	0.53	-1.39	0.54

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
$P_t = 1$	-1.60	0.67	-2.36	0.77	***-3.29***	0.89	***-3.30***	0.95	***-3.53***	0.91
$P_t = 0$	-1.39	0.58	-1.96	0.66	***-2.38***	0.66	***-1.72***	0.70	***-1.42***	0.72

Note. Boldfaced italicized estimates are statistically significant, $p < .05$. Odds ratios and standard errors for the grade level interaction effects were estimated separately for each combination of yearly participation.

School violations and disciplinary actions. Table 36 shows the number of students included in analyses comparing the weighted school violation scores (WSV) of and weighted disciplinary action scores (WDA) for yearly participants and non-participants. Disciplinary actions analyses for Years 1 through 3 were omitted because there were not sufficient data available for the model to converge[38]

Tables 37 and 38 show goodness of fit statistics (χ^2), the significance levels, and the effect sizes (**w**) for the GEE results regressing WSV and WDA, respectively, on yearly participation (main effects and interactions) controlling for demographic characteristics. A power analysis shows power to detect each of the main effects to be greater than or equal to .995 (**f** = .15, α = .05) for both analyses of behavior.

Results demonstrate that there were no significant differences in weighted school violation scores or weighted disciplinary action scores for those participating and those not participating in the MYS in any year. Further, there were no consistent departures from representativeness for any main effect or any interactions for WSV scores. However, consistent departures from representativeness are evident in the yearly participation × grade interaction for WDA in Years 4, 5, and 8. Table 39 shows the odds ratios for this interaction that produced statistically significant results for WDA.

These odds ratios show that for the specified years, there was a stronger positive relationship between WDA and grade level for those participating yearly than for those not participating yearly. Even though statistically significant, effect sizes for all significant effects are very small ($.02 \leq \mathbf{w} \leq .07$) and the statistical significance was not found across measures; thus, there is support for ignorable missingness.

[38] Variables could have been systematically removed from the model to force convergence, but then results for these years would not have been comparable to other years.

While it is important to describe the statistically significant results, it is also important to note the lack of significance in all but one of the interactions that include yearly participation as a component. There is no consistent statistical significance in any interaction across measures (WSV and WDA) and across years. This lack of statistical significance suggests that those who are yearly participants in the MYS are representative of the population on behavioral characteristics and that missingness may be ignorable.

Table 36

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods Participating Yearly in the MYS with School Violation and Disciplinary Action Data

	Year									
	1	2	3	4	5	6	7	8	9	10
n	2,770	2,722	2,276	3,138	2,606	2,361	2,244	10,600	10,830	8,994

Note. For Years 1 through 7, only students living in target neighborhoods were included in the analyses. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Table 37

Significance and Effect Size for MYS Yearly Participation and Demographic Characteristics on Weighted School Violation Scores

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
P_t	1	0.22	.639	.01	1.38	.240	.02	1.80	.179	.03	0.64	.423	.01	1.15	.283	.02
NT	1															
Gr	1	68.29	<.001	.16	131.31	<.001	.22	61.69	<.001	.16	50.06	<.001	.13	97.15	<.001	.19
G	1	25.39	<.001	.10	51.43	<.001	.14	47.27	<.001	.14	82.69	<.001	.16	34.57	<.001	.12
R	1	3.31	.069	.03	3.36	.067	.04	6.47	.011	.05	6.30	.012	.04	11.24	.001	.07
LE	2	2.54	.281	.03	2.06	.356	.03	2.40	.301	.03	11.70	.003	.06	7.36	.025	.05
N	[a]12	166.29	<.001	.25	103.85	<.001	.20	43.32	<.001	.14	65.55	<.001	.14	67.78	<.001	.16
LE × NT	2															
G × NT	1															
P_t × G	1	2.71	.100	.03	7.17	.007	.05	2.00	.157	.03	1.18	.277	.02	0.83	.364	.02
P_t × Gr	1	0.09	.763	.01	1.61	.204	.02	0.27	.606	.01	4.11	.043	.04	9.19	.002	.06
P_t × LE	2	2.45	.294	.03	0.40	.819	.01	0.27	.873	.01	3.25	.197	.03	3.83	.147	.04
P_t × NT	1															
P_t × R	1	0.83	.362	.02	1.61	.205	.02	3.72	.054	.04	0.04	.839	.00	1.20	.274	.02
P_t × NT × Gr	2															
P_t × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
P_t	1	0.27	.602	.01	0.01	.939	.00	3.03	.082	.02	0.87	.352	.01	3.97	.046	.02
NT	1							0.72	.397	.01	0.03	.852	.00	1.12	.290	.01
Gr	1	0.62	.430	.02	8.10	.004	.06	29.40	<.001	.05	60.92	<.001	.08	124.20	<.001	.12
G	1	49.43	<.001	.14	72.11	<.001	.18	103.94	<.001	.10	116.95	<.001	.10	71.65	<.001	.09
R	1	6.37	.012	.05	3.39	.066	.04	0.02	.895	.00	1.08	.299	.01	0.62	.430	.01
LE	2	10.72	.005	.07	0.28	.869	.01	49.74	<.001	.07	24.10	<.001	.05	13.68	.001	.04
N	12	84.95	<.001	.19	66.45	<.001	.17	303.71	<.001	.08	226.68	<.001	.17	235.82	<.001	.16
LE × NT	2							0.69	.708	.01	1.06	.587	.01	1.27	.530	.01
G × NT	1							1.45	.229	.01	0.02	.889	.00	0.70	.404	.01
P_t × G	1	0.90	.343	.02	1.07	.301	.02	1.87	.171	.01	1.41	.235	.01	0.32	.574	.01
P_t × Gr	1	1.58	.208	.03	0.05	.821	.00	2.05	.153	.01	0.63	.427	.01	0.05	.818	.00
P_t × LE	2	4.05	.132	.04	0.62	.735	.02	4.77	.092	.02	0.27	.872	.00	1.84	.398	.01
P_t × NT	1							2.67	.102	.02	0.61	.435	.01	0.07	.795	.00
P_t × R	1	0.32	.575	.01	2.21	.137	.03	0.96	.327	.01	0.70	.401	.01	1.57	.210	.01
P_t × NT × Gr	2							1.32	.517	.01	0.34	.844	.01	1.39	.499	.01
P_t × G × NT	1							0.20	.659	.00	0.16	.685	.00	0.53	.468	.01

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. Pt = yearly participation; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status; N = neighborhood.

[a]Because not all neighborhoods existed at all times for all analyses, *df* for N is inconsistent across years. Further, because only neighborhoods in the target neighborhoods were used in the Year 1 – 7 analyses, there are further inconsistencies in degrees of freedom. In Years 8 and 9, *df* = 45; in Year 10, *df* = 46.

Table 38

Significance and Effect Size for MYS Yearly Participation and Demographic Characteristics on Weighted Disciplinary Action Scores

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
P_t	1										0.81	.368	.02	1.60	.207	.02
NT	1															
Gr	1										108.49	<.001	.19	121.07	<.001	.22
G	1										66.13	<.001	.15	32.02	<.001	.11
R	1										3.99	.046	.04	10.39	.001	.06
LE	2										12.23	.002	.06	7.85	.020	.05
N	[a]12										77.25	<.001	.16	59.22	<.001	.15
LE × NT	2															
G × NT	1															
P_t × G	1										1.39	.239	.02	0.90	.344	.02
P_t × Gr	1										4.38	.036	.04	11.16	.001	.07
P_t × LE	2										1.58	.454	.02	1.35	.510	.02
P_t × NT	1															
P_t × R	1										0.19	.666	.01	1.93	.165	.03
P_t × NT × Gr	2															
P_t × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
P_t	1	0.99	.320	.02	0.02	.902	.00	4.24	.040	.02	1.03	.309	.01	2.65	.104	.02
NT	1							2.27	.132	.01	0.64	.424	.01	1.18	.277	.01
Gr	1	29.17	<.001	.11	21.17	<.001	.10	50.32	<.001	.07	13.79	.000	.04	76.83	<.001	.09
G	1	47.50	<.001	.14	74.01	<.001	.18	99.21	<.001	.10	128.69	<.001	.11	81.64	<.001	.10
R	1	6.17	.013	.05	4.84	.028	.05	0.09	.765	.00	1.04	.309	.01	1.94	.164	.01
LE	2	9.61	.008	.06	0.25	.882	.01	41.79	<.001	.06	19.29	<.001	.04	13.86	.001	.04
N	12	71.31	<.001	.17	46.11	<.001	.14	306.19	<.001	.17	143.06	<.001	.11	173.24	<.001	.14
LE × NT	2							0.75	.686	.01	0.02	.989	.00	1.04	.593	.01
G × NT	1							1.43	.232	.01	0.08	.778	.00	0.07	.799	.00
P_t × G	1	0.44	.507	.01	2.79	.095	.04	2.98	.085	.02	0.64	.423	.01	0.64	.424	.01
P_t × Gr	1	1.78	.182	.03	0.03	.871	.00	4.14	.042	.02	3.04	.081	.02	0.89	.345	.01
P_t × LE	2	4.13	.127	.04	0.88	.644	.02	5.61	.061	.02	0.04	.980	.00	1.96	.376	.01
P_t × NT	1							5.16	.023	.02	0.81	.369	.01	0.13	.720	.00
P_t × R	1	0.01	.906	.00	3.11	.078	.04	1.75	.186	.01	0.67	.411	.01	1.13	.288	.01
P_t × NT × Gr	2							4.61	.100	.02	1.27	.529	.01	3.14	.208	.02
P_t × G × NT	1							0.70	.403	.01	0.16	.691	.00	2.79	.095	.02

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. P_t = yearly participation; NT = neighborhood type; Gr = grade level, G = gender; LE = free lunch eligibility status; N = neighborhood.

[a]Because not all neighborhoods existed at all times for all analyses, *df* for N is inconsistent across years. Further, because only neighborhoods in the target neighborhoods were used in the Year 1 – 7 analyses, there are further inconsistencies in degrees of freedom. In Years 8 and 9, *df* = 45; in Year 10, *df* = 46.

Table 39

Odds Ratios and Standard Errors for MYS Yearly Participation × Grade Level on Weighted Disciplinary Action Scores

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
$P_t = 1$	1.09	0.04	2.26	0.12	2.16	0.14	***2.45***	0.16	***5.71***	0.21
$P_t = 0$	1.07	0.02	1.51	0.11	2.15	0.11	***2.36***	0.11	***3.14***	0.13

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
$P_t = 1$	4.18	0.23	2.92	0.27	***4.80***	0.29	6.39	0.34	4.17	0.34
$P_t = 0$	3.70	0.17	2.45	0.17	***2.29***	0.17	2.90	0.21	3.29	0.23

Note. Boldfaced italicized estimates are statistically significant, $p < .05$. Odds ratios and standard errors for the grade interaction effects were estimated separately for each combination of yearly participation.

Research question 2: Longitudinal Participation ($P_{t-1,\,t}$)

The second research question is concerned with longitudinal representativeness and whether the between-wave missing data mechanisms in the MYS data can be considered to be missing at random. Here, similar analyses to Research Question 1 were completed. Like Research Question 1, both demographic (grade level, gender, race, free lunch eligibility status) and functional (cognitive and behavioral) characteristics are analyzed and those students who dropped out of the MYS at $Time_t$ are compared to those students who participated at $Time_{t-1}$. Because these analyses use pairs of years, Year 1 is undefined. Table 40 shows, for each pair of years (t – 1, t), the number of students from the MYS neighborhoods that participated both years ($P_{t-1,\,t} = 1$) and that dropped out ($P_{t-1,\,t} = 0$). Notably, between 10.4% and 32.8% of those participants at $Time_t$ then dropped out in the next year of possible MYS participation.

Table 40

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods by MYS Longitudinal Participation Status

	Year									
	1	2	3	4	5	6	7	8	9	10
$P_{t-1,\,t} = 1$		652	626	662	734	645	581	811	965	844
$P_{t-1,\,t} = 0$		76	149	111	215	140	129	271	470	180

Note. For Years 1 through 7, analyses were limited to those living in target neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Demographics

Table 41 shows the goodness of fit statistics (χ^2), significance levels, and effect sizes (**w**) for the GEE results regressing longitudinal participation on demographic variables. A power analysis shows power to detect each of the main effects to be greater than or equal to .995 (**f** = .15, α = .05).

Results demonstrate that there was a significant departure from longitudinal representativeness in gender for Year 5 (that is the gender-makeup of those who dropped out of the MYS in Year 5 is significantly different from gender-makeup of those who participated in Year 4)[39], with females being more likely to drop out than males. Consistent departures from representativeness were found for free lunch eligibility status (Years 2, 4, 5, and 8) and for neighborhood type (Years 8 and 9).

Table 42 shows the estimates for those variables that produced statistically significant results. For free lunch eligibility status and neighborhood type, these estimates represent proportions for each category that are longitudinal participants in the MYS. The rates of longitudinal participation are higher for those who qualify for some form of federal assistance for school lunch are higher than for those who are not eligible. Further, the rates of longitudinal participation are higher for those in target neighborhoods than for those in expansion neighborhoods. Even though statistically significant, effect sizes for all significant effects are small, with the exception of free lunch eligibility status in Year 4, which was small to moderate: (a) for free lunch eligibility status: $.02 \leq \mathbf{w} \leq .2$; (b) for neighborhood type: $.08 \leq \mathbf{w} \leq .15$.

It is important to also consider the main effects and interactions that are not statistically significant. There were no consistent departures from representativeness (in three or more years) for gender or either interaction effect (free lunch status × neighborhood type and gender × neighborhood type) included in this analysis of demographic representativeness. That is, there are no consistent differences in the dropout rates for males and females. Next, while there were significant departures in representativeness for free lunch eligibility status, when included as a component of interaction with neighborhood type, there are no statistically significant results. This indicates that across neighborhood types, rates of dropout for students qualifying (or not)

[39] For this result, because it only occurred in one year, estimates are not reported in any table.

for free lunch are similar, and that this interaction does not contribute anything beyond the main effects. Finally, while there were statistically significant differences in neighborhood type, the rates of dropout for males were similar in both neighborhood types, and that this interaction does not contribute anything beyond the main effects. This lack of significance does suggest that those who drop out of the MYS are similar to those who do not drop out on these characteristics and fails to suggest that missingness is non-ignorable.

Table 41

Significance and Effect Size for Demographic Characteristics on MYS Longitudinal Participation

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
G	1				0.06	.813	.01	1.24	.265	.04	0.44	.508	.03	10.66	.001	.12
LE	2				8.58	.003	.03	0.87	.350	.04	27.26	<.001	.20	5.00	.025	.08
NT	1															
LE × NT	2															
G × NT	1															
Gr	1				7.21	.007	.03	0.05	.815	.01	5.87	.015	.09	0.26	.611	.02

		Year														
		6			7			8			9			10		
		χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
G	1	0.05	.815	.01	0.00	.975	.00	0.48	.488	.02	0.72	.397	.03	0.19	.665	.02
LE	2	0.29	.589	.02	0.15	.697	.02	5.53	.019	.08	1.77	.183	.04	3.85	.050	.07
NT	1							19.13	<.001	.15	21.44	<.001	.15	5.90	.670	.08
LE × NT	2							1.90	.168	.05	0.58	.448	.02	0.01	.921	.00
G × NT	1							2.87	.090	.06	0.43	.510	.02	0.09	.761	.01
Gr	1	0.04	.834	.01	0.04	.850	.01	0.72	.398	.03	1.71	.191	.04	0.18	.670	.01

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. G = gender; LE = free lunch eligibility status; NT = neighborhood type; Gr = grade level.

Table 42

Adjusted Estimates and Standard Errors for Neighborhood Type and Free Lunch Eligibility Status on MYS Longitudinal Participation

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Expansion Neighborhood										
Target Neighborhood										
Paid Lunch			***0.64***	0.35	0.75	0.34	***0.68***	0.18	***0.66***	0.25
Free or Reduced Lunch			***0.91***	0.14	0.81	0.10	***0.90***	0.14	***0.79***	0.09

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Expansion Neighborhood					***0.57***	0.16	***0.56***	0.12	0.69	0.21
Target Neighborhood					***0.81***	0.19	***0.75***	0.15	0.83	0.27
Paid Lunch	0.85	0.42	0.80	0.34	***0.63***	0.24	0.63	0.18	0.67	0.33
Free or Reduced Lunch	0.82	0.10	0.82	0.10	***0.77***	0.08	0.69	0.07	0.84	0.09

Note. Boldfaced italicized estimates are statistically significant, $p < .05$.

SAT percentile ranks

Table 43 shows the number of students included in these analyses: those aged 10 through 15 who are in grades 3 through 8 and have either a total reading percentile rank (SAT_{TR}) or a total math percentile rank (SAT_{TM}). Tables 44 and 45 show goodness of fit statistics (χ^2), the significance levels, and the effect sizes (**w**) for the GEE results regressing SAT_{TR} and SAT_{TM}, respectively, on longitudinal participation (main effects and interactions) controlling for demographic characteristics. A power analysis shows power to detect each of the main effects to be greater than or equal to .87 (**f** = .15, α = .05) for both analyses of cognitive ability.

Results demonstrate no significant differences in SAT_{TR} or SAT_{TM} for those participating and those not participating longitudinally in the MYS in any year. Further, there were no consistent departures from representativeness for any main effect or any interaction involving longitudinal participation for either SAT_{TR} or SAT_{TM}.

In this case, it is important to note the lack of significance in any of the interactions that include longitudinal participation as a component. There is no consistent statistical significance in any interaction across measures (SAT_{TR} and SAT_{TM}) and across years. This lack of statistical significance suggests that those who drop out of the MYS are representative of those who do not drop out on cognitive characteristics and that missingness may be ignorable.

Table 43

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods Participating Longitudinally in the MYS with SAT Data

	Year									
	1	2	3	4	5	6	7	8	9	10
With SAT_{TR} Percentile Ranks		344	410	419	590	471	445	612	806	630
With SAT_{TM} Percentile Ranks		345	403	442	593	469	442	606	800	623

Note. For Years 1 through 7, analyses were limited to those living in target neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Table 44

Significance and Effect Size for MYS Longitudinal Participation and Demographic Characteristics on Total Reading Percentile Ranks

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
$P_{t-1,\,t}$	1				1.02	.314	.05	0.01	.918	.00	0.19	.664	.02	0.01	.904	.00
NT	1															
Gr	1				6.17	.013	.13	0.50	.479	.03	0.17	.681	.02	0.52	.471	.03
G	1				1.94	.164	.08	4.63	.032	.11	4.53	.033	.10	4.66	.031	.09
LE	1				0.29	.593	.03	1.05	.306	.05	2.83	.093	.08	0.33	.565	.02
LE × NT	1															
G × NT	1															
$P_{t-1,\,t}$ × G	1				0.23	.633	.03	0.46	.499	.03	0.00	.989	.00	0.01	.934	.00
$P_{t-1,\,t}$ × Gr	1				1.86	.173	.07	0.25	.619	.02	0.31	.577	.03	0.00	.997	.00
$P_{t-1,\,t}$ × LE	1				0.77	.380	.05	1.42	.233	.06	0.81	.369	.04	0.41	.525	.03
$P_{t-1,\,t}$ × NT	1															
$P_{t-1,\,t}$ × NT × Gr	2															
$P_{t-1,\,t}$ × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
$P_{t-1,\,t}$	1	0.18	.670	.02	0.52	.471	.03	0.02	.879	.01	0.57	.449	.03	1.11	.291	.04
NT	1							0.77	.381	.04	0.24	.626	.02	0.15	.703	.02
Gr	1	3.65	.056	.08	1.89	.169	.06	1.35	.246	.05	0.07	.791	.01	3.70	.055	.08
G	1	10.42	.001	.13	7.31	.007	.12	12.34	<.001	.14	11.08	.001	.12	7.26	.007	.11
LE	1	1.22	.269	.05	1.96	.161	.06	0.34	.560	.02	1.37	.242	.04	0.01	.925	.00
LE × NT	1							4.61	.032	.09	0.02	.889	.00	3.02	.083	.07
G × NT	1							0.43	.514	.03	1.61	.204	.04	1.07	.301	.04
$P_{t-1,\,t}$ × G	1	0.89	.345	.04	0.06	.802	.01	1.52	.218	.05	0.19	.667	.02	0.03	.868	.01
$P_{t-1,\,t}$ × Gr	1	0.83	.362	.04	0.28	.598	.02	0.15	.694	.02	2.31	.129	.05	0.58	.445	.03
$P_{t-1,\,t}$ × LE	1	0.52	.469	.03	0.31	.580	.03	1.61	.205	.05	5.59	.018	.08	2.34	.126	.06
$P_{t-1,\,t}$ × NT	1							0.78	.377	.04	0.48	.490	.02	2.25	.134	.06
$P_{t-1,\,t}$ × NT × Gr	2							1.46	.482	.05	0.53	.767	.03	5.84	.054	.10
$P_{t-1,\,t}$ × G × NT	1							0.84	.360	.04	0.39	.534	.02	0.54	.463	.03

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. $P_{t-1,\,t}$ = longitudinal participation; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status.

Table 45

Significance and Effect Size for MYS Longitudinal Participation and Demographic Characteristics on Total Math Percentile Ranks

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
$P_{t-1,\,t}$	1				0.82	.364	.05	0.16	.688	.02	0.00	.977	.00	3.05	.081	.07
NT	1															
Gr	1				3.06	.081	.09	5.88	.015	.12	0.30	.582	.03	6.10	.014	.10
G	1				0.97	.324	.05	0.34	.562	.03	0.44	.509	.03	6.16	.013	.10
LE	1				0.49	.486	.04	0.16	.688	.02	0.94	.333	.05	0.96	.328	.04
LE × NT	1															
G × NT	1															
$P_{t-1,\,t}$ × G	1				0.61	.437	.04	0.12	.730	.02	0.00	.946	.00	2.22	.137	.06
$P_{t-1,\,t}$ × Gr	1				1.15	.283	.06	0.43	.512	.03	0.12	.730	.02	3.02	.082	.07
$P_{t-1,\,t}$ × LE	1				0.00	.976	.00	0.14	.704	.02	0.48	.489	.03	0.03	.866	.01
$P_{t-1,\,t}$ × NT	1															
$P_{t-1,\,t}$ × NT × Gr	2															
$P_{t-1,\,t}$ × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
$P_{t-1,\,t}$	1	0.55	.457	.03	0.13	.720	.02	0.46	.499	.03	0.47	.494	.02	0.04	.833	.01
NT	1							0.40	.528	.03	1.53	.216	.04	0.17	.680	.02
Gr	1	0.40	.527	.03	0.01	.903	.00	3.64	.056	.08	1.62	.203	.05	5.11	.024	.09
G	1	7.92	.005	.13	2.29	.130	.07	4.53	.033	.09	5.05	.025	.08	7.98	.005	.11
LE	1	0.40	.529	.03	0.67	.413	.04	4.75	.029	.09	0.98	.323	.04	0.37	.541	.02
LE × NT	1							2.46	.117	.06	0.05	.829	.01	0.90	.344	.04
G × NT	1							1.36	.243	.05	0.71	.401	.03	1.94	.164	.06
$P_{t-1,\,t}$ × G	1	0.79	.373	.04	0.08	.781	.01	0.31	.581	.02	0.32	.569	.02	0.91	.339	.04
$P_{t-1,\,t}$ × Gr	1	0.91	.341	.04	0.00	.954	.00	0.83	.363	.04	0.71	.400	.03	0.01	.919	.00
$P_{t-1,\,t}$ × LE	1	0.37	.541	.03	1.54	.215	.06	0.89	.346	.04	0.07	.794	.01	0.68	.410	.03
$P_{t-1,\,t}$ × NT	1							0.52	.471	.03	0.03	.868	.01	1.59	.207	.05
$P_{t-1,\,t}$ × NT × Gr	2							0.67	.715	.03	1.94	.379	.05	3.16	.206	.07
$P_{t-1,\,t}$ × G × NT	1							0.11	.743	.01	0.38	.537	.02	0.01	.918	.00

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. $P_{t-1,\,t}$ = longitudinal participation; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status.

School violations and disciplinary actions

Table 46 shows the number of students included in analyses comparing the weighted school violation scores (WSV) of and weighted disciplinary action scores (WDA) for dropouts and non-dropouts. Tables 47 and 48 show goodness of fit statistics (χ^2), the significance levels, and the effect sizes (**w**) for the GEE results regressing WSV and WDA, respectively, on longitudinal participation (main effects and interactions) controlling for demographic characteristics. A power analysis shows power to detect each of the main effects to be greater than or equal to .995 (**f** = .15, α = .05) on both analyses of behavior.

Results demonstrate no significant main effects for longitudinal participation on either WSV or WDA in any year; however, results demonstrate a significant longitudinal participation × free lunch eligibility status interaction for WSV scores in Year 8[40]. Because this effect was only statistically significant for one year, it is not consistent across measures, and the effect size is small (**w** = .10), it is not indicative of non-ignorable missingness. Results demonstrate no significant interactions for WDA.

While it is important to describe the statistically significant results, it is also important to describe the lack of significance in the main effect of yearly participation or in most interaction effects with enrollment as a component across years and across measures (WSV and WDA). Even in the case of the statistical significance found in the longitudinal participation × free lunch interaction, significance was not found across years and measures. These results suggest that with respect to these behavioral characteristics, those who drop out of the MYS are similar to those who do not drop out, and fails to supports the conclusion of non-ignorable missingness.

[40] For this result, because it only occurred in one year, estimates are not reported in any table.

Table 46

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods Participating Longitudinally in the MYS with School Violation and Disciplinary Action Data

	Year									
	1	2	3	4	5	6	7	8	9	10
n		728	775	773	949	785	710	1,082	1,435	1,024

Note. For Years 1 through 7, analyses were limited to students living in target neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Table 47

Significance and Effect Size for MYS Longitudinal Participation and Demographic Characteristics on Weighted School Violation Scores

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
$P_{t-1,\,t}$	1				0.00	.994	.00	0.04	.835	.01	5.88	.015	.09	3.61	.057	.06
NT	1															
Gr	1				9.67	.002	.12	13.19	<.001	.13	4.55	.033	.08	12.32	<.001	.11
G	1				5.04	.025	.08	11.48	.001	.12	7.66	.006	.10	8.87	.003	.10
LE	1				0.16	.685	.01	0.66	.415	.03	4.05	.044	.07	2.26	.132	.05
LE × NT	1															
G × NT	1															
$P_{t-1,\,t}$ × G	1				1.33	.249	.04	0.61	.434	.03	0.21	.648	.02	0.29	.592	.02
$P_{t-1,\,t}$ × Gr	1				0.01	.923	.00	0.18	.669	.02	2.49	.115	.06	3.77	.052	.06
$P_{t-1,\,t}$ × LE	1				0.37	.544	.02	1.06	.302	.04	1.05	.305	.04	0.01	.928	.00
$P_{t-1,\,t}$ × NT	1															
$P_{t-1,\,t}$ × NT × Gr	2															
$P_{t-1,\,t}$ × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
$P_{t-1,\,t}$	1	0.59	.441	.03	0.09	.760	.01	7.01	.008	.08	0.10	.752	.01	2.51	.113	.05
NT	1							4.64	.031	.07	1.34	.247	.03	2.47	.116	.05
Gr	1	1.62	.203	.05	1.71	.191	.05	0.15	.694	.01	0.32	.572	.01	8.14	.004	.09
G	1	18.88	<.001	.16	14.10	<.001	.14	7.68	.006	.08	23.35	<.001	.13	7.63	.006	.09
LE	1	0.75	.386	.03	0.86	0.353	.03	22.30	<.001	.14	0.86	.353	.02	1.12	.290	.03
LE × NT	1							0.03	.853	.01	0.10	.752	.01	0.73	.392	.03
G × NT	1							0.19	.666	.01	0.03	.864	.00	0.63	.428	.02
$P_{t-1,\,t}$ × G	1	1.29	.256	.04	1.22	.270	.04	0.53	.469	.02	0.13	.716	.01	0.23	.630	.01
$P_{t-1,\,t}$ × Gr	1	0.17	.684	.01	0.25	.616	.02	0.55	.458	.02	0.53	.468	.02	0.89	.344	.03
$P_{t-1,\,t}$ × LE	1	0.45	.501	.02	0.00	.977	.00	9.93	.002	.10	0.28	.594	.01	1.16	.281	.03
$P_{t-1,\,t}$ × NT	1							2.24	.135	.05	0.08	.772	.01	3.16	.075	.06
$P_{t-1,\,t}$ × NT × Gr	2							2.68	.262	.05	1.43	.489	.03	2.56	.278	.05
$P_{t-1,\,t}$ × G × NT	1							0.61	.437	.02	1.82	.177	.04	1.35	.245	.04

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. $P_{t-1,\,t}$ = longitudinal participation; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status.

Table 48

Significance and Effect Size for MYS Longitudinal Participation and Demographic Characteristics on Weighted Disciplinary Action Scores

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
$P_{t-1,\,t}$	1				0.05	.818	.01	0.30	.582	.02	2.37	.124	.06	3.28	.070	.06
NT	1															
Gr	1				8.92	.003	.11	24.39	<.001	.18	5.36	.021	.08	13.86	<.001	.12
G	1				6.43	.011	.09	14.45	<.001	.14	9.80	.002	.11	8.99	.003	.10
LE	1				0.08	.772	.01	1.07	.301	.04	2.59	.108	.06	2.29	.130	.05
LE × NT	1															
G × NT	1															
$P_{t-1,\,t}$ × G	1				2.58	.108	.06	2.08	.150	.05	1.41	.235	.04	0.11	.744	.01
$P_{t-1,\,t}$ × Gr	1				0.32	.574	.02	0.08	.777	.01	0.35	.556	.02	3.42	.065	.06
$P_{t-1,\,t}$ × LE	1				0.02	.894	.01	0.35	.555	.02	1.31	.252	.04	0.02	.883	.00
$P_{t-1,\,t}$ × NT	1															
$P_{t-1,\,t}$ × NT × Gr	2															
$P_{t-1,\,t}$ × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
$P_{t-1,\,t}$	1	0.42	.515	.02	0.10	.758	.01	2.92	.087	.05	0.04	.848	.01	0.64	.424	.03
NT	1							2.37	.123	.05	1.38	.240	.03	1.04	.307	.03
Gr	1	0.40	.525	.02	1.19	.276	.04	0.00	.983	.00	5.18	.023	.06	0.55	.457	.02
G	1	17.54	<.001	.15	13.44	<.001	.14	5.80	.016	.07	28.96	<.001	.14	8.56	.003	.09
LE	1	1.45	.229	.04	0.06	.812	.01	12.58	<.001	.11	1.23	.268	.03	3.55	.060	.06
LE × NT	1							0.01	.923	.00	0.14	.711	.01	0.37	.544	.02
G × NT	1							0.03	.869	.01	1.09	.296	.03	1.23	.267	.03
$P_{t-1,\,t}$ × G	1	1.07	.301	.04	2.08	.149	.05	0.34	.558	.02	0.57	.450	.02	1.07	.300	.03
$P_{t-1,\,t}$ × Gr	1	0.24	.626	.02	1.01	.314	.04	0.00	.980	.00	0.13	.722	.01	0.07	.795	.01
$P_{t-1,\,t}$ × LE	1	0.02	.886	.01	0.28	.597	.02	7.64	.006	.08	0.01	.906	.00	0.32	.572	.02
$P_{t-1,\,t}$ × NT	1							0.41	.520	.02	0.04	.833	.01	1.56	.212	.04
$P_{t-1,\,t}$ × NT × Gr	2							1.35	.508	.04	3.02	.221	.05	1.17	.558	.03
$P_{t-1,\,t}$ × G × NT	1							0.66	.417	.02	3.91	.048	.05	4.92	.027	.07

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. $P_{t-1,\,t}$ = longitudinal participation; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status.

Missing data patterns (MP)

The first sub-analysis for Research Question 2 concerns missing data patterns of dropouts. Table 49 shows the number of students enrolled in the MYS each year who did not have complete data (i.e., dropouts) and the patterns of their dropout. As described in Chapter 2 and in Table 2, monotonic missingness can be described as dropping out of a study and never participating again (when eligible to participate). Non-monotonic missingness can be described as dropping out and then dropping back into a study.

Table 49

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods with MYS Between-Wave Missingness

	Year									
	1	2	3	4	5	6	7	8	9	10
With Monotone Missingness	1,039	997	853	833	788	645	503	826	772	284
With Non-monotone Missingness	518	640	503	667	687	582	458	773	665	375

Note. For Years 1 through 7, analyses were limited to those living in target neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Demographics. Table 50 shows the goodness of fit statistics (χ^2), significance levels, and effect sizes (**w**) for the GEE results regressing missing data pattern type on demographic variables. A power analysis shows power to detect each of the main effects to be greater than or equal to .993 (**f** = .15, α = .05).

Results demonstrate that there was a significant departure from representativeness in grade level for multiple years (Years 1 through 4 and Years 8 through 10). Table 51 shows an odds ratio for grade level, which are, for Years 1 through 4, less than 1.0 indicating a negative relationship between missing data pattern and grade level. That is, in the earlier years of the MYS, older students were more likely to drop out and stay dropped out of the MYS. In contrast, the odds ratio for grade level in Years 8 through 10 is greater than 1.0 indicating a positive

relationship between missing data pattern and grade level. That is, in the later years of the MYS, older students were more likely to drop out and drop back into the MYS. The inconsistency in direction in the results raises questions about the meaning of these findings as they relate to ignorable missingness. However, because effect sizes are primarily in the medium range ($.10 \leq \mathbf{w} \leq .32$), the effect should not be ignored.

It is important to also consider the main effects and interactions that are not statistically significant. There were no consistent departures from representativeness (in three or more years) for gender, free lunch eligibility status, neighborhood type, or either interaction effect (free lunch status × neighborhood type and gender × neighborhood type) included in this analysis of demographic representativeness. This lack of significance does suggest that those who drop out of the MYS, but have different patterns of dropout (monotonic and non-monotonic) are similar in terms of demographic characteristics. These results suggest that missing data patterns (monotone vs. non-monotone) are not related to demographic characteristics.

Table 50

Significance and Effect Size for Demographic Characteristics on MYS Missing Data Patterns

		Year														
	df	1			2			3			4			5		
		χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
G	1	0.67	.413	.02	3.77	.052	.05	0.90	.343	.03	6.64	.010	.07	3.81	.051	.01
LE	1	1.64	.200	.03	0.91	.341	.02	0.15	.701	.01	0.04	.839	.01	0.55	.457	.02
NT	1															
LE × NT	2															
G × NT	1															
Gr	1	162.11	<.001	.32	140.49	<.001	.29	91.45	<.001	.26	40.98	<.001	.17	0.70	.402	.02

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
G	1	2.62	.105	.05	1.01	.314	.03	0.00	.993	.00	0.41	.523	.02	0.00	.999	.00
LE	1	0.79	.375	.03	3.74	.053	.06	0.97	.325	.02	0.40	.526	.02	1.26	.262	.04
NT	1							1.86	.173	.03	0.65	.420	.02	1.47	.225	.05
LE × NT	2							0.99	.320	.02	0.07	.798	.01	0.37	.544	.02
G × NT	1							0.77	.379	.02	0.34	.560	.02	0.04	.846	.01
Gr	1	0.01	.935	.00	0.85	.356	.03	16.41	<.001	.10	25.49	<.001	.13	23.01	<.001	.19

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. G = gender; LE = free lunch eligibility status; NT = neighborhood type; Gr = grade level.

Table 51

Odds Ratios and Standard Errors for Grade Level on MYS Missing Data Patterns

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Grade Level	***0.63***	0.04	***0.69***	0.03	***0.73***	0.03	***0.83***	0.03	0.97	0.03

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Grade Level	1.0	0.03	1.04	0.04	***1.13***	0.03	***1.22***	0.04	***1.39***	0.07

Note. Boldfaced italicized estimates are statistically significant, $p < .05$.

SAT percentile ranks. Table 52 shows the number of students included in these analyses: those aged 10 through 15 who are in grades 3 through 8 and have either a total reading percentile rank (SAT_{TR}) or a total math percentile rank (SAT_{TM}) and have dropped out of the MYS. Tables 53 and 54 show goodness of fit statistics (χ^2), the significance levels, and the effect sizes (**w**) for the GEE results regressing SAT_{TR} and SAT_{TM}, respectively, on missing data pattern type (main effects and interactions) controlling for demographic characteristics. A power analysis shows power to detect each of the main effects to be greater than or equal to .82 (**f** = .15, α = .05) for both analyses of cognitive ability.

Results demonstrate no significant difference in SAT_{TR} or SAT_{TM} for those who dropped out monotonically versus those who dropped out non-monotonically. Further, there were no consistent departures from representativeness for any main effect or any interaction involving monotonic or non-monotonic missingness either SAT_{TR} or SAT_{TM}.

It is important to note that missing data pattern as a main effect is not significant. Further, there is a lack of significance in any of the interactions that include missing data pattern as a component. There is no statistical significance in any interaction across measures (SAT_{TR} and SAT_{TM}) and across years. This lack of statistical significance suggests that those who drop out and have different missing data patterns (monotonic and non-monotonic) are similar on cognitive characteristics. These results suggest that missing data patterns (monotone vs. non-monotone) are not related to cognitive outcomes.

Table 52

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods with MYS Between Wave Missing Data and SAT Data

	Year									
	1	2	3	4	5	6	7	8	9	10
With SAT_{TR} Percentile Ranks	889	881	799	892	979	748	578	845	675	304
With SAT_{TM} Percentile Ranks	898	893	789	929	988	743	573	838	672	300

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Table 53

Significance and Effect Size for MYS Missing Data Patterns and Demographic Characteristics on Total Reading Percentile Ranks

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
MP	1	3.27	.071	.06	0.00	.987	.00	0.73	.394	.03	0.61	.434	.03	0.81	.369	.03
NT	1															
Gr	1	0.71	.399	.03	3.28	.070	.06	1.11	.292	.04	1.74	.188	.04	1.92	.166	.04
G	1	8.18	.004	.10	7.81	.005	.09	4.81	.028	.08	6.69	.010	.09	15.33	<.001	.13
LE	1	2.86	.091	.06	0.38	.540	.02	0.14	.706	.01	0.70	.403	.03	0.26	.613	.02
LE × NT	1															
G × NT	1															
MP × G	1	0.67	.414	.03	1.31	.252	.04	0.07	.792	.01	0.32	.572	.02	0.45	.504	.02
MP × Gr	1	1.20	.273	.04	0.18	.668	.01	0.40	.529	.02	1.41	.235	.04	0.94	.332	.03
MP × LE	1	2.32	.127	.05	0.28	.594	.02	0.22	.639	.02	2.47	.116	.05	0.11	.738	.01
MP × NT	1															
MP × NT × Gr	2															
MP × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
MP	1	0.04	.838	.01	0.10	.757	.01	4.48	.034	.07	2.81	.094	.06	1.75	.186	.08
NT	1							0.04	.849	.01	3.80	.051	.08	0.77	.381	.05
Gr	1	0.12	.732	.01	0.78	.377	.04	0.84	.361	.03	0.01	.931	.00	0.23	.632	.03
G	1	20.75	<.001	.17	20.24	<.001	.19	14.96	<.001	.13	16.01	<.001	.15	13.08	<.001	.21
LE	1	3.54	.060	.07	2.10	.148	.06	0.36	.549	.02	0.37	.544	.02	0.82	.366	.05
LE × NT	1							1.45	.229	.04	3.27	.071	.07	0.29	.588	.03
G × NT	1							0.10	.757	.01	1.50	.221	.05	1.83	.177	.08
MP × G	1	2.29	.131	.06	0.93	.334	.04	0.38	.538	.02	0.55	.458	.03	0.15	.699	.02
MP × Gr	1	0.33	.565	.02	0.06	.805	.01	4.79	.029	.08	1.32	.251	.04	1.67	.196	.07
MP × LE	1	1.67	.197	.05	0.20	.652	.02	0.08	.771	.01	1.95	.162	.05	0.13	.718	.02
MP × NT	1							0.01	.905	.00	0.06	.805	.01	0.44	.505	.04
MP × NT × Gr	2							0.04	.981	.01	2.25	.325	.06	1.28	.528	.06
MP × G × NT	1							0.20	.656	.02	0.38	.536	.02	0.10	.753	.02

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. MP = missing data pattern; NT = neighborhood type; Gr = grade level, G = gender; LE = free lunch eligibility status.

Table 54

Significance and Effect Size for MYS Missing Data Patterns and Demographic Characteristics on Total Math Percentile Ranks

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
MP	1	1.40	.237	.04	0.02	.885	.00	0.50	.480	.03	0.00	.957	.00	0.03	.874	.01
NT	1															
Gr	1	1.54	.215	.04	0.53	.469	.02	24.11	<.001	.17	20.12	<.001	.15	7.65	.006	.09
G	1	5.08	.024	.08	0.84	.360	.03	0.40	.525	.02	3.69	.055	.06	8.05	.005	.09
LE	1	1.04	.308	.03	0.12	.727	.01	0.14	.713	.01	0.08	.777	.01	0.37	.542	.02
LE × NT	1															
G × NT	1															
MP × G	1	0.02	.897	.00	0.59	.441	.03	0.06	.804	.01	2.24	.135	.05	0.68	.409	.03
MP × Gr	1	0.28	.598	.02	0.86	.354	.03	0.13	.720	.01	0.10	.753	.01	0.44	.506	.02
MP × LE	1	1.75	.185	.04	0.63	.429	.03	0.43	.512	.02	3.92	.048	.06	3.08	.079	.06
MP × NT	1															
MP × NT × Gr	2															
MP × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
MP	1	0.96	.328	.04	0.96	.328	.04	0.75	.386	.03	0.36	.550	.02	0.76	.384	.05
NT	1							0.02	.894	.00	0.03	.863	.01	0.67	.413	.05
Gr	1	9.74	.002	.11	9.74	.002	.13	14.23	<.001	.13	1.12	.290	.04	0.40	.528	.04
G	1	8.71	.003	.11	8.71	.003	.12	8.36	.004	.10	19.46	<.001	.17	16.11	<.001	.23
LE	1	0.00	.985	.00	0.00	.985	.00	0.98	.321	.03	0.45	.504	.03	0.33	.563	.03
LE × NT	1							2.26	.133	.05	0.08	.777	.01	0.55	.458	.04
G × NT	1							0.18	.669	.01	0.81	.368	.03	2.55	.111	.09
MP × G	1	2.14	.144	.05	2.14	.144	.06	2.47	.116	.05	0.01	.942	.00	0.12	.732	.02
MP × Gr	1	0.44	.508	.02	0.44	.508	.03	0.21	.645	.02	0.09	.761	.01	0.53	.468	.04
MP × LE	1	0.83	.361	.03	0.83	.361	.04	1.55	.214	.04	7.66	.006	.11	0.00	.973	.00
MP × NT	1							0.51	.474	.02	0.24	.622	.02	0.71	.399	.05
MP × NT × Gr	2							0.64	.726	.03	0.88	.643	.04	0.89	.642	.05
MP × G × NT	1							0.14	.711	.01	0.01	.916	.00	0.01	.927	.01

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. MP = missing data pattern; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status.

School violations and disciplinary actions. Table 55 shows the number of students included in analyses comparing the weighted school violation scores (WSV) of and weighted disciplinary action scores (WDA) for dropouts and non-dropouts. Tables 56 and 58 show goodness of fit statistics (χ^2), the significance levels, and the effect sizes (**w**) for the GEE results regressing WSV and WDA, respectively, on missing data pattern type (main effects and interactions) controlling for demographic characteristics. A power analysis shows power to detect each of the main effects to be greater than or equal to .993 (**f** = .15, α = .05).

Results demonstrate that there were significant differences in WSV for those who had monotonic missing patterns and those who had non-monotonic missing patterns in Year 10[41]. Because this effect was only statistically significant for one year, was not significant across measures, and the effect size is small (**w** = .13), it does not suggest that differences in dropout pattern are related to behavioral characteristics. Further, results demonstrate a significant missing data pattern type × free lunch eligibility status interaction in WSV (Table 57) for years 4, 6, and 10. In this interaction, those with monotonic missingness and who qualify for free lunch have lower WSV scores than those who with monotonic missingness who do not qualify for free lunch. In contrast, those with non-monotonic missingness and who qualify for free lunch have higher WSV scores than those with non-monotonic missingness and who do not qualify for free lunch. While the general pattern is consistent across years, the slopes of these effects are not consistent. Further, this interaction was not significant across measures. Finally, effect sizes, even for the specified years, were small to very small ($.05 \leq \mathbf{w} \leq .09$). Results demonstrate no significant differences (main effects or interactions) in WDA.

While it is important to describe the statistically significant results, it is important to note that missing data pattern as a main effect is not consistently significant across years or across

[41] For this result, because it only occurred in one year, estimates are not reported in any table.

measures. Further, there is a lack of significance in many of the interactions that include missing data pattern as a component. There is also no statistical significance in any interaction across measures (WSV and WDA) and across years. This lack of statistical significance suggests that those who drop out and have different missing data patterns (monotonic and non-monotonic) are similar on behavioral characteristics. These results suggest that missing data patterns (monotone vs. non-monotone) are not related to behavioral outcomes.

Table 55

Sample Size: MCPSS Students (Ages 10 through 15) Living in MYS Neighborhoods with MYS Between Wave Missing Data and with School Violation and Disciplinary Action Data

	Year									
	1	2	3	4	5	6	7	8	9	10
n		1,637	1,356	1,500	1,475	1,227	961	1,599	1,437	659

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Table 56

Significance and Effect Size for MYS Missing Data Patterns and Demographic Characteristics on Weighted School Violation Scores

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
MP	1				1.88	.171	.03	0.21	.650	.01	1.56	.211	.03	2.67	.102	.04
NT	1															
Gr	1				83.82	<.001	.23	35.28	<.001	.16	13.40	<.001	.09	39.49	<.001	.16
G	1				22.12	<.001	.12	26.36	<.001	.14	44.00	<.001	.17	19.83	<.001	.12
LE	1				0.00	.968	.00	2.73	.099	.04	2.12	.146	.04	0.93	.335	.03
LE × NT	1															
G × NT	1															
MP × G	1				0.05	.827	.01	0.00	.989	.00	2.18	.140	.04	5.85	.016	.06
MP × Gr	1				4.88	.027	.05	0.19	.663	.01	1.98	.159	.04	2.98	.084	.04
MP × LE	1				0.01	.935	.00	0.00	.986	.00	3.93	.047	.05	0.43	.511	.02
MP × NT	1															
MP × NT × Gr	2															
MP × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
MP	1	0.00	.998	.00	0.02	.886	.00	0.62	.431	.02	3.13	.077	.05	11.01	.001	.13
NT	1							4.84	.028	.06	2.39	.122	.04	0.56	.455	.03
Gr	1	2.34	.126	.04	1.10	.294	.03	1.15	.284	.03	4.77	.029	.06	0.49	.484	.03
G	1	33.84	<.001	.17	41.66	<.001	.21	34.11	<.001	.15	24.61	<.001	.13	14.43	<.001	.15
LE	1	0.37	.541	.02	0.06	.802	.01	12.67	<.001	.09	1.42	.234	.03	3.12	.077	.07
LE × NT	1							0.03	.855	.00	2.83	.093	.04	1.54	.215	.05
G × NT	1							1.17	.279	.03	0.45	.504	.02	0.11	.744	.01
MP × G	1	3.11	.078	.05	1.40	.237	.04	1.67	.196	.03	0.95	.330	.03	0.63	.427	.03
MP × Gr	1	0.37	.543	.02	0.20	.654	.01	1.47	.225	.03	0.56	.455	.02	6.15	.013	.10
MP × LE	1	4.10	.043	.06	1.30	.254	.04	0.62	.433	.02	2.52	.112	.04	5.92	.015	.09
MP × NT	1							0.55	.459	.02	0.00	.988	.00	6.15	.013	.10
MP × NT × Gr	2							4.68	.096	.05	0.92	.633	.03	7.09	.029	.10
MP × G × NT	1							1.82	.178	.03	1.45	.228	.03	0.25	.616	.02

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. MP = missing data pattern; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status.

Table 57

Adjusted Estimates and Standard Errors for MYS Missing Data Patterns × Free Lunch Eligibility Status on Weighted School Violation Scores

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Monotone Missingness, Paid Lunch			2.04	0.24	3.75	0.21	***3.39***	0.11	5.03	0.17
Monotone Missingness, Free/Reduced Lunch			2.06	0.07	2.64	0.08	***3.55***	0.07	4.10	0.07
Non-Monotone Missingness, Paid Lunch			2.38	0.26	3.89	0.28	***4.16***	0.12	4.47	0.17
Non-Monotone Missingness, Free/Reduced Lunch			2.34	0.08	2.75	0.10	***3.00***	0.07	4.28	0.07

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Monotone Missingness, Paid Lunch	***4.43***	0.19	5.26	0.19	4.15	0.16	5.52	0.11	***4.11***	0.34
Monotone Missingness, Free/ Reduced Lunch	***5.35***	0.06	5.95	0.10	5.78	0.05	6.91	0.06	***8.34***	0.07
Non-Monotone Missingness, Paid Lunch	***7.00***	0.17	6.64	0.20	3.45	0.23	8.94	0.12	***11.73***	0.17
Non-Monotone Missingness, Free/Reduced Lunch	***4.95***	0.06	5.48	0.10	5.95	0.06	8.74	0.05	***10.88***	0.07

Note. Boldfaced italicized estimates are statistically significant, $p < .05$.

Table 58

Significance and Effect Size of MYS Missing Data Patterns and Demographic Characteristics on Weighted Disciplinary Action Scores

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
MP	1				0.13	.720	.01	0.78	.376	.02	0.68	.411	.02	4.71	.030	.06
NT	1															
Gr	1				77.76	<.001	.22	58.70	<.001	.21	44.54	<.001	.17	45.33	<.001	.18
G	1				19.37	<.001	.11	23.55	<.001	.13	37.04	<.001	.16	17.81	<.001	.11
LE	1				0.01	.916	.00	3.10	.078	.05	0.76	.384	.02	0.91	.340	.02
LE × NT	1															
G × NT	1															
MP × G	1				0.01	.931	.00	0.18	.673	.01	1.50	.220	.03	5.94	.015	.06
MP × Gr	1				1.82	.177	.03	1.00	.316	.03	1.06	.304	.03	6.32	.012	.07
MP × LE	1				0.54	462	.02	0.00	.993	.00	2.75	.097	.04	0.00	.999	.00
MP × NT	1															
MP × NT × Gr	2															
MP × G × NT	1															

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
MP	1	0.07	.798	.01	0.06	.809	.01	0.48	.489	.02	3.99	.046	.05	9.57	.002	.12
NT	1							4.78	.029	.05	3.16	.075	.05	0.27	.605	.02
Gr	1	2.71	.100	.05	2.54	.111	.05	1.05	.306	.03	15.76	<.001	.10	2.05	.152	.06
G	1	32.18	<.001	.16	37.40	<.001	.20	31.92	<.001	.14	30.42	<.001	.15	8.49	.004	.11
LE	1	0.50	.482	.02	0.12	.729	.01	8.71	.003	.07	2.55	.111	.04	2.06	.151	.06
LE × NT	1							0.00	.982	.00	0.73	.392	.02	0.70	.403	.03
G × NT	1							0.97	.326	.02	0.21	.645	.01	0.16	.686	.02
MP × G	1	3.29	.070	.05	3.17	.075	.06	2.81	.094	.04	0.51	.474	.02	0.85	.355	.04
MP × Gr	1	0.09	.765	.01	0.01	.942	.00	1.32	.250	.03	1.35	.245	.03	5.28	.022	.09
MP × LE	1	4.60	.032	.06	2.01	.156	.05	1.05	.305	.03	2.19	.139	.04	3.73	.054	.08
MP × NT	1							0.28	.594	.01	0.23	.635	.01	2.08	.149	.06
MP × NT × Gr	2							4.31	.116	.05	2.99	.225	.05	2.84	.242	.07
MP × G × NT	1							0.54	.464	.02	0.81	.368	.02	0.22	.641	.02

Note. Analysis is limited to those aged 10 through 15 living in MYS Neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. MP = missing data pattern; NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status.

Alternative Proxy Measures

Thus far, in Research Question 2, school records have been used as a proxy to measure characteristics of dropouts. In this analysis, MYS responses at $Time_{t-1}$ were used as a proxy to determine characteristics of dropouts. There were two parts of this analysis. First, dropout was regressed on demographic characteristics and the previously listed MYS variables measured at $Time_{t-1}$ using 10 through 18 year olds who dropped out of the MYS (living in MYS neighborhoods) as the sample. A power analysis shows power to detect each of the main effects to be greater than or equal to .995 (**f** = .15, α = .05). Second, dropout was regressed on demographic characteristics and the previously listed MYS variables measured at $Time_{t-1}$ using 16 through 18 year olds who dropped out of the MYS (living in MYS neighborhoods) as the sample. A power analysis shows power to detect each of the main effects to be greater than or equal to .61 (**f** = .15, α = .05).

Analysis 1: 10 through 18 year olds. Table 59 shows the number of students (ages 10 through 18) who did not and who did drop out of the MYS in two consecutive years. Table 60 shows the results of these analyses for each year. Results demonstrate that there was a significant departure from representativeness in length of time spent in residence (residential tenure) for multiple years (Years 3, 6, and 7). Table 61 shows odds ratios for this variable, which are, for all years, greater than 1.0, indicating a positive relationship between longitudinal participation and residential tenure. That is, those who have spent more time living in a neighborhood are less likely to drop out. Effect sizes, however, are small ($.07 \leq \mathbf{w} \leq .11$). Notably, no risk behaviors are associated with dropout. These results taken together indicate that there those who drop out of the MYS and those who do not drop out of the MYS are not different

with respect to these variables, failing to suggest that missingness is non-ignorable (i.e., not related to outcomes).

Table 59

Sample Size: MYS Participants (Ages 10 through 18) by Longitudinal Participation Status

	Year									
	1	2	3	4	5	6	7	8	9	10
$P_{t-1,\,t} = 1$		772	732	849	877	751	724	1,065	1,244	1,059
$P_{t-1,\,t} = 0$		100	181	165	288	182	167	381	644	293

Note. Analysis in Years 1 through 7 is limited to those living in target neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Table 60

Significance and Effect Size for Selected MYS Variables and Demographic Characteristics on MYS Longitudinal Participation (Ages 10 through 18)

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
NT	1															
Gr	1				0.34	.558	.02	1.99	.158	.05	0.08	.772	.01	2.42	.119	.05
G	1				0.02	.899	.00	0.55	.458	.02	0.38	.535	.02	9.06	.003	.09
LE	1				9.87	.002	.11	1.84	.176	.04	41.04	<.001	.20	7.96	.005	.08
LE × NT	1															
G × NT	1															
Residence	1				2.79	.095	.06	11.82	.001	.11	1.91	.167	.04	3.28	.070	.05
Arrest	1				1.62	.203	.04	0.00	.970	.00	0.07	.789	.01	0.21	.647	.01
Gun	1				2.23	.135	.05	0.02	.881	.00	0.23	.635	.02	0.08	.780	.01
Marijuana	1				0.63	.426	.03	0.06	.801	.01	0.36	.546	.02	0.06	.810	.01

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*P*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
NT	1							33.28	<.001	.15	34.24	<.001	.13	3.70	.054	.05
Gr	1	1.52	.217	.04	1.34	.247	.04	3.44	.064	.05	6.04	.014	.06	8.01	.005	.08
G	1	1.14	.287	.04	0.00	.985	.00	2.19	.139	.04	0.06	.800	.01	0.18	.676	.01
LE	1	0.34	.560	.02	1.27	.259	.04	17.98	<.001	.11	2.16	.141	.03	8.50	.004	.08
LE × NT	1							7.13	.008	.07	0.65	.420	.02	2.42	.120	.04
G × NT	1							2.56	.109	.04	5.64	.018	.05	1.85	.174	.04
Residence	1	5.14	.023	.07	3.91	.048	.07	2.43	.119	.04	0.08	.781	.01	3.73	.054	.05
Arrest	1	0.75	.385	.03	0.00	.961	.00	0.31	.579	.01	0.18	.674	.01	0.42	.518	.02
Gun	1	0.09	.765	.01	0.50	.478	.02	0.95	.330	.03	0.52	.470	.02	0.57	.452	.02
Marijuana	1	0.16	.688	.01	0.28	.595	.02	4.33	.037	.05	0.18	.668	.01	0.03	.864	.00

Note. Analysis is limited to those aged 10 through 18 living in MYS Neighborhoods. Analysis in Years 1 through 7 is limited to those living in target neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status.

Table 61

Odds Ratios and Standard Errors for Residential Tenure on MYS Longitudinal Participation

	Year									
	1		2		3		4		5	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Residential Tenure			1.11	0.06	***1.18***	0.04	1.08	0.05	1.08	0.04

	Year									
	6		7		8		9		10	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Residential Tenure	***1.12***	0.05	***1.11***	0.05	1.06	0.04	1.01	0.03	1.08	0.04

Note. Boldfaced italicized estimates are statistically significant, $p < .05$.

Analysis 2: 16 through 18 year olds. Table 62 shows the number of students (ages 16 through 18) who did not and who did drop out of the MYS in two consecutive years. Table 63 shows the results of these analyses for each year. Results demonstrate no significant differences in any of the MYS variables for those who did and those who did not longitudinally participate in the MYS. This lack of significance indicates that there are no differences in these characteristics among dropouts and non-dropouts. First, this lack of significance indicates that for those older adolescents who participate in the MYS, residential tenure is not associated with dropout. Next, this lack of significance indicates that those older adolescents who drop out of the MYS and those who do not drop out of the MYS are not different with regard to risk behaviors, residential tenure, or time spent in paid employment. While a difference in dropouts and non-dropouts for time spent in paid employment might suggest a missing at random mechanism (i.e., the missingness is not related to risk outcomes), the lack of difference does not undermine the assumption of an ignorable missing data mechanism.

Table 62

Sample Size: MYS Participants (Ages 16 through 18) by Longitudinal Participation Status

	Year									
	1	2	3	4	5	6	7	8	9	10
$P_{t-1,\,t} = 1$		158	152	210	180	147	189	280	311	247
$P_{t-1,\,t} = 0$		30	46	58	84	48	51	131	198	118

Note. Analysis in Years 1 through 7 is limited to those living in target neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2.

Table 63

Significance and Effect Size for Selected MYS Variables and Demographic Characteristics on MYS Longitudinal Participation (Ages 16 through 18)

		Year														
		1			2			3			4			5		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
NT	1															
Gr	1				1.32	.251	.08	5.23	.022	.16	0.60	.439	.05	0.03	.866	0.01
G	1				0.17	.682	.03	0.01	.922	.01	0.01	.915	.01	1.20	.273	0.07
LE	1				3.67	.055	.14	3.47	.063	.13	12.88	<.001	.22	2.07	.151	0.09
LE × NT	1															
G × NT	1															
Residence	1				0.13	.714	.03	2.40	.121	.11	1.30	.254	.07	1.28	.257	0.07
Arrest	1				0.14	.711	.03	1.19	.275	.08	0.67	.413	.05	0.41	.523	0.04
Work	1				0.32	.574	.04	1.78	.183	.09	0.01	.930	.01	0.07	.791	0.02
Gun	1				0.02	.897	.01	0.00	.972	.00	0.60	.438	.05	0.87	.351	0.06
Marijuana	1				0.01	.918	.01	0.13	.716	.03	0.88	.348	.06	0.49	.482	0.04

		Year														
		6			7			8			9			10		
	df	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**	χ^2	*p*	**w**
NT	1							19.49	<.001	.22	8.08	.005	.13	0.21	.649	.02
Gr	1	0.05	.819	.02	1.88	.170	.09	0.40	.528	.03	2.74	.098	.07	0.01	.921	.01
G	1	6.96	.008	.19	0.49	.483	.05	3.46	.063	.09	0.29	.590	.02	0.95	.329	.05
LE	1	3.00	.083	.12	1.38	.240	.08	13.34	<.001	.18	1.06	.303	.05	6.38	.012	.13
LE × NT	1							2.18	.140	.07	0.08	.775	.01	4.43	.035	.11
G × NT	1							0.43	.510	.03	7.47	.006	.12	4.37	.037	.11
Residence	1	1.18	.277	.08	2.39	.122	.10	0.52	.471	.04	0.08	.784	.01	1.95	.162	.07
Arrest	1	0.04	.847	.01	0.22	.640	.03	0.37	.544	.03	0.10	.747	.01	0.89	.346	.05
Work	1	0.17	.682	.03	0.92	.339	.06	1.33	.249	.06	0.03	.868	.01	0.01	.921	.01
Gun	1	0.03	.869	.01	0.16	.688	.03	0.05	.816	.01	0.02	.884	.01	0.65	.420	.04
Marijuana	1	0.37	.546	.04	0.04	.844	.01	2.01	.157	.07	0.10	.758	.01	5.73	.017	.13

Note. Analysis is limited to those aged 16 through 18 living in MYS Neighborhoods. Analysis in Years 1 through 7 is limited to those living in target neighborhoods. Year 2 analysis excludes those living in neighborhood 48 and Years 8 and 9 analyses exclude those living in neighborhood 2. NT = neighborhood type; Gr = grade level; G = gender; LE = free lunch eligibility status.

CHAPTER 5

DISCUSSION

Justification for Analysis

Any research involving humans requires that a sample of research participants be selected from the population. The process of selecting a sample can have important implications for the validity of the conclusions that are drawn from the research. Even if the sample is selected randomly, low response rates can threaten that validity. Response rates that deviate from unity introduce missing data into the study, which can reduce the representativeness of the sample. It is, therefore, very important to explore sample representativeness in all research studies. When non-random sampling strategies are used in a research study (whether by convenience or by snowball effect or by simply being eligible) certain characteristics may be oversampled and consequently, certain characteristics may be undersampled, resulting in a non-representative sample. Further, even in random samples with less than 100% participation, representativeness should be explored because there may be certain characteristics (related to outcomes) leading a group of recruited individuals to not participate in the study. In both of these cases, when important characteristics in the sample differ from the population, the missing data that result can be said to be missing not at random, or missing for non-ignorable reasons.

Sometimes, but not often, authors point out that missing data or nonrepresentativeness is a potential weakness in their studies. More frequently, this is left to authors of review articles; but, seldom do they suggest how sampling limitations may affect conclusions the articles may draw. This study, rather than simply reporting that results generated from the Mobile Youth

Survey (MYS) data are potentially limited by missing data, assesses the extent to which the MYS sample deviates from the population of adolescents living in MYS neighborhoods, and therefore, the extent to which the missing data may be non-ignorable. Further, this study developed a protocol for using auxiliary data to determine whether non-randomness in the MYS sampling strategy led to biased estimates. This, in turn, leads to a discussion about missing data mechanisms that exist in the MYS.

Before discussing the results with implications specific to the MYS, it is important to remember that in any dataset, there may be both ignorable and non-ignorable missingness (Daniels & Hogan, 2008). If there is evidence to support the conclusion that the majority of the missing data patterns represent a missing at random (MAR) mechanism, then overall, the missing data within the dataset can be considered MAR contaminated with missing not at random (MNAR) data. Even with moderate amounts of contamination, nonparametric statistics (and in some cases, even parametric statistics) are robust estimators of location and dispersion (Rey, 1983). Though this study does not seek to implement techniques to manage missing data for subsequent analyses, it is important to note that the amount of missing data, categorized by mechanism, cannot be definitively quantified.

Because the categories of missing data mechanisms cannot be definitively quantified, and because the statistical procedures used to study missing data are often beyond the grasp of the typical researcher, a method of studying missing data that is within the grasp of typical researchers is necessary. In this study, a procedure for studying missing data mechanisms, that is, assessing sample representativeness, has been presented and demonstrated to be effective.

Here, the representativeness was examined in a longitudinal community-based study. In longitudinal studies, it is important to examine representativeness in two ways: cross-sectionally

and longitudinally. That is, it is important to look at each year's data to determine whether each wave (and aggregately, the entire sample) is representative, and to determine whether the follow-up sample each year is representative of the previous year's sample. This two-pronged approach is necessary, because cross-sectional representativeness does not guarantee longitudinal representativeness and vice-versa.

Researchers often conclude whether a sample is representative based exclusively on demographic characteristics. Demographic characteristics are important in any study to determine representativeness of a sample; however, in many studies, they are not enough. When a study's research questions are centered around something like beliefs, attitudes, and behaviors of at-risk adolescents living in poverty (as in the MYS), it is also necessary to determine the representativeness of the sample in terms of functional characteristics like cognitive abilities and behavioral characteristics, which may be more directly related to study outcomes.

Because the MYS project (the youth survey, in particular, but also the other projects and data sources associated with it) has such a wealth of data, the quality of the data is directly relevant to millions of dollars in grant funding that can be obtained for studies based on the data. The quality of the data is also directly relevant to articles that have been written, and that will be written in the future, using the data. Further, the quality of these data also is indirectly related to research that others may conduct using survey methods to study vulnerable populations, and the potential bias that might characterize these studies.

Summary of Analyses

When data are missing, there is no way to definitively determine sample representativeness, at least with respect to the outcomes being measured in the study (e.g., MYS measures of risk behaviors). A number of techniques have been proposed to handle missing

data, as identified in Chapter 2. But, these methods are not easily accessible to most researchers because of their complex statistical nature and programming requirements (e.g., writing code for an Expectation Maximization algorithm). Alternatively, an auxiliary dataset, if available, can be used to examine whether differences exist between those who are surveyed in the primary study and those in the population. In the case of the present study, an auxiliary dataset in the form of student records from the Mobile County Public School System (MCPSS) is available which was used to measure population characteristics for those aged 10 through 15 in the neighborhoods represented in the MYS. Characteristics assessed for representativeness in this study include: demographic characteristics (gender, race, free lunch eligibility status (as a proxy for socio-economic status), neighborhood, and neighborhood type) and functional characteristics (Stanford Achievement Test percentile ranks, weighted school violation scores, and weighted school disciplinary action scores).

First, cross-sectional representativeness was assessed for 10 years of the MYS on demographic and functional characteristics. Each year, between 1,317 and 10,967 students aged 10 through 15 who were living in 48 MYS neighborhoods were used in these cross-sectional analyses. Using a Generalized Estimating Equations approach, MYS enrollees were compared to non-enrollees and each year, MYS participants were compared to non-participants.

Second, longitudinal representativeness was assessed in order to determine whether dropout mechanisms in the MYS sample were ignorable or non-ignorable. To assess the longitudinal representativeness in the MYS, pairs of adjacent years were used, comparing participants at $Time_{t-1}$ with participants at $Time_t$, to ensure that the latter are representative of the former. In each of these analyses, between 344 and 1,435 MYS participants aged 10 through 15 who were living in the 48 neighborhoods, were analyzed, comparing those who dropped out and

those who did not. Also, as a part of the analysis of longitudinal representativeness, characteristics of these dropouts were examined. First, patterns of missingness were examined: those with monotonic missingness patterns were compared to those with non-monotonic missingness patterns to determine whether they differed on demographic and functional characteristics. Finally, to extend the ability to interpret dropout mechanisms, data from the MYS at $Time_{t-1}$ was used to compare participants and dropouts at $Time_t$ (in terms of residential tenure, paid employment hours, and risky behaviors).

Summary of Findings

The goals of this study are to (a) identify the representativeness of the MYS and thus, identify missing data mechanisms that exist in the MYS, (b) address concerns that selection bias is inherent in sampling hard-to-reach populations, (c) test whether non-random sampling procedures necessarily lead to biased estimates, and (d) provide an accessible diagnostic tool for researchers to assess the representativeness, and thus the missing data mechanisms in their own research studies and are addressed through two research questions.

In order to draw conclusions about whether missingness is non-ignorable, several criteria should be met: (a) statistical significance in three or more years; (b) significance in both measures of cognitive characteristics or in both measures of behavior; (c) significance across analyses; (d) consistency in functional form; (e) $\mathbf{w} > .15$.

When considering representativeness in terms of ***enrollment***, deviations from representativeness (either demographic or functional) must be consistent across time. Recall that enrollees are youths who have participated in the MYS during any of the 10 years considered in the study: as such, the enrollment measure does not vary by time. If non-representativeness is found for only a single year, only two explanations could be entertained. One explanation is that

there is a difference in school demographics or outcomes of the MYS enrollees and non-MYS enrollees, but only for that year. This finding would require further investigation to determine if significant differences are localized in different schools. If there are localized school differences, the differences may well not be related to the representativeness of the MYS sample, but rather to school policies. Second, neighborhood composition with respect to MYS and non-MYS enrollees may have changed (e.g., the MYS enrollees moved to non-MYS neighborhoods that year, but moved back into MYS neighborhoods the following year). This situation is not very likely, but could be examined further to rule it out definitively. Thus, the only instance that suggests non-ignorable missingness for the enrollment analyses, on both demographic and functional measures, would be consistency in trends across years.

When considering ***yearly participation*** and ***longitudinal participation***, annual deviations from representativeness may be more meaningful, providing evidence of non-ignorable missingness; that is, the people who are missing, but eligible, from the MYS are different from those who participated in the MYS during that year. There are three possible reasons for deviations that are non-ignorable: first, in any year, there may have been fewer recruiters (or less active or skilled recruiters) available for MYS recruitment, and thus there may not have been as substantial an effort to recruit potential participants who were hardest to reach. Second, in vulnerable populations, there may be an inherent mistrust of outsiders. In any given year, events (related to or unrelated to research) may have exacerbated this mistrust and led to either lower rates of participation or higher rates of dropout. Third, a non-representative new cohort may have enrolled in any given year but did not participate in subsequent years. The first two reasons presented should affect yearly participation and longitudinal participation. However, deviations because of the any event leading to mistrust in a specific year or because of new-cohort

differences would only manifest themselves in yearly participation. Before deviations from representativeness in yearly participation are used to support conclusions of non-ignorable missingness, however, there should be an examination of whether these deviations also occur in the enrollment analyses.

While the analyses were conducted by type of participation in the MYS, that is, enrollment, yearly participation, and longitudinal participation, a discussion of the results by type of measure (i.e., demographic and functional) is more logical. The discussions of these results will first identify what is required to suggest non-ignorable missingness and second, describe the results from the analyses of the MYS.

Demographic Characteristics

Results support non-ignorable differences in demographic effects if (a) there are consistent statistically significant differences for enrollment across years, (b) consistent statistically significant differences across participation analyses (i.e., yearly and longitudinal); (c) consistent statistically significant differences across enrollment and yearly participation analyses; and/or (d) consistent trends across years.

Results indicate consistent differences in the demographic composition of the neighborhoods containing MYS participants in the enrollment and yearly participation analyses (Table 64): African Americans enrolled and participated at higher rates than non-African Americans; those eligible for free lunch enrolled and participated at higher rates than those who were not; those in target neighborhoods enrolled and participated at higher rates than those in expansion neighborhoods. Also, during initial years, MYS enrollees and participants in lower grades over-represented the population; while in later years, enrollees and participants in higher grades over-represented the population. Similar patterns were demonstrated in the longitudinal

analyses for free lunch eligibility status and neighborhood type. While statistically significant, these results typically show very small to small effect sizes. There were no consistent interactions involving demographic variables.

These results are not unexpected. The MYS study design may account for the racial differences in enrollment and participation of the MYS sample. The initial neighborhoods selected to participate in the MYS (the poorest 13 neighborhoods in the Mobile MSA) were almost exclusively African American. As a result, most of the people surveyed in the early years of the MYS were African Americans. Perhaps due to the initial racial distribution in the MYS, the MYS study as a whole may have been perceived to be only for African Americans, which may have discouraged non-African Americans from enrolling even when eligible. An additional explanation for the racial distribution in the MYS is that the racial distribution within neighborhoods is not uniform, which is discussed in subsequent sections. In a similar way, the MYS study design can perhaps account for differences in free lunch eligibility status. In the United States, there is a correlation between race and socio-economic status (SES) (Do, Frank, & Finch, 2012). The two explanations for deviations from racial representativeness may equally apply to deviations from representativeness in free lunch eligibility status.

While free lunch eligibility status is treated categorically in these analyses, it perhaps should be viewed as an ordinal measure. Free and reduced lunch eligibility is determined by income levels and number of people in the household ("NSLP," 2012). Thus, one would expect enrollment and participation rates by free lunch eligibility status to have the same ordinality, which was not consistently the case in the yearly participation analyses (see Tables 18 and 31). In each of the discrepant analyses, the functional form of the reduced-lunch eligibility category is inconsistent, raising questions about the validity of including the reduced-cost lunch status as a

category. If this variable were dichotomized, it is unclear whether reduced-cost lunch status should be combined with not eligible status or with free lunch status.

Finally, explanations for deviations in grade level representativeness may be contextual. Specifically, during the initial years of the survey, older students, who may have been more likely to participate in illegal behaviors and therefore were more concerned about anonymity, may have been hesitant to participate in the MYS (Bolland et al., 2007). Younger adolescents were less likely to have these concerns, so their participation rates were higher. But over time, two things happened: the study became more established in the neighborhoods, and participants aged. The former may have reduced the anonymity concerns, particularly for those youths who were familiar with the MYS study (i.e., enrollees). In the latter case, MYS data show that the likelihood of participation during any given year increase as the number of years of prior participation increases; thus, as participants aged, any concern about anonymity may have been counterbalanced by the positive effects of prior participation.

While it is important to discuss the significant results or departures from representativeness, it is also important to discuss the lack of significance in some of the demographic analyses and in some of the main effects and interactions. Rarely are there significant differences in gender in any of the demographic analyses, and never are there consistent differences in gender across years in any of the demographic analyses. This result suggests that the MYS sample is representative of the population as far as the proportion of males and females. Next, there are never any significant interactions of free lunch eligibility status × neighborhood type, indicating that the interaction does not contribute anything beyond the significant main effects (when they occurred). Finally, there are never any significant

interactions of neighborhood type × gender, indicating that the interaction does not contribute anything beyond the significant main effect of neighborhood type (when it occurred).

Table 64

Summary of Significant Effects for Demographic Analyses

	Year									
	1	2	3	4	5	6	7	8	9	10
Effects										
R	•	•◘	•◘	•	•	•	•	•◘	•◘	•◘
G	•		•		*					
NT	•	•	•		•	•	•	•◘*	•◘*	•◘
LE	•◘	•◘*	•◘	•◘*	•◘*	•	•◘	•◘*	•◘	•◘
Gr	•	•◘	•◘	◘	•◘	•◘	•◘	•◘	•◘	•◘
LE × NT				•						
SATTR						•				
SATTM						•				
WSV					•		•			•
WDA										

Note. R = race; G = gender; NT = neighborhood type; LE = free lunch eligibility status; Gr = grade level; SAT_{TR} = SAT reading percentile ranks; SAT_{TM} = SAT math percentile ranks; WSV = weighted school violation scores; WDA = weighted school disciplinary scores; • = significance in enrollment; ◘ = significance in yearly participation; * = significance in longitudinal participation.

Table 65

Summary of Significant Interaction Effects for Functional Analyses

	Year									
	1	2	3	4	5	6	7	8	9	10
Interaction Effects										
SAT_{TR}										
× LE	•	•	•			•				
SAT_{TM}										
× Gr								◘	◘	◘
× NT		•								
× G × NT		•	•	•						
WSV										
× G		◘								
× Gr				◘	◘					
× LE		•			•		•	*		
× R		•								
× NT × Gr		•		•	•					
WDA										
× Gr				◘	◘			◘	•	
× LE		•			•			•		
× NT								◘		
× R		•								

Note. R = race; G = gender; NT = neighborhood type; LE = free lunch eligibility status; Gr = grade level; SAT_{TR} = SAT reading percentile ranks; SAT_{TM} = SAT math percentile ranks; WSV = weighted school violation scores; WDA = weighted school disciplinary scores; • = significance in enrollment; ◘ = significance in yearly participation; * = significance in longitudinal participation. Only significant interactions are presented in this table.

Functional Characteristics

Table 65 demonstrates that unlike the consistency found in demographic deviations from representativeness, there was little consistency in deviations from representativeness for functional outcomes. First, with respect to main effects, the only consistent result across years was for weighted school violation scores (WSV) for enrollment (enrollees had higher WSV scores than non-enrollees in each significant year). Three factors limit the importance of this finding. While statistically significant, the effect sizes for these results were very small. Further, this significant result was not found for either of the participation analyses. Finally, this significant result was not found in any other functional outcome. Second, the only consistent results across measures within one year were for Year 6 SAT percentile ranks for enrollment (those enrolled had lower percentile ranks than those un-enrolled on both cognitive measures), but this result was not replicated across years or type of analysis. Further, while statistically significant, the effect sizes for these results were very small.

Interaction effects are more complicated in terms of what they reveal. For example, enrollment × free lunch eligibility status was a significant predictor for total math percentile ranks (SAT_{TM}) in four years, for WSV in four years, and for weighted disciplinary action scores (WDA) in three years. Although these significant results occurred across time and across outcome measure, the trends in outcome were not consistent within measure or across time (Figures 2a, 4a, and 5a and b). Further, while statistically significant, effect sizes in these results were small.

Next, there is a consistent significant interaction of yearly participation × grade level on SAT_{TM}, WSV, and WDA. There is also one significant interaction of enrollment × grade level on WDA. The patterns are consistent across measures with older students having worse

outcomes (i.e., lower SAT_{TM} percentile ranks and higher WSV and WDA scores). These patterns were consistently more pronounced for yearly participants than for non-participants. While these results were significant across time and measure and patterns were consistent both within and between measures, the effect sizes are still in the very small to small range.

Another example is the number of significant interactions involving enrollment in Year 2 (see Table 64). However, this appears to be a contextual anomaly occurring in that year. That is, since all but one of these significant results occurred for the enrollment analyses, rather than the participation analyses, it is unlikely that this was a consequence of sampling. However, the sheer number of significant results indicates that a further examination into school context in Year 2 is warranted.

While deviations in functional representativeness were unanticipated, significant results occurred only for enrollment and yearly participation analyses. Longitudinal analyses of functional characteristics failed to consistently result in significance, with no significant main effects and only one significant interaction effect (Table 65). Thus, the results support the conclusion that MYS dropouts at $Time_t$ are representative of the sample of participants at $Time_{t-1}$ and fail to suggest that missingness is non-ignorable. One possible explanation for the lack of significant results is the lower sample sizes used in these analyses (ranging from 344 to 1,435). However, an analysis of the statistical power available for these analyses in a sample of at least 344 with $\alpha = .05$ and $\mathbf{f} = .15$ is greater than or equal to .87 (Cohen, 1988).

Supplemental Analyses

Even though no meaningful results were obtained for MYS dropout, analyses of dropout patterns were conducted. Results indicated that in the earlier years of the MYS, older students were more likely to have monotonic patterns of missingness. In contrast, in the later years, older

students were more likely to have non-monotonic patterns of missingness. The results from the earlier years are more intuitive, because older students have less opportunity to drop out and then drop back into a study, because they will age out of eligibility. The results from the later years seem more counter-intuitive. There were also significant consistent interactions for missing data pattern × free lunch eligibility status, but these results were neither consistent across years or across measures. The lack of significant results could possibly be accounted for in because of the categorization of missing patterns. Only two categories of pattern missingness were used in these analyses: monotonic missingness and non-monotonic missingness. In reality, there are many patterns of missingness that could occur in 10 years of data collection. Dichotomizing these patterns may limit the potential to find significance in complex patterns.

Although the lack of significant results is inconclusive with respect to missing data and missing data patterns in this study, this is not the case with respect to the analysis of alternative proxy measures, i.e., MYS measures at $Time_{t-1}$. Using 10 through 18 year olds, after controlling for demographic factors, the only consistent predictor of longitudinal participation was residential tenure. Not surprisingly, greater residential stability is associated with lower dropout rates. The power of these analyses (**f** = .15, α = .05) is greater than or equal to .995 (Cohen, 1988). This analysis is especially useful because it is an analysis that can be done absent an auxiliary dataset.

Approaches to Remediation

There are four possible approaches to handling missing data associated with demographic deviations from representativeness: weighting, deletion, modeling, and ignoring.

The first approach to remediation is weighting the sample to make it more representative. In these analyses, because of the low prevalence of non-African Americans (ranging from 0.6%

to 1.8% across years) in the MYS sample, weighting could be used; however, the resulting estimates would be quite unstable. Although to a lesser extent, this logic also holds for weighting the free lunch eligibility status variable.

Second in cases where there are sample deviations from representativeness and where distribution among categories is severely unequal, deletion of categories from analyses could be used. The analyses presented here suggest that the removal of non-African Americans and/or the removal of non-free lunch recipients may be more appropriate. This method of remediation is only advisable, however, after a determination that the differences in the categories to be deleted and the categories to be kept (e.g., African Americans and non-African Americans) are related to the outcomes of interest. In the event that differences in categories are not related to outcomes of interest, but deleted from analyses, inappropriate conclusions could be drawn (e.g., if only African Americans were retained, the present data would suggest that they have low test percentile ranks because of their race). In contrast, if there are no differences in outcomes between categories, using a full dataset allows for conclusions about differences (or non-differences) between categories (e.g., African Americans and non-African Americans living in the same poor neighborhoods; Bolland et al., 2007).

Third, modeling can be used to remediate deviations between sample and population demographics. An example of modeling is what was done in the analyses of functional characteristics in this study by including demographic variables as covariates in the analyses. Modeling these characteristics in subsequent analyses, however, may not completely remove the effects of deviations from representativeness, although, it does allow for those effects to be somewhat controlled.

Finally, in the event that deviations are largely ignorable, but contaminated with small amounts of non-ignorable missingness, one might choose to simply ignore those deviations. Because amounts of ignorable and non-ignorable missingness cannot be quantified, a possible proxy for estimating the amount of non-ignorable missingness might come from effect size. That is, as effect size increases, the amount of non-ignorable contamination also increases. Rey (1983) demonstrates that even in the face of modest contamination, estimates are still robust (particularly if non-parametric estimates of location and dispersion are used).

When considering how to remediate demographic deviations from representativeness for substantive analyses, there are two final considerations. First, the researcher should revisit the initial goals of the project. In this case, the goals of the MYS are reconsidered: to describe the (a) characteristics, circumstances, and behaviors of its participants (disadvantaged urban adolescents, aged, 10 through 18) and (b) the etiology of these characteristics, circumstances, and behaviors. Then, decisions should be made about the deviations from population representativeness.

Next, demographic deviations from representativeness may ***not*** indicate non-ignorable missingness; therefore, before remediation, researchers should investigate whether and/or how those deviations affect further substantive analyses. When auxiliary datasets are available, this step is not necessary because relationships between demographics and proxy outcomes can be determined empirically. However, if an auxiliary dataset is not available, researchers must use logic and results from other studies to determine the plausibility of demographics being related to outcomes and therefore, non-ignorable.

Deviations from functional representativeness create a more complex problem for researchers because they seldom have access to auxiliary datasets with which to empirically

measure proxy outcomes. Therefore, the analyses of functional representativeness using an auxiliary dataset can rarely be replicated (which is not to say that researchers should not attempt to gain access to an auxiliary dataset for these purposes). Even so, there are recommendations that are relevant for all researchers, and especially for those who intend on using the MYS data for substantive analyses.

In the present study, where an auxiliary data set is available, there does seem to be a weak bias in the sample, as indicated by the effect sizes and inconsistency of significant results across analyses, measures, and years. This bias indicates that enrollees in the MYS are at greater risk for poor outcomes than non-enrollees. Second, there seems to be a potential bias introduced by grade level for yearly participation. This suggests that some remedial action may be necessary with respect to grade level in the MYS, since grade level had a significant main effect for both enrollment and yearly participation; in addition, it suggests that researchers using non-MYS data should closely examine the representativeness of their samples with respect to grade level. Similarly, age should also be considered and examined in the same way.

If a longitudinal analysis does uncover significant (both statistically and practically) results, an examination of missing data patterns should be conducted (as discussed previously). Finally, all researchers conducting longitudinal studies can use their datasets to examine dropout as a function of participant responses at $Time_{t-1}$.

Risk Status and Neighborhoods Revisited

Researchers have argued that the hardest-to-reach and most vulnerable populations are under-sampled in studies of risk (e.g., Hatchett, Holmes, Duran, & Davis, 2000; Kerkorian, Traub, & McKay, 2007; Pottick & Lerman, 1991), resulting in biased conclusions about these populations. At the very least, there are questions about whether the reported prevalence of

some risky behaviors is reflective of true levels of that behavior in the population. One of the goals of the MYS is to capture low prevalence risky behaviors among a very vulnerable population. Results presented in this dissertation show that there are not many functional differences between the sampled and the un-sampled residents of high poverty neighborhoods in Mobile and between MYS dropouts and non-dropouts. If anything, these results show that that typically, those sampled have poorer outcomes than those who were un-sampled, even though they are living in the same neighborhoods. While determining representativeness of the sample to the population is an important component of this study, one might argue that the population should be redefined as the most vulnerable and highest risk adolescents living in these neighborhoods, and that the sample of enrollees and participants in the MYS is reflective of this new population. Any deviations from representativeness may therefore be even less problematic than what is suggested in the analyses (and it was small even considering the entire neighborhood student population). In other words, missingness may not be ignorable as it relates to the original neighborhood population, but, it becomes largely ignorable as it relates to a redefined population.

This justification, of course, is post-hoc, but it also has a geographical basis. Distributions of demographic characteristics may not be uniformly distributed in neighborhoods, and in fact, are not in many of the MYS neighborhoods. Within MYS neighborhoods, there are pockets of relative affluence and relative deprivation. For example (Figures 6a and 6b), one expansion neighborhood, according to the 2000 census, was approximately 88% Caucasian. However, because MYS enrollees from originally targeted neighborhoods moved to this neighborhood, it was identified as an expansion neighborhood and active recruitment of participants occurred in this neighborhood. Further, all of the African Americans in this

neighborhood were concentrated in 16 census blocks out of 126 (12.7%) located in the southern part of the neighborhood. None the less, however, all students from the MCPSS who lived in the entire neighborhood were counted, even though the MYS sample was limited to a few blocks (Lian & Bolland, 2009).

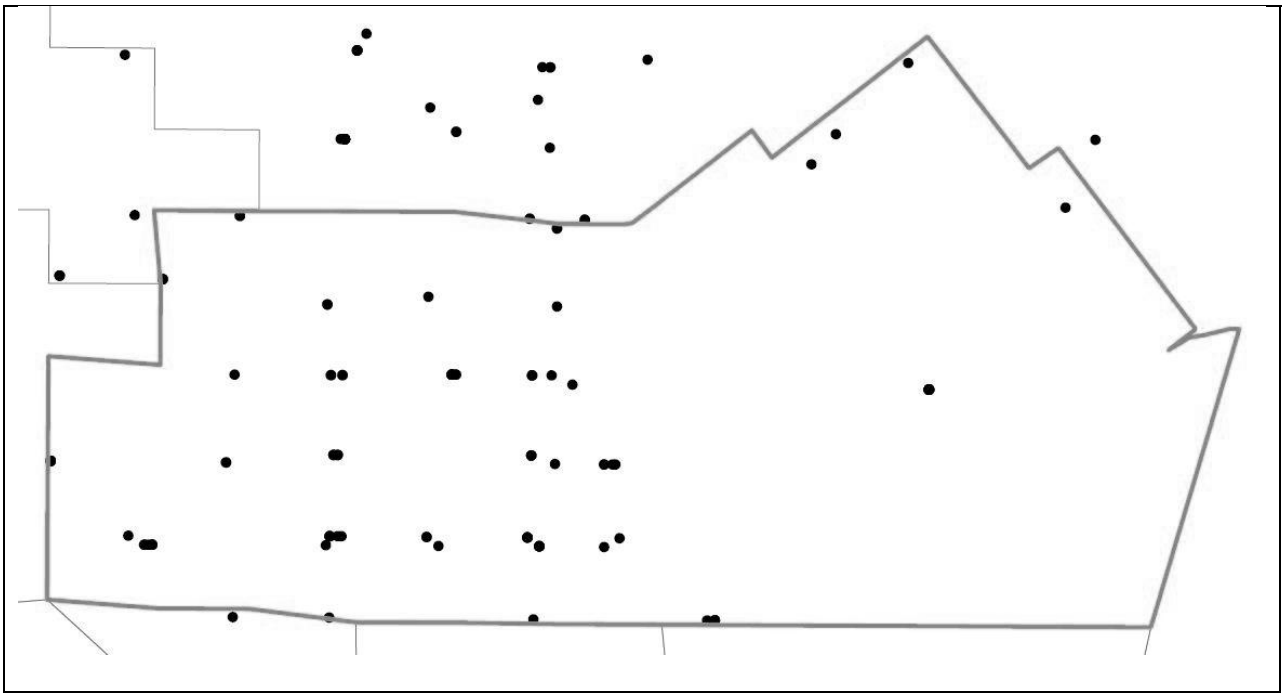

Figure 6a. Neighborhood with MYS Enrollees.
Note. The bold line indicates the neighborhood border. A dot indicates that at some time between summer 1998 and summer 2007 at least one, but not limited to one, MYS enrollee indicated residence.

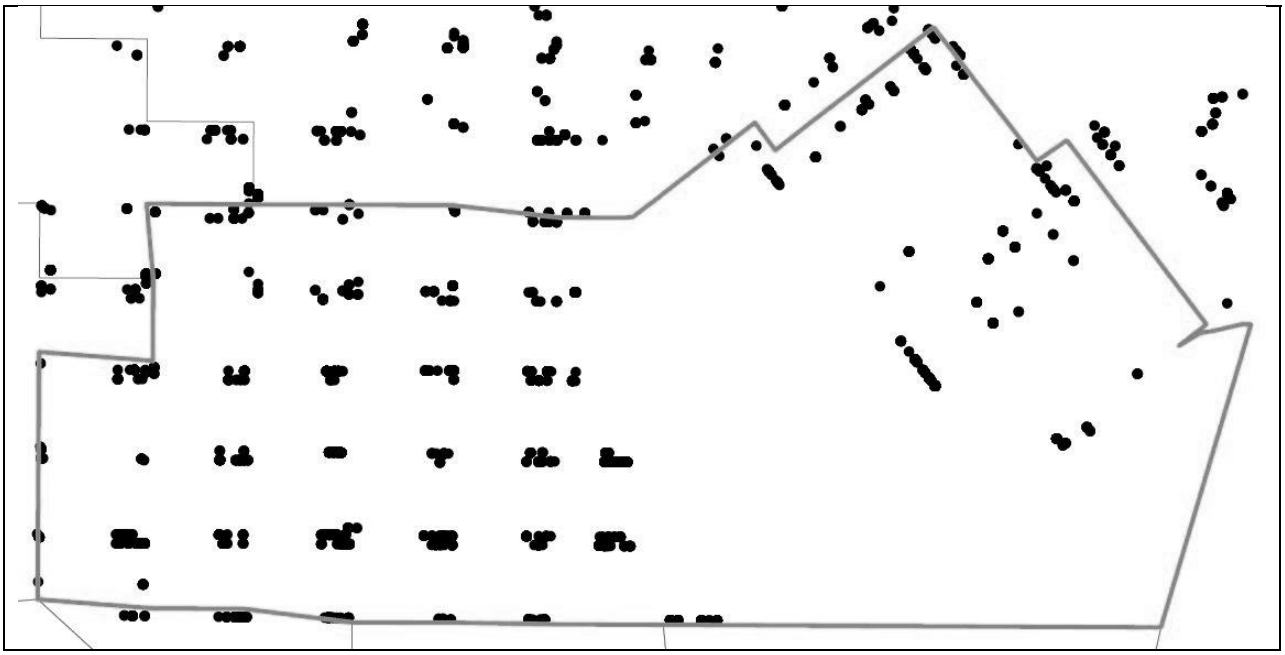

Figure 6b. Neighborhood with MYS Enrollees and MCPSS Students.
Note. The bold line indicates the neighborhood border. A dot indicates that at least one, but not limited to one, MCPSS student lived at address between 1998-1999 and 2007-2008 academic school years.

Strengths and Limitations

Strengths

There are several strengths of this study. First, because for most analyses the sample is large, statistical power is substantial for most analyses. Second, an auxiliary dataset was available for use. Third, the analyses of functional representativeness used two measures of cognitive ability and two measures of risky behavior, allowing consistency to be assessed across measures. Fourth, both the MYS dataset and the auxiliary MCPSS dataset are longitudinal, which allows representativeness to be assessed with respect to enrollment, yearly participation, and longitudinal participation. Moreover, the longitudinal nature of the data allows the consistency of effects (both their existence across years and their functional forms across years) to be considered. Fifth, the auxiliary dataset used in this study does not contain self-reported

measures, but rather is based on actual records (e.g., birth certificates and SAT percentile ranks). Thus, these datasets are not limited by recall biases or self-report biases.

Sixth, these samples are drawn from reasonably homogenous neighborhoods, that is, nearly everyone sampled is living in some level of poverty and all neighborhood income and poverty rates are below the median for the Mobile MSA ("Census," 2012). Socioeconomic status (SES) is a multi-dimensional concept, and it is difficult (if not impossible) to measure (e.g., Oakes & Rossi, 2003). When the population studied is heterogeneous with respect to SES, non-measurement or mismeasurement of SES creates a substantial residual confounding effect, and results are almost always uninterpretable (e.g., Geronimus & Korenman, 1992). In the present study, free lunch eligibility status is used as a proxy for SES. Admittedly, it is a poor measure of this multidimensional concept. However, since the sample is relatively homogeneous with respect to SES, any deficiencies associated with free lunch eligibility status are minimized. Additionally, substantive findings from the MYS are unlikely to be confounded by unmeasured or mismeasured SES.

Similarly, because this study is geographically confined to one area, the question of generalizability is validly raised. If this study cannot be generalized, then perhaps the representativeness and approach to handling missing data is not as important as it might be for other, geographically dispersed studies. Moreover, as Bolland (2008) explains, being geographically confined increases internal validity. This study measures risk behaviors in minority adolescents and as it has been described, the neighborhoods that make up this population consist of an overwhelming percentage of minorities living in poverty. Again, socioeconomic status is difficult to conceptualize and measure (e.g., Oakes & Rossi, 2003), however, this concept appears to be linked to behavioral and developmental variables.

Therefore, if socioeconomic heterogeneity was maximized in this study, the effects of behavior and development would not be easily sorted from effects of SES. This sacrifice of generalizability, for the researchers and for these purposes, has been noted and accepted.

Limitations

As with its strengths, the study also has several limitations. First, researchers involved in the MYS worked extremely hard to gain the trust of the community and worked with other agencies within the community to gain access to additional data (e.g., MCPSS records, court records, housing records). The access to additional data may not be feasible for all researchers to gather for a research study. Admittedly, it often takes resources to gain access to additional data sources; however, it would be advantageous for researchers and evaluators to seek additional data sources when possible, especially if the data sources are complete.

Second, and related, the problem with missing data is that because the data are missing, the researchers cannot definitively determine why those data are missing. In the MYS, the researchers have the ability to consult additional available data as a way to measure the representativeness of the MYS, but only to the extent that the available data sources are complete. In this case, representativeness can be measured for those aged 10 through 15 because that is the extent that the MCPSS records are complete. Therefore, the sample of 16 through 18 year olds represented in the MYS has not been concluded to be representative because data are not available to use to assess representativeness.

Third, the proxy measures used to measure functional outcomes do not directly correspond to risk behaviors in the MYS. While there is reason to believe that people who score low on achievement tests or commit more school violations are more likely to engage in risky behaviors, the correspondence is not complete.

Fourth, grade level was used as a demographic characteristic in these analyses, and while statistically significant, it is unclear what the results would have occurred if age were used in its place.

Fifth, as stated in Chapter 1, a research study is only as good as the quality of its data. The representativeness of the MYS data was assessed with auxiliary data that was defined as a measure of population characteristics. The MCPSS dataset, not unlike any other datasets, does contain error, and it is not unreasonable to expect some errors in the large amount of data contained in those records. The MCPSS, however, is as good as any other available dataset that could be used to compare the sample characteristics with the population characteristics of those aged 10 through 15 living in 48 pre-determined neighborhoods. Even though the MCPSS records are the best available, drawing conclusions based on this single source of data should not be done hastily. It is because of the measurement error in the MCPSS dataset that representativeness in the MYS was examined in so many ways.

Conclusions and Recommendations for the MYS

This study potentially benefits methodologists in allowing them to re-conceptualize non-ignorable missingness as non-representativeness, a more accessible concept for most researchers. This study also benefits substantive researchers conducting studies of high-poverty adolescents by suggesting what they should examine to assess bias in their results. Maybe most relevant, this study also benefits those using the MYS dataset by pointing out potential sources of bias due to non-ignorable missing data.

When considering recommendations, attention should be paid to the criteria established in Chapter 3. While results in this study may meet individual criteria, none of the results in this study meet all of the criteria needed to conclude that missingness is non-ignorable. That is, when

assessing the missing data mechanisms in the MYS, in light of its intended sampling frame, one can conclude that data missingness is ignorable. Still, however, in terms of demographic characteristics, remediation may be considered for some applications. This section suggests those remedial steps.

1. With respect to race, low effect sizes suggest that even though deviations in representativeness are statistically significant, they pose little threat to the validity of conclusions derived from the MYS study. Therefore, any missing data associated with race should be ignored. This recommendation is supported by the lack of any consistent significant results of functional outcomes coupled with the fact that the MYS is focused on the most vulnerable populations in selected neighborhoods. While the sample could be weighted, doing so would not be advisable because of the instability of resulting estimates. The lack of differences between races on most functional measures suggests there is no reason to delete non-African Americans from this sample. In substantive analyses, however, race should be modeled as a covariate.
2. Gender was not found to be consistently significant in any analysis, and therefore any missing data with respect to gender can be ignored as potential source of bias.
3. Free lunch eligibility status was used in these analyses as an indicator of SES, and there were consistent significant deviations from representativeness across years, measures, and analyses. However, the patterns of these trends were not consistent. One recommendation is to dichotomize this variable; however, more research should be conducted to determine the most justifiable way of combining categories. If a researcher is primarily interested in the most vulnerable population, without comparison between levels of vulnerability, removing those qualifying for reduced-cost lunch or not eligible at

all from the analysis would be justified. This recommendation is consistent with the MYS goal of studying the most vulnerable adolescents. However, not everyone who is eligible for free-lunch status actually fills out the forms to qualify, and those are automatically included in the paid-lunch category; therefore, eliminating some of the sample because they do not qualify for free lunch would result in potential bias as well. Because of this limitation, and the fact that free lunch eligibility status for many students changes over time, researchers perhaps should consider only eliminating those who never qualified for free lunch.

4. For neighborhood type, though consistently significant, these deviations from representativeness are at least partially a function of the MYS design. Elimination of cases would therefore, not be appropriate. While the sample could be weighted by neighborhood type to control for these differences, this would effectively reduce the representation of the most vulnerable in the study. Weighting by neighborhood is also inadvisable because neighborhood, like SES, is a multi-dimensional construct. Thus, before weighting was implemented, defining the neighborhood as it relates to the MYS sample would first be necessary. Until more information is gathered missing data for neighborhood and neighborhood type should be ignored.
5. Grade level consistently deviated from representativeness, but in inconsistent ways. Because of its statistical significance, it is likely that the differences in sample and population are non-ignorable. Moreover, grade level cannot logically be related to underlying vulnerability (e.g., low SES); therefore, differences cannot be attributed to non-homogeneity of vulnerability within neighborhoods. The inconsistency of trends in grade level and enrollment, though, raises questions about using any elimination strategy

for grade level. Rather, the sample should be weighted to reflect population grade level (and age). The challenge comes in determining the population of 16 through 18 year olds in the geographically relevant portions of the neighborhoods, since this would have to be used for weighting the older segments of the sample. Census data perhaps could be best used for this purpose, with annual population characteristics estimated from the 1990, 2000, and 2010 censuses.

6. A question inevitably rises about how to treat substantive analyses using MYS data that have been published with no remediation of sample characteristics. If the missing data in the MYS are non-ignorable, then the assumptions used in these published substantive analyses are incorrect and the resulting conclusions are also incorrect. However, the analyses reported here do not tend to support this concern.

Implications Beyond

This study has implications not only for the MYS project, but also beyond the MYS project. First, regarding the MYS study, results show that while demographically, the MYS sample (ages 10 through 15) is not strictly representative of the population, these deviations from randomness do not suggest that those who (a) were eligible, but did not participate or (b) dropped out did so for non-ignorable reasons. Further, results show that functionally, the MYS sample is representative of the population. This representativeness confirms that the results found in published articles using MYS data are not biased and the conclusions drawn from those results are not incorrect. This representativeness also allows the MYS data to be analyzed without fear of biased of estimates and incorrect conclusions.

Second, this study has implications for studying vulnerable populations. Results show that vulnerable populations can be studied without bias. Overall, deviations from

representativeness in the MYS sample occurred most consistently for race and free lunch eligibility status. However, in both cases, the "hardest to reach" groups (African Americans, those eligible for free lunch) participated at higher rates than the "easier to reach" groups. Few functional differences were found between participants and nonparticipants; but those that did occur tended to show lower test scores and higher rates of school violations and school discipline for enrollees/participants than for non-enrollees/nonparticipants.

More generally, this study also shows that random sampling procedures are not necessary to achieve representativeness. The MYS used a combination active and passive sampling strategies, with the sampling frame initially established as all youths living in the 13 target neighborhoods. As youths moved, however, and expansion neighborhoods were added, previous enrollees continued to be actively recruited, while new enrollees (at least those not living in MYS households) were, for the most, passively recruited. Yet, even given this lack of random sampling procedures, the MYS resulted in a largely representative sample.

Although the design and implementation of the MYS was not directly studied here, it is worth noting that the MYS involved a strong community presence, even in the expansion neighborhoods. While the importance of such a presence is suggested in the community-based participatory research literature (Israel, Schulz, Parker, & Becker, 1998; Minkler & Wallerstein, 1997), this literature is silent on exactly how to implement this type of research, particularly in the most impoverished neighborhoods where community-based organizations are largely absent. The results presented here provide indirect evidence that the methodology used in the MYS may be an effective way to accomplish this. Thus, while establishing a strong community presence may not be feasible for every study, it can be an effective way to achieve a representative sample when studying vulnerable populations.

Third, and even more generally, this study suggests the importance of establishing an appropriate sampling frame in research involving vulnerable populations. Neighborhoods, which are often used as sampling frames, are typically defined by a contiguous set of census blocks; however, this study has highlighted a problem with using neighborhoods as sampling frames and has also highlighted the importance of examining neighborhood context beyond demographic characteristics. Especially when studying the most vulnerable populations, this study suggests that using a neighborhood sampling frame may not be appropriate; rather, using blocks or a set of blocks as a sampling frame may be more appropriate.

Fourth, this study establishes an approachable and accessible way of conceptualizing and exploring missing data, that is, in terms of sample representativeness. Missing data techniques are statistically quite complex, and beyond the reach of the typical researcher. As a result, most published research studies that even address the issue of missing data simply assert that missing data are likely missing at random, and make no attempt to test the assumption. This study provides an alternative framework for assessing whether missing data are ignorable, based on the concept of sample representativeness. In every case, this framework should allow the researcher to assess, on a rudimentary basis, whether missing data can be ignored, by comparing sample characteristics with known population characteristics (e.g., those available through the census or aggregated school data). In cases where the researcher has access to an auxiliary dataset, this dissertation suggests a procedure for examining the correspondence between the sample and the population on a case-level. Along these lines, this study also establishes the importance of acquiring an auxiliary dataset when conducting a study or an evaluation.

Further Research

This study has shed light on some directions for future research that could and should be conducted with this dataset in order to more accurately and fully describe the MYS sample and the population it represents.

One such study could be the development of a scale to more precisely measure the school violations and disciplinary actions in the school system records, using techniques associated with attitude scale construction (e.g., Guttman scaling, Thurston scaling) (Edwards, 1979).

Knowing that error exists in all datasets, a second possible study would be an examination of the MCPSS records' consistency and error. In addition, the concept of consistency can be explored further. One such exploration would include an examination of the robustness of consistency and inconsistency over time or across measures, to determine whether results could be described as meaningful patterns or random error.

Another direction for future research is an examination of special education in the school records and how a specification of special education may affect data missingness. Especially interesting may be an examination of how labeling of students with special needs may affect SAT percentile ranks, particularly because a disproportionate number of students in the MYS neighborhoods have been given special education status. In addition, it would be useful to examine whether these special needs would be sufficient to disqualify a student from participating in the MYS, and in turn redefining the population to exclude students who would not be able to provide valid or reliable responses to MYS questions.

Also, neighborhood identification was almost always significant in the demographic analyses of enrollment and yearly participation. That is, enrollment and yearly participation rates were different by neighborhood. An examination of the contextual characteristics of

neighborhoods would be an important study to determine what characteristics of these neighborhoods make them different from each other, and whether demographic differences, functional differences, or more contextual differences account for different rates of enrollment and yearly participation.

Finally, while results from this study to fail to suggest non-ignorable missingness, still results from published MYS studies should be reexamined to determine whether weighting based on grade level and elimination of students who have never qualified for free lunch change the conclusions that were drawn.

REFERENCES

Aber, J. L. & Bennett, N. G. (1997). The effects of poverty on child health and development. *Annual Review of Public Health*, 18, 463-483.

Add Health (2012). Social, Behavioral, and Biological Linkages Across the Life Course. Retrieved from http://www.cpc.unc.edu/projects/addhealth.

Alabama Department of Education (n.d.). Reports. Retrieved from http://www.alsde.edu/html/reports1.asp?systemcode=049&schoolcode=0000.

Allison, P. D. (2000). Multiple imputation for missing data: A cautionary tale. *Sociological Methods and Research*, 28, 301-309.

Allison, P. D. (2002). *Missing Data*. Thousand Oaks, CA: Sage University Paper No. 136.

Aarts, B., Close, J., & Wallis, S. (2009). *Choices over time: Some methodological issues in research into current change*. Presented at the symposium on Current Change in the English Verb Phrase at the third International Conference on the Linguistic s of Contemporary English (ICLCE 3), London, 14 July 2009.

Baron, G., Ravaud, P., Samson, A., & Giraudeau, B. (2008). Missing data in randomized controlled trials of rheumatoid arthritis with radiographic outcomes: A simulation study. *Arthritis & Rheumatism (Arthritis Care & Research),* 59, 25-31.

Barry, A. E. (2005). How attrition impacts the internal and external validity of longitudinal research. *Journal of School Health*, 75, 267-270.

Bolland, J. M. (2003). Hopelessness and risk behaviour among adolescents living in high-poverty inner-city neighbourhoods. *Journal of Adolescence, 26,* 145-158.

Bolland, J. M. (2008). *Testing methodological assumptions about the Mobile Youth Survey*. Unpublished grant proposal submitted to National Institute for Child Health and Human Development.

Bolland, J. M., Bryant, C. M., Lian, B. E., McCallum, D. M., Vazsonyi, A. T., & Barth, J. M. (2007). Development and risk behavior among African American, Caucasian, and mixed-race adolescents living in high poverty inner-city neighborhoods. *American Journal of Community Psychology, 40,* 230-249.

Bolland, J. M., Lian, B. E., & Formichella, C. M. (2005). The origins of hopelessness among inner-city African-American adolescents. *American Journal of Community Psychology, 36,* 293-305.

Bolland, K.A., Bolland, J.M., Tomek, S., Devereaux, R., Mrug, S., & Wimberly, J. C. (n.d.). Trajectories of adolescent alcohol use among early and non-early alcohol initiators. "Manuscript submitted for publication."

Bor,W., Najman, J., O'Callaghan, M., Williams, G. M., & Anstey, K. (2001). Aggression and the development of delinquent behavior in children. *Australian Institute of Criminology Trends and Issues in Crime and Criminal Justice*, 207, 1-6.

Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. *American Educational Research Journal*, 5, 437-474.

Braveman, P.A., Cubbin, C., Egerter, S., Williams, D.R. & Pamuk, E. (2010). Socioeconomic disparities in health in the United States: What the patterns tell us. *American Journal of Public Health*, 100, S186-S196.

Buhi, E. R. & Goodson, P. (2006). Predictors of adolescent sexual behavior and intention: A theory-guided systematic review. *Journal of Adolescent Health*, 40, 4-21.

Buhi, E. R., Goodson, P., & Neilands, T. B. (2008) Out of sight, not out of mind: Strategies for handling missing data. *American Journal of Health Behavior,* 32, 83-92.

Carpenter, J. (2009) Statistical modeling with missing data using multiple imputation. Lecture.

Cayo, M. R. & Talbot, T. T. (2003). Positional error in automated geocoding of residential addresses. *International Journal of Health Geographics*, 2, 2-10.

Cleveland, W. S. (1993). *Visualizing Data*. Summit, NJ: Hobart Press.

Cohen J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.), Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Collins, L. M., Schafer, J. L., & Kam, C. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods*, 6, 330-351.

Daniels, M. J. & Hogan, J. W. (2008).*Missing data in longitudinal studies: Strategies for Bayesian modeling and sensitivity analysis*. Boca Raton, FL: Chapman & Hall/CRC.

deLeeuw, E. D. (2001). Reducing missing data in surveys: An overview of methods. *Quality & Quantity*, 35, 147-160.

deMorais, S. R., & Aussem, A. (2009). Exploiting data missingness in Bayesian network modeling. In N. Adams, et al (Eds.), *Lecture Notes in Computer Science* (pp. 35-46). London, England: Springer.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39, 1-38.

Diggle, P., & Kenward, M. G. (1994). Informative drop-out in longitudinal data analysis. *Journal of the Royal Statistical Society*, 43, 49-93.

Diggle, P., Heagerty, P., Liang, K., & Zeger, S. (2002) *Analysis of Longitudinal Data*. Oxford, England: Oxford University Press.

Do, D. P., Frank, R., & Finch, B. K. (2012). Does SES explain more of the black/white health gap than we thought? Revisiting our approach toward understanding racial disparities in health. *Social Science and Medicine*, 74, 1385-1393.

Drummond, H., Bolland, J. M., & Harris, W. A. (2011). Becoming violent: Evaluating the mediating effect of hopelessness on the code of the street thesis. *Delinquent Behavior, 32,* 191-233.

Edwards, A. L. (1979) *Techniques of Attitude Scale Construction*. NY, NY: Irvington Publishers, Inc.

Ellen, I.G. & Turner, M.A. (1997). Does neighborhood matter? Assessing recent evidence. *Housing Policy Debate*, 8, 833-866.

Enders, C. K. (2001). A primer on maximum likelihood algorithms available for use with missing data. *Structural Equation Modeling*,8, 128-141.

Enders, C. K., & Bandalos, D. L. (2001). The relative performance of full information maximum likelihood estimation for missing data in structural equation models. Structural *Equation Modeling: A Multidisciplinary Journal*, 8, 430-457.

Frees, E. W. (2004). *Longitudinal and Panel Data: Analysis and Applications in the Social Sciences*. Cambridge, UK: Cambridge University Press.

Geronimus, A. T. & Korenman, S. (1992). The socioeconomic consequences of teen childbearing reconsidered. *The Quarterly Journal of Economics*, 107, 1187-1214.

Gheyas, I. A., & Smith, L. S. (2009) Reconstruction of cross-sectional missing data using neural networks. In D. Palmer-Brown, C. Draganova, E. Pimenidis, & H. Mouratidis (Eds.), *Engineering Applications of Neural Networks* (pp. 28-34). London, England: Springer-Verlag Berlin Heidelberg.

Graham, J. W. (2009). Missing data analysis: Making it work in the real world *Annual Review of Psychology*, 60: 549-576.

Graham, J. W., Cumsille, P. E., & Elek-Fisk, E. (2003). Handling missing data. In J. A. Schinka, & W. F. Velicer (Eds.), *Handbook of Psychology: Research Methods in Psychology* (pp. 87-114). Hoboken, NJ: Wiley & Sons, Inc.

Groves, R. M. & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias. *Public Opinion Quarterly, 72,* 167-189.

Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, *70*, 646-675.

Hardin, J. W. & Hilbe, J. M. (2003). Generalized Estimating Equations. Boca Raton, FL. Chapman & Hal/CRC.

Hatchett, B. F., Holmes, K., Duran, D. A., & Davis, C. (2000). African Americans and research participation: The recruitment process. *Journal of Black Studies, 30,* 664-675.

Hay, G., Kypri, K, Whigham, P, Langley, J. (2008). Potential biases due to geocoding error in spatial analyses of official data. *Health & Place*, 15, 562-567.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47, 53-161.

Heckman, J., Ichimura, H., Smith, J., & Todd, P. (1998). Characterizing selection bias using experimental data. *Econometrica*, 66, 1017-1098.

Heitjan, D. F., &Basu, S. (1996). Distinguishing "missing at random" and "missing completely at random." *The American Statistician*, 50, 207-213.

Hogan, J. W., Roy, J., & Korkontelou, C. (2004). Handling drop-out in longitudinal studies. *Statistics in Medicine*, 23, 1455-1497.

Honaker, J. & King, G. (2007). What to do about missing values in time series cross-sectional data.

Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A.W., & Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions*, 11:3, 133-144.

Howell, D. C. (2002). *Statistical Methods for Psychology*. Belmont, CA: Wadsworth Publishing.

Ibrahim, J. G., Lipsitz, S. R., & Horton, N. (2001). Using auxiliary data for parameter estimation with non-ignorably missing outcomes. *Journal of the Royal Statistical Society*, 50, 361-373.

Israel, B.A., Schulz, A.J., Parker, E.A., & Becker, A.B. (1998). Review of community-based research: Assessing partnership approaches to improve public health. *Annual Review of Public Health*, 19, 173-202.

Jackson, D., White, I. R., & Leese, M. (2010). How much can we learn about missing data?: An exploration of a clinical trial in psychiatry. *Journal of the Royal Statistical Society*, 173, 1-20.

Jamshidian, M., & Mata, M. (2008).Postmodeling sensitivity analysis to detect the effect of missing data mechanisms. *Multivariate Behavioral Research*, 43, 432-452.

Jargowsky, P. A. (1997). *Poverty and Place: Ghettos, Barrios, and the American City*. New York, NY: Russell Sage Foundation.

Jargowsky, P. A. (2003). *Stunning progress, hidden problems: The dramatic decline of concentrated poverty in the 1990s*.Washington, DC: Center on Urban and Metropolitan Policy, Brookings Institution.

Johnson, J. E., Berry, K. J. & Mielke, P. W. (2006). Measures of effect size for chi-squared and likelihood-ratio goodness-of-fit tests. *Perceptual and Motor Skills*, 103, 412-414.

Kenward. M. G. (1998). Selection models for repeated measurements with non-random dropout: An illustration of sensitivity. *Statistics in Medicine*, 17, 2723-2732.

Kerkorian, D., Traube, D. E., & McKay, M. M. (2007). Understanding the African American research experience (KAARE): Implications for HIV prevention. *Social Work in Mental Health, 5,* 295-312.

Kim, K. H., & Bentler, P. M. (2002).Tests of homogeneity of means and covariance matrices for multivariate incomplete data. *Psychometrika*, 67, 609-624.

Krieger, N., Waterman, P., Lemieux, K., Zierler, S., & Hogan, J. W. (2001). On the wrong side of the tracts? Evaluating the accuracy of geocoding in public health research. *American Journal of Public Health*, 91, 1114-1116.

Lian, B. E. & Bolland, J. M. (2009) *Residential mobility and youth violence in underclass neighborhoods.* Unpublished grant proposal to National Institute of Child Health and Human Development.

Liang, K.Y. & Zeger, S.L. (1986). "Longitudinal Data Analysis Using Generalized Linear Models," *Biometrika*, 73, 13-22.

Lin, I., & Schaeffer, N. C. (1995). Using survey participants to estimate the impact of nonparticipation. *The Public Opinion Quarterly*, 59, 236-258.

Little, R. J. (1993). Pattern-mixture models for multivariate incomplete data. *Journal of the American Statistical Association*, 88, 125-134.

Little, R. J. A. (1988). Missing-data adjustments in large surveys. *Journal of Business & Economic Statistics*, 6, 287-296.

Little, R. T. (1995). Modeling the drop-out mechanism in repeated-measures studies. *Journal of the American Statistical Association*, 90, 1112-1121.

Little, R.J.A., & Rubin, D.B. (2002).*Statistical Analysis with Missing Data* (2nd ed.), New York, NY: Wiley-Interscience.

May, M., Pasquier, M., Korner, C., Hofmann, U., Hecker, D., Mende, F. (2009). Handling missing values in GPS surveys using survival analysis. Proceedings from the Third

International Workshop on Data Mining and Audience Intelligence for Advertising. New York, NY: ACM.

McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing Data: A Gentle Introduction*. New York, NY: The Guilford Press.

Minkler, M. & Wallerstein, N. (1997). Improving health through community organization and community building. In M. Minkler (Editor), *Community Organizing & Community Building for Health*. New Brunswick, NJ: Rutgers University Press.

Molenberghs, G. (2008). Every missing not at random model has got a missing at random counterpart with equal fit. Lecture presented at the Graybill Conference.

Molenberghs, G., & Kenward, M. (2007).*Missing Data in Clinical Studies*. New York, NY: Wiley & Sons, Inc.

Molenberghs, G., Thijs, H., Kenward, M. G., & Verbeke, G. (2002). Sensitivity analysis for continuous incomplete longitudinal outcomes. *Statistica Neerlandica*, 57, 112-135.

Nakagawa, S. & Cutthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological Review*, 82, 591-605.

Nakagawa, S., & Freckleton, R. P. (2008). Missing inaction: The dangers of ignoring missing data. *Trends in Ecology and Evaluation*, 23, 592-596.

O'Rourke, T. W. (2003). Methodological techniques for dealing with missing data. *American Journal of Health Studies*, 18, 165-168.

Oakes, J. M. & Rossi, P. H. (2003). The measurement of SES in health research: Current practice and steps toward a new approach. *Social Science & Medicine, 56,* 769-784.

Office of Disease Prevention and Health Promotion, U.S. Department of Health and Human Services. (n.d.) *Healthy People 2010*. Retrieved from http://healthypeople.gov.

Oliver, M. N., Matthews, K. A., Siadaty, M., Hauck, F. R., & Pickle, L. W. (2005). Geographic bias related to geocoding in epidemiologic studies. *International Journal of Health Geographics*, 4, 4-29.

Park, N. S., Lee, B. S., Bolland, J. M., Vazsonyi, A. T., & Fei, S. (2008). Early adolescent pathways of antisocial behaviors in poor, inner-city neighborhoods. *Journal of Early Adolescence, 28,* 185-205.

Park, N. S., Lee, B. S., Sun, F., Vazsonyi, A. T., & Bolland, J. M. (2010). Pathways and predictors of antisocial behaviors in African American adolescents from poor neighborhoods. *Child Youth.Services Review, 32,* 409-415.

Patrician, P. A. (2002). Multiple imputation for missing data. *Research in Nursing & Health*, 25, 76-84.

Pearson (2012). Assessment and Information, Stanford Achievement Test. Retrieved from http://education.pearsonassessments.com/haiweb/cultures/en-us/productdetail.htm?pid=SAT10C.

Piggot, T. D. (2001). A review of methods for missing data. *Educational Research and Evaluation*, 7, 353-383.

Pottick, K. J. & Lerman, P. (1991). Maximizing survey response rates for hard-to-reach inner-city populations. *Social Science Quarterly, 72,* 172-180.

Raymond, M.R., & Roberts, D.M. (1987) A comparison of methods for treating incomplete data in selection research. *Educational and Psychological Measurement*, 47, 13–26.

Rey, W. J. J. (2008). *Introduction to robust and quasi-robust statistical methods*. Berlin: Springer-Verelag.

Roth, P. L., & Switzer III, F. S. (1995). A Monte Carlo analysis of missing data techniques in a HRM setting. *Journal of Management*, 21, 1003-1023.

Rubin, D. B. (1976). Inference and missing data. *Biometrika*, 63, 581-592.

Rubin, D. B., & Schenker, N. (1986). Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *Journal of the American Statistical Association*,81, 366-374.

Schafer, J. L. & Olsen, M. K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavior Research,* 33, 545-571.

Schafer, J. L. (1997). *Analysis of Incomplete Multivariate Data*. London, England: Chapman & Hall.

Schafer, J. L. (1999). Multiple imputation: A primer. *Statistical Methods in Medical Research*, 8, 3-15.

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147-177.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002).*Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.

Shaffer J. P. (1995). Multiple hypothesis testing. *Annual Review of Psychology*, 46, 561-584.

Spano, R., Freilich, J. D., & Bolland, J. (2008). Gang membership, gun carrying, and employment: Applying routine activities theory to explain violent victimization among inner city, minority youth living in extreme poverty. *Justice Quarterly, 25,* 381-410.

Spano, R., Rivera, C., & Bolland, J. (2006). The impact of timing of exposure to violence on violent behavior in a high poverty sample of inner city African American youth. *Journal of Youth and Adolescence, 35,* 681-692.

Spano, R., Rivera, C., & Bolland, J. M. (2010). Does parenting shield youth from exposure to violence during adolescence? A 5-year longitudinal test in a high-poverty sample of minority youth. *Journal of Interpersonal Violence,* 930-949.

Spano, R., Rivera, C., Vazsonyi, A. T., & Bolland, J. M. (2008). The impact of exposure to violence on a trajectory of (declining) parental monitoring: A partial test of the ecological--transactional model of community violence. *Criminal Justice and Behavior, 35,* 1411-1428.

Stewart, C. & Bolland, J. M. (2002). Parental style as a possible mediator of the relationship between religionsity and substance use in Africa- American adolescents. *Journal of Ethnicity in Substance Abuse, 1,* 63-81.

Stoddard, S. A., Henly, S. J., Sieving, R. E., & Bolland, J. (2010). Social connections, trajectories of hopelessness, and serious violence in impoverished urban youth. *Journal of Youth and Adolescence, 40,* 278-295.

Streiner, D. L. (2002). The case of missing data: Methods of dealing with dropouts and other research vagaries. *The Canadian Journal of Psychiatry*, 47, 68-75.

Strickland, M. J., Siffel, C., Gardner, B. R., Berzen, A. K. & Correa, A. (2007). Quantifying geocode location error using GIS methods. *Environmental Health*, 6, 6-10.

Tanguma, J. (2000). A review of the literature on missing data. Paper presented at the Annual Meeting of the Mid-South Educational Research Association.

Teitler, J. O., Reichman, N. E., & Sprachman, S. (2003). Costs and benefits of improving response rates for a hard-to-reach population. *The Public Opinion Quarterly*, 67, 126-138.

U.S. Census Bureau. (2012). 1990 Census. Retrieved from http://www.census.gov/.

U.S. Census Bureau. (2012). 2000. Retrieved from http://www.census.gov/.

U.S. Census Bureau. (2012). *Statistical Abstract of the United States*. Retrieved from http://www.census.gov/prod/2004pubs/03statab/educ.pdf.

U.S. Department of Agriculture: Food and Nutrition Service. (2012). National School Lunch Program. Retrieved from http://www.fns.usda.gov/cnd/lunch/.

Vazsonyi, A. T., Pickering, L. E., & Bolland, J. M. (2006). Growing up in a dangerous developmental milieu: The effects of parenting processes on adjustment in inner-city African American adolescents. *Journal of Community Psychology*, 34, 47-73.

Way, N. & Robinson, M. G. (2003). A longitudinal study of the effects of family, friends, and school experiences on the psychological adjustment of ethnic minority, low-SES adolescents. *Journal of Adolescent Research*, 18, 324-346.

Weisberg, H. F. (2005). *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago: The University of Chicago Press.

Yang, L. & Tsiatis, A. A. (2001). Efficiency study of estimators for a treatment effect in a pretest-postest trial. *The American Statistician*, 55, 314-321.

Appendix A

Items used from the Mobile Youth Survey

1. How long have you lived in your neighborhood
    a. Less than one year
    b. About one year
    c. About two years
    d. About three years
    e. About four years
    f. Five years or longer

2. During the past year (12 months), were you arrested?
    a. No
    b. Yes

3. How many hours each week do you work at a paid job?
    a. None; I don't have a job
    b. 1 to 5 hours each week
    c. 6 to 10 hours each week
    d. 11 to 20 hours each week
    e. More than 20 hours each week

4. In the past 3 months (90 days), did you carry a gun?
    a. No
    b. Yes, just once
    c. Yes, more than once

5. In the past year (12 months), did you use marijuana?
    a. No
    b. Yes, just once
    c. Yes, more than once

Appendix B

Office for Research

Institutional Review Board for the
Protection of Human Subjects

THE UNIVERSITY OF
ALABAMA
RESEARCH

September 28, 2011

John M. Bolland, Ph.D.
The University of Alabama
Human Environmental Sciences
Box 870158

Re: IRB Protocol # 09-015-R2: "Testing Methodological Assumptions about the Mobile Youth Survey"

Dear Dr. Bolland:

The University of Alabama Non-Medical IRB recently met to consider your renewal application. The IRB voted to approve your protocol for one year.

Your application will expire on September 27, 2012. If your research will continue beyond this date, please complete FORM: IRB Renewal Application. If you need to modify the study, please submit FORM: Modification of an Approved Protocol. Changes in this study cannot be initiated without IRB approval, except when necessary to eliminate apparent immediate hazards to participants. When the study closes, please complete FORM: Request for Study Closure (Investigator).

Should you need to submit any further correspondence regarding this application, please include the above application number.

Good luck with your research.

Sincerely,

Stuart Usdan, Ph.D.
Chair, Non-Medical IRB
The University of Alabama

358 Rose Administration Building
Box 870127
Tuscaloosa, Alabama 35487-0127
(205) 348-8461
FAX (205) 348-7189
TOLL FREE (877) 820-3066

IRB#: 09-015-R2

UNIVERSITY OF ALABAMA INSTITUTIONAL REVIEW BOARD FOR THE PROTECTION OF HUMAN SUBJECTS
REQUEST FOR APPROVAL OF RESEARCH INVOLVING HUMAN SUBJECTS

I. Identifying information

	Principal Investigator	Second Investigator	Third Investigator
Name:	John M. Bolland		
Department:			
College:	Human Environmental Sciences		
University:	University of Alabama		
Address:	208 CDRC Box 870158		
Telephone:	8-9953		
FAX:	8-8153		
E-mail:	jbolland@ches.ua.edu		

Title of Research Project: Testing Methodological Assumptions about the Mobile Youth Survey

Date Printed: 9/14/2011 Funding Source: NICHD

Type of Proposal: __New | X Revision X Renewal __Completed | __Exempt

Attach a renewal application

Attach a continuing review of studies form

Please enter the original IRB # at the top of the page

UA faculty or staff member signature:

II. NOTIFICATION OF IRB ACTION (to be completed by IRB):

Type of Review: ✓ Full board ____ Expedited

IRB Action:

__Rejected Date:_____

__Tabled Pending Revisions Date:_____

__Approved Pending Revisions Date:_____

✓ Approved—this proposal complies with University and federal regulations for the protection of human subjects.

Approval is effective until the following date: 9/27/2012

Items approved:

✓ Research protocol: dated

__ Informed consent: dated

__ Recruitment materials: dated

__ Other: dated

Approval signature _Date 9/28/2011

AAHRPP DOCUMENT # 140

THE UNIVERSITY OF ALABAMA
HUMAN RESEARCH PROTECTION PROGRAM

IRB Application Study Personnel Sheet (Insert after Face Sheet

Study Personnel and Study Responsibility

Name and Degree(s) or student status (e.g., master's student)	Study Position Title (PI, Interviewer, Data Analyst, etc.)	Study Responsibilities	Date of Certificate of Investigator /staff Human Subjects Training
Cassandra Coddington	Investigator	Analyse data, prepare manuscripts	
Brad Lian	Investigator	Analyse data, prepare manuscripts	
Anneliese Bolland	GRA	Analyse data, prepare manuscripts	
Shannon Hitchcock	GRA	Analyse data, prepare manuscripts	
Nicole du Maine	GRA	Analyse data, prepare manuscripts	
Unita (Shanette) Granstaff	GRA	Analyse data, prepare manuscripts	
John Bolland	Investigator	Analyse data, prepare manuscripts	