UNDERSTANDING ADOLESCENT SURVEY RESPONSES: IMPACT OF MODE AND OTHER CHARACTERISTICS ON DATA OUTCOMES AND QUALITY

by

ERIKA SHAUN TRAPL

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Dissertation Adviser: Elaine A. Borawski, Ph.D.

Department of Epidemiology and Biostatistics

CASE WESTERN RESERVE UNIVERSITY

May, 2007
We hereby approve the dissertation of

Erika Shaun Trapl

candidate for the Ph.D. degree *

(signed) Elaine A. Borawski
(Chair of the committee)

H. Gerry Taylor

Natalie Colabianchi

David Litaker

(date) April 6, 2007

*We also certify that written approval has been obtained for any proprietary material contained therein.
I grant to Case Western Reserve University the right to use this work, irrespective of any copyright, for the University’s own purposes without cost to the University or to its students, agents and employees. I further agree that the University may reproduce and provide single copies of the work, in any format other than in or from microforms, to the public for the cost of reproduction.

___________________________Erika Shaun Trapl__________________________
DEDICATION

This work is dedicated to my parents, Ian and Julie, and my husband, Jeremy.
<table>
<thead>
<tr>
<th>CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES..................................................................................7</td>
</tr>
<tr>
<td>LIST OF FIGURES................................................................................10</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS.......................................................................11</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS.............................................................................12</td>
</tr>
<tr>
<td>ABSTRACT.............................................................................................13</td>
</tr>
<tr>
<td>CHAPTER 1: INTRODUCTION....................................................................15</td>
</tr>
<tr>
<td>CHAPTER 2: BACKGROUND AND LITERATURE REVIEW.................................21</td>
</tr>
<tr>
<td>1. Modes of Data Collection..............................................................21</td>
</tr>
<tr>
<td>1.1 Self-Administered Approaches to Data Collection..........................21</td>
</tr>
<tr>
<td>1.2. Computer Assisted Self Interview &amp; Audio-Computer Assisted Self Interview 24</td>
</tr>
<tr>
<td>1.3 Using Personal Digital Assistants for Data Collection..................26</td>
</tr>
<tr>
<td>2. Previous APDA Study.......................................................................30</td>
</tr>
<tr>
<td>3. Cognitive Burden: Literacy, Language and Attention.......................34</td>
</tr>
<tr>
<td>4. Confidentiality, Privacy and Environmental Factors.......................36</td>
</tr>
<tr>
<td>5. Demographic Confounders..................................................................38</td>
</tr>
<tr>
<td>6. Conceptual Framework......................................................................39</td>
</tr>
<tr>
<td>CHAPTER 3: SPECIFIC AIMS AND HYPOTHESES.....................................42</td>
</tr>
<tr>
<td>1. Specific Aim 1...............................................................................42</td>
</tr>
<tr>
<td>2. Specific Aim 2...............................................................................43</td>
</tr>
<tr>
<td>3. Specific Aim 3...............................................................................46</td>
</tr>
<tr>
<td>4. Specific Aim 4...............................................................................47</td>
</tr>
<tr>
<td>CHAPTER 4: RESEARCH DESIGN &amp; METHODS.........................................49</td>
</tr>
<tr>
<td>1. Overview.........................................................................................49</td>
</tr>
<tr>
<td>2. School Sites....................................................................................49</td>
</tr>
</tbody>
</table>
3. Student Recruitment and Randomization ........................................................ 50

4. Data Collection Mode ...................................................................................... 54

5.0 Data Sources .................................................................................................. 55

5.1. Self-Administered Survey Content ............................................................ 55
5.2. Group Reading Assessment and Diagnostic Evaluation (GRADE) .......... 56
5.3. Teacher Assessment of Students ................................................................. 56
5.4. Data Collection Observer Checklist ............................................................. 57
5.5. Debriefer Questionnaire .............................................................................. 57

6.0 Outcomes ....................................................................................................... 58

6.1 Number of Questions Answered ................................................................. 58
6.2 Missing Data ................................................................................................. 58
6.3 Failed Skip Patterns & Inconsistent Answers ................................................ 58
6.4 Exposure to sensitive questions .................................................................... 58
6.5 Reporting of Behaviors ................................................................................ 59
6.6 Internal Consistency and Factorial Invariance ............................................ 59
6.7 Student Evaluation ....................................................................................... 59

7.0 Independent Variables .................................................................................. 60

7.1 Reading Comprehension .............................................................................. 60
7.2 Language Barriers ....................................................................................... 61
7.3 Reading Enjoyment and Reading Difficulty ................................................ 62
7.4 Inattentiveness ............................................................................................. 63
7.5 Perceptions of Privacy and Confidentiality ................................................ 64

8.0 Data Analytic Plan ....................................................................................... 64

8.1 Analytic Approach for Specific Aim 1 ........................................................ 65
8.2 Factorial Invariance ..................................................................................... 65
8.3 Analytic Approach for Specific Aim 2 ........................................................ 67
8.4 Analytic Approach for Specific Aim 3 ........................................................ 68
8.5 Power Analyses .......................................................................................... 69

CHAPTER 5: RANDOMIZATION, SAMPLE DESCRIPTION, AND METHODOLOGICAL ISSUES ................................................................. 70

1. Summary of Recruitment and Randomization .............................................. 70

2. Sample Description ....................................................................................... 72

2.1 Summary of Demographic Variables ......................................................... 72
2.2 Summary of Individual Characteristics ...................................................... 73
2.3 Summary of Outcome Measures ................................................................. 75
3. Survey Environment ........................................................................................................ 78

4. Methodological Issues .................................................................................................. 79

4.1. Teacher Academic Assessment ............................................................................... 79
4.2. Reading Enjoyment and Reading Difficulty ............................................................. 80

CHAPTER 6: RESULTS FOR SPECIFIC AIM 1 .................................................................. 81

1. Impact of Mode on Number of Questions Answered ................................................. 81
2. Impact of Mode on Data Quality: Missing Data ......................................................... 82
3. Impact of Mode on Data Quality: Inconsistent or Missed Skips ............................... 86
4. Impact of Mode on Data Quality: Construct Reliability and Factorial Invariance .... 88

4.1. Results of Factorial Invariance Testing .................................................................. 89
4.1.1. Reading Enjoyment ......................................................................................... 90
4.1.2. Reading Difficulty ......................................................................................... 92
4.1.3. Sex Beliefs ..................................................................................................... 93
4.1.4. Consequences of Early Sexual Behavior ....................................................... 95
4.1.5. Parental Sex Beliefs ..................................................................................... 97
5. Impact of Mode on Data Quality: Reporting of Behavior ........................................... 99
6. Impact of Mode on Student Evaluation of Survey Experience ............................... 100

6.1. Overall Honesty, Discomfort, and Enjoyment ....................................................... 100
6.2. Mode Preference .................................................................................................... 101

CHAPTER 7: RESULTS FOR SPECIFIC AIM 2 .............................................................. 103

1. Description of the Individual Characteristics in the Analyses ................................. 103

1.1. Sentence Comprehension .................................................................................... 103
1.2. Listening Comprehension .................................................................................. 104
1.3. Language Mastery ............................................................................................ 104
1.4. Reading Enjoyment ......................................................................................... 105
1.5. Reading Difficulty ............................................................................................ 105
1.6. Inattentiveness .................................................................................................. 106
1.7. Perceived Confidentiality ................................................................................ 106
1.8. Perceived Privacy ............................................................................................. 107
1.9. Consideration of Confounders in Multivariate Models ...................................... 107

2. Impact of Individual Characteristics on Number of Questions Answered ............ 108
2.1 Multivariate Results for Number of Questions Answered............................... 110
2.1.1 Sentence Comprehension.............................................................................. 111
2.1.2 Listening Comprehension ........................................................................... 111
2.1.3 Perceived Reading Difficulty................................................................. 112
2.1.4 Inattentiveness......................................................................................... 113
2.1.5 Perceived Privacy.................................................................................... 113

3. Impact of Individual Characteristics on Missing Data ................................. 114
3.1 Multivariate Models for Missing Data......................................................... 116
3.1.1 Inattentiveness......................................................................................... 117
3.1.2 Perceived Privacy.................................................................................... 117

4. Impact of Individual Characteristics on Student Evaluation ......................... 118
4.1 Multivariate Models for Overall Reported Honesty ..................................... 120
4.1.1 Sentence Comprehension......................................................................... 121
4.1.2 Language Barriers:.................................................................................. 122
4.1.3 Reading Difficulty................................................................................. 123
4.1.4 Inattentiveness....................................................................................... 124
4.1.5 Perceived Confidentiality....................................................................... 124
4.2 Multivariate Models for Overall Discomfort .............................................. 125
4.2.1 Sentence Comprehension....................................................................... 126
4.2.2 Inattentiveness...................................................................................... 126
4.2.3 Perceived Confidentiality..................................................................... 127
4.3 Multivariate Models for Overall Enjoyment Completing the Survey .......... 128
4.3.1 Reading Enjoyment............................................................................... 129
4.3.2 Perceived Confidentiality................................................................... 129

5.0 Factorial Invariance ..................................................................................... 130
5.1 Factorial Invariance by Sentence Comprehension....................................... 131
5.1.1 Factorial Invariance by Sentence Comprehension: Reading Enjoyment...... 131
5.1.2 Factorial Invariance by Sentence Comprehension: Reading Difficulty ...... 132
5.1.3 Factorial Invariance by Sentence Comprehension: Sex Beliefs.............. 133
5.1.4 Factorial Invariance by Sentence Comprehension: Consequences of Early Sexual Behavior.............................................................. 133
5.1.5 Factorial Invariance by Sentence Comprehension: Parental Sex Beliefs .... 134
5.2 Factorial Invariance by Listening Comprehension....................................... 135
5.2.1 Factorial Invariance by Listening Comprehension: Reading Enjoyment...... 135
5.2.2 Factorial Invariance by Listening Comprehension: Reading Difficulty...... 135
5.2.3 Factorial Invariance by Listening Comprehension: Sex Beliefs.............. 136
5.2.4 Factorial Invariance by Listening Comprehension: Consequences of Early Sexual Behavior.............................................................. 137
5.2.5 Factorial Invariance by Listening Comprehension: Parental Sex Beliefs .... 137
5.3 Factorial Invariance by Language/US Index............................................... 138
5.3.1 Factorial Invariance by Language/US Index: Reading Enjoyment........... 138
LIST OF TABLES

Table 2.1: Scale Reliabilities: Regular Education, English as a Second Language, and Special Education Students ................................................................. 32

Table 5.1: Stratified Randomization of Schools to Mode ................................................. 71
Table 5.3: Individual Characteristics ................................................................................ 74
Table 5.4: Sample Outcomes ............................................................................................ 76
Table 5.5: Survey Environment Characteristics ................................................................ 78
Table 5.6: Correlations Among Teacher Measures .......................................................... 80

Table 6.1: Questions Answered by Mode ......................................................................... 82
Table 6.2: Frequency of Missing Data by Mode ............................................................. 84
Table 6.3: Missing Data by Mode: Percentage of Students with any Missing Data .......... 86
Table 6.4: Missed Skips and Inconsistent Responses Among SAQ .................................. 88
Table 6.5: Construct Reliabilities by Mode ...................................................................... 89
Table 6.6: Model Fit for Reading Enjoyment: Overall and By Mode ................................ 91
Table 6.7: Model Fit for Reading Difficulty: Overall and By Mode ................................ 93
Table 6.9: Model Fit for Consequences of Early Sexual Behavior: Overall and By Mode . 97
Table 6.10: Model Fit for Parental Sex Beliefs: Overall and By Mode ............................ 98
Table 6.12: Student Evaluation by Mode ....................................................................... 100
Table 6.13: Mode Preference By Mode .......................................................................... 101

Table 7.1: Number of Questions Answered by Characteristic ....................................... 109
Table 7.2: Regression Parameters for Questions Answered and Demographic Variables .................................................................................................................. 110
Table 7.3: Regression Parameters for Questions Answered and Sentence Comprehension Stanine .............................................................................................................. 111
Table 7.4: Regression Parameters for Questions Answered and Listening Comprehension ...................................................................................................................... 112
Table 7.5: Regression Parameters for Questions Answered and Reading Difficulty ....... 112
Table 7.6: Regression Parameters for Questions Answered and Perceived Inattentiveness ......................................................................................................................... 113
Table 7.8: Missing Data by Characteristic ....................................................................... 115
Table 7.9: Regression Parameters for Missing Data and Demographics ......................... 116
Table 7.10: Regression Parameters for Missing Data and Inattentiveness ....................... 117
Table 7.11: Regression Parameters for Missing Data and Perceived Privacy ................. 118
Table 7.12: Student Evaluation by Individual Characteristics ......................................... 120
Table 7.13: Regression Parameters for Overall Reported Honesty and Demographics .. 121
Table 7.14: Regression Parameters for Overall Reported Honesty and Sentence Comprehension Stanine ........................................................................................................ 122
Table 7.15: Regression Parameters for Overall Reported Honesty and Language/US Index ............................................................... 123
Table 7.16: Regression Parameters for Overall Reported Honesty and Reading Difficulty ............................................................... 123
Table 7.17: Regression Parameters for Overall Reported Honesty and Perceived Confidentiality .................................................. 124
Table 7.18: Regression Parameters for Overall Reported Honesty and Perceived Confidentiality .................................................. 125
Table 7.19: Regression Parameters for Discomfort and Demographics ......................................................................................... 125
Table 7.20: Regression Parameters for Discomfort and Sentence Comprehension Stanine ............................................................... 126
Table 7.21: Regression Parameters for Discomfort and Inattentiveness ......................................................................................... 127
Table 7.22: Regression Parameters for Discomfort and Perceived Confidentiality ................................................................. 128
Table 7.23: Regression Parameters for Enjoyment and Demographics ......................................................................................... 128
Table 7.24: Regression Parameters for Enjoyment and Reading Enjoyment ................................................................................... 129
Table 7.25: Regression Parameters for Enjoyment and Perceived Confidentiality ................................................................. 130
Table 7.26: Model Fit for Reading Enjoyment: Overall and By Sentence Comprehension ................................................................. 132
Table 7.27: Model Fit for Reading Difficulty: Overall and By Sentence Comprehension ................................................................. 132
Table 7.28: Model Fit for Sex Beliefs: Overall and By Sentence Comprehension ................................................................. 133
Table 7.29: Model Fit for Consequences of Early Sexual Behavior: Overall and By Sentence Comprehension ................................................. 134
Table 7.30: Model Fit for Parental Sex Beliefs: Overall and By Sentence Comprehension ................................................................. 134
Table 7.31: Model Fit for Reading Enjoyment: Overall and By Listening Comprehension ................................................................. 135
Table 7.32: Model Fit for Reading Difficulty: Overall and By Listening Comprehension ................................................................. 135
Table 7.33: Model Fit for Sex Beliefs: Overall and By Listening Comprehension ................................................................. 136
Table 7.34: Model Fit for Consequences of Early Sexual Behavior: Overall and By Listening Comprehension ................................................. 137
Table 7.35: Model Fit for Parental Sex Beliefs: Overall and By Listening Comprehension ................................................................. 137
Table 7.36: Model Fit for Reading Enjoyment: Overall and By Language/US Index ................................................................. 139
Table 7.37: Model Fit for Reading Difficulty: Overall and By Language/US Index ................................................................. 139
Table 7.38: Model Fit for Sex Beliefs: Overall and By Language/US Index ................................................................. 140
Table 7.39: Model Fit for Consequences of Early Sexual Behavior: Overall and By Language/US Index ................................................. 141
Table 7.40: Model Fit for Parental Sex Beliefs: Overall and By Language/US Index.. 141
Table 8.1: Regression Parameters for Questions Answered and Sentence Comprehension
Table 8.2: Regression Parameters for Questions Answered and Listening Comprehension
Table 8.3: Regression Parameters for Questions Answered and Reading Difficulty
Table 8.4: Regression Parameters for Questions Answered and Inattentiveness
Table 8.5: Regression Parameters for Missing Data and Inattentiveness
Table 9.1: Regression Parameters for Questions Answered and Privacy
Table 9.2: Regression Parameters for Missing Data and Privacy
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Conceptual Model</td>
<td>18</td>
</tr>
<tr>
<td>2.1</td>
<td>Tourangeau &amp; Smith’s Model</td>
<td>39</td>
</tr>
<tr>
<td>2.2</td>
<td>Conceptual Model</td>
<td>40</td>
</tr>
<tr>
<td>4.1</td>
<td>Sample Confirmatory Factor Analytic Model in AMOS</td>
<td>67</td>
</tr>
<tr>
<td>6.1</td>
<td>Example of Failed Skip Pattern and Logically Inconsistent Response</td>
<td>87</td>
</tr>
<tr>
<td>6.2</td>
<td>Measurement Model for Reading Enjoyment</td>
<td>91</td>
</tr>
<tr>
<td>6.3</td>
<td>Measurement Model for Reading Difficulty</td>
<td>92</td>
</tr>
<tr>
<td>6.4</td>
<td>Measurement Model for Sex Beliefs</td>
<td>94</td>
</tr>
<tr>
<td>6.5</td>
<td>Measurement Model for Consequences of Early Sexual Behavior</td>
<td>96</td>
</tr>
<tr>
<td>6.6</td>
<td>Measurement Model for Parental Sex Beliefs</td>
<td>98</td>
</tr>
<tr>
<td>8.1</td>
<td>Plot of Moderating Effect of Reading Difficulty on Questions Answered</td>
<td>148</td>
</tr>
<tr>
<td>9.1</td>
<td>Model for Testing for Mediation</td>
<td>155</td>
</tr>
</tbody>
</table>
LIST OF ABBREVIATIONS

A-CASI: Audio computer-assisted self-interviewing
APDA: Audio-enhanced personal digital assistant
CASI: Computer-assisted self-interviewing
CMSD: Cleveland Municipal School District
GRADE: Group Reading Assessment and Diagnostic Evaluation
LC: Listening comprehension
PDA: Personal digital assistant
SAQ: Paper-pencil self-administered questionnaire
SC: Sentence comprehension
SWAN: Strengths and Weakness of ADHD-symptoms and Normal Behavior
ACKNOWLEDGEMENTS

First and foremost, I would like to thank to Elaine Borawski, my dissertation advisor, mentor, and friend. Over the past several years, Elaine has guided and encouraged me to grow as an academic and a researcher. I am quite proud to have been one of her doctoral students. She has led me well through my academic career, and I look forward to continuing our professional and personal relationship.

I would like to acknowledge the members of my dissertation committee for their guidance and advice, and for making this such an enjoyable process: Drs. David Litaker, H. Gerry Taylor, and Natalie Colabianchi.

Many colleagues, new and old, have greatly contributed to my success along the way. Paul Stork made surveying via audio-enhanced personal digital assistant not just an idea, but my new reality. M. Cecil Smith, an expert in adult literacy, welcomed me into the area of literacy research. Finally, my work could not have been as successful without the everyday support and assistance from the staff (past and present) of the Center for Health Promotion Research at Case Western Reserve University: LaTylia Clemmons, Danyel Savarda, Melissa Hanahan, Manish Jhawar, Ramakrishnan Vijayakumar, Kathryn Przepyszny, Matthew Russell, Ashley Brooks, Dana Jankus, Sarah Boyle, Alida Walker and Nital Subhas. I also would like to thank and acknowledge the students, teachers, and staff of Cleveland Municipal School District who participated in this study.

This work was supported through funds from the National Institute of Child Health and Human Development: HD-R01-41364-S2.
Understanding Adolescent Survey Responses: Impact of Mode and Other Characteristics on Data Outcomes and Quality

Abstract

by

ERIKA SHAUN TRAPL

BACKGROUND: Adolescents spend a majority of time at school, making the classroom a natural venue not only for risk behavior interventions, but for administration of surveys assessing the effectiveness of these interventions. Computerized approaches provide valuable advantages over self-administered questionnaires (SAQ); however, few comparative studies have included the personal digital assistant (PDA) or the PDA with audio-enhancement (APDA). Even less is known about the role of cognitive burden (e.g. reading ability and preferences, language mastery, attentiveness) on survey completion by different mode.

PURPOSE: This study has four aims: (1) to understand the impact of three data collection modes (SAQ, PDA, APDA) on the number of survey questions answered, missing data, data consistency, and student evaluation of the survey experience; (2) examine associations between cognitive burden and outcomes; (3) assess moderating effects of cognitive burden on relationships between mode and outcomes; (4) assess the role (mediating or moderating) of perceptions of confidentiality/privacy on relationships between mode and outcomes.
METHODS: Two-hundred seventy-five students were recruited from seven urban K-8 schools in the Midwest. Consented participants were stratified based on reading scores and randomized to complete a survey by SAQ, PDA, or APDA. Upon completion, students completed a paper-based debriefing survey, assessing student survey experience. Academic and behavioral assessments were completed by each student’s teacher.

RESULTS: APDA respondents completed significantly more questions compared to SAQ and PDA. Both PDA and APDA had significantly less missing data than SAQ. No differences were found for student evaluation by mode. Several measures of cognitive burden were related to outcomes. While data inconsistency (e.g., factorial variance) was found across modes, data from students at varying levels of cognitive burden was not significantly different. One significant moderator was found with the relationship between mode and number of questions varying by perceived reading difficulty. Finally, neither perceived confidentiality nor privacy acted as a moderator or mediator.

CONCLUSIONS: This study indicates strong benefits to be gained by the use of APDA for school-based data collection with adolescents, particularly those with reading, language or attention issues. While students with greater cognitive burden answered fewer questions, their data was found to reliable.
Chapter 1: INTRODUCTION

Early adolescence (ages 10-14) is increasingly seen as an appropriate time to introduce interventions aimed at preventing risky behaviors such as drug and alcohol use, violence and unprotected sexual activity (Peterson and Leffert, 1995; Resnick et al., 1997). As adolescents spend a majority of their time at school, the classroom becomes a natural venue for implementing these interventions (Gans and Brindis, 1995). Researchers often use survey methodology to assess the effectiveness of these interventions. Due to the sensitive nature of the targeted behavior, researchers must ask highly sensitive questions that often involve sophisticated branching patterns that student must navigate. Furthermore, these assessments usually require reading competency and command of the English language, neither of which is guaranteed in today’s diverse urban school settings.

While the most common and economical method for collecting data from a large number of students is the paper/pencil self-administered questionnaire (SAQ), there are many limitations to this method. SAQs require moderate reading skills, often require students to navigate sophisticated skip patterns, and necessitate large testing areas to guarantee privacy. Research, mostly on adults, has shown that the administration of and data quality from the SAQ is largely influenced by the testing environment, as well as literacy and attention demands of the respondent (Beebe et al, 1998; Couper, Singer, and Tourangeau, 2003; Tourangeau and Smith; 1998; Couper et al, 1998). Moreover, the test environment can have an effect on the respondents’ perceptions of privacy, which can indirectly affect administration and data quality (Couper, Singer, and Tourangeau, 2003; Brener, Billy and Grady, 2003; Kann et al, 2002).
A second self-administered data collection option that addresses many of the limitations of the SAQ is the computer-assisted self-interview (CASI) system, with or without audio-enhancement (i.e., questions and/or responses are read to student via headphones). CASI systems are used with both desktop and laptop computer systems, allowing for transportability of the system (Beebe et al, 1997; Couper et al, 1998). However, the cost, resource and staffing requirements of this method reduce its feasibility in school-based research where space and resources (e.g. electrical) are limited.

A third self-administered data collection option is the small, handheld personal digital assistant (PDA), which has many of the advantages of the CASI system, but is cheaper and more portable than the desktop or laptop based systems. However, there are very few published studies on the use of PDAs in school-based research and until recently (Trapl et al, 2005), there were no studies using PDAs with audio-enhancement.

Trapl (2004) sought to develop and pilot a PDA-based data collection system with audio-enhancement (APDA), completed in October, 2003. This method was successfully used to collect baseline data for a large, longitudinal NIH-funded study involving middle school adolescents (Trapl et al, 2005). While this research answered questions regarding feasibility of using the APDA system, the study was unable to answer questions of comparability or improvement over existing methods of data collection. Further, there was no way to validate empirically observations made in the field, such as students being less distracted and willing to tolerate a much longer survey, or reduction in the exposure to questions with sensitive content without loss of data. Most importantly, students who spoke English as a second language (ESL) and special education students were not only included in the process, but completed the survey with no help and with similar time-to-
completion rates as their peers. However, closer examination of data collected in this study dampened investigators initial enthusiasm as it appeared that the internal consistency of several measures among the ESL and special education students was lower than that of their regular education peers, thus threatening the appropriateness of their inclusion in further data analyses.

This previous work underscores several important knowledge gaps. It is unclear whether audio-enhancement removes the burden of reading, resulting in improved completion rates or if completion rates may be improved as a result of reduction in distractions due to the audio-enhancement, allowing students to stay focused and on task. Further, while it appears that a broader range of students may be included in survey-based data collection using APDA, it remains unclear if students with limited reading skills and language mastery are appropriately comprehending and responding to the survey questions. Lastly, while some parallels may be drawn between APDA and desktop or laptop ACASI, there has been no comparative study published in the literature examining the potential benefits or limitations of the use of audio-enhanced PDAs in any survey research.

Tourangeau and Smith (1996) propose that data quality, such as level of reporting, accuracy, reliability, and rate of missing item responses, is impacted by such characteristics as cognitive burden, perceptions of privacy and legitimacy of the research or survey. These characteristics themselves are affected by the survey mode. For example, their model implies that the computerization of a survey influences the cognitive burden of the respondent, which in turn impacts the rate of missing data. By
reducing missing data, the overall quality of the data is improved, providing for a more complete dataset and more reliable results.

Figure 1.1: Conceptual Model

Our study builds upon the work of Tourangeau and Smith (1996) and expands the model to include other factors including demographics (e.g. gender, socioeconomic status), cognitive burden (i.e. reading comprehension, listening comprehension, English mastery, reading enjoyment and difficulty, inattentiveness), and perceptions of confidentiality and privacy may have direct effects on outcomes (number of questions answered, data quality and student evaluation)(see Figure 1.1). Further, we suspect that
these effects could modify the relationship between data collection mode and outcomes and that perceptions of confidentiality and privacy of the survey environment may actually act as mediators.

This exploratory study will address the following aims:

- To examine the differential effects of three different data collection modes (SAQ, PDA, APDA) on the number of questions answered, data quality, and student evaluation of the process (Path A);

- To assess the impact of cognitive burden (i.e. reading competency, language barriers, inattentiveness, perceived burden of reading) on the number of questions answered, data quality and student evaluation (Path B);

- To determine whether and how these characteristics modify the relationship between data collection mode and data outcomes (Path C).

- To clarify the role of perceived confidentiality and perceived privacy on the number of questions answered, data quality and student evaluation as either a moderator (Path D1) or a mediator (Paths D2).

This comparative and exploratory study was designed to inform future studies, and that of others utilizing this technology with school-age children by comparing data derived from three different data collection modes. From this aim alone, our study will contribute to a literature that has been found to be remarkably sparse. In addition, we sought to increase understanding of how adolescents with different cognitive and behavioral characteristics respond to different aspects of the different data collection
modes by exploring a variety of possible relationships between individual characteristics, data collection mode, and data outcomes.
Chapter 2: Background and Literature Review

Chapter Overview

The contents of this chapter will provide the reader with a brief summary of the impact of survey mode on data outcomes, as well as an understanding and rationale of current survey methodology used to collect sensitive data from adolescents. Further, it will highlight and discuss the individual characteristics of adolescents that may have great potential to impact reliable responses and inclusion in self-administered survey data collection.

1. Modes of Data Collection

1.1 Self-Administered Approaches to Data Collection

Self administered questionnaires are one of the most efficient approaches to survey research, allowing researchers to reach the largest number of respondents in the most economical way, requiring one survey administrator for a large number of respondents and minimal investment in equipment and software to produce surveys (Bourque and Felder, 1995). Further, research has shown that removal of personal interaction with an interviewer provides a more private environment for both adults and adolescents and reduces reporting bias (Aquilino 1994; Couper et al, 1998; Gribble et al, 1999; Jones and Forrest, 1992; Tourangeau and Smith; 1996; Tourangeau and Smith, 1998).

However, the SAQ does not account for respondent literacy and ability to navigate the survey, thereby potentially reducing data quality among all respondents (Gribble et al, 1999; Lessler and O-Reilly, 1997), especially those with reading limitations or cognitive deficiencies. In examining non-response in the National
Educational Longitudinal Study of 1994 administered in a paper-and-pencil format, Wolfe (2003) found that academically lower achieving students were more likely to have missing data than those with higher achievement scores. In addition to missing data, improperly navigated skip patterns, logically inconsistent answers, and other recording errors (e.g. choosing two responses to a question requiring a single response) have also been found to be linked to low literacy (Al-Tayyib et al, 2002), with one study finding that only 56.5% of adults with a 7th/8th grade literacy level were able to complete a paper-and-pencil health survey on alcohol use without errors. Further, even among adults with high literacy levels, the ability to follow form instructions and navigate skip patterns was somewhat limited, with errors detected on almost 30% of an alcohol use survey. The rate of error detection (e.g. identifying logically inconsistent responses or two responses provided to a question requiring a single response) was even higher (46.2%) among high literacy adults completing a same-sex sexual intercourse and masturbation survey, possibly due to the sensitive nature of the survey content.

Despite these limitations, researchers still consider self-administered, paper-and-pencil questionnaires to be the preferred approach when compared to interviewer-administered survey (Catania, McDermott and Pollack, 1986), including face-to-face interviews or telephone surveying. While these latter approaches may better address many of the limitations of SAQ, the pervasiveness of self-presentation bias (e.g. respondents report behaviors or beliefs consistent with social norms and underreport information that is embarrassing, socially stigmatizing, or illegal) under conditions of interviewer-administered questionnaires, especially when respondents are asked highly sensitive questions, is of much greater concern (Sudman and Bradburn, 1974).
Several studies in adults have identified increased reporting of sensitive behaviors among self-administered approaches compared interviewer-administered approaches in the areas of sexual behavior (Boekeloo et al, 1994; Hewett, 2002; Jobe et al, 1997; Metzger et al, 2000; Robinson and West, 1992; Tourangeau et al, 1997), illicit drug use (Aquilino, 1994, Aquilino and LoSciuto, 1990, Schober et al, 1992; Tourangeau and Smith, 1996; and Turner, Lessler and DeVore, 1992), alcohol consumption (Aquilino and LoSciuto, 1990; Waterton and Duffy, 1984), reporting on abortion (London and Williams, 1990; Mosher and Duffer, 1994; Mott, 1985) and mental health (Epstein, Barker and Kroutil, 2001). As there is not always a gold standard to validate reporting of behavior (e.g. sexual intercourse), it is assumed that self-presentation and social desirability bias are reduced in the absence of an interviewer, thus increased reporting of these sensitive behaviors is assumed to be more accurate (Brener et al, 2003; Tourangeau and Smith, 1996, Boekeloo et al, 1994, Turner et al, 1998b; Des Jarlais et al, 1999). Similarly, extensive research with adolescents has shown that self-administered techniques are much more likely to yield increased reports of sensitive behaviors when compared with interviewer-administered methods (Gribble et al, 1999; Beebe et al, 1997; Ellen et al, 2002; Romer et al, 1997; Turner et al, 1992; Turner et al, 1998; Turner et al 1998b).

Still, in survey research with youth, researchers oftentimes encounter resistance from school administration and parents due to the fact that the SAQ cannot limit exposure of detailed and sensitive questions (e.g., types of sexual behavior, condom use frequency, number and types of partners) to only those for whom it applies (i.e., students who have
ever engaged in sexual behavior), potentially exposing students to developmentally inappropriate questions.

### 1.2. Computer Assisted Self Interview & Audio-Computer Assisted Self Interview

Computerized data collection methods were developed to address many of the limitations of paper-pencil surveys and interviewer-administered surveys. Computer-assisted self-interviewing (CASI) provides computer-controlled navigation of sophisticated branching patterns (to skip past non-applicable and developmentally inappropriate questions), programmed consistency checks, and automatic data entry (Beebe et al, 1997; Ramos, Sedivi and Sweet 1998). Because most programs bring unanswered questions to the attention of the respondent before allowing them to advance, missing data are reduced (Couper, Singer, and Tourangeau, 2003; Couper, et al, 1998; Hallfors et al, 2000). CASI also has the benefit of standardizing survey administration across the study population (Supple, Aquilino, and Wright, 1999; Lessler and O’Reilly, 1997). Audio-enhancement features were added to CASI programs (i.e. A-CASI), allowing the survey respondent to be read the questions through headphones, thereby potentially reducing issues related to literacy and comprehension (Romer et al, 1997; Turner et al, 1998b).

Survey researchers have carried out extensive comparative experiments, mostly with adults, to better understand the impact of computerization and audio-enhancement on self-administered questionnaires, although the literature indicates mixed results. Surveys completed by adolescents in their homes or in a public space (e.g. a recreation room) using CASI or A-CASI elicited increased reporting of sensitive behaviors (i.e.
sexual behaviors, drug or alcohol use) when compared to face-to-face interviews, phone interviews and self-administered paper-pencil surveys (Le et al, 2006; Romer et al, 1997; Ellen et al, 2002; Supple, Aquilino, and Wright, 1999; Turner, Ku, et al, 1998; Turner et al, 1992; Wright, Aquiline, and Supple, 1998). Other studies have found quite different results. Beebe et al (1998) found a decrease in reporting of sensitive behaviors using A-CASI when conducting a comparative study with alternative high school students, but attributed this finding to the impact of distance between students and perceived privacy of the survey environment of students using A-CASI. Alternatively, Hallfors et al (2000) found no difference between reporting of substance use between SAQ and CASI implemented in a school setting. In a clinical setting, Webb et al (1999) found differential reporting by mode based on gender, with females reporting greater alcohol use on CASI than SAQ, and males reporting greater alcohol use on SAQ than CASI.

In spite of the benefits of the A-CASI systems, there are also significant limitations to their use, particularly in school-based research, including cost, transport, and security of a portable data collection system, venue limitations, and privacy issues (Beebe et al, 1998). To facilitate A-CASI data collection, researchers either rely on currently existing school resources, such as computer labs, or build their own portable data collection system using laptop computers. Utilization of existing school resources requires less expenditure on the part of the researcher, but also restricts the researcher to the availability of those resources, and may even eliminate prospective study sites due to lack of resources (Beebe et al, 1998), potentially increasing selection bias. Use of school resources may also be perceived as less confidential to students who may fear that their responses will be accessible to school administrators and teachers (Beebe et al, 1998).
While creating a portable data collection system with laptops, software, and programming costs requires a substantial initial investment by the researcher, it reduces selection bias of schools and standardizes equipment use across all study participants. A remaining concern with this approach is the need to ensure the availability of adequate testing sites with sufficient flat-top space and electrical outlets. The transportation and security of this system also requires substantial research staff, particularly if surveying a large number of students at one time.

Traditional A-CASI features (i.e. aural and visual presentation, programming capabilities) common to desktop and laptop computers can easily be extended to personal digital assistants (PDA), or handheld computers, whose technological qualities offer many practical advantages over both desktop and laptop systems.

1.3 Using Personal Digital Assistants for Data Collection

Technology of PDAs

Personal digital assistants (PDA) became widely available in the 1990s, although handheld electronic personal organizers were available in the 1980s. The original Palm Pilot handheld electronic organizer, introduced in 1996 and built on its own operating system platform (Palm OS), became an instant staple for the busy professional, with new programs continuing to appear on the market to enhance personal and business applications. Major computer manufacturers, including Dell, IBM, Compaq, Toshiba, Sony and Hewlett-Packard, quickly followed suit with their own versions of the handheld computer, complete with a Windows-based operating system.
PDAs fit easily into an individual’s hand and have a screen size approximately the size of one’s palm. A stylus, an instrument like an inkless pen, is used to navigate the touch-sensitive screen of the PDA. PDAs are outfitted with a range of memory capabilities, and most offer memory extension through the use of a memory card. In addition to providing multimedia support, color screens and volume control are standard features. A rechargeable battery can generally run for up to several hours before requiring charging, although a lithium back-up battery is generally included to prevent loss of data.

In considering the cost of hardware, PDAs cost substantially less per unit than a laptop or desktop computer ($300 vs $1000+ respectively), allowing researchers to buy a greater number of PDAs or more efficiently redirect their data collection expenses to other areas of the research (Gravlee, 2002). Further, PDAs are easier to transport than laptop computers, placing less physical burden on research staff. For example, 25 PDAs and the accompanying accessories (e.g. charging cords) can easily fit into a carrying case that would accommodate only two laptops computers (Trapl et al, 2005). As with A-CASI, researchers have found that the PDA is acceptable to adolescents (Bobula et al 2004, Johnson et al, 1996, Trapl et al, 2005). Additionally, PDAs allows for greater flexibility in the testing venue requirements when compared to laptops, which require flat surfaces and ample electrical supply (Gravlee 2002). PDAs are charged prior to data collection and are able to run continuously for up to six hours before needing to be recharged, thus allowing researchers to use them in a variety of settings with and without tabletop access, including classrooms, libraries, cafeterias, media centers, and auditoriums. In addition, PDAs create a natural interface for the student as they fit easily
into the palm of one’s hand, allowing students physical control over the privacy of their answers by moving the PDA or physically shifting positions. This sense of control may increase students’ perceptions of privacy, even when fellow students are sitting in close proximity.

Survey Software for PDAs

There are currently a number of survey design and data collection packages commercially available that contain PDA-based data collection modules (SurveyView Survey Software; Mercator Research Group, Snap Surveys; PocketPC PocketSurvey; Adesso Systems; Perseus MobileSurvey; Global Bay Technologies; Apian Software, Inc., SurveyHost; DigSee SURE; Raosoft EZSurvey). However, these systems are typically built on a marketing research platform with the broad assumptions that a single PDA will be distributed to a designated interviewer who then goes out into the field and interviews multiple respondents using several different surveys. While these packages have many features widely sought in other types of research, specifically marketing research, options that would be desirable for a self-administered survey are not in place, such as auditory enhancement (i.e. reading of the questions and/or responses) or disabling menu bars and equipment buttons that might otherwise distract the respondent.

Research Using PDAs

While most journal articles discuss the use of PDAs in clinical setting by a variety of health professionals for non-research purposes, some studies describe the use of PDAs by adult participant of research projects in areas such as drug and alcohol use (Bernhardt, Usdan, and Burnett, 2005; van Griensven et al, 2005), assessment of circadian rhythms (Varkevisser and Kerkhof, 2003) and collection of pain diaries (Gaertner et al, 2004;
Roeleofs, et al, 2006; Stone et al, 2002) and health-related quality of life measures (Saleh et al, 2002). Through these projects, researchers identified benefits of PDA-based data collection including utilization of electronic features such as auditory prompts to enhance compliance (Stone et al, 2002), and potentially decreased cost and increased convenience compared to paper-based tools (Saleh et al, 2002). Further, work by Bernhardt and colleagues (2001) found that most of the adult participants were comfortable using the PDA and stated a preference for this method over paper-based methods.

A body of literature has also developed around the use of PDAs by adolescents for data collection. Whalen and colleagues have had great success using PDAs to collect daily life diary data among older adolescents (Henker et al, 2002; Whalen et al, 2001; Whalen, Henker, Ishikawa et al, 2006; Whalen, Henker, Jamner et al, 2006; Whalen et al, 2004). In their studies, adolescents were provided with PDAs programmed to prompt the adolescent at certain intervals throughout the day to collect such information as adolescent social setting, moods, activities, dietary intake, and cigarette and alcohol use, with the questions typically taking only a few minutes to complete. Palermo and colleagues (2004) and Chang and colleagues (2003) have used a similar approach of the PDA-based diary with youth. Bobula and colleagues (2004) reported on their use of PDAs for in-home survey data collection with youth between the ages of nine and eleven. Based on skip patterns, youth could have answered up to 491 questions on the PDA, and among 57 respondents, the average time to complete the PDA-based survey was 57 minutes.

Only one rigorous comparison of data collection modes including PDA among youth and young adults has been published to date (van Griensven et al, 2005). Thai
students aged 15 to 21 were randomized to complete a health behavior survey assessing both sensitive (e.g. sexual behavior) and non-sensitive (e.g. dietary practices) behaviors by either PDA-based survey, audio computer-assisted self interview (A-CASI), and paper-pencil self-administered questionnaire (SAQ), or interviewer-administered survey. Findings from this study indicated that PDA was acceptable and not generally different from A-CASI and SAQ for the collection of sensitive behavioral data. Further, investigators noted the portability and ease of use in settings where other modes of data collection would be more cumbersome.

In spite of the many perceived “common-sense” benefits of this mode of data collection, few studies have examined the value of the PDA as a data collection tool in school-based research or with adolescents generally (Chang, Omery, and Mayo, 2003; Henker et al, 2002; Johnson et al, 1996; van Griensven et al, 2005; Whalen et al, 2001; Whalen, Henker, Ishikawa et al, 2006; Whalen, Henker, Jamner et al, 2006; Whalen et al, 2004), and due to the limited availability of the audio enhancement feature, there has been only one study using the APDA (Trapl et al, 2005).

2. Previous APDA Study

In 2005, Trapl and colleagues published a study describing their implementation of an audio-enhanced PDA (APDA) data collection system with 645 7th grade students in a diverse, urban school district in the Midwest. As part of their participation in a larger study, students completed a health belief and health behavior survey, with content focusing on nutrition, physical activity, and sexual behavior. The survey, which included
203 to 243 questions, depending on branching patterns, was implemented using the APDA.

The average time to complete the survey was 52.33 minutes, with no significant difference in total minutes by gender or age. There was also no difference in time to completion by sexual experience, important due to the increased number of questions asked of sexually experienced students. The authors’ experience with APDA was similar to that of Romer et al. (1997) with computers; while students generally were unfamiliar with the use of APDAs, they quickly learned how to use the technology, exhibited little fatigue during the hour-long survey, and were very respectful of the equipment (e.g. none of the twenty-five PDAs were damaged or stolen in the course of the data collection). In regards to data quality, the proportion of missing data was quite low. On a potentially sensitive question assessing sexual intercourse, the proportion of non-response using the APDA was 1.2%.

Of note, the study sample specifically included students speaking English as a second language (ESL) and special education students who might not have been able to participate in the survey had it been conducted using a paper-and-pencil survey. Nearly 13% of the sample was identified as having special individual education plans (IEPs) by their respective schools due to behavioral or cognitive deficits, yet these students were able to complete the survey in approximately the same amount of time as the regular education students (50.48 vs. 52.41 minutes respectively). Less acculturated students (i.e. those living in the US for 6 years or less and speaking mostly or only a language other than English with their families) were found to take longer to complete the survey than the rest of the sample.
To explore the association of quality of data and individual characteristics among these subgroup of students (regular education, students receiving special education classes and students new(er) the US who spoke another language at home), the authors examined the internal consistency of the items assessing four different latent constructs within the survey that vary in the level of their abstraction: (a) consequences of early sexual behavior; (b) condom-related prevention beliefs; (c) beliefs about sexual behavior in adolescents; and, (d) peer social support.

Table 2.1: Scale Reliabilities: Regular Education, English as a Second Language, and Special Education Students

<table>
<thead>
<tr>
<th>CONSTRUCT (# items)</th>
<th>Regular Educ (REG)</th>
<th>English as 2nd lang (ESL)</th>
<th>F statistic</th>
<th>p value</th>
<th>Special Educ (SE)</th>
<th>F statistic</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consequences of Early Sexual Behavior (4)</td>
<td>0.800</td>
<td>0.725</td>
<td>1.375</td>
<td>0.107</td>
<td>0.651</td>
<td>1.745</td>
<td>0.009</td>
</tr>
<tr>
<td>Condom-Related Prevention Beliefs (3)</td>
<td>0.862</td>
<td>0.756</td>
<td>1.768</td>
<td>0.014</td>
<td>0.733</td>
<td>1.935</td>
<td>0.003</td>
</tr>
<tr>
<td>Beliefs about Sexual Behavior (4)</td>
<td>0.833</td>
<td>0.696</td>
<td>1.820</td>
<td>0.010</td>
<td>0.545</td>
<td>2.725</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Peer Social Support (4)</td>
<td>0.770</td>
<td>0.722</td>
<td>1.209</td>
<td>0.230</td>
<td>0.638</td>
<td>1.573</td>
<td>0.027</td>
</tr>
</tbody>
</table>

As shown in the above table, internal consistency in all four subscales was the highest among regular education students followed by less acculturated students (ESL),
and lowest among the special education students. This suggests that even though they answered the survey questions, there was increased measurement error introduced in the data from less acculturated students and even more so from special education students, suggesting that comprehension may be limited in spite of the questions being read to the respondents. Unclear from this table is whether these differences would be observed among other modes of administration.

A second important finding from this study was high engagement of the students in the survey. The multi-sensory approach of the APDA appeared to remove the impact of external distractions on students’ abilities to remain focused. The ability to move the PDA and hide it under a desk, or to turn sideways in a chair, may have increased students’ sense of privacy of their responses. However, there were no measures collected to assess these anecdotal findings.

Although the findings of the initial study were interesting, conclusions on the improvement of administration (e.g. time to answer a certain number of questions), data quality, and student evaluation compared to SAQ could not be drawn due to the descriptive nature of the study and lack of a comparison group. Further, questions remained to be answered as to the impact of literacy skills, English mastery, and attention issues on administration (e.g. time to answer a certain number of questions), data quality, and student evaluation. The analytic and anecdotal results indicated that a comparative study was imperative.
3. Cognitive Burden: Literacy, Language and Attention

Literacy is an important consideration in the use of self-administered data collection approaches. Literacy among urban school youth is quite low (US Department of Education, 2003), and these students are often targeted for health behavior intervention evaluations and tapped for youth surveillance systems. This is both important and appropriate as research has shown that populations most at risk for risky health behaviors and poor health have much lower literacy levels than the general population (Baker et al, 1997; Davis et al, 1999; Kirsch, et al, 1993; Weiss, Hart, and Pust, 1991; Weiss, et al, 1992). Urban schools may also have a larger proportion of students who speak English as a second language as immigration patterns indicate those new to the US tend to settle in urban centers (US Department of Homeland Security, 2003). Lastly, as with regular test-taking, a lengthy survey requires sustained focus and attention, a task not easily achieved by students with attention problems (Barkley, 1996). Even for the skilled reader, a student with a limited attention span is more likely to skim over the instructions or to read the questions less thoroughly (Mosse, 1982).

Research has shown that low-literacy adults are more likely to provide logically inconsistent survey responses and have higher rates of missing data due to limited understanding of survey questions and instructions (Coyne, 1999). Al-Tayyib and colleagues (2002) found that as adult literacy decreased, rates of errors in navigation of complex surveys increased, with extremely high rates of error-prone data among the lowest literacy groups of adults.

It has long been a concern of survey researchers that literacy issues impact not only the reliability and quality of the data (Aday, 1996; Cooley et al, 2001; Fowler, 1995)
but also introduce a potential bias by excluding individuals who are unable to complete a survey due to literacy limitations. McGrew et al (1993) described the systematic exclusion from national data collection programs of youth with a wide array of disabilities, including learning disabilities, emotional disabilities, speech and language impairments, sensory disabilities, and those with multiple or severe disabilities. Investigation of nine national databases by McGrew et al (1993) indicated that while 40% to 50% of students with disabilities were excluded from the national education data collection programs (e.g. National Assessment of Education Progress), exclusion from non-educational data collection programs (e.g. Longitudinal Study of American Youth) was minimal. Points of exclusion included separate specialized schools for those with disabilities, students without a school grade designation (often occurs with special education students), and identification during the data collection process of students who appear to struggle with completing the data collection instrument independently.

One of the recommendations to come from this study was the need for modifications to the assessments, including the use of computers and audio-enhancement, to reduce selection bias by including vulnerable subgroups often under-represented in research. Over the last thirteen years since McGrew et al’s (1993) article was published, great strides have been made in the educational testing and assessment literature to develop and study technological accommodations for special needs populations (Butler and Stevens, 2001; Stock et al, 2004). This literature has developed parallel to the development of modified, self-administered questionnaire techniques, such as CASI and A-CASI.
In the survey literature, CASI itself has been shown to reduce the cognitive burden of literacy by eliminating the need for respondents to navigate through branching patterns on their own. However, it is the audio enhancement that most greatly reduces cognitive burden (Couper, Singer, and Tourangeau, 2003). Studies with adults have shown that use of the A-CASI system reduces the amount of missing data and reduces the interview time significantly among those with less education and lower reading levels (Coyne, 1999). A-CASI has been used successfully with non-English speaking populations, many of whom rely on the audio support (Hewett, 2002). Still, Hewett (2002) found that Spanish-speaking Hispanic women had difficulty navigating use of the keyboard to provide responses, and many minority group members were more likely to prefer personal interviews. As Hewett points out, this may be due to the limited exposure to computers among this population at the time of the study.

4. Confidentiality, Privacy and Environmental Factors

As discussed above, self-administered modes of data collection, including paper-and-pencil surveys as well as computerized approaches, remove the personal interaction with an interviewer and are believed to provide a more private environment for the respondent, which has been shown to be important in reducing reporting bias (Couper et al., 1998; Gribble et al, 1999; Beebe et al, 1998; Romer et al, 1997). Data collection setting and elements of the environment can also impact perceptions of privacy of the survey environment and confidentiality of responses. In school-based data collection, Gans and Brindis (1995) point out that physical changes within the environment, such as moving desks far apart or providing an adolescents with paper to cover their responses,
can add to an adolescent’s sense of privacy of their responses from other students or teachers. Gans and Brindis (1995) also advocate discouraging teachers from walking amongst students while they are completing their surveys and to prohibit talking. Use of school resources may also be perceived as less confidential to students who may fear that their responses will be accessible to school administrators and teachers (Beebe et al, 1998).

Many studies have found that adolescents have higher reports of sensitive behavior in a school setting as compared to the home environments (Gfroerer, Wright and Kopstein, 1997; Hedges and Jarvis 1998; Kann et al, 2002; Needle et al, 1983), although some have found no difference (O’Malley et al, 1983; Zanes et al, 1979). Gfroerer, Wright & Kopstein (1997) and Needle et al (1983) found increased reports of substance use, while the work of Hedges & Jarvis (1998) found increased reports of cigarette smoking. Kann and colleague’s (2002) comparison of the school-administered Youth Risk Behavior Survey (YRBS) compared to the in-home administered National Health Interview Survey (NHIS) and YRBS supplement illustrated increased reporting across a number of sensitive risky behaviors, including tobacco use, alcohol use, and sexual behavior; differential reporting of non-sensitive topics, such as physical activity and diet, were not found.

While schools have clearly been shown to be better settings for survey administration when compared to the home, there is a paucity of the literature in understanding the differential impact of mode on survey response among youth in schools. As evidenced above, the literature suggests schools may provide a better setting
in which to survey youth (Gfroerer, Wright and Kopstein, 1997; Hedges and Jarvis 1998; Kann et al, 2002; Needle et al, 1983).

In addition, the literature indicates that youth provide increased reporting of sensitive behaviors, assumed to be more accurate, using CASI or ACASI compared to SAQ (Le et al, 2006; Romer et al, 1997; Ellen et al, 2002; Turner et al, 1998b; Turner et al, 1992; Wright, Aquilino, and Supple, 1998). However, there is little evidence of mode effects in the reporting of sensitive behaviors among school-based studies and further research is needed.

Beebe et al (1998) found a decrease in reporting of sensitive behaviors using A-CASI when conducting a comparative study with alternative high school students, but attributed this finding to the impact of distance between students and perceived privacy of the survey environment of students using A-CASI. Alternatively, Hallfors et al (2000) found no difference between reporting of substance use between SAQ and CASI implemented in a school setting. These findings indicate a need to further clarify the role of mode on students’ perceptions of response confidentiality, survey environment privacy, and reporting of sensitive behaviors.

5. Demographic Confounders

While the adolescent literature is sparse, there are some studies that indicate that age, gender, race and socioeconomic status may be associated with data outcomes, such as differential reporting of sensitive behaviors by gender, race, or ethnicity (Aquilino, 1994; Schober et al., 1992; Rogers et al, 1998; Webb et al, 1999; van Griensven et al, 2005), or differential reporting of non-sensitive academic measures by socioeconomic
status (MacCann, 2006, Pomplun, Ritchie, and Custer, 2006). Further, Boone and colleagues (1997) have documented gender and race differences on completion rates and missing data among academic assessment tools, which could have implications related to survey research in adolescents. Finally, given that some outcomes of interest are developmental (e.g. sexual initiation) (Santelli et al, 2000), age was also considered a confounder.

6. Conceptual Framework

Tourangeau and others have conducted numerous studies exploring the best ways in which to obtain accurate information to sensitive questions, such as sexual behavior or other high risk behavior (Couper et al, 2003; Coyne, 1999; Tourangeau and Smith, 1996; Tourangeau and Smith, 1996; Turner et al, 1998). In their conceptual model, Tourangeau and Smith (1996) (Figure 2.1 below) proposed that both computerization and auditory presentation can affect data quality, but this relationship is mediated through the individual’s perception of privacy, legitimacy of the research and survey, and cognitive burden related to the survey experience. That is, the mediating variables are viewed as being responsive to the mode, rather than characteristics of the individual.

Figure 2.1: Tourangeau & Smith’s Model
In contrast, as shown in our conceptual model (Figure 2.2), we propose that individuals bring to the survey environment a set of abilities, limitations, preferences, and perceptions that can affect data quality as well, such as their level of sentence comprehension, listening comprehension, English mastery, perceived reading enjoyment and burden, inattentiveness, and perceptions of confidentiality and privacy of the survey environment. That is, the impact of data collection mode is likely to have differential effects on data outcomes, depending upon the set of characteristics the individual brings to the data collection setting.

Figure 2.2: Conceptual Model

The model as proposed above indicates four paths of interest. First, drawing from the literature on data collection mode, we suspected that data collection mode in a school setting has direct impact on data outcomes (Path A), including the number of questions
answered, data quality (e.g. missing data, missed skips, logically inconsistent responses, internal consistency, reporting of behaviors), and student evaluation (e.g. honesty, discomfort, and enjoyment). Second, we were interested in better understanding the relationships between respondent characteristics (e.g. gender, cognitive burden) and data outcomes (Path B), contributing to a gap in the literature. Additionally, we sought to explore the conditions under which the relationship between data collection mode and data outcomes was modified by these respondent characteristics (Path C). Lastly, as the literature suggests that perceptions of privacy and confidentiality could potentially act as either a moderator or a mediator, we investigated the direct effect of these variables as well as the potential for both moderation (Path D1) and mediation (Path D2).

In summary, this comparative and exploratory study aims to address many of the issues outlined in this chapter, with the intention to inform future studies, and that of others utilizing this technology with school-age children by comparing data derived from three different data collection modes. From this aim alone, our study will contribute to a literature that has been found to be remarkably sparse. In addition, this study seeks to increase understanding of how adolescents with different cognitive and behavioral characteristics respond to different aspects of the different data collection modes by exploring a variety of possible relationships between individual characteristics, data collection mode, and data outcomes.
Chapter 3: SPECIFIC AIMS AND HYPOTHESES

Chapter Overview

This research has four aims, in addition to specific questions related to each aim, which are described in detail below. Because no study to date has comparatively evaluated audio-enhanced PDAs for data collection, nor has any study directly evaluated the relationships of individual characteristics on data outcomes, the questions within each aim are exploratory in nature with few exceptions. In cases where the parallel literature is firm, questions are written in a hypothesis format and indicated by the “H” notation. In cases where the literature is sparse or non-existent, the exploratory question is indicated by the “Q” notation.

1. Specific Aim 1

To examine the differential effects of three different data collection modes (SAQ, PDA, APDA) on the number of questions answered, data quality, and student’s evaluation of the process (Path A).

Q1.1: Will the number of survey questions answered by students vary by mode, with incremental effects of both hardware (PDA) and audio-enhancement? That is, will students responding to APDA answer more items than those responding to questions administered by PDA, and will those responding to PDA answer more than those responding to SAQ?

H1.2: The amount of missing data will vary by mode, but only due to the impact of the PDA software (e.g. programmed navigation and alerts); thus, SAQ respondents
will have more missing data than PDA and APDA respondents, but there will be no difference in missing data between PDA and APDA respondents.

H1.3: Missed skips and logical inconsistencies will be present among SAQ respondents. (Note: this is not applicable to the PDA and APDA owing to the programmed skips within the survey application).

Q1.4: Will individuals respond to questions assessing a single construct (e.g., scale) in a similar way across modes? Specifically, will the measurement models of five different constructs (i.e. reading enjoyment, reading difficulty, consequences of early sexual behavior, sex beliefs, and parental sex beliefs) exhibit factorial invariance (i.e. indicate similar factor loadings) by mode?

Q1.5: Will reporting of sensitive behaviors (e.g. marijuana use) and non-sensitive behaviors (e.g. fast food consumption) vary by mode?

Q1.6: Will overall student honesty vary by mode; specifically, will students report higher levels of honesty on PDA and APDA when compared to SAQ?

Q1.7: Will student discomfort vary by mode?

Q1.8: Will student enjoyment vary by mode?

2. Specific Aim 2

To assess the impact of individual characteristics on number of survey questions answered, data quality, and student evaluation (Path B).

As the literature is sparse or non-existent for many of the relationships examined in this aim, this exploratory aim will ask the following questions:

Q2.1: Will the number of questions answered vary by:
1. sentence comprehension?
2. listening comprehension?
3. Language/US index?
4. self-assessed reading enjoyment?
5. self-assessed reading difficulty?
6. inattentiveness?
7. perceptions of confidentiality?
8. perceptions of survey privacy?

Q2.2: Will missing data vary by:

1. sentence comprehension?
2. listening comprehension?
3. Language/US index?
4. self-assessed reading enjoyment?
5. self-assessed reading difficulty?
6. inattentiveness?
7. perceptions of confidentiality?
8. perceptions of survey privacy?

Q2.3: Does overall honesty vary by sentence comprehension?

1. sentence comprehension?
2. listening comprehension?
3. Language/US index?
4. self-assessed reading enjoyment?
5. self-assessed reading difficulty?
Based on the questions of comprehension left unanswered in our previous work, we were specifically interested in better understanding if students’ responses to construct-specific questions (e.g. latent constructs) would be similar, regardless of literacy skills, including sentence comprehension, listening comprehension, and English mastery as measured by the Language/US index.

Q2.1e: Will the measurement models of five different constructs (i.e. reading enjoyment, reading difficulty, consequences of early sexual behavior, sex beliefs, and parental sex beliefs) exhibit factorial invariance (i.e. indicate similar factor loadings) by sentence comprehension?
Q2.2e: Will the measurement models of five different constructs (i.e. reading enjoyment, reading difficulty, consequences of early sexual behavior, sex beliefs, and parental sex beliefs) exhibit factorial invariance (i.e. indicate similar factor loadings) by *listening comprehension*?

Q2.3e: Will the measurement models of five different constructs (i.e. reading enjoyment, reading difficulty, consequences of early sexual behavior, sex beliefs, and parental sex beliefs) exhibit factorial invariance (i.e. indicate similar factor loadings) by *English mastery as measured by the Language/US index*?

### 3. Specific Aim 3

To determine whether and how demographics and cognitive burden modify the relationship between data collection mode and data outcomes (Path C).

This aim is a post-hoc analysis guided by the results of Aim 1 and Aim 2. When significant relationships were found between an outcome of interest and both mode and a respondent characteristic, that characteristic was probed for a modifying effect (Cohen, Cohen, West and Aiken, 2003). If modification was indicated by a significant interaction term, post-hoc testing was conducted to more fully understand the impact of the modifying variable (Holmbeck, 2002; Cohen, Cohen, West and Aiken, 2003). While specific hypotheses cannot be outlined, potential questions are provided below.

1. If data collection mode and reading comprehension are both significantly related to the number of questions answered, is the relationship between the number of questions answered and data collection mode modified by reading comprehension?
2. If data collection mode and self-assessed reading difficulty are both significantly related to missing data, is the relationship between data collection mode and missing data modified by self-assessed reading difficulty?

3. If data collection mode and gender are both significantly related to overall honesty, is the relationship between data collection mode and overall honesty modified by gender?

4. Specific Aim 4

To clarify the role of perceived confidentiality and perceived privacy on the number of questions answered, data quality and student evaluation as either a moderator (Path D1) or a mediator (Paths D2).

This aim is also a post-hoc analysis based on the relationships established between data collection mode and outcomes in Aim 1 and perceived confidentiality or privacy and outcomes in Aim 2. If significant relationships were established in both Aim 1 and Aim 2 for the same outcome, then cross-terms were created and entered into the model. If modification was indicated by a significant interaction term, post-hoc testing was conducted to more fully understand the impact of the modifying variable. Similarly, if significant relationships were established in Aim 1 and Aim 2 for the same outcome, steps were taken to assess mediation as outlined by Baron and Kenny (1986). While specific hypotheses cannot be provided, potential questions are provided below.
1. If data collection mode and perceived confidentiality are both significantly related to overall honesty, is the relationship between data collection mode and overall honesty modified by perceptions of confidentiality?

2. If data collection mode and perceived privacy are related to number of questions answered, and data collection mode is related to perceived privacy, are the other necessary criteria met to establish that perceived privacy mediates the relationship between data collection mode and the number of questions answered?
Chapter 4: Research Design & Methods

Chapter Overview

This chapter provides detail on the design and implementation of the study, as well as the measure and methods used to carry out the study.

1. Overview

This study involved a cross-sectional survey of 7th grade students using a stratified random sampling design. Students were first stratified based on their scores on sentence and listening comprehension assessments conducted prior to the day of randomization. Based on their scores, students were placed into one of three groups: (1) moderate/high scoring on both; (2) low scoring on only one; (3) low scoring on both. Then, students were individually randomized to complete a self-administered survey by one of three methods: paper-and-pencil (SAQ); personal digital assistant (PDA); or audio-enhanced personal digital assistant (APDA), with a goal of 96 students in each mode. Seventh grade students were recruited from seven K-8 schools. Students completed the survey in small groups by mode (e.g. all SAQ), and were placed at tables with controlled seating.

2. School Sites

Students were recruited from seven of seventy-eight Cleveland Municipal School District (CMSD) K-8 schools. The non-random selection of the schools was based on their ethnic and cultural heterogeneity (42% Hispanic, 29% Caucasian, 26% African American, and 3% Other). Moreover, a large percentage of the students spoke a language other than English at home or with friends at least some of the time. Reading
proficiencies were low (~30% passing the 6th grade proficiency exam), although similar to the rates across the school district in general. This rich environment was highly beneficial for this study as it provided diversity with regards to literacy and language barriers.

3. Student Recruitment and Randomization

The seven participating schools were neighborhood schools and ranged in size, with anywhere between one and five classes, each containing between 20 and 30 seventh grade students, determined by the size of the school building. Classes in the seven schools were comprised of students with a range of abilities, including students with cognitive, behavioral, and/or health deficits. While students with specified deficits might not remain in regular education classes for an entire day, the general policy of the district was that all students were included in core classes when appropriate as determined by classroom teachers as a means of maximizing learning potential. Additionally, bilingual education classes existed at five of the seven schools. Students speaking languages other than English who had not mastered a minimal level of the English language were provided with teachers and aides who taught in the student’s native language, predominantly Spanish. Student assignment to bilingual education was at the discretion of each school.

The study and all recruitment and consent materials were approved by the Institutional Review Board (IRB), and active parental consent and student assent were required for participation in the study. Informational packets including a letter of support from the principal, a “frequently asked questions” informational sheet, and two copies of the consent form were distributed during language arts to each seventh grade student two
months prior to data collection. Although written parental consent was the ideal, the IRB also approved an alternative approach to recruitment that allowed the research team to obtain parental consent by phone or e-mail and subsequently confirm the consent decision by a letter to each student’s home.

Students identified as seventh grade students based on school rosters were recruited in spring semester of their seventh grade, with the final data collection taking place in May and June. Students were recruited into the study through their language arts teacher. Based on power calculations (See Section 8.5 in this chapter), active parental consent was sought for a minimum of 288 students across the seven schools. To account for an average daily absentee rate of 15-20%, observed in other studies in these schools, the investigators sought consent from more students than were expected to be surveyed. Bilingual students were oversampled as were students in special education classes. Students with extreme cognitive, behavioral, and physical disabilities as identified by their special need teachers (n=2) or those with limited English proficiency (n=11) were excluded from the study.

Prior to randomization into the study, students completed the Sentence Comprehension and Listening Comprehension sections of the Group Reading and Diagnostic Evaluation (GRADE). As the investigators were concerned about having adequate numbers of students with low reading ability across each of the data collection modes, stanine scores on these two sections were used to split students into low scoring and moderate/high scoring groups. Stanines, or standard nines, assign raw scores to a value of one through nine, assuming a normal distribution with a mean of five and standard distribution of two. The bottom third of each section’s stanine (i.e. students
scoring a stanine of one, two, or three) was identified as low, and the top two-thirds (i.e. students scoring four through nine) were identified as moderate/high. Students were stratified into one of three groups based on a combination of their stanine scores on the Sentence Comprehension (SC) and Listening Comprehension (LC) sections of the GRADE: (1) moderate/high SC and moderate/high LC; (2) moderate/high SC and low LC or low SC and moderate/high LC; and (3) low SC and low LC. Following this stratification, students within each of these stratifications in each school were individually randomized to one of three data collection modes (i.e. SAQ, PDA, or APDA).

Survey administration groups were expected to range in size from 9 to 18, depending on the student enrollment and recruitment at each school, with up to two data collection sessions for each mode in larger schools. To determine whether the PDA and APDA increased one’s sense of privacy even when sitting in close proximity of others, students were seated three to a table, placing students 2-3 feet from one another. Even in small data collection sessions, students were always placed at a table with other students rather than placing them at tables by themselves to control for proximity to other students.

Prior to the survey administration, the survey administration team prepared each testing station with an assent form, a prepared survey, and a survey debriefer. In PDA or APDA administration, the PDA was prepared individually by the survey administration team in order to ensure inclusion of the study identification number. Both PDA-based and paper-based surveys were placed facedown once prepared by the survey administration team. The students were led through the assent form by the primary survey administrator, who also maintained responsibility for timekeeping.
In determining the best way to measure the impact of data collection mode on survey administration, two options were available: (1) allow all students to work through the entirety of the survey and document time to completion for each student; or (2) count the number of questions completed in a specific period of time. The latter option was chosen for several reasons. First, this approach seemed to be most in line with our efforts to be considerate in minimizing the classroom disruption caused by the study. Instead of students trickling back to class one by one as they completed the survey, the entire group of students was returned to their classroom at the same time. Also, it was known from previous data collection experiences that students occasionally neglect to inform data collection administrators that they have completed the survey because they do not want to return to class, which could allow the data collection session to continue for over an hour. Additionally, the chosen approach reduced the data collection burden on the staff and created a specific timeline for the data collection protocol, allowing several data collection sessions to be conducted in one day. Lastly, researchers are often limited to a single class period for data collection efforts, and our entire data collection protocol was designed to be completed in less than one hour, including the assent process, data collection, and a brief survey which assessed the student’s data collection experience (i.e. debriefer).

Once all students completed the assent form, they were given twenty-five minutes to work through the survey. The number of students completing the survey within the twenty-five minute period was recorded. After twenty-five minutes, the primary survey administrator advised students to stop working on the survey, and instructed the students on completion of the survey debriefer, which asked questions about the student’s data
collection experience. Each successfully enrolled student was given a t-shirt upon completion of the data collection session as a token of appreciation for participating in the study.

4. Data Collection Mode

The surveys designed for this study were identical in content and only varied in regards to instruction on the data collection method (SAQ, PDA, APDA). All attempts were made to assure that the length of mode instruction was similar across all three modes.

*Teleform*, an optical character recognition software package, was used to design and manage data from the multi-page, paper-pencil based surveys (SAQ). *Teleform* is equipped with a survey design module, *Teleform Designer*, which was used to construct the sixteen-page, paper-pencil survey. Surveys were hand-edited by trained research assistants prior to optical scanning. *Teleform Reader* and *Teleform Verifier* modules import the scanned image, digitally recognize the responses which are validated by a research team member, and export values for each item response to an SPSS file.

*Surveyor*, a survey development software package developed by Paul Stork at Case Western Reserve University for use with PDAs, was used to design and execute the PDA-based surveys. *Surveyor* creates electronic surveys with features similar to CASI software, including built-in skip patterns and out-of-range checks. Survey results were saved on the PDA as a *.txt* file as a function of the *Surveyor* PDA module and were transferred to the secure database after each data collection.
All methods for APDA were identical to the PDA-based survey, except for the audio-enhancement. *Surveyor* was designed to allow the linking of voice files (i.e. *.wav files) to each section introduction, question and response for use with the survey. A voice professional produced the voice files used with the survey in a professional recording studio. Voice files were created for questions only and were not created for response options. Students were provided with plug-in headphones (that they kept) to facilitate the audio-enhancement.

### 5.0 Data Sources

While the majority of data used in this project were derived from the SAQ and PDA-based systems, a few measures came from external sources. All data sources are described below. All data collection surveys, assessments, and forms are included in the appendix.

#### 5.1. Self-Administered Survey Content

Although the primary purpose of the study was to assess the variables in our conceptual model, the survey was presented to the students as one assessing “general health beliefs and behavior”, with a focus on sexuality, physical activity and nutrition. The survey was designed with more questions than students should have been able to complete in the allotted 25 minutes, although it was suspected that some students might finish before time was called.

In addition to the measures in the model, other measures included:

- Demographics, including age, gender, self-identified race and ethnicity, and current living arrangement (socioeconomic status proxy);
• Attitudinal and belief-based measures: reading enjoyment and reading difficulty (discussed in more detail in Section 7.3 below); sex beliefs (alpha= .56) and consequences of early sexual behavior (alpha=.66) (Mathematica Policy Research, Inc); parental sex beliefs (alpha=.72) (ETR Associates);

• Sensitive and non-sensitive behaviors including: fast food and breakfast consumption, alcohol, tobacco, and substance use; dishonest in-school behavior; stealing and property damage; pre-coital behavior; sexual behavior. All behavioral measures were drawn from standardized national population-based surveys (e.g. Youth Risk Behavior Survey (YRBS), National Longitudinal Study of Adolescents (AddHealth)).

The items outlined above comprised the first seventy items of the survey. The remainder of the survey was built to contain questions related to academic achievement not pertinent to this study other than to provide additional survey length.

5.2 Group Reading Assessment and Diagnostic Evaluation (GRADE)

Two subtests of the Group Reading Assessment and Diagnostic Evaluation (GRADE), developed by AGS Publishing, are used: Sentence Comprehension and Listening Comprehension. This brief (estimated 30 minutes for both subtests), easy to use, highly reliable, group administered test was given to the students to enable stratified randomized sampling by sentence and listening comprehension. Validity and reliability of the GRADE subtests are provided below in Section 7.1.

5.3 Teacher Assessment of Students

Language arts teachers in each of the schools were asked to complete the Strengths and Weakness of ADHD-symptoms and Normal Behavior (SWAN) Scale
teacher form and Academic Performance Assessment for each of their recruited students. Content of each of these assessment tools is described in detail below in 7.1, 7.2, and 7.4. Teachers were compensated with a $5 gift card per child for the completion of both assessment instruments.

5.4 Data Collection Observer Checklist

All survey administration staff were trained to oversee each data collection session and complete an Observer Checklist identifying the data collection method (SAQ, PDA, APDA) and characteristics of the data collection session, including quality of the survey environment (i.e. chaos, distraction/interruption, noise level) and the individual students’ affect while completing the survey (i.e. bored, restless, rejecting) rated on seven-point scales and number of individuals in the survey area (i.e. students completing survey, research/school personnel). A minimum of three survey administration staff were present for each survey administration session. These data were collected to provide a description of the data collection sessions and were not considered to be part of the conceptual model.

5.5 Debrief Questionnaire

All students were asked to complete a paper-based Debrief Questionnaire immediately following the 25-minute survey administration. The Debrief Questionnaire assessed students’ impressions of their survey and mode experience. The Debrief Questionnaire used in this study was adapted from a Debriefing Questionnaire used by Couper, Singer, and Tourangeau (2003) in their study on mode effects of A-CASI with an adult population.
6.0 Outcomes

Descriptions and sources of study outcomes are provided below.

6.1 Number of Questions Answered

The number of questions with validly answered out of the possible 178 questions during the 25-minute survey period was calculated for each student. Questions skipped as part of a skip pattern were not considered “answered” questions. However, among SAQ respondents, if a student provided a response to a question which should have been skipped, this response was considered an answered question.

6.2 Missing Data

Proportion of missing data by question and survey section (e.g., sensitive vs. non-sensitive questions; questions at the beginning vs. the end of the survey) were assessed. Additionally, missing data was examined as a dichotomous variable, grouping students as either having no missing data or having any missing data among the first seventy questions.

6.3 Failed Skip Patterns & Inconsistent Answers

Primarily an issue with the SAQ, any skip pattern where students answered questions they should not have answered was deemed a failed skip pattern. Logically inconsistent responses (e.g. student answers “never drank alcohol” but reports having been drunk) were also assessed. Proportion of failed skip patterns and inconsistent answers were calculated for each of the eight skip patterns examined.

6.4 Exposure to sensitive questions

It is known from previous work that one is able to reduce exposure to sensitive questions using the PDA/APDA (Trapl et al, 2005). The proportion of students who were
unnecessarily exposed to sensitive questions (e.g. in-depth questions about sexual behavior) on the SAQ was calculated following the same logic pattern used in programming the PDA and APDA surveys (i.e. responding “no” to ever having sex).

6.5 Reporting of Behaviors

Reporting of both sensitive and non-sensitive behaviors was calculated. Missing responses were recoded as a “no” response in order to more accurately represent reported (i.e. yes) behavior prevalence.

6.6 Internal Consistency and Factorial Invariance

Internal consistency refers to the reliability of observed variables contributing to a latent construct (e.g., do respondents answers in a similar fashion across items intended to tap a single latent construct). Construct alpha reliabilities were examined, as well as the factor structure invariance across mode among five constructs measured in the survey (reading enjoyment, reading difficulty, consequences of early sexual behavior, sex beliefs, parental sex beliefs). The methodology of factorial invariance is discussed in-depth below in section 8.1.

6.7 Student Evaluation

Students were asked to evaluate their overall honesty, enjoyment, comfort, and mode preference via the paper-based Debriefer Questionnaire. Perceived enjoyment was assessed by asking students to assess how much they enjoyed taking the survey. Students using the PDA or APDA were also asked in the Debriefer Questionnaire if they had ever used a PDA and asked to list likes and dislikes. Students were also asked which survey mode they preferred for completing the survey.
To measure self-reported honesty, students were asked to report how honest they were when answering the survey using a single item with a 4-point response scale ranging from completely honest to not at all honest (Siegel et al, 1998). Students were also asked to report how honest they were on specific sections of the survey (e.g. questions about sexual behavior; questions about physical activity) using items with a 4-point response scale ranging from completely honest to not at all honest (Siegel et al, 1998).

7.0 Independent Variables

A brief description of the independent variables is provided below. For those measures not previously validated, psychometric analyses are provided in Chapter 5, section 4.

7.1 Reading Comprehension

Sentence and Listening Comprehension were assessed using the two subtests (of the same names) of the Group Reading Assessment and Diagnostic Evaluation (GRADE) completed by each student. Subtests were scored using the GRADE scoring system. Scores were assigned to a stanine, ranging from 1 to 9, as set forth in the GRADE Technical Manual (Williams, 2003) and were used both as continuous measures and as well as dichotomous measures (i.e. low versus moderate/high stanine). As described elsewhere, the stanines were split into thirds, and the lower third of stanine scores was grouped together to represent low scores and the upper two-thirds of stanine scores (i.e. four through nine) was group together to represent moderate/high scores.
Internal reliabilities were calculated for each GRADE subtest and academic grade as part of its development (Williams, 2003). Coefficient alphas for the seventh grade sentence comprehension subtest and listening comprehension subtests were 0.86 and 0.72 respectively. To assess validity of the GRADE, the GRADE was correlated with the Gates-MacGintie Reading Tests, a set of nationally standardized, group-administered reading tests (0.87) and the Iowa Test of Basic Skills (0.83).

Overall reading comprehension was also assessed via the Teacher Academic Assessment. Teacher ratings of student’s reading and listening comprehension were assessed by four separate composite scales. Two of the scales assessed the students current academic performance in the area of reading comprehension (e.g. accurately follows written instructions) and listening comprehension (e.g. can follow simple oral instructions) following seventh grade achievement standards. The remaining two scales assessed the same academic performance in the same areas, but compared the student’s achievement to that of other students in the Cleveland Municipal School District. Each scale was comprised of 4 items that were measured on a 5-point scale from (1) far below grade level to (5) far above grade level for current assessment. The measures from the Teacher Academic Assessment had not previously been used in any other studies, and psychometric properties of the items were assessed prior to data analysis to determine the value of the data.

7.2 Language Barriers

Language Barriers were assessed by the student and by their core teacher. The self-assessment is a composite of three items, which include: the frequency with which
the student spoke a language other than English with their friends; the frequency with which the student spoke a language other than English at home; and the number of years the student had been living in the United States. The language variables were dichotomized as speaking mostly or only a language other than English (1) versus at least half English (0); number of years in the US is dichotomized as less than six (1) versus six years or more (0). These items are summed to create the Language/US index, which ranges from zero to three.

Additionally, three items in the teacher assessment asked teachers to assess the student’s oral and written proficiency with the English language compared to other students their age. As with the teacher-assessed reading and listening skills, the items had not been used previously, and psychometric properties of the items were assessed prior to data analysis to determine the value of the data.

7.3 Reading Enjoyment and Reading Difficulty

Reading enjoyment and reading difficulty were both assessed using six items adapted for adolescents from the Adult Survey of Reading Attitudes (Smith, 1990). Six items of the 12-item reading enjoyment scale and six items of the 16-item reading anxiety/difficulty scales were adapted for use with adolescents. In collaboration with Dr. M. Cecil Smith, developer of the Adult Survey of Reading Attitudes, the items were selected based on content relevance in an adolescent population. Following data collection, psychometric properties of the items were assessed and used to determine the final scale content.
7.4 Inattentiveness

Inattentiveness was assessed using a nine-item sub-scale of the Strengths and Weaknesses of ADHD-symptoms and Normal Behavior (SWAN) Scale Teacher form (Swanson, unpublished). The SWAN Scale is adapted from Swanson, Nolan and Pelham (SNAP) rating scale, an 18-item scale which assess two domains of Attention Deficit Hyperactivity Disorder (ADHD), Inattention and Hyperactivity/Impulsivity. The two nine-item subscales are based on the nine criteria outlined in the Diagnostic and Statistical Manuals (DSM) of the American Psychiatric Association and rate behavior on a scale ranging from never (0) to very often (3). In order to more fully capture a range of behaviors (e.g. strengths and weaknesses) as opposed to clinical diagnosis, Swanson and colleagues extended the four-point response categories of the SNAP and reworded the questions to capture behaviors both far above average and far below average. Psychometrics conducted on the SWAN Scale Teacher form data indicated a two-factor structure in line with the DSM-IV criteria for both inattentiveness and hyperactivity/impulsivity (Swanson et al, unpublished).

Using the SWAN subscale for inattentiveness, statements were provided and teachers rated the student’s behavior on a range from “far below average” (1) to “far above average” (7). This provided for a broader and more continuous measure of student inattentiveness as compared to the more traditional attention measures that capture only the low end of the scale (Swanson et al, unpublished). The mean of the nine items provided a single score for each student assessed ranging from one to seven.
7.5 Perceptions of Privacy and Confidentiality

Perceived confidentiality was assessed with two items asking about the student’s understanding of the confidentiality of his/her responses. Items included: how likely do you think it is that your answers will remain confidential; and how likely do you think it is that your answers will fall into the wrong hands. Perceived privacy of survey responses in the survey environment was assessed using a single question that asks the student if they believe that students sitting next to them could see their answers.

8.0 Data Analytic Plan

SAQ data was entered via the Teleform optical scanning software system; Surveyor produced data files in ASCII format. All data were compiled into an SPSS data file and analyzed using SPSS 13.0 and AMOS 7.0. Factor analysis was used to assess the factor structure of items not previously used or published in the literature and bivariate correlations were assessed to better understand the relationships among these variables.

As an overview, different analytic approaches were used, dependent upon the outcome. First, bivariate statistics established the presence of associations between independent variables and the outcomes. If the bivariate evidence was indicative of an association at a significance level of 0.05, multivariate regression models were run to account for potential confounders (e.g. age) and demographic variables found to be significantly different across mode group (see Chapter 5, Section 2, Sample Description). Second, to assess data consistency, confirmatory factor analyses (CFA) were used to test for invariance across factorial structures from different groups (Byrne, 1994).
8.1 Analytic Approach for Specific Aim 1

The first set of analyses addressed Path A of the model, the number of questions answered, data quality and student evaluations across the three modes. First, simple bivariate statistics (p<.05), including analysis of variance and chi-square tests, were performed to determine differences across the three modes. When these associations were found to be significant, multivariate linear or logistic regression models were run to account for potential confounders identified in the literature (e.g. age, gender, two parents in the home (SES proxy)) as well as demographic variables found to vary by mode despite the randomization scheme. Multicollinearity within these variables was examined and the variables were deleted if deemed problematic using two approaches. First, the tolerance and variance inflation factor (VIF) statistics were examined for small (tolerance close to zero) and large (VIF>4) values (Mason and Perrault, 1991). Additionally, significant changes of at least one standard deviation of the parameter estimates of the model after addition of the variable of interest (i.e. changes in gender after addition of sentence comprehension) were examined.

Internal consistently, measured using Cronbach’s alpha, as well as group comparisons using confirmatory factor analysis were also assessed.

8.2.1 Factorial Invariance

Measurement models of specific latent constructs were simultaneously examined across mode groups to test whether they were in fact the same (i.e., null hypothesis). Using AMOS 7.0, confirmatory factor analysis was conducted to assess the factor structure of a construct using data from all modes. Chi-square statistic, normed fit index (NFI), Tucker-Lewis index (TLI or NNFI), and the root mean square error of
approximation (RMSEA) were used to assess the fit of the overall model (Browne and Cudeck, 1993; Byrne, 1994; Hu and Bentler, 1999). After assessing the fit of the overall data, group comparisons of the measurement model were tested. First, the confirmatory factor analytic model was defined in AMOS, represented by Figure 4.1. Group membership for each group was identified within AMOS. Once the comparative model was built, the system first analyzed the unconstrained model; that is, the system determined the measurement weights for each variable of the construct by mode group. Once the best fitting model was identified by AMOS, a nested model constraining the measurement paths of the model was tested. In essence, a measurement model was determined for the first group, and these measurement weights were applied to the other groups. By running nested models where factor loadings across the groups were constrained to be equal, models were assessed for statistical difference using goodness-of-fit statistics (e.g. chi-square, TLI) (Byrne, 1994). As the chi-square test can be biased in large sample sizes larger (Marsh, 1994), the change in the fit statistics was examined. This same approach was used to assess factor invariance for the groups of interest in Aim 2.
8.3 Analytic Approach for Specific Aim 2

The second set of analyses addresses the direct effects of the independent variables (sentence comprehension, listening comprehension, Language/US Index, reading enjoyment, reading difficulty, inattentiveness, and perceived privacy and confidentiality) on the outcomes, assessed separately. First, descriptive statistics were run on each variable to assess assumptions of linearity (i.e. skewness). Similar to Specific Aim 1, bivariate analyses were conducted to determine the associations between independent variables and outcome measures. Correlations were used to assess the association between two continuous variables (e.g. sentence comprehension stanine and number of questions answered). T-tests and analysis of variance were used to assess mean score differences among categorical variables (e.g. Language/US index and overall honesty). Lastly, chi-square tests were used to assess relationships between two
categorical variables (e.g. Language/US index and missing data). Next, to describe the relationships of the independent variables and the outcome measures, the independent variables were dichotomized and outcome means or proportions were reported. Finally, if the bivariate tests indicated an association between an independent variable and the outcome measure, a multivariate regression model (linear or logistic regression) was run to account for potential confounders (e.g. age, gender, two parents in the home (SES proxy)) and any demographic variables found to vary by mode. Multicollinearity within these variables was examined and the variables were deleted if deemed problematic using two approaches. First, the tolerance and variance inflation factor (VIF) statistics were examined for small (tolerance close to zero) and large (VIF>4) values (Mason and Perrault, 1991). Additionally, significant changes of at least one standard deviation of the parameter estimates of the model after addition of the variable of interest (i.e. changes in gender after addition of sentence comprehension) were examined. In cases where inattentiveness was found to be significantly associated with an outcome, the multivariate analysis was also run through STATA to adjust for the clustering of the inattentiveness measure by teacher.

8.4 Analytic Approach for Specific Aim 3

The final set of analyses examined whether any of the significant associations established in the first aim are modified by any of the respondent characteristics found to be significant in the second aim for corresponding outcomes. These hypotheses were tested using multivariate linear and logistic regression models with the addition of the main mode-by-independent variable interaction terms in the model. The main effects were entered into the model, followed by the interaction term (i.e. mode, sentence
comprehension, mode*sentence comprehension). Significant interactions were explored more completely in post-hoc analysis to determine the direction and strength of the interaction, informing us as to how the impact of the direct effects varies by respondent characteristic.

8.5 Power Analyses

The differences in outcomes across groups had to be fairly substantive in order to meaningfully inform decisions regarding selection of data collection mode. A minimal change of 25% should be observed in outcome differences across mode. For example, based on data from over 25,000 middle school surveys, missing data from the SAQ was estimated to be 8-10% on attitudinal questions and 13-15% on sexual behavior questions. A 25% change would mean a reduction of 2-2.5% for attitudinal questions and 3.25-3.75% for behavior. When considering that the missing rate for sexual behavior in Trapl et al (2005) using the APDA was less than 2% (nearly 90% reduction), this was a reasonable level of expected change. Power calculations indicated that a final sample size of 96 per group (total 288) was adequate to detect a 25% difference, with a p<.05 and 80% power.
Chapter 5: Randomization, Sample Description, and Methodological Issues

Chapter Overview

This chapter first summarizes the recruitment and randomization efforts. Next, it provides descriptions of the demographics variables, individual characteristics (e.g. sentence comprehension), and outcome measures (e.g. number of questions answered) of the sample. A discussion of the survey environment as rated by the observer checklists is presented. Lastly, methodological issues that arose during analysis are discussed.

1. Summary of Recruitment and Randomization

Using seventh grade rosters from each of the seven schools, 521 students were identified as eligible for inclusion in the study. Of these 521 students, teachers identified 11 students with limited English proficiency and 2 students with cognitive deficiencies which would prohibit the students from participating in the study. Among the remaining 508 students, 385 (75.8%) provided “yes” consents, 48 (9.4%) provided “no” consents, and 75 (14.8%) did not respond. Of the 385 students with parental consent for participation, 275 students (71.4%), or 54.1% of the eligible student population, participated in the study.

Prior to randomization into the study, students completed the Sentence Comprehension and Listening Comprehension sections of the Group Reading and Diagnostic Evaluation (GRADE). These two sections were scored and students were divided into three groups based on their stanine scores on the Sentence Comprehension (SC) and Listening Comprehension (LC) sections of the GRADE: (1) moderate/high SC
and moderate/high LC; (2) moderate/high SC and low LC or low SC and moderate/high LC; and (3) low SC and low LC. As discussed elsewhere, stanine scores of one through three (lower third of scores, not distribution) were considered to be low, while scores of four through nine were considered moderate to high. Students in attendance on the day of data collection were individually stratified based on stanine group designation and randomized into one of three modes on the day of data collection to ensure equal mode groups. The school-by-school stratified randomization breakdown is shown below in Table 5.1.

<table>
<thead>
<tr>
<th>School</th>
<th>Hi RC and LC</th>
<th>Total</th>
<th>SAQ</th>
<th>PDA</th>
<th>APDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>School 1</td>
<td></td>
<td>21</td>
<td>5</td>
<td>7</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Low RC or LC</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Low RC and LC</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>School 2</td>
<td></td>
<td>24</td>
<td>8</td>
<td>9</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Low RC or LC</td>
<td>21</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Low RC and LC</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>School 3</td>
<td></td>
<td>21</td>
<td>7</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Low RC or LC</td>
<td>12</td>
<td>3</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Low RC and LC</td>
<td>7</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>School 4</td>
<td></td>
<td>16</td>
<td>6</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Low RC or LC</td>
<td>15</td>
<td>6</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Low RC and LC</td>
<td>8</td>
<td>3</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>School 5</td>
<td></td>
<td>24</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Low RC or LC</td>
<td>7</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Low RC and LC</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>School 6</td>
<td></td>
<td>41</td>
<td>14</td>
<td>14</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Low RC or LC</td>
<td>12</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Low RC and LC</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>School 7</td>
<td></td>
<td>13</td>
<td>5</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Low RC or LC</td>
<td>7</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Low RC and LC</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>TOTAL</td>
<td>Hi RC and LC</td>
<td>160</td>
<td>52</td>
<td>53</td>
<td>55</td>
</tr>
<tr>
<td></td>
<td>Low RC or LC</td>
<td>82</td>
<td>28</td>
<td>27</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Low RC and LC</td>
<td>33</td>
<td>10</td>
<td>13</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>TOTAL</td>
<td>275</td>
<td>90</td>
<td>93</td>
<td>92</td>
</tr>
</tbody>
</table>
While all attempts were made to strictly follow the randomization protocol, there were some problematic recruitment issues that led to a breakdown in the randomization process in two schools. Due to low consent response at two schools, some students were recruited the day of data collection and had not yet completed the GRADE assessment necessary for stratification. In order to utilize the cases, we employed an additional randomization scheme by entering students without the GRADE into a fourth “unknown GRADE” column of randomization; that is, the first student was randomized to SAQ, the second to PDA, the third to APDA, and so on. Following the students’ participation in data collection, GRADE data were collected. This deviation further impacted the distribution of students to the different data collection modes in other schools as we attempted to maintain equal numbers across mode assignment within the stratification. Due to a lack of systematic deviation from the stratified randomization, it is not believed that this deviation as described should have had any impact on the overall success of the randomization in producing comparable mode groups.

2. Sample Description

This section provides an overview of the study sample, including demographics, individual characteristics, and outcome measures.

2.1 Summary of Demographic Variables

As shown in Table 5.2, the sample was 48.4% female with a mean age of 13.15 years (sd=0.75); these characteristics were similar across the three modes. Students were similar across the three modes in frequency of spoken English with family, living in the
US for more than six years, and living with two parents. However, in spite of the individual randomization, there was a significantly different racial/ethnic distribution across the three modes, with Caucasians overrepresented in the SAQ mode, African-Americans over-represented in the PDA mode, and Hispanics over-represented in the APDA mode.

Table 5.2: Sample Demographics

<table>
<thead>
<tr>
<th>DEMOGRAPHICS</th>
<th>TOTAL n=275</th>
<th>SAQ n=90</th>
<th>PDA n=93</th>
<th>APDA n=92</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (% Female)</td>
<td>48.4</td>
<td>47.8</td>
<td>47.3</td>
<td>50.0</td>
<td>NS</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.021</td>
</tr>
<tr>
<td>% Caucasian</td>
<td>23.0</td>
<td>34.8</td>
<td>17.2</td>
<td>17.4</td>
<td></td>
</tr>
<tr>
<td>% African-American</td>
<td>29.6</td>
<td>25.8</td>
<td>38.7</td>
<td>23.9</td>
<td></td>
</tr>
<tr>
<td>% Hispanic</td>
<td>37.6</td>
<td>30.3</td>
<td>37.6</td>
<td>44.6</td>
<td></td>
</tr>
<tr>
<td>% Other</td>
<td>10.8</td>
<td>9.0</td>
<td>6.5</td>
<td>14.2</td>
<td></td>
</tr>
<tr>
<td>Age in Years: Mean(SD)</td>
<td>13.15(.75)</td>
<td>13.20(.78)</td>
<td>13.14(.76)</td>
<td>13.10 (.71)</td>
<td>NS</td>
</tr>
<tr>
<td>Speak Mostly/Only English with Family</td>
<td>60.8</td>
<td>62.9</td>
<td>64.1</td>
<td>55.4</td>
<td>NS</td>
</tr>
<tr>
<td>Lived in US &gt;6 years</td>
<td>89.1</td>
<td>88.9</td>
<td>93.5</td>
<td>84.8</td>
<td>NS</td>
</tr>
<tr>
<td>Live with 2 parents</td>
<td>49.8</td>
<td>43.3</td>
<td>52.7</td>
<td>53.3</td>
<td>NS</td>
</tr>
</tbody>
</table>

2.2 Summary of Individual Characteristics

Table 5.3 below details the individual characteristics of the sample. Sentence comprehension ranged from a low of one to a high of nine, with a mean score for the sample of 3.97 and a standard deviation of 1.38. This deviated from the intended mean of 5 and standard deviation of two implied by a stanine score. Listening comprehension also had a range of one to nine, and students scored higher here, with a mean score of 5.19, standard deviation of 1.98, in line with the intended stanine mean and standard deviation.
Table 5.3: Individual Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>% or Mean (SD)</th>
<th>Skewness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence Comprehension (1=low, 9=high)</td>
<td>3.97 (1.38)</td>
<td>.171</td>
</tr>
<tr>
<td>Listening Comprehension (1=low, 9=high)</td>
<td>5.19 (1.98)</td>
<td>.001</td>
</tr>
<tr>
<td>Language/US Index (%) *</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No indicators</td>
<td>76.9%</td>
<td></td>
</tr>
<tr>
<td>One indicator</td>
<td>14.7%</td>
<td></td>
</tr>
<tr>
<td>Two indicators</td>
<td>7.0%</td>
<td></td>
</tr>
<tr>
<td>Three indicators</td>
<td>1.5%</td>
<td></td>
</tr>
<tr>
<td>Perceived Reading Enjoyment (1=low, 5=high)</td>
<td>2.80 (0.87)</td>
<td>.241</td>
</tr>
<tr>
<td>Perceived Reading Difficulty (1=low, 5=high)</td>
<td>2.19 (0.88)</td>
<td>.693</td>
</tr>
<tr>
<td>Inattentiveness (1=low, 7=high)</td>
<td>4.15 (1.23)</td>
<td>.090</td>
</tr>
<tr>
<td>Perceived Confidentiality (1=low, 5=high)</td>
<td>3.99 (0.98)</td>
<td>-.744</td>
</tr>
<tr>
<td>Perceived Privacy of Responses (1=low, 4=high)</td>
<td>2.85 (1.06)</td>
<td>-.368</td>
</tr>
</tbody>
</table>

* Indicators include <6 years living in US and mostly/only speaking English with family or friends

As a measure of English mastery, the Language/US Index was created based on students’ responses to three questions: the number of years living in the US; frequency of speaking a language other than English with family; and the frequency of speaking a language other than English with friends. Number of years was dichotomized as <6 years (1) or more than 6 years living in the US. Frequency of speaking a language other than English for both family and friends was dichotomized as mostly/only another language (1) or half or more of English (0). The indicators were summed to create an index ranging from zero to three, with the score indicating the number of indicators each student had. As shown in Table 5.3, almost 77% of students indicated living in the US for at least six years and speaking English at least half of the time with their family and friends. However, almost one quarter of the sample reported one or more indicators. Only 1.5% of the sample indicated living in the US for less than six years and speaking
mostly or only another language with both family and friends. Based on skewness statistics of less than one, none of the measures significantly deviated from normality.

Following psychometric assessment (discussed in more detail below in section 4), perceived reading enjoyment was a composite of six items (Cronbach’s alpha of 0.886) and perceived reading difficulty was a composite of five items (Cronbach’s alpha of 0.830). Both scale composite scores ranged from one to five, with higher scores indicating higher reading enjoyment and higher reading difficulty. Students had a mean reading enjoyment score of 2.80 (sd=0.87) and a mean reading difficulty score of 2.19, indicating relatively low perceptions of reading difficulty. Inattentiveness was based a composite of nine items from the SWAN Teacher Assessment (Cronbach’s alpha of 0.979), where students were rated far above grade level (1) to far below grade level (7). As indicated in Table 5.3, students scored a mean of 4.15, very close to the “at grade level” designation. Perceived confidentiality was also a composite based on student responses to three items and ranging from one to five, with one indicating low perceptions of confidentiality and five indicating high perceptions of confidentiality. Overall, students rated perceptions of confidentiality fairly high, with a mean score for the sample of 3.99. Student perceptions of the privacy of their responses was also quite high, with a mean score on the single item of 2.85 on a scale of one to four, with four representing the response that other students definitely could not see their survey responses.

2.3 Summary of Outcome Measures

Table 5.4 below summarized outcome variables for the sample. All students were allotted 25 minutes in which to work through the 178-question survey. To have answered
all 178 questions, students in the PDA or APDA groups would have had to “opt-in” to all skip patterns; however, due to the nature of SAQ, it may have been possible to ignore skip patterns and answer all 178 questions. Overall, students completed an average of 127.6 questions (se=33.2), median of 130 questions, with completion ranging from a low of 28 questions to a high of 178 questions. Anecdotally, 33 students finished prior to the end of the allotted 25 minutes, and were distributed by mode as follows: 5 SAQ; 6 PDA; 22 APDA (p=.001).

In order to differentiate true missing data from data missing due to a respondent not getting far enough in the survey, missing data was examined dichotomously as having any missing responses among the first 70 questions in the survey compared to having none. As indicated in Table 5.4, 17.8% of all students had at least one missing response within the first 70 questions of the survey.

<table>
<thead>
<tr>
<th>Table 5.4: Sample Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome</td>
</tr>
<tr>
<td>----------------------------</td>
</tr>
<tr>
<td>Number of Questions Answered (max=178)</td>
</tr>
<tr>
<td>Students with Any Missing Data (%)</td>
</tr>
<tr>
<td>Students with Any Failed Skips*</td>
</tr>
<tr>
<td>Students with Any Inconsistencies*</td>
</tr>
<tr>
<td>Unnecessary Exposure to Sensitive Questions*</td>
</tr>
<tr>
<td>Reporting of Behaviors (%)</td>
</tr>
<tr>
<td>Ate Fastfood</td>
</tr>
<tr>
<td>Ever drank alcohol</td>
</tr>
<tr>
<td>Ever used marijuana</td>
</tr>
<tr>
<td>Ever cheated in school</td>
</tr>
<tr>
<td>Ever shopped</td>
</tr>
<tr>
<td>Ever had sexual intercourse</td>
</tr>
<tr>
<td>Overall Honesty (1=low, 5=high)</td>
</tr>
<tr>
<td>Overall Discomfort (1=low, 5=high)</td>
</tr>
<tr>
<td>Overall Enjoyment (1=low, 5=high)</td>
</tr>
</tbody>
</table>

* Among SAQ respondents only
Failed skip patterns and logically inconsistent responses were examined among SAQ respondents only due to the programmed skips of the PDA and APDA. Among SAQ respondents, 35.6% failed to appropriately complete a skip pattern, and 12.2% provided logically inconsistent responses. Further, 62.2% of the SAQ respondents were unnecessarily exposed to additional questions related to sexual intercourse.

Reporting of behaviors among the sample was also assessed for behaviors considered sensitive (e.g. sexual intercourse) and non-sensitive (e.g. eating fast food). Missing responses were recoded to represent lack of prevalence (i.e. “no” response) in order to more accurately represent the reported prevalence. Seventy-eight percent of all students reported eating fast food at least once in the previous week. While 61.4% of students admitted to ever using alcohol, only 16.4% reported ever trying marijuana, which was the least prevalent of the behaviors examined. Over 52% of students reported cheating in school, 28.3% of students reported shoplifting, and 22.6% reported ever having sex.

Student self-reported overall honesty was quite high with a mean of 4.51 (standard deviation of 0.89) on a scale of “not at all honest” (1) to “completely honest” (5). Students’ overall discomfort was fairly low with a mean of 2.10 (SD=1.12) on a scale ranging from “very comfortable” (1) to “not at all comfortable” (5). Overall enjoyment appeared to be quite high given the mean score of 3.96 (SD=1.12) on the scale ranging from “didn’t like it at all” (1) to “liked it a lot” (5). Interestingly, among PDA and APDA respondents, 57.1% of students reported never having used a PDA prior to the data collection session.
3. Survey Environment

Twenty-three unique data collection sessions were held across the seven schools. Test environment was rated by up to eight trained data collection observers. Observers scored qualities of the test environment, such as whether the environment was orderly or chaotic, had no distractions or many distractions, or were quiet or loud. Observers also scored student affect during the data collection session, indicating if the data collection group was enthusiastic or bored, attentive or restless, or accepting or rejecting. Data describing the survey environment across the 23 sessions is provided below in Table 5.5.

<table>
<thead>
<tr>
<th>Table 5.5: Survey Environment Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Characteristic</td>
</tr>
<tr>
<td>Number of Students</td>
</tr>
<tr>
<td>General Survey Environment (1=poor, 7=ideal)</td>
</tr>
<tr>
<td>Distraction in the Survey Environment (1=minimal distraction; 7=many distractions)</td>
</tr>
<tr>
<td>Student Engagement (1=not engaged; 7=engaged)</td>
</tr>
</tbody>
</table>

As shown in Table 5.5, the mean number of students for each data collection session was 13.4. Session size ranged from 3 to 19, however, only three of the twenty-three sessions had six or fewer students. The general survey environment, representing the general nature of the room as far as lighting, temperature, and resources, was rated fairly high on a scale ranging from “unacceptable” (1) to “ideal” (7), with a mean across observers of 5.22. Distraction of the survey environment was rated fairly low by observers on a one to seven scale, with a mean score of 2.42. This was likely due to the fact that all data collections sessions took place in each school’s media center, which is typically not an area with many distractions. Student engagement was also rated highly by observers, with a mean of 5.73 on a scale of one to seven.
4. Methodological Issues

During the course of the analyses, several measurement issues were discovered and changes were made. These processes are discussed in detail below.

4.1. Teacher Academic Assessment

All items from the reading and listening comprehension questions as well as English mastery questions comparing students to others in Cleveland Municipal School District from the Teacher Academic Assessment were factor analyzed using principal axis factor extraction with a varimax rotation in SPSS 13.0. This analysis yielded only 1 factor among the eleven items. Next, the same items were factor analyzed using the same techniques but forcing three factors. Using this approach, the data approximately reproduced the assumed factor structure, with the exception of one item. Additionally, all items had very large secondary factor loadings, reinforcing the single factor structure of the initial factor analysis.

Anecdotally, as research assistants had reviewed each teacher assessment, it became apparent that teachers seemed to choose the same response category for all assessment statement, seemingly choosing a “rating” for a student, ranging from far below to far above grade level, and circling that rating regardless of the skill being assessed. This anecdotal assessment was further reinforced by the Cronbach’s alphas which were greater than .95 for each of the three conceptualized scales. Further, as shown in Table 5.6, the correlations among the three conceptualized scales were quite high.
Table 5.6: Correlations Among Teacher Measures

<table>
<thead>
<tr>
<th></th>
<th>Reading</th>
<th>Listening</th>
<th>English Mastery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>1.000</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Listening</td>
<td>0.939</td>
<td>1.000</td>
<td>---</td>
</tr>
<tr>
<td>English Mastery</td>
<td>0.887</td>
<td>0.862</td>
<td>1.000</td>
</tr>
</tbody>
</table>

In considering measures to assess reading and listening comprehension, the teacher assessment was just one of the two measures proposed and collected. Given the stability and validation of the GRADE sub-tests and the lack of variation in the teacher assessment data, it was decided that the teacher assessments of reading and listening would not be used in the analyses.

4.2 Reading Enjoyment and Reading Difficulty

In order to assess the success of the adaptation of the items from the Adult Survey of Reading Attitudes for use in an adolescent population, the twelve items were factor analyzed using principal axis factor extraction with a varimax rotation in SPSS 13.0. Upon initial assessment of the communalities, one item stood out as being very low in comparison to the other items. Based on eigenvalues, three factors were extracted. However, using the “elbow” test, the scree plot indicated only two factors. Thus, factor analysis was re-run with the single item removed. Upon removal of this item, the eleven-item structure yielded two unique factors. Reading enjoyment was constructed as a six-item composite ranging from one to five, with an internal reliability of 0.88, and reading difficulty was constructed as a five-item composite ranging from one to five, with an internal reliability of 0.83. Bivariate correlation analysis indicated that these two scales were not significantly correlated.
Chapter 6: Results for Specific Aim 1

Specific Aim 1: To examine the differential effects of three different data collection modes (SAQ, PDA, APDA) on survey administration, data quality and student evaluation of the process.

Chapter Overview

This chapter describes the relationship between outcomes and mode. Results within this chapter are discussed by outcome, with bivariate and descriptive results preceding multivariate results.

1. Impact of Mode on Number of Questions Answered

Q1.1: Will the number of survey questions answered by students vary by mode, with incremental effects of both hardware (PDA) and audio-enhancement? That is, will students responding to APDA answer more items than those responding to questions administered by PDA, and will those responding to PDA answer more than those responding to SAQ?

As mentioned previously, all students were allotted 25 minutes in which to work through the 178-question survey. To have answered all 178 questions, students in the PDA or APDA groups would have had to “opt-in” to all skip patterns; however, due to the nature of SAQ, it may have been possible to ignore skip patterns and answer all 178 questions. Overall, students completed an average of 127.6 questions (se=33.2), median of 130 questions, with completion ranging from a low of 28 questions to a high of 178 questions. The number of questions answered varied significantly by mode, with
students completing the survey on the APDA completing significantly more questions (148.7) than both SAQ respondents (116.2) and PDA respondents (117.8). Anecdotally, 33 students finished prior to the end of the allotted 25 minutes, and were distributed by mode as follows: 5 SAQ; 6 PDA; 22 APDA (p=.001). Most students finished between one and five minutes early, but one student was able to complete the survey eight minutes prior to the end of the survey period, completing the survey in just 17 minutes on the PDA.

<table>
<thead>
<tr>
<th></th>
<th>Unadjusted</th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode</td>
<td>n</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>SAQ</td>
<td>90</td>
<td>116.2 (33.2)</td>
</tr>
<tr>
<td>PDA</td>
<td>93</td>
<td>117.8 (32.8)</td>
</tr>
<tr>
<td>APDA</td>
<td>92</td>
<td>149.7 (21.8)</td>
</tr>
</tbody>
</table>

The analysis of the number of questions answered was then adjusted for race of participants, which was found to be significantly different by mode despite randomization. Results remained similar to the unadjusted results, with APDA respondents answering significantly more questions than both SAQ (p<.001) and PDA respondents (p<.001).

2. Impact of Mode on Data Quality: Missing Data

H1.2: The amount of missing data will vary by mode, but only due to the impact of the PDA software (e.g. programmed navigation and alerts); thus, SAQ respondents will have more missing data than PDA and APDA respondents, but there will be no difference in missing data between PDA and APDA respondents.
In this study, missing data was measured in two distinct ways. First, missing data were examined among individual questions to better understand the characteristics of questions left unanswered. Second, missing data were examined by student. That is, students were broken into two groups: those with no missing data among the first seventy questions, and those with at least one missing response among the first seventy questions.

In considering the first approach, missing data was examined among variables chosen at random among the first four sections of the survey in order to provide a variety of question type (i.e. sensitive versus non-sensitive) and placement (i.e. beginning versus middle), and all behavior questions were examined to determine the extent of missing data. In order to differentiate true missing data separate from data missing due to a respondent not getting far enough in the survey, missing data was examined for respondents who answered at least 70 questions in the survey. This yielded a subset of 261 respondents. Overall, prevalence of missing data was quite low. Among the demographic questions examined, only self-reported grades yielded any missing data. Interestingly, this variable had the highest rate of missing data, and all missing cases were among those completing the survey via SAQ. The reading self-assessment questions included a variety of questions assessing reading enjoyment and difficulty and could be considered non-sensitive. Within the variables examined in this section, there was only one missing response. Student beliefs were composed of questions regarding students’ beliefs around sexual behavior and condom use and could be considered sensitive in nature. Among the four sensitive variables examined, there were five total missing responses in the entire sample, with a low of one missing response to a high of three missing responses among the variables examined.
Table 6.2: Frequency of Missing Data by Mode

<table>
<thead>
<tr>
<th>Variables with possible missing values</th>
<th>Total</th>
<th>SAQ</th>
<th>PDA</th>
<th>APDA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>261</td>
<td>81</td>
<td>88</td>
<td>92</td>
</tr>
</tbody>
</table>

**Demographics**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total</th>
<th>SAQ</th>
<th>PDA</th>
<th>APDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gender</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Self-report Grades in School</td>
<td>7</td>
<td>7</td>
<td>0</td>
<td>0*</td>
</tr>
</tbody>
</table>

**Reading**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total</th>
<th>SAQ</th>
<th>PDA</th>
<th>APDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>I love to read.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>I quickly forget what I have read even if I have just read it.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Reading is one of my favorite activities.</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Beliefs around sexual behavior**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total</th>
<th>SAQ</th>
<th>PDA</th>
<th>APDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>I believe people my age should wait until they are older to have sex</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Having sexual intercourse as a teenager makes it harder for someone to develop a loving relationship in the future.</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>When used properly, a condom prevents girls from getting pregnant.</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sexual intercourse is better--more enjoyable, intense and satisfying--when two people are married to each other.</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

**Sensitive (S) and Non-sensitive (NS) Behaviors**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Total</th>
<th>SAQ</th>
<th>PDA</th>
<th>APDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>In the past two weeks, how many times did you eat at a fast food restaurant? (NS)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>In the past seven days, how many times did you eat breakfast? (NS)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Please describe your experience with smoking. (NS)</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Have you ever had a drink of alcohol, even just a sip? (NS)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Have you ever smoked marijuana? (S)</td>
<td>4</td>
<td>3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Have you ever cheated on a test in school? (NS)</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Have you ever copied from someone else's homework? (NS)</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Have you ever shoplifted/stolen something from a store? (S)</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Have you ever beaten up another person? (S)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Have you ever kissed someone on the lips, other than a parent or relative? (S/NS)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Have you ever French kissed? (S/NS)</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Have you ever touched a girl's breasts, or if you are a girl, has anyone ever touched your breasts? (S)</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Have you ever touched someone else's private parts below the waist under their clothes? (S)</td>
<td>3</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Has someone touched your private parts below the waist under your clothes? (S)</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Have you ever had sexual intercourse? (S)</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>0*</td>
</tr>
</tbody>
</table>

* Among Respondents who answered at least 70 questions.

* Indicates p<.05 across mode.
Behavior questions included a variety of non-sensitive (e.g. eating fast food) and sensitive (e.g. sexual intercourse) questions. As one of the goals of this study was to examine differences in reporting of non-sensitive and sensitive behaviors, all behavior variables, excluding those within skip patterns, were examined for rate of missing data. As shown in Table 6.2, missing responses across these variables ranged from zero to four. Questions regarding breakfast and fast food consumption yielded no missing data, while questions regarding marijuana use, shoplifting, and sexual experimentation yielded more, yet still minimal, missing data. While there appears to be a trend indicating more missing data among SAQ respondents, missing data was only significantly higher among SAQ respondents for self-reported grades (p<.001) and sexual intercourse (p<.05). However, the rates of missing were still quite low, with 7.8% of SAQ respondents missing self-reported grades (compared to 0% among PDA and APDA respondents) and 3.3% of SAQ respondents missing sexual intercourse (compared to 0% among PDA and APDA respondents).

To further understand the characteristics of students with missing data, the number of missing responses among the first seventy questions was summed for each student and then dichotomized to create a variable indicating (0) no missing data versus (1) any missing data. There was a significant difference in the prevalence of missing data by mode, with the APDA group (5.4%) and the PDA group (15.1%) having significantly fewer students with missing data than the SAQ group (33.3%) (p<.001 for both comparisons). While there appears to be a difference between the PDA and APDA groups, this difference is not statistically significant.
### Table 6.3: Missing Data by Mode: Percentage of Students with any Missing Data

<table>
<thead>
<tr>
<th>Mode</th>
<th>n</th>
<th>Unadjusted</th>
<th>Adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAQ</td>
<td>90</td>
<td>33.3%</td>
<td>33.6%</td>
</tr>
<tr>
<td>PDA</td>
<td>93</td>
<td>15.1%***</td>
<td>14.5%***</td>
</tr>
<tr>
<td>APDA</td>
<td>92</td>
<td>5.4%***</td>
<td>5.7%***</td>
</tr>
</tbody>
</table>

*** Indicates significant difference when compared to SAQ; p<.001

3. Impact of Mode on Data Quality: Inconsistent or Missed Skips

*H1.3: Missed skips and logical inconsistencies will be present among SAQ respondents.*

Missed skip patterns and logically inconsistent responses were examined among only SAQ respondents owing to the programmed skips of the PDA and APDA. A skip pattern is missed when a student responds to a question which should have been skipped based on the response to a feeder question. In Figure 6.1 below, the missed skip is represented by the shaded circle. A skipped response is considered logically inconsistent if the response to the skipped question is in direct conflict with the stem question, as shown in the shaded square in Figure 6.1 below. For example, a student responding no to ever drinking alcohol is instructed in the response to skip ahead in the survey (e.g. No→ Go to #48). However, after choosing this response, the student still answers the subsequent question assessing “ever been drunk”, which should have been skipped. This is considered a missed skip. If the student answers “no” to this question, it is consistent; however, answering “yes” to “ever been drunk” is in direct conflict with the initial response of “no” to ever drinking alcohol.
Table 6.4 first provides the number of students whose response to the stem question would have indicated a need to skip. Indented questions below the stem question indicate first the number of students who missed the skip, followed by the number of students who provided inconsistent responses. While significant numbers of students missed skip patterns (7.1% to 32.8%), many fewer provided inconsistent responses (0% to 8.9%). The greatest number of inconsistencies was reported among questions regarding sexual experience. It is important to note that age of initiation, number of partners, and use of a condom at last sex did not provide response options that would allow the student to be logically consistent (e.g. “I’ve never had sex”).
Table 6.4: Missed Skips and Inconsistent Responses Among SAQ

<table>
<thead>
<tr>
<th>Questions</th>
<th>Missed Skip</th>
<th>Inconsistencies</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Never Had A Drink of Alcohol</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ever got drunk</td>
<td>6</td>
<td>18.8%</td>
</tr>
<tr>
<td>Had a drink of alcohol in past month</td>
<td>7</td>
<td>21.9%</td>
</tr>
<tr>
<td>Never smoked marijuana</td>
<td>69</td>
<td>24.6%</td>
</tr>
<tr>
<td>Smoked marijuana in past month</td>
<td>17</td>
<td>29.6%</td>
</tr>
<tr>
<td>Never kissed on the lips</td>
<td>27</td>
<td>29.6%</td>
</tr>
<tr>
<td>Kissed on the lips in past 3 months</td>
<td>8</td>
<td>29.6%</td>
</tr>
<tr>
<td>Never French kissed</td>
<td>43</td>
<td>18.6%</td>
</tr>
<tr>
<td>French kissed in past 3 months</td>
<td>8</td>
<td>29.6%</td>
</tr>
<tr>
<td>Never touch breasts/breasts touched</td>
<td>49</td>
<td>30.6%</td>
</tr>
<tr>
<td>Touched breasts/breasts touched in past 3 months</td>
<td>15</td>
<td>30.6%</td>
</tr>
<tr>
<td>Never touched other's private parts</td>
<td>64</td>
<td>32.8%</td>
</tr>
<tr>
<td>Touched private parts in past 3 months</td>
<td>21</td>
<td>32.8%</td>
</tr>
<tr>
<td>Never had private parts touched</td>
<td>60</td>
<td>32.8%</td>
</tr>
<tr>
<td>Private parts touched in past 3 months</td>
<td>17</td>
<td>28.3%</td>
</tr>
<tr>
<td>Never had sex</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>Valid age of initiation</td>
<td>4</td>
<td>7.1%</td>
</tr>
<tr>
<td>Valid number of partners</td>
<td>4</td>
<td>7.1%</td>
</tr>
<tr>
<td>Valid gender of sexual partners</td>
<td>5</td>
<td>8.9%</td>
</tr>
<tr>
<td>Valid response to condom at last sex</td>
<td>3</td>
<td>5.4%</td>
</tr>
<tr>
<td>Sex in the past 3 months</td>
<td>8</td>
<td>14.3%</td>
</tr>
</tbody>
</table>

4. Impact of Mode on Data Quality: Construct Reliability and Factorial Invariance

Internal consistency was examined for the first five scales present in the survey using Cronbach’s alpha. These scales included Reading Enjoyment (5 items), Reading Difficulty (6 items), Sex Beliefs (5 items), Consequences of Early Sexual Involvement (4 items), and Parental Sex Beliefs (4 items). Overall, Cronbach’s alpha for these five scales was quite high, ranging from 0.784 (Parental Sex Beliefs) to 0.886 (Reading Enjoyment). Among four of the five scales, APDA respondents yielded the highest
internal consistency, with parental sex beliefs as the exception. There was no consistent pattern evident between SAQ and PDA respondents.

<table>
<thead>
<tr>
<th>Table 6.5: Construct Reliabilities by Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Construct</strong></td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>Group</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td>SAQ</td>
</tr>
<tr>
<td>PDA</td>
</tr>
<tr>
<td>APDA</td>
</tr>
</tbody>
</table>

4.1 Results of Factorial Invariance Testing

Q1.4: Will the measurement models of five different constructs (i.e. reading enjoyment, reading difficulty, consequences of early sexual behavior, sex beliefs, and parental sex beliefs) exhibit factorial invariance (i.e. indicate similar factor loadings) by mode?

The constructs examined above were also examined for factorial invariance. That is, the measurement models of each of the constructs were simultaneously examined across mode groups to test whether the measurement weights were the same. Results are presented below by construct.

Testing of all models followed the same approach. One by one, measurement models for each construct were created in AMOS 7.0 and overall fit was assessed based on chi-square, NFI, TLI, and RMSEA. Next, comparison groups were defined using the “manage groups” interface and multi-group analysis was run in AMOS 7.0. As described earlier, AMOS first fits an unconstrained model, allowing the measurement weights for
all variable paths in the model to be determined for each identified group. Next, the measurement weights for one group are applied to the other identified group(s), effectively constraining the paths in these other groups to be equal to the paths in the first group. The fit statistics from the constrained measurement weights model are compared to the unconstrained model, allowing for nested model comparisons. The model comparisons are presented below for each of the five latent constructs.

4.1.1 Reading Enjoyment

First, the overall fit of the model was assessed by defining the latent construct as composed of six unique, observed variables, each with its own unobserved error variance, as shown in Figure 6.2. As shown in Table 6.6, the chi-square statistic was significant, indicating that the model was not a good fit for the data. However, this statistic is very sensitive to sample size, and other measures of fit are generally a better indicator of model fit in the case of a large sample size. Examining other fit statistics, the overall fit of the model was good as indicated by the NFI and TLI close to one and the RMSEA less than .08 (Browne and Cudeck, 1993, Hu and Bentler, 1999).
Next, the mode groups were defined in AMOS, and models were tested to assess the fit for the unconstrained group analysis (Model 2) and the constrained measurement weights group analysis (Model 3). Fit statistics, provided in Table 6.6, indicate that both Model 2 and Model 3 fit the data well. A chi-square difference test between Model 2 and Model 3 indicated that there was no significant difference between the fit for the unconstrained model and the constrained model, implying that the measurement weights across mode groups were not significantly different for reading enjoyment.

<table>
<thead>
<tr>
<th>Table 6.6: Model Fit for Reading Enjoyment: Overall and By Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Chi-square</strong></td>
</tr>
<tr>
<td>Model 1: Overall Model Fit</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
</tr>
</tbody>
</table>
4.1.2 Reading Difficulty

As was done with reading enjoyment, the overall fit of the model was assessed by defining the latent construct as composed of five unique, observed variables, each with its own unobserved error variance, as shown in Figure 6.3. As shown in Table 6.7, the chi-square statistic is significant, indicating a poor fitting model, although this is likely driven by the large sample size. Still, the overall fit of the model was not very good as indicated by the NFI of less than 0.90, TLI of .51, and the large RMSEA value.

Next, the mode groups were defined in AMOS, and models were tested to assess the fit for the unconstrained group analysis (Model 2) and the constrained measurement weights group analysis (Model 3). Fit statistics, provided in Table 6.7, indicate that both Model 2 and Model 3 fit the data well. A chi-square difference test between Model 2 and Model 3 indicated that there was no significant difference between the fit for the...
unconstrained model and the constrained model, implying that the measurement weights across mode groups were not significantly different for reading difficulty.

| Table 6.7: Model Fit for Reading Difficulty: Overall and By Mode |
|-----------------|-----------------|-----------------|
|                 | Chi-            |                 |
|                 | square          | df              | p       | NFI   | TLI   | RMSEA |
| Model 1: Overall Model Fit | 97.933          | 5               | 0.000   | 0.831 | 0.506 | 0.26  |
| Model 2: Unconstrained Group Comparison | 104.468         | 15              | 0.000   | 0.831 | 0.53  | 0.148 |
| Model 3: Constrained Measurement Weights | 118.794         | 23              | 0.000   | 0.807 | 0.672 | 0.124 |
| Model 2 vs. Model 3 | 14.326          | 8               | 0.074   | -0.024| 0.142 |

4.1.3 Sex Beliefs

Following the steps outlined above, the overall fit of the measurement model for the latent construct “sex beliefs” was assessed by defining the latent construct as composed of five unique, observed variables, each with its own unobserved error variance, as shown in Figure 6.4. As shown in Table 6.8, the overall fit of the model was good as indicated by the non-significant chi-square statistics, NFI and TLI close to one, and the RMSEA less than .08.
Next, the mode groups were defined in AMOS, and models were tested to assess the fit for the unconstrained group analysis (Model 2) and the constrained measurement weights group analysis (Model 3). Fit statistics, provided in Table 6.8, indicate that both Model 2 and Model 3 fit the data well. A chi-square difference test between Model 2 and Model 3 indicated that there was a significant difference between the fit for the unconstrained model and the constrained model, implying that the measurement weights across mode groups were significantly different for sex beliefs.
Table 6.8: Model Fit for Sex Beliefs: Overall and By Mode

<table>
<thead>
<tr>
<th></th>
<th>Chi-square</th>
<th>df</th>
<th>p</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>6.606</td>
<td>5</td>
<td>.252</td>
<td>.986</td>
<td>.990</td>
<td>.034</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>30.995</td>
<td>15</td>
<td>.009</td>
<td>.943</td>
<td>.904</td>
<td>.063</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td>51.888</td>
<td>23</td>
<td>.001</td>
<td>.904</td>
<td>.887</td>
<td>.068</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
<td>20.893</td>
<td>8</td>
<td>.007</td>
<td>-0.039</td>
<td>-0.017</td>
<td></td>
</tr>
</tbody>
</table>

4.1.4 Consequences of Early Sexual Behavior

Following the steps outlined above, the overall fit of the measurement model for the latent construct consequences of early sexual behavior was assessed by defining the latent construct as composed of four unique, observed variables, each with its own unobserved error variance, as shown in Figure 6.5. As shown in Table 6.9, the overall fit of the model was good as indicated by the NFI, but not ideal based on the TLI and the RMSEA. Additionally, the chi-square statistic is significant, indicating a poor fitting model, likely due to sample size.
Next, the mode groups were defined in AMOS, and models were tested to assess the fit for the unconstrained group analysis (Model 2) and the constrained measurement weights group analysis (Model 3). Fit statistics, provided in Table 6.9, indicate that both Model 2 and Model 3 fit the data well. A chi-square difference test between Model 2 and Model 3 indicated that there was a significant difference between the fit for the unconstrained model and the constrained model, implying that the measurement weights across mode groups were significantly different for this latent construct.
Table 6.9: Model Fit for Consequences of Early Sexual Behavior: Overall and By Mode

<table>
<thead>
<tr>
<th></th>
<th>Chi-square</th>
<th>df</th>
<th>p</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>19.346</td>
<td>2</td>
<td>0.000</td>
<td>0.943</td>
<td>0.736</td>
<td>0.178</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>29.046</td>
<td>6</td>
<td>0.000</td>
<td>0.925</td>
<td>0.679</td>
<td>0.119</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td>54.709</td>
<td>12</td>
<td>0.000</td>
<td>0.859</td>
<td>0.703</td>
<td>0.114</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
<td>25.663</td>
<td>6</td>
<td>0.000</td>
<td>-0.066</td>
<td>0.024</td>
<td></td>
</tr>
</tbody>
</table>

4.1.5 Parental Sex Beliefs

Following the steps outlined above, the overall fit of the measurement model for the latent construct consequences of early sexual behavior was assessed by defining the latent construct as composed of four unique, observed variables, each with its own unobserved error variance, as shown in Figure 6.6. As shown in Table 6.10, the overall fit of the model was good as indicated by the NFI, the TLI and the RMSEA. Further, the chi-square statistic is not significant, indicating a good fitting model.

Next, the mode groups were defined in AMOS, and models were tested to assess the fit for the unconstrained group analysis (Model 2) and the constrained measurement weights group analysis (Model 3). Fit statistics, provided in Table 6.10, indicate that both Model 2 and Model 3 fit the data reasonably well. A chi-square difference test between Model 2 and Model 3 indicated that there was a significant difference between the fit for the unconstrained model and the constrained model, implying that the measurement weights across mode groups were significantly different for this latent construct.
Figure 6.6: Measurement Model for Parental Sex Beliefs

![Diagram of Parental Sex Views]

Table 6.10: Model Fit for Parental Sex Beliefs: Overall and By Mode

<table>
<thead>
<tr>
<th>Model</th>
<th>Chi-square</th>
<th>df</th>
<th>p</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>2.988</td>
<td>2</td>
<td>.224</td>
<td>0.988</td>
<td>0.980</td>
<td>0.042</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group</td>
<td>18.599</td>
<td>6</td>
<td>0.005</td>
<td>0.947</td>
<td>0.803</td>
<td>0.088</td>
</tr>
<tr>
<td>Comparison</td>
<td>38.692</td>
<td>12</td>
<td>0.003</td>
<td>0.889</td>
<td>0.791</td>
<td>0.090</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
<td>20.093</td>
<td>6</td>
<td>0.003</td>
<td>-0.058</td>
<td>-0.012</td>
<td></td>
</tr>
</tbody>
</table>
5. Impact of Mode on Data Quality: Reporting of Behavior

Q1.5: Will reporting of sensitive behaviors (e.g. marijuana use) and non-sensitive behaviors (e.g. fast food consumption) vary by mode?

All behavior stem questions were examined for overall reporting and reporting by mode; questions within skip patterns were not examined as estimates for these questions may be inflated or deflated depending on students “opting-in” on the stem question.

In order to more accurately assess differences in reporting of behaviors, general linear modeling analyses were used to adjust for mode differences of racial distribution. All missing responses were coded as a “no” response to more accurately represent true reporting of behavior. Variables in the adjusted model include: gender, age, race/ethnicity, two parents in the home.

<table>
<thead>
<tr>
<th>% Reported</th>
<th>TOTAL</th>
<th>SAQ</th>
<th>PDA</th>
<th>APDA</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ate Fast food</td>
<td>77.4</td>
<td>71.1</td>
<td>84.7</td>
<td>76.3</td>
<td>0.094</td>
</tr>
<tr>
<td>Ever drank alcohol</td>
<td>60.7</td>
<td>62.9</td>
<td>63.7</td>
<td>55.6</td>
<td>0.480</td>
</tr>
<tr>
<td>Ever used marijuana</td>
<td>16.0</td>
<td>17.4</td>
<td>15.7</td>
<td>15.0</td>
<td>0.903</td>
</tr>
<tr>
<td>Ever cheated in school</td>
<td>51.2</td>
<td>46.8</td>
<td>52.7</td>
<td>54.2</td>
<td>0.591</td>
</tr>
<tr>
<td>Ever shoplifted</td>
<td>27.6</td>
<td>29.2</td>
<td>27.2</td>
<td>26.5</td>
<td>0.915</td>
</tr>
<tr>
<td>Ever had sexual intercourse</td>
<td>21.9</td>
<td>30.6</td>
<td>19.1</td>
<td>16.0</td>
<td>0.046</td>
</tr>
</tbody>
</table>

* Adjusted for race

Overall, prevalence of behaviors was reported similarly and without any consistent pattern across modes with the exception of sexual experience. SAQ respondents were much more likely to report ever engaging in sexual activity compared to PDA and APDA respondents.
6. Impact of Mode on Student Evaluation of Survey Experience

6.1 Overall Honesty, Discomfort, and Enjoyment

Q1.6: Will overall student honesty vary by mode; specifically, will students report higher levels of honesty on PDA and APDA when compared to SAQ?

Q1.7: Will student discomfort vary by mode?

Q1.8: Will student enjoyment vary by mode?

In the Debrief Questionnaire, students were asked to rate their overall honesty, discomfort, and enjoyment while completing the survey. Table 6.12 below provides students’ mean responses by mode adjusted for race. Overall honesty, rated on a scale from very honest (one) to not at all honest (five), was rated highly by students in all three modes, with the lowest reported honesty among PDA respondents (4.43) and highest reported honesty among APDA respondents (4.58). These scores were not significantly different.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Overall Honesty Mean (SD)</th>
<th>Overall Discomfort Mean (SD)</th>
<th>Enjoyment Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTAL</td>
<td>275</td>
<td>4.51 (.89)</td>
<td>2.10 (1.12)</td>
<td>3.96 (1.12)</td>
</tr>
<tr>
<td>SAQ</td>
<td>90</td>
<td>4.52 (0.91)</td>
<td>2.08 (1.07)</td>
<td>3.80 (1.06)</td>
</tr>
<tr>
<td>PDA</td>
<td>93</td>
<td>4.43 (1.00)</td>
<td>2.03 (1.10)</td>
<td>4.13 (1.09)</td>
</tr>
<tr>
<td>APDA</td>
<td>92</td>
<td>4.58 (0.75)</td>
<td>2.20 (1.21)</td>
<td>3.93 (1.21)</td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>0.517</td>
<td>0.590</td>
<td>0.146</td>
</tr>
</tbody>
</table>

* Adjusted for race

Overall discomfort, ranging from very comfortable (one) to not at all comfortable (five), was rated as low by students, with the lowest discomfort among PDA respondents (2.03) and the highest reported discomfort among the APDA respondents (2.20). These scores were not significantly different. Overall enjoyment, rated from one to five, was
rated highly, with the lowest enjoyment score among the SAQ respondents (3.80) and the highest score among the PDA respondents (4.13). Again, these scores were not significantly different.

### 6.2 Mode Preference

Students were also asked to indicate their preferred mode to complete the survey among SAQ (By a written questionnaire that I could fill out myself), PDA (By PDA—handheld computer), APDA (By PDA—handheld computer—with the questions read to me through earphones), in-person with an interview, on the phone with an interviewer, or no preference/doesn’t matter. The table below provides a summary of the students’ responses overall and by variables of interest. Overall, students’ top choice of mode was APDA (29.2%), closely followed by PDA (27.7%), which together accounted for almost 57% of students’ preference. Interestingly, the next most popular choices were “no preference” and interviewer-administered survey (16.2%). SAQ (7.0%) and phone interview (3.0%) rounded out the remainder of the responses.

<table>
<thead>
<tr>
<th>Preference (%)/Mode</th>
<th>SAQ</th>
<th>PDA</th>
<th>APDA</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAQ</td>
<td>13.6</td>
<td>3.3</td>
<td>4.4</td>
</tr>
<tr>
<td>PDA</td>
<td>14.8</td>
<td>46.7</td>
<td>20.9</td>
</tr>
<tr>
<td>APDA</td>
<td>15.9</td>
<td>23.9</td>
<td>47.3</td>
</tr>
<tr>
<td>Interview</td>
<td>28.4</td>
<td>7.6</td>
<td>13.2</td>
</tr>
<tr>
<td>Phone Interview</td>
<td>1.1</td>
<td>3.3</td>
<td>4.4</td>
</tr>
<tr>
<td>No Preference</td>
<td>26.1</td>
<td>15.2</td>
<td>9.9</td>
</tr>
</tbody>
</table>

When examining student preference by mode, it seems that a majority of students in the PDA and APDA groups chose the mode they had just used, 46.7% and 47.3%
respectively. This was followed by a large group within each mode choosing the remaining PDA mode, with 23.9% of PDA-completers choosing APDA, and 20.9% of APDA-completers choosing PDA. The trend was quite different for students completing the survey on SAQ. These students were approximately evenly distributed across SAQ (13.6%), PDA (14.8%), and APDA (15.9%). The most popular choice was interviewer-administered survey (28.4%), possibly due to the burden of reading.
Chapter 7: Results for Specific Aim 2

Impact of Individual Characteristics on Data Outcomes

Specific Aim 2: To assess the impact of individual characteristics on administration, data quality and student evaluation.

Chapter Overview

The next set of analyses focus on the relationship between the eight identified individual characteristics (e.g. sentence comprehension, listening comprehension, English mastery, reading enjoyment, reading difficulty, inattentiveness, perceived confidentiality and privacy) with outcomes. Similar to the previous chapter, this chapter is organized by outcomes, which include: number of questions answered, missing data, and student evaluation, including overall honesty, discomfort, and enjoyment. Lastly, the factorial invariance of five constructs will be examined for sentence comprehension, listening comprehension, and the Language/US index.

A brief description of the measures and the analyses are repeated here to refresh the reader’s memory; however, details are provided in Chapter 4.

1. Description of the Individual Characteristics in the Analyses

1.1 Sentence Comprehension

Sentence comprehension was measured by the GRADE sentence comprehension subtest stanine, which standardizes students’ raw scores to a value ranging from one to nine. Correlation analysis and t-tests were run to first establish bivariate associations
between this variable and the outcomes. To descriptively compare lower sentence comprehension stanines to moderate and higher sentence comprehension stanines, the measure was dichotomized, with a stanine of less than four representing lower sentence comprehension and a stanine of four or more representing moderate/higher sentence comprehension. If bivariate associations were indicated, sentence comprehension stanine was entered into the multivariate linear or logistic model as a continuous variable.

1.2 Listening Comprehension

Listening comprehension was measured by the GRADE listening comprehension subtest stanine, which standardizes students’ raw scores to a value ranging from one to nine. Correlation analysis and t-tests were run to first establish bivariate associations between this variable and the outcomes. To descriptively compare lower listening comprehension stanines to moderate and higher listening comprehension stanines, the measure was dichotomized, with a stanine of less than four representing lower sentence comprehension and a stanine of four or more representing moderate/higher sentence comprehension. If bivariate associations were indicated, listening comprehension stanine was entered into the multivariate linear or logistic model as a continuous variable.

1.3 Language Mastery

English language mastery is measured by the Language/US index which is constructed based on students’ responses to three questions: (1) how long have you lived in the US; (2) what language do you usually speak with your friends; (3) what language do you usually speak with your family. To create the index, the language variables were dichotomized as speaking mostly or only a language other than English (1) versus at least half English (0); number of years in the US is dichotomized as less than six (1) versus six
years or more (0). These items are summed to create the Language/US index, which ranges from zero to three. Few students have all three indicators, so students with two or three indicators were collapsed into one group.

Analysis of variance and chi-square tests were used to assess the bivariate relationship between this variable and outcomes. In comparing low Language/US index scores to high Language/US index scores, low is represented by those with no indicators and high is represented by those with at least one of the three indicators. If bivariate associations were indicated, Language/US index was entered into the multivariate linear or logistic model as a categorical variable.

**1.4 Reading Enjoyment**

Reading enjoyment is measured by a six-item scale with responses ranging from strongly disagree (one) to strongly agree (five). Correlation analysis and t-tests were run to first establish bivariate associations between this variable and the outcomes. To descriptively compare students with lower reading enjoyment to students with higher reading enjoyment, the measure was dichotomized with low represented by those scoring less than three and high represented by those with a score of at least three. If bivariate associations were indicated, reading enjoyment was entered into the multivariate linear or logistic model as a continuous variable.

**1.5 Reading Difficulty**

Reading difficulty is measured by a five-item scale with responses ranging from strongly disagree (one) to strongly agree (five). Correlation analysis and t-tests were run to first establish bivariate associations between this variable and the outcomes. To descriptively compare students with lower reading difficulty to students with higher
reading difficulty, the measure was dichotomized with low represented by those scoring less than three and high represented by those with a score of at least three. If bivariate associations were indicated, reading difficulty was entered into the multivariate linear or logistic model as a continuous variable.

1.6 Inattentiveness

Inattentiveness was measured by a nine-item scale ranging from one to seven. Correlation analysis and t-tests were run to first establish bivariate associations between this variable and the outcomes. To descriptively compare students with low inattentiveness scores to students with high inattentiveness scores, inattentiveness was dichotomized with low representing scores less than four (i.e. average) and high representing scores of four and higher. If bivariate associations were indicated, inattentiveness was entered into the multivariate linear or logistic model as a continuous variable.

1.7 Perceived Confidentiality

Perceived confidentiality is measured by taking the mean of two items ranging from low (one) to high (five). Correlation analysis and t-tests were run to first establish bivariate associations between this variable and the outcomes. To descriptively compare students with low perceived confidentiality to students with high perceived confidentiality, scores were dichotomized with low representing scores of three or less and high representing those scoring higher than three. If bivariate associations were indicated, perceived confidentiality was entered into the multivariate linear or logistic model as a continuous variable.
1.8 Perceived Privacy

Perceived privacy was measured by a single item which assessed the extent to which a student believed that others could not see their survey responses, and ranged from not at all (1) to definitely could see their responses (4). Correlation analysis and t-tests were run to first establish bivariate associations between this variable and the outcomes. To descriptively compare students with low perceived privacy of responses to students with high perceived privacy of responses, scores were dichotomized with low representing the responses “definitely could” and “probably could” see responses and high representing those reporting “definitely could not” and “probably could not” see responses. If bivariate associations were indicated, perceived privacy was entered into the multivariate linear or logistic model as a continuous variable.

1.9 Consideration of Confounders in Multivariate Models

When bivariate associations were indicated, the individuals characteristics were then entered into a multivariate model that included age, gender and number of parents in the household (as a proxy for socioeconomic status) as potential confounders identified in the literature (see Chapter 2) and race due to the differences in racial composition by mode. In assessing the multivariate models, we assumed the tradition of retaining all variables in the model, regardless of their statistical significance (i.e. p value), based on the analytic approach that all variables considered to contribute to the model should be maintained in the model regardless of the level of statistical significance.
2. Impact of Individual Characteristics on Number of Questions Answered

Q2.1: Will the number of questions answered vary by:

1. sentence comprehension?
2. listening comprehension?
3. Language/US index?
4. self-assessed reading enjoyment?
5. self-assessed reading difficulty?
6. inattentiveness?
7. perceptions of confidentiality?
8. perceptions of survey privacy?

The relationship between the eight individual characteristics described above and the number of questions answered during the twenty-five minute survey administration session was examined. As shown in the table below, sentence comprehension, listening comprehension, inattentiveness, and privacy were found to have a significant, positive correlation with the number of questions answered. Reading difficulty had a significant, negative correlation with the number of questions answered. However, no relationship existed between the number of questions answered and language/US index, reading enjoyment, and perceived confidentiality.
Descriptively, an overall trend indicates that students with lower levels of cognitive burden were able to answer more survey questions in the allotted time period. Students with lower sentence comprehension completed significantly fewer questions (118.1) than students with higher sentence comprehension (133.0). Similarly, students
with lower listening comprehension answered significantly fewer questions (111.1) than students with higher listening comprehension (131.1). Lower scores on perceived reading difficulty were related to answering more questions, 129.9 questions compared to 118.7 questions. Students with lower inattentiveness scores (i.e. high attentiveness) completed significantly more questions (134.0) in the allotted time when compared to students with high inattentiveness scores (121.7).

2.1 Multivariate Results for Number of Questions Answered

To further assess the relationship of the sentence comprehension, listening comprehension, reading difficulty, inattentiveness, and perceived privacy on the number of questions answered, variables were entered individually into a multivariate linear regression model that adjusted for gender, age, race, and living arrangement (two parents vs. other). The results of this adjusted model prior to the inclusion of the variables of interest are shown in Table 7.2. As indicated by the results, age, gender, race, and having two parents in the home (SES proxy) were not significantly associated with the number of questions answered on the survey.

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.143</td>
<td>2.820</td>
<td>0.960</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>5.971</td>
<td>4.070</td>
<td>0.144</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-3.149</td>
<td>5.674</td>
<td>0.579</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-5.488</td>
<td>5.358</td>
<td>0.307</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-4.683</td>
<td>7.640</td>
<td>0.540</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>5.453</td>
<td>4.143</td>
<td>0.189</td>
</tr>
</tbody>
</table>
2.1.1 Sentence Comprehension

Entered into the linear regression model, sentence comprehension stanine was significantly related to the number of questions answered (see Table 7.3). As indicated by the bivariate correlation, higher sentence comprehension was positively related to answering more questions on the survey. For every one stanine increase (raw sentence comprehension score increase varies for one stanine increase), the average number of questions answered increased by 7.56.

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>2.100</td>
<td>2.725</td>
<td>0.442</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>5.143</td>
<td>3.888</td>
<td>0.187</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-0.328</td>
<td>5.443</td>
<td>0.952</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-0.545</td>
<td>5.201</td>
<td>0.917</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-0.035</td>
<td>7.346</td>
<td>0.996</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>3.187</td>
<td>3.978</td>
<td>0.424</td>
</tr>
<tr>
<td>Sentence Comprehension</td>
<td>7.560</td>
<td>1.450</td>
<td>0.000</td>
</tr>
</tbody>
</table>

2.1.2 Listening Comprehension

Entered into the linear regression model, listening comprehension stanine was significantly related to the number of questions answered (see Table 7.4). As indicated by the bivariate correlation, higher listening comprehension was positively related to answering more questions on the survey. For every one stanine increase (raw listening comprehension score increase varies for one stanine increase), the average number of questions answered increased by 4.02.
Table 7.4: Regression Parameters for Questions Answered and Listening Comprehension

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.662</td>
<td>2.754</td>
<td>.810</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>6.557</td>
<td>3.968</td>
<td>.100</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-2.631</td>
<td>5.529</td>
<td>.635</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-1.631</td>
<td>5.312</td>
<td>.759</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-2.676</td>
<td>7.461</td>
<td>.720</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>3.305</td>
<td>4.073</td>
<td>.418</td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>4.023</td>
<td>1.026</td>
<td>.000</td>
</tr>
</tbody>
</table>

2.1.3 Perceived Reading Difficulty

A linear regression model was run to examine the impact of reading difficulty on number of questions answered. As indicated in Table 7.5 below, reading difficulty was significantly related to the number of questions answered (p=.006). That is, for each one point increase on the reading difficulty scale, students answered approximately 6.4 fewer questions on the survey.

Table 7.5: Regression Parameters for Questions Answered and Reading Difficulty

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.931</td>
<td>2.812</td>
<td>0.741</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>6.969</td>
<td>4.037</td>
<td>0.085</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-4.265</td>
<td>5.620</td>
<td>0.449</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-4.885</td>
<td>5.297</td>
<td>0.357</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-5.069</td>
<td>7.549</td>
<td>0.502</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>4.477</td>
<td>4.108</td>
<td>0.277</td>
</tr>
<tr>
<td>Reading difficulty</td>
<td>-6.409</td>
<td>2.318</td>
<td>0.006</td>
</tr>
</tbody>
</table>
2.1.4 Inattentiveness

A linear regression model was run to assess the relationship between inattentiveness and the number of questions answered. As indicated in the table below, inattentiveness was significantly and negatively related to the number of questions answered (p<.001). For each one point increase on the inattentiveness scale, students answered approximately 8 fewer questions on the survey.

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.683</td>
<td>2.716</td>
<td>0.802</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>1.007</td>
<td>4.032</td>
<td>0.803</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-2.301</td>
<td>5.461</td>
<td>0.674</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-5.974</td>
<td>5.144</td>
<td>0.246</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-8.111</td>
<td>7.368</td>
<td>0.272</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>3.406</td>
<td>4.000</td>
<td>0.395</td>
</tr>
<tr>
<td>Inattentiveness</td>
<td>-8.019</td>
<td>1.643</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Due to the clustering of the inattentiveness measure by teacher (ICC=.133), the multivariate regression analysis was also run through STATA, yielding adjusted robust standard errors and corresponding p-value. Analysis in STATA indicated no impact of the clustering on the relationship between inattentiveness and the number of questions answered.

2.1.5 Perceived Privacy

A linear regression model was to assess the relationship between perceived privacy of responses and the number of questions answered (see Table 7.7). Despite the significant bivariate relationship between privacy and the number of questions answered,
this relationship did not remain after the addition of age, gender, race, and two parents in the home in the linear regression model.

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.848</td>
<td>2.867</td>
<td>.768</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>6.558</td>
<td>4.079</td>
<td>.109</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-3.620</td>
<td>5.641</td>
<td>.522</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-7.161</td>
<td>5.379</td>
<td>.184</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-7.331</td>
<td>7.707</td>
<td>.342</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>5.463</td>
<td>4.161</td>
<td>.190</td>
</tr>
<tr>
<td>Perceived privacy</td>
<td>3.802</td>
<td>1.944</td>
<td>.052</td>
</tr>
</tbody>
</table>

3. Impact of Individual Characteristics on Missing Data

Q2.2. Will missing data vary by:

1. sentence comprehension?
2. listening comprehension?
3. Language/US index?
4. self-assessed reading enjoyment?
5. self-assessed reading difficulty?
6. inattentiveness?
7. perceptions of confidentiality?
8. perceptions of survey privacy?

In this analysis, students were broken into two groups: those with no missing data among the first seventy questions, and those with at least one missing response among
the first seventy questions. The number of missing responses among the first seventy questions was summed for each student and then dichotomized to create a variable indicating (0) no missing data versus (1) any missing data. The summary table below provides the bivariate results.

<table>
<thead>
<tr>
<th>Table 7.8: Missing Data by Characteristic</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Proportion Miss Data*</td>
<td>t-test or Chi-square (df)</td>
</tr>
<tr>
<td>Low Sentence Comp (≤ 3)</td>
<td>100</td>
<td>14.9%</td>
<td>.866</td>
</tr>
<tr>
<td>Hi Sentence Comp (&gt;3)</td>
<td>175</td>
<td>23.0%</td>
<td>.398</td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>0.106</td>
<td></td>
</tr>
<tr>
<td>Low Listening Comp (≤ 3)</td>
<td></td>
<td>15.4%</td>
<td>1.782</td>
</tr>
<tr>
<td>High Listening Comp (&gt;3)</td>
<td></td>
<td>29.2%</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>.056</td>
<td></td>
</tr>
<tr>
<td>Low Language/US Score (No indicators)</td>
<td>63</td>
<td>17.6%</td>
<td>6.508 (3)</td>
</tr>
<tr>
<td>High Language/US Score (1+ indicators)</td>
<td>210</td>
<td>15.9%</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>0.749</td>
<td></td>
</tr>
<tr>
<td>Low Reading Enjoyment (&lt;3)</td>
<td>153</td>
<td>20.3%</td>
<td>-.108</td>
</tr>
<tr>
<td>High Reading Enjoyment (≥3)</td>
<td>121</td>
<td>14.1%</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>0.173</td>
<td></td>
</tr>
<tr>
<td>Low Reading Difficulty (≥3)</td>
<td>219</td>
<td>16.4%</td>
<td>-.466</td>
</tr>
<tr>
<td>High Reading Difficulty (≥3)</td>
<td>56</td>
<td>23.2%</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>0.279</td>
<td></td>
</tr>
<tr>
<td>Low Inattentiveness (≤4)</td>
<td>135</td>
<td>12.6%</td>
<td>3.143</td>
</tr>
<tr>
<td>High Inattentiveness (≥4)</td>
<td>139</td>
<td>22.3%</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>0.034</td>
<td>.002</td>
</tr>
<tr>
<td>Low perceived Confidentiality (≤3)</td>
<td>66</td>
<td>13.6%</td>
<td>-.922</td>
</tr>
<tr>
<td>High perceived Confidentiality (≥3)</td>
<td>207</td>
<td>19.3%</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>0.264</td>
<td></td>
</tr>
<tr>
<td>Low perceived Privacy (&lt;3)</td>
<td>104</td>
<td>21.2%</td>
<td>2.342</td>
</tr>
<tr>
<td>High perceived Privacy (≥3)</td>
<td>166</td>
<td>16.3%</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>.324</td>
<td>.020</td>
</tr>
</tbody>
</table>

* Proportion of students with at least one missing response among the first 70 Qs

T-tests were conducted among the continuous variables to determine the association with students’ missing data. Chi-square test was run to assess the relationship
between Language/US index and missing data. Inattentiveness (p=.002) and perceived privacy (p=.02) were the only variables found to be significantly related to a student having any missing data. Descriptively, students in the highly inattentive group were more likely to have missing data (22.30%) when compared to less inattentive (i.e. more attentive) students (12.59%); students with lower perceptions of privacy were more likely to have missing data (21.2%) when compared to students with higher perceptions of privacy (16.3%).

3.1 Multivariate Models for Missing Data

To examine missing data in a multivariate approach, the number of missing values among the first 70 questions was calculated for all students completing at least 70 questions in the survey. Due to the non-linear nature of the resulting variable, the variable was dichotomized into (0) no missing data and (1) any missing data. As inattentiveness and perceived privacy were the only characteristic found to have a bivariate relationship with missing data, these variables were entered into a multivariate logistic regression model that adjusted for gender, age, race, and living arrangement (two parents vs. other). This adjusted model is shown in Table 7.9 below. As indicated in the table, none of the demographic variables were significantly associated with missing data.

<table>
<thead>
<tr>
<th>Table 7.9: Regression Parameters for Missing Data and Demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Exp(B)</strong></td>
</tr>
<tr>
<td><strong>Gender: Female (1) vs. Male (0)</strong></td>
</tr>
<tr>
<td><strong>Age</strong></td>
</tr>
<tr>
<td><strong>Race: Black (1) vs. White (0)</strong></td>
</tr>
<tr>
<td><strong>Race: Hispanic (1) vs. White (0)</strong></td>
</tr>
<tr>
<td><strong>Race: Other (1) vs. White (0)</strong></td>
</tr>
<tr>
<td><strong>Two parents in home (1) vs. Other (0)</strong></td>
</tr>
</tbody>
</table>
3.1.1 Inattentiveness

A logistic regression model adjusting for gender, age, race and two parents in the home was run to examine the impact of inattentiveness on students’ missing data. As indicated in the table below, inattentiveness was significantly related to missing data (p=.002). A single point increase on the one to seven inattentiveness scale increased the odds of missing data by just over one and a half times.

<table>
<thead>
<tr>
<th>Table 7.10: Regression Parameters for Missing Data and Inattentiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
</tr>
<tr>
<td>Inattentiveness</td>
</tr>
</tbody>
</table>

Due to the clustering of the inattentiveness measure by teacher (ICC=.133), the multivariate logistic regression analysis was also run through STATA, yielding adjusted robust standard errors and corresponding p-value. Analysis in STATA indicated no impact of the clustering on the relationship between inattentiveness and missing data.

3.1.2 Perceived Privacy

A logistic regression model adjusting for gender, age, race and two parents in the home was run to examine the impact of perceived privacy on students’ missing data. As indicated in the table below, inattentiveness was significantly related to missing data (p=.044). A single point increase in perceived privacy (ranging from one to four) reduced the odds of missing data by just over 25%.
Table 7.11: Regression Parameters for Missing Data and Perceived Privacy

<table>
<thead>
<tr>
<th></th>
<th>Exp(B)</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Lower</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>1.232</td>
<td>.648</td>
</tr>
<tr>
<td>Age</td>
<td>.919</td>
<td>.587</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>1.131</td>
<td>.486</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>0.723</td>
<td>.304</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>0.962</td>
<td>.292</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>0.625</td>
<td>.322</td>
</tr>
<tr>
<td>Perceived Privacy</td>
<td>.733</td>
<td>.542</td>
</tr>
</tbody>
</table>

4. Impact of Individual Characteristics on Student Evaluation

Q2.3: Does overall honesty vary by:

1. sentence comprehension?
2. listening comprehension?
3. Language/US index?
4. self-assessed reading enjoyment?
5. self-assessed reading difficulty?
6. inattentiveness?
7. perceptions of confidentiality?
8. perceptions of survey privacy?

Q2.4: Will overall discomfort and enjoyment vary:

1. sentence comprehension?
2. listening comprehension?
3. Language/US index?
4. self-assessed reading enjoyment?
5. self-assessed reading difficulty?
6. inattentiveness?
Students completed a Debriefing Survey which provided feedback on students’ perceptions of their survey experience. Students reported on honesty, comfort, and enjoyment related to completing the survey. Students were also asked to indicate a mode preference if they had been given a choice.

As shown in Table 7.12, self-reported honesty was found to vary significantly by the following variables: sentence comprehension; language/US index score, perceived reading difficulty, inattentiveness, and perceived confidentiality. Students with less cognitive burden (i.e. higher sentence comprehension stanine, higher language/US index score, lower reading difficulty, lower inattentiveness score) reported higher levels of honesty, as did students with higher perceptions of confidentiality.

Among the variables examined in the table below, sentence comprehension, inattentiveness and perceived confidentiality were significantly associated with reporting of overall discomfort completing the survey. Students with lower inattentiveness and higher perceptions of confidentiality had lower scores of perceived discomfort completing the survey.

Reading enjoyment and perceptions of confidentiality were the only individual characteristics significantly related to a student reporting liking the survey. Students with higher reading enjoyments scores and higher perceptions of confidentiality reported higher self-reported enjoyment (liking) of the survey.
Table 7.12: Student Evaluation by Individual Characteristics

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Overall Honesty Mean(SD)</th>
<th>Corr or F</th>
<th>Overall Discomfort Mean(SD)</th>
<th>Corr or F</th>
<th>Like Survey Mean(SD)</th>
<th>Corr or F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Sentence Comp</td>
<td>100</td>
<td>4.21 (.17)</td>
<td>.305</td>
<td>2.28 (1.23)</td>
<td>.120</td>
<td>3.93 (1.25)</td>
<td>.016</td>
</tr>
<tr>
<td>High Sentence Comp</td>
<td>175</td>
<td>4.67 (.64)</td>
<td></td>
<td>2.01 (1.05)</td>
<td></td>
<td>3.98 (1.05)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>0.068</td>
<td>.049</td>
<td>0.746</td>
<td>.790</td>
</tr>
<tr>
<td>Low Listening Comp</td>
<td>48</td>
<td>4.27 (1.22)</td>
<td>.121</td>
<td>2.36 (1.29)</td>
<td>.082</td>
<td>3.67 (1.37)</td>
<td>.078</td>
</tr>
<tr>
<td>High Listening Comp</td>
<td>226</td>
<td>4.56 (.80)</td>
<td>.045</td>
<td>2.05 (1.08)</td>
<td>.177</td>
<td>4.02 (1.06)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>&lt;.028</td>
<td>.001</td>
<td>0.126</td>
<td>.096</td>
<td>0.191</td>
<td>.112</td>
</tr>
<tr>
<td>Low Lang/US Score</td>
<td>63</td>
<td>4.26 (1.09)</td>
<td></td>
<td>2.03 (1.07)</td>
<td>.181</td>
<td>4.13 (1.17)</td>
<td></td>
</tr>
<tr>
<td>High Lang/US Score</td>
<td>210</td>
<td>4.59 (.80)</td>
<td></td>
<td>2.13 (1.14)</td>
<td></td>
<td>3.92 (1.09)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>&lt;.028</td>
<td>.001</td>
<td>0.550</td>
<td>.834</td>
<td>0.191</td>
<td>.112</td>
</tr>
<tr>
<td>Low Read Enjoyment</td>
<td>153</td>
<td>4.46 (.88)</td>
<td>.094</td>
<td>2.18 (1.23)</td>
<td>.055</td>
<td>3.79 (1.19)</td>
<td>.202</td>
</tr>
<tr>
<td>High Read Enjoyment</td>
<td>121</td>
<td>4.59 (.88)</td>
<td>.123</td>
<td>2.01 (.97)</td>
<td></td>
<td>4.20 (.96)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>.241</td>
<td>.123</td>
<td>0.203</td>
<td>.366</td>
<td>0.191</td>
<td>.112</td>
</tr>
<tr>
<td>Low Read Difficulty</td>
<td>219</td>
<td>4.63 (.71)</td>
<td>-.299</td>
<td>2.09 (1.12)</td>
<td>-.063</td>
<td>4.00 (1.07)</td>
<td>-.101</td>
</tr>
<tr>
<td>High Read Difficulty</td>
<td>56</td>
<td>4.00 (1.29)</td>
<td></td>
<td>2.16 (1.14)</td>
<td></td>
<td>3.80 (1.30)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>0.666</td>
<td>.299</td>
<td>0.245</td>
<td>.097</td>
</tr>
<tr>
<td>Low Inattentiveness</td>
<td>135</td>
<td>4.67 (.73)</td>
<td>.133</td>
<td>1.90 (.93)</td>
<td>.112</td>
<td>4.09 (.97)</td>
<td>.086</td>
</tr>
<tr>
<td>High Inattentiveness</td>
<td>139</td>
<td>4.35 (1.01)</td>
<td></td>
<td>2.28 (1.24)</td>
<td></td>
<td>3.85 (1.23)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>&lt;.001</td>
<td>.028</td>
<td>0.005</td>
<td>.006</td>
<td>0.063</td>
<td>.160</td>
</tr>
<tr>
<td>Low Confidentiality</td>
<td>66</td>
<td>3.92 (1.23)</td>
<td>.403</td>
<td>2.43 (1.30)</td>
<td>.187</td>
<td>3.67 (1.27)</td>
<td>.162</td>
</tr>
<tr>
<td>High Confidentiality</td>
<td>207</td>
<td>4.70 (.66)</td>
<td></td>
<td>2.00 (1.05)</td>
<td></td>
<td>4.04 (1.06)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>&lt;.001</td>
<td>&lt;.001</td>
<td>0.017</td>
<td>.002</td>
<td>0.037</td>
<td>.008</td>
</tr>
<tr>
<td>Low Privacy</td>
<td>104</td>
<td>4.55 (.82)</td>
<td>.019</td>
<td>2.24 (1.12)</td>
<td>.077</td>
<td>4.00 (1.05)</td>
<td>.006</td>
</tr>
<tr>
<td>High Privacy</td>
<td>166</td>
<td>4.49 (.94)</td>
<td></td>
<td>2.01 (1.12)</td>
<td></td>
<td>3.92 (1.23)</td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td></td>
<td>.592</td>
<td>.752</td>
<td>.105</td>
<td>.206</td>
<td>.554</td>
<td>.922</td>
</tr>
</tbody>
</table>

4.1 Multivariate Models for Overall Reported Honesty

To further assess the relationship of sentence comprehension, language/US index, reading difficulty, inattentiveness, and perceived confidentiality on overall reported honesty, variables were entered into a multivariate linear regression model that adjusted for gender, age, race, living arrangement (two parents vs. other). The results of this
adjusted model prior to the inclusion of the variables of interest are shown in Table 7.13. As indicated by the results, age, gender, race were not significantly associated with self-reported overall honesty. However, having two parents in the home (SES proxy) was positively associated with self-reported overall honesty.

**Table 7.13: Regression Parameters for Overall Reported Honesty and Demographics**

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-.060</td>
<td>.075</td>
<td>.424</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>.176</td>
<td>.109</td>
<td>.106</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-.098</td>
<td>.151</td>
<td>.518</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-.235</td>
<td>.143</td>
<td>.101</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-.174</td>
<td>.207</td>
<td>.399</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>.237</td>
<td>.111</td>
<td>.033</td>
</tr>
</tbody>
</table>

### 4.1.1 Sentence Comprehension

This analysis examined the effect of sentence comprehension stanine on overall reported honesty. A linear regression model examining the relationship between overall honesty and sentence comprehension was adjusted for gender, age, race and two parents in the home. As indicated in the table below, sentence comprehension was significantly related to overall honesty (p<.001), with higher sentence stanine scores related to greater overall honesty. As indicated by the model, for every one point increase in sentence comprehension, overall honesty increases by 0.181.
4.1.2 Language Barriers:

This analysis examined the effect of Language/US index on overall reported honesty. A linear regression model examining the relationship between overall reported honesty and Language/US index was adjusted for gender, age, race and two parents in the home. Language/US index was entered into the model as a categorical variable, with students with no indicators serving as the reference category. Students with two or three indicators were collapsed into one category due to the small number of students with all three indicators. As indicated in Table 7.15 below, Language/US index was significantly related to overall honesty, with students with one indicator reporting lower overall honesty when compared to students with no indicators (p=.001); there was no difference between students with two or three indicators when compared to students with no indicators.

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.006</td>
<td>0.073</td>
<td>0.935</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>0.158</td>
<td>0.105</td>
<td>0.134</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-0.030</td>
<td>0.146</td>
<td>0.836</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-0.116</td>
<td>0.140</td>
<td>0.406</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-0.067</td>
<td>0.200</td>
<td>0.736</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>0.184</td>
<td>0.107</td>
<td>0.088</td>
</tr>
<tr>
<td>Sentence Comp. Stanine</td>
<td>0.181</td>
<td>0.039</td>
<td>0.000</td>
</tr>
</tbody>
</table>
4.1.3 Reading Difficulty

This analysis examined the effect of reading difficulty on overall reported honesty. A linear regression model examining the relationship between overall reported honesty and reading difficulty was adjusted for gender, age, race and two parents in the home. As indicated in Table 7.16 below, reading difficulty was significantly related to overall honesty (p<.001), with increasing reading difficulty related to lower reporting of overall reported honesty.

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-.045</td>
<td>.074</td>
<td>.544</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>.150</td>
<td>.106</td>
<td>.159</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-.085</td>
<td>.148</td>
<td>.566</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-.081</td>
<td>.152</td>
<td>.593</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-.118</td>
<td>.204</td>
<td>.564</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>.196</td>
<td>.109</td>
<td>.073</td>
</tr>
<tr>
<td>Language/US Index: 1 vs. 0 indicators</td>
<td>-.543</td>
<td>.166</td>
<td>.001</td>
</tr>
<tr>
<td>Language/US Index: 2/3 vs. 0 indicators</td>
<td>.030</td>
<td>.103</td>
<td>.767</td>
</tr>
</tbody>
</table>

Table 7.16: Regression Parameters for Overall Reported Honesty and Reading Difficulty
4.1.4 Inattentiveness

This analysis examined the relationship between inattentiveness and overall reported honesty. A linear regression model examining the relationship between overall reported honesty and inattentiveness was adjusted for gender, age, race and two parents in the home. As indicated in the table below, despite the bivariate association, inattentiveness was not significantly related to overall reported honesty.

Table 7.17: Regression Parameters for Overall Reported Honesty and Perceived Confidentiality

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.049</td>
<td>0.075</td>
<td>0.519</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>0.137</td>
<td>0.112</td>
<td>0.223</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-0.102</td>
<td>0.151</td>
<td>0.501</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-0.239</td>
<td>0.142</td>
<td>0.094</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-0.209</td>
<td>0.207</td>
<td>0.313</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>0.225</td>
<td>0.111</td>
<td>0.044</td>
</tr>
<tr>
<td>Inattentiveness</td>
<td>-0.077</td>
<td>0.046</td>
<td>0.092</td>
</tr>
</tbody>
</table>

Due to the clustering of the inattentiveness measure by teacher (ICC=.133), the multivariate regression analysis was also run through STATA, yielding adjusted robust standard errors and corresponding p-value. Analysis in STATA indicated no impact of the clustering on the relationship between inattentiveness and overall reported honesty.

4.1.5 Perceived Confidentiality

This analysis examined the relationship between perceived confidentiality and overall reported honesty. A linear regression model examining the relationship between overall reported honesty and perceived confidentiality was adjusted for gender, age, race and two parents in the home. As indicated in the table below, perceived confidentiality
was significantly related to overall reported honesty (p<.001), with higher scores of perceived confidentiality related to higher reporting of overall honesty.

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.048</td>
<td>0.070</td>
<td>0.495</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>0.127</td>
<td>0.101</td>
<td>0.209</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-0.037</td>
<td>0.140</td>
<td>0.789</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-0.159</td>
<td>0.132</td>
<td>0.231</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-0.024</td>
<td>0.192</td>
<td>0.899</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>0.201</td>
<td>0.103</td>
<td>0.051</td>
</tr>
<tr>
<td>Perceived Confidentiality</td>
<td>0.354</td>
<td>0.051</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### 4.2 Multivariate Models for Overall Discomfort

To further assess the relationship of the independent variables of interest on students’ discomfort completing the survey, sentence comprehension, inattentiveness, and perceived confidentiality were entered into multivariate linear regression models that adjusted for gender, age, race, living arrangement (two parents vs. other). The results of this adjusted model prior to the inclusion of the variables of interest are shown in Table 7.19. As indicated by the results, age, gender, race, and having two parents in the home (SES proxy) were not significantly associated with students’ overall discomfort.

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-.025</td>
<td>.097</td>
<td>.797</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>-1.188</td>
<td>.139</td>
<td>.177</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>.225</td>
<td>.192</td>
<td>.241</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>.101</td>
<td>.182</td>
<td>.578</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>.001</td>
<td>.262</td>
<td>.998</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>-.144</td>
<td>.141</td>
<td>.306</td>
</tr>
</tbody>
</table>
4.2.1 Sentence Comprehension

This analysis examined the relationship between sentence comprehension and overall discomfort. A linear regression model examining the relationship between overall discomfort and sentence comprehension was adjusted for gender, age, race and two parents in the home. As indicated in the table below, despite the bivariate association, sentence comprehension was not significantly related to discomfort.

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.052</td>
<td>0.098</td>
<td>0.594</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>-0.177</td>
<td>0.138</td>
<td>0.201</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>0.191</td>
<td>0.192</td>
<td>0.320</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>0.040</td>
<td>0.184</td>
<td>0.828</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-0.054</td>
<td>0.263</td>
<td>0.839</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>-0.116</td>
<td>0.141</td>
<td>0.411</td>
</tr>
<tr>
<td>Sentence Comp. Stanine</td>
<td>-0.092</td>
<td>0.051</td>
<td>0.075</td>
</tr>
</tbody>
</table>

4.2.2 Inattentiveness

This analysis examined the relationship between inattentiveness and overall discomfort. A linear regression model examining the relationship between overall discomfort and inattentiveness was adjusted for gender, age, race and two parents in the home. As indicated in the table below, despite the bivariate relationship, inattentiveness was not significantly related to discomfort.
Table 7.21: Regression Parameters for Discomfort and Inattentiveness

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.024</td>
<td>0.096</td>
<td>0.806</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>-0.116</td>
<td>0.142</td>
<td>0.414</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>0.188</td>
<td>0.190</td>
<td>0.325</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>0.108</td>
<td>0.180</td>
<td>0.549</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>0.030</td>
<td>0.261</td>
<td>0.908</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>-0.106</td>
<td>0.140</td>
<td>0.450</td>
</tr>
<tr>
<td>Inattentiveness</td>
<td>0.083</td>
<td>0.057</td>
<td>0.148</td>
</tr>
</tbody>
</table>

Due to the clustering of the inattentiveness measure by teacher (ICC=.133), the multivariate regression analysis was also run through STATA, yielding adjusted robust standard errors and corresponding p-value. Analysis in STATA indicated no impact of the clustering on the relationship between inattentiveness and discomfort completing the survey.

4.2.3 Perceived Confidentiality

This analysis examined the relationship between perceived confidentiality and overall discomfort while completing the survey. A linear regression model examining the relationship between overall discomfort and perceived confidentiality was adjusted for gender, age, race and two parents in the home. As indicated in the table below, perceived confidentiality was significantly related to discomfort (p<.003), with higher scores of perceived confidentiality related to lower reporting of discomfort.
Table 7.22: Regression Parameters for Discomfort and Perceived Confidentiality

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.033</td>
<td>0.096</td>
<td>0.728</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>-0.158</td>
<td>0.138</td>
<td>0.253</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>0.190</td>
<td>0.190</td>
<td>0.318</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>0.064</td>
<td>0.180</td>
<td>0.724</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-0.087</td>
<td>0.261</td>
<td>0.738</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>-0.124</td>
<td>0.139</td>
<td>0.376</td>
</tr>
<tr>
<td>Confidentiality</td>
<td>-0.208</td>
<td>0.070</td>
<td>0.003</td>
</tr>
</tbody>
</table>

4.3 Multivariate Models for Overall Enjoyment Completing the Survey

To assess the relationship of reading enjoyment and perceived confidentiality on enjoyment completing the survey, variables were entered into multivariate linear regression models that adjusted for gender, age, race, living arrangement (two parents vs. other). The results of this adjusted model prior to the inclusion of the variables of interest are shown in Table 7.23. As indicated by the results, gender and race were not significantly associated with students’ overall enjoyment completing the survey. However, gender and having two parents in the home (SES proxy) were found to be associated with enjoyment completing the survey; with females and youth living with two parents enjoying the process more than males and those not living with two parents.

Table 7.23: Regression Parameters for Enjoyment and Demographics

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.208</td>
<td>.095</td>
<td>.029</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>.196</td>
<td>.138</td>
<td>.156</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>.014</td>
<td>.191</td>
<td>.944</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>.047</td>
<td>.180</td>
<td>.793</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-.068</td>
<td>.260</td>
<td>.794</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>.322</td>
<td>.140</td>
<td>.022</td>
</tr>
</tbody>
</table>
4.3.1 Reading Enjoyment

This analysis examined the relationship between reading enjoyment and overall enjoyment completing the survey. A linear regression model examining the relationship between enjoyment completing the survey and reading enjoyment was adjusted for gender, age, race and two parents in the home. As indicated in the table below, reading enjoyment was significantly related to enjoyment completing the survey (p=.001), with higher reading enjoyment related to greater enjoyment completing the survey.

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.226</td>
<td>0.092</td>
<td>0.015</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>0.051</td>
<td>0.139</td>
<td>0.714</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>0.051</td>
<td>0.186</td>
<td>0.782</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>0.111</td>
<td>0.176</td>
<td>0.530</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-0.078</td>
<td>0.253</td>
<td>0.759</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>0.303</td>
<td>0.136</td>
<td>0.027</td>
</tr>
<tr>
<td>Reading enjoyment</td>
<td>0.266</td>
<td>0.079</td>
<td>0.001</td>
</tr>
</tbody>
</table>

4.3.2 Perceived Confidentiality

This analysis examined the relationship between perceived confidentiality and enjoyment completing the survey. A linear regression model examining the relationship between enjoyment completing the survey and perceived confidentiality was adjusted for gender, age, race and two parents in the home. As indicated in the table below, perceived confidentiality was significantly related to enjoyment (p=.008), with higher scores of perceived confidentiality related to higher reporting of enjoyment completing the survey.
Table 7.25: Regression Parameters for Enjoyment and Perceived Confidentiality

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.223</td>
<td>0.094</td>
<td>0.019</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>0.164</td>
<td>0.137</td>
<td>0.233</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>0.040</td>
<td>0.189</td>
<td>0.834</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>0.074</td>
<td>0.179</td>
<td>0.680</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>0.009</td>
<td>0.259</td>
<td>0.972</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>0.298</td>
<td>0.139</td>
<td>0.033</td>
</tr>
<tr>
<td>Perceived Confidentiality</td>
<td>0.185</td>
<td>0.070</td>
<td>0.008</td>
</tr>
</tbody>
</table>

5.0 Factorial Invariance

In thinking about the inclusion of students with varying literacy levels, it was important to assess the consistency of data collected among students with a range of abilities. Because we were able to include students with a range of reading and language skills, factor invariance was assessed for sentence comprehension, listening comprehension, and language/US index.

Five constructs examined for factorial invariance by mode in Chapter 6 (i.e. reading enjoyment, reading difficulty, sex beliefs, consequences of early sex, and parental sex beliefs) were also examined for factorial invariance by sentence comprehension, listening comprehension, and language/US index. That is, the measurement models of each of the constructs were simultaneously examined across defined groups to test whether the measurement weights were the same. Sentence comprehension, listening comprehension, and language/US index were dichotomized to provide adequate group sample size. Dichotomization of each variable followed previous groupings: sentence comprehension and listening comprehension were split based on stanine score, with low defined as a stanine score of three or less, and moderate/high
defined as a stanine score of four and higher; language/US index was divided as no indicators versus at least one indicator.

Testing of all models followed the same approach. One by one, models for each construct were created in AMOS 7.0. Comparison groups were defined using the “manage groups” interface within AMOS 7.0. As described earlier, AMOS first fits an unconstrained model, allowing the measurement weights for all variable paths in the model to be determined for each identified group. Next, the measurement weights for one group are applied to the other identified group(s), effectively constraining the paths in these other groups to be equal to the paths in the first group. The fit statistics from the constrained measurement weights model are compared to the unconstrained model, allowing for nested model comparisons. The nested model comparisons testing the constrained measurement weights model against the unconstrained model are presented below.

5.1 Factorial Invariance by Sentence Comprehension

5.1.1 Factorial Invariance by Sentence Comprehension: Reading Enjoyment

The measurement model for reading enjoyment was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.1 for a graphical representation of the measurement model. After defining the two sentence comprehension groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.26 below. As indicated by the non-significant p-value in the nested model comparison, the measurement weights for reading enjoyment did not vary significantly by sentence comprehension.
### Table 7.26: Model Fit for Reading Enjoyment: Overall and By Sentence Comprehension

<table>
<thead>
<tr>
<th></th>
<th>Chi-square</th>
<th>df</th>
<th>p</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>22.593</td>
<td>9</td>
<td>0.007</td>
<td>.972</td>
<td>.959</td>
<td>0.074</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>40.269</td>
<td>18</td>
<td>.002</td>
<td>.951</td>
<td>.934</td>
<td>0.067</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td>43.594</td>
<td>23</td>
<td>.006</td>
<td>.947</td>
<td>.952</td>
<td>0.057</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
<td>3.325</td>
<td>5</td>
<td>.650</td>
<td>.004</td>
<td>.004</td>
<td></td>
</tr>
</tbody>
</table>

### 5.1.2 Factorial Invariance by Sentence Comprehension: Reading Difficulty

The measurement model for reading difficulty was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.2 for a graphical representation of the measurement model. After defining the two sentence comprehension groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.27 below. As indicated by the non-significant p-value in the nested model comparison, the measurement weights for reading difficulty did not vary significantly by sentence comprehension.

### Table 7.27: Model Fit for Reading Difficulty: Overall and By Sentence Comprehension

<table>
<thead>
<tr>
<th></th>
<th>Chi-square</th>
<th>df</th>
<th>P</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>97.933</td>
<td>5</td>
<td>0.000</td>
<td>.831</td>
<td>.506</td>
<td>0.26</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>100.049</td>
<td>10</td>
<td>.000</td>
<td>.819</td>
<td>.483</td>
<td>.182</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td>104.013</td>
<td>14</td>
<td>.000</td>
<td>.812</td>
<td>.631</td>
<td>.153</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
<td>3.964</td>
<td>4</td>
<td>.411</td>
<td>-.007</td>
<td>-.148</td>
<td></td>
</tr>
</tbody>
</table>
5.1.3 Factorial Invariance by Sentence Comprehension: Sex Beliefs

The measurement model for sex beliefs was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.3 for a graphical representation of the measurement model. After defining the two sentence comprehension groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.28 below. The significant p-value of .049 in the nested model comparison would seem to indicate that the measurement weights for sex beliefs varied significantly by sentence comprehension. However, due to the potential inflation of the chi-square test in large samples (Marsh, 1994) and the minimal changes in the NFI and TLI, it was concluded that the measurement weights were, in fact, invariant.

| Table 7.28: Model Fit for Sex Beliefs: Overall and By Sentence Comprehension |
|-----------------------------|------------------|-----------------|-----|-----|-----|
|                             | Chi-square | df  | P   | NFI  | TLI  | RMSEA |
| Model 1: Overall Model Fit  | 6.606      | 5   | .252| .986 | .990 | .034  |
| Model 2: Unconstrained Group Comparison | 18.001 | 10  | .055| .966 | .952 | .054  |
| Model 3: Constrained Measurement Weights | 27.561 | 14  | .016| .948 | .942 | .060  |
| Model 2 vs. Model 3        | 9.560      | 4   | .049| -.018| -.010|       |

5.1.4 Factorial Invariance by Sentence Comprehension: Consequences of Early Sexual Behavior

The measurement model for consequences of early sexual behavior was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.4 for a graphical representation of the measurement model. After defining the two sentence comprehension groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.29 below. As indicated by the non-significant
p-value in the nested model comparison, the measurement weights for consequences of early sexual behavior do not vary significantly by sentence comprehension.

<table>
<thead>
<tr>
<th>Table 7.29: Model Fit for Consequences of Early Sexual Behavior: Overall and By Sentence Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
</tr>
</tbody>
</table>

5.1.5 Factorial Invariance by Sentence Comprehension: Parental Sex Beliefs

The measurement model for parental sex beliefs was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.5 for a graphical representation of the measurement model. After defining the two sentence comprehension groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.30 below. As indicated by the non-significant p-value in the nested model comparison, the measurement weights for parental sex beliefs do not vary significantly by sentence comprehension.

<table>
<thead>
<tr>
<th>Table 7.30: Model Fit for Parental Sex Beliefs: Overall and By Sentence Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
</tr>
</tbody>
</table>
5.2 Factorial Invariance by Listening Comprehension

Factorial invariance was next assessed for the five constructs among low vs. moderate/high listening comprehension. Results are provided below.

5.2.1 Factorial Invariance by Listening Comprehension: Reading Enjoyment

The measurement model for reading enjoyment was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.1 for a graphical representation of the measurement model. After defining the two listening comprehension groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.31 below. As indicated by the non-significant p-value in the nested model comparison, the measurement weights for reading enjoyment did not vary significantly by listening comprehension.

<table>
<thead>
<tr>
<th>Table 7.31: Model Fit for Reading Enjoyment: Overall and By Listening Comprehension</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>Model 1: Overall Model Fit</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
</tr>
</tbody>
</table>

5.2.2 Factorial Invariance by Listening Comprehension: Reading Difficulty

The measurement model for reading difficulty was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.2 for a graphical representation of the measurement model. After defining the two listening comprehension groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.32 below. As indicated by the non-significant p-value in the nested model
comparison, the measurement weights for reading difficulty did not vary significantly by listening comprehension.

Table 7.32: Model Fit for Reading Difficulty: Overall and By Listening Comprehension

<table>
<thead>
<tr>
<th>Model</th>
<th>Chi-square</th>
<th>df</th>
<th>P</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>97.933</td>
<td>5</td>
<td>0.000</td>
<td>0.831</td>
<td>0.506</td>
<td>0.26</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>100.044</td>
<td>10</td>
<td>.000</td>
<td>0.827</td>
<td>0.507</td>
<td>0.182</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td>103.786</td>
<td>14</td>
<td>.000</td>
<td>0.820</td>
<td>0.649</td>
<td>0.153</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
<td>3.742</td>
<td>4</td>
<td>.442</td>
<td>-.007</td>
<td>-.142</td>
<td></td>
</tr>
</tbody>
</table>

5.2.3 Factorial Invariance by Listening Comprehension: Sex Beliefs

The measurement model for sex beliefs was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.3 for a graphical representation of the measurement model. After defining the two listening comprehension groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.33 below. As indicated by the non-significant p-value in the nested model comparison, the measurement weights for sex beliefs did not vary significantly by listening comprehension.

Table 7.33: Model Fit for Sex Beliefs: Overall and By Listening Comprehension

<table>
<thead>
<tr>
<th>Model</th>
<th>Chi-square</th>
<th>df</th>
<th>p</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>6.606</td>
<td>5</td>
<td>.252</td>
<td>.986</td>
<td>.990</td>
<td>.034</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>15.923</td>
<td>10</td>
<td>.102</td>
<td>.968</td>
<td>.962</td>
<td>.047</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
<td>4.87</td>
<td>4</td>
<td>.301</td>
<td>-.010</td>
<td>.007</td>
<td></td>
</tr>
</tbody>
</table>
5.2.4 Factorial Invariance by Listening Comprehension: Consequences of Early Sexual Behavior

The measurement model for consequences of early sexual behavior was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.4 for a graphical representation of the measurement model. After defining the two listening comprehension groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.34 below. As indicated by the non-significant p-value in the nested model comparison, the measurement weights for consequences of early sexual behavior did not vary significantly by listening comprehension.

| Table 7.34: Model Fit for Consequences of Early Sexual Behavior: Overall and By Listening Comprehension |
|-------------------------------------------------|------|--------|--------|--------|
| Model 1: Overall Model Fit                      | 19.346 | 2     | 0.000  | 0.943  | 0.736  | 0.178 |
| Model 2: Unconstrained Group Comparison         | 25.074 | 4     | 0.000  | 0.927  | 0.676  | 0.139 |
| Model 3: Constrained Measurement Weights        | 27.127 | 7     | 0.000  | 0.922  | 0.823  | 0.103 |
| Model 2 vs. Model 3                             | 2.053  | 3     | .561   | -.005  | .147   |

5.2.5 Factorial Invariance by Listening Comprehension: Parental Sex Beliefs

The measurement model for parental sex beliefs was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.5 for a graphical representation of the measurement model. After defining the two listening comprehension groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.35 below. As indicated by the non-significant p-value in the nested model comparison, the measurement weights for parental sex beliefs did not vary significantly by listening comprehension.
Table 7.35: Model Fit for Parental Sex Beliefs: Overall and By Listening Comprehension

<table>
<thead>
<tr>
<th>Model</th>
<th>Chi-square</th>
<th>df</th>
<th>p</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>2.988</td>
<td>2</td>
<td>.224</td>
<td>0.988</td>
<td>0.980</td>
<td>0.042</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>3.472</td>
<td>4</td>
<td>.482</td>
<td>.987</td>
<td>1.000</td>
<td>.000</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td>4.446</td>
<td>7</td>
<td>.727</td>
<td>.984</td>
<td>1.000</td>
<td>.000</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
<td>0.974</td>
<td>3</td>
<td>.808</td>
<td>-.003</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

5.3 Factorial Invariance by Language/US Index

Factorial invariance was assessed for five constructs for students with no indicators of the Language/US index compared to students with at least one indicator of the Language/US index. Results are provided below.

5.3.1 Factorial Invariance by Language/US Index: Reading Enjoyment

The measurement model for reading enjoyment was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.1 for a graphical representation of the measurement model. After defining the two Language/US index groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.36 below. As indicated by the non-significant p-value in the nested model comparison, the measurement weights for reading enjoyment did not vary significantly by Language/US index.
Table 7.36: Model Fit for Reading Enjoyment: Overall and By Language/US Index

<table>
<thead>
<tr>
<th></th>
<th>Chi-square</th>
<th>df</th>
<th>p</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>22.593</td>
<td>9</td>
<td>0.007</td>
<td>0.972</td>
<td>0.959</td>
<td>0.074</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>27.248</td>
<td>18</td>
<td>.074</td>
<td>.983</td>
<td>.987</td>
<td>.044</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td>35.154</td>
<td>23</td>
<td>.050</td>
<td>.979</td>
<td>.986</td>
<td>.044</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
<td>7.906</td>
<td>5</td>
<td>.162</td>
<td>-0.004</td>
<td>-0.001</td>
<td></td>
</tr>
</tbody>
</table>

5.3.2 Factorial Invariance by Language/US Index: Reading Difficulty

The measurement model for reading difficulty was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.2 for a graphical representation of the measurement model. After defining the two Language/US index groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.37 below. As indicated by the non-significant p-value in the nested model comparison, the measurement weights for reading difficulty did not vary significantly by Language/US index.

Table 7.37: Model Fit for Reading Difficulty: Overall and By Language/US Index

<table>
<thead>
<tr>
<th></th>
<th>Chi-square</th>
<th>df</th>
<th>p</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>97.933</td>
<td>5</td>
<td>0.000</td>
<td>0.831</td>
<td>0.506</td>
<td>0.26</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>80.825</td>
<td>10</td>
<td>.000</td>
<td>.937</td>
<td>.831</td>
<td>.162</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td>85.460</td>
<td>14</td>
<td>.000</td>
<td>.934</td>
<td>.879</td>
<td>.137</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
<td>4.635</td>
<td>4</td>
<td>.327</td>
<td>-.003</td>
<td>.048</td>
<td></td>
</tr>
</tbody>
</table>

5.3.3 Factorial Invariance by Language/US Index: Sex Beliefs

The measurement model for sex beliefs was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.3 for a graphical representation of the measurement model. After
defining the two Language/US index groups within AMOS, the unconstrained model and
the constrained measurement weights model were run. Fit statistics are provided in Table
7.38 below. As indicated by the highly significant p-value of less than .001 in the nested
model comparison, the measurement weights for sex beliefs varied significantly by
Language/US index.

<table>
<thead>
<tr>
<th>Model 1: Overall Model Fit</th>
<th>Chi-square</th>
<th>df</th>
<th>p</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>40.008</td>
<td>10</td>
<td>.000</td>
<td>.918</td>
<td>.803</td>
<td>.105</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td>135.354</td>
<td>14</td>
<td>.000</td>
<td>.723</td>
<td>.432</td>
<td>.179</td>
</tr>
</tbody>
</table>

Model 2 vs. Model 3 | 95.346 | 4   | .000   | .195  | .371  |

5.3.4 Factorial Invariance by Language/US Index: Consequences of Early Sexual Behavior

The measurement model for consequences of early sexual behavior was defined
in AMOS 7.0; refer to Chapter 6, Section 4.1.4 for a graphical representation of the
measurement model. After defining the two Language/US index groups within AMOS,
the unconstrained model and the constrained measurement weights model were run. Fit
statistics are provided in Table 7.39 below. As indicated by the non-significant p-value
in the nested model comparison, the measurement weights for consequences of early
sexual behavior did not vary significantly by Language/US index.
Table 7.39: Model Fit for Consequences of Early Sexual Behavior: Overall and By Language/US Index

<table>
<thead>
<tr>
<th>Model</th>
<th>Chi-square</th>
<th>df</th>
<th>P</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>19.346</td>
<td>2</td>
<td>0.000</td>
<td>0.943</td>
<td>0.736</td>
<td>0.178</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>23.299</td>
<td>4</td>
<td>0.000</td>
<td>0.930</td>
<td>0.693</td>
<td>0.133</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td>24.484</td>
<td>7</td>
<td>0.000</td>
<td>0.927</td>
<td>0.841</td>
<td>0.096</td>
</tr>
</tbody>
</table>

Model 2 vs. Model 3 1.185 3 .757 -.003 .148

5.3.5 Factorial Invariance by Language/US Index: Parental Sex Beliefs

The measurement model for parental sex beliefs was defined in AMOS 7.0; refer to Chapter 6, Section 4.1.5 for a graphical representation of the measurement model. After defining the two Language/US index groups within AMOS, the unconstrained model and the constrained measurement weights model were run. Fit statistics are provided in Table 7.40 below. As indicated by the significant p-value of 0.016 in the nested model comparison, the measurement weights for parental sex views varied significantly by Language/US index.

Table 7.40: Model Fit for Parental Sex Beliefs: Overall and By Language/US Index

<table>
<thead>
<tr>
<th>Model</th>
<th>Chi-square</th>
<th>df</th>
<th>P</th>
<th>NFI</th>
<th>TLI</th>
<th>RMSEA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1: Overall Model Fit</td>
<td>2.988</td>
<td>2</td>
<td>.224</td>
<td>0.988</td>
<td>0.980</td>
<td>0.042</td>
</tr>
<tr>
<td>Model 2: Unconstrained Group Comparison</td>
<td>3.203</td>
<td>4</td>
<td>.525</td>
<td>0.988</td>
<td>1.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Model 3: Constrained Measurement Weights</td>
<td>13.477</td>
<td>7</td>
<td>.061</td>
<td>0.951</td>
<td>0.927</td>
<td>0.058</td>
</tr>
<tr>
<td>Model 2 vs. Model 3</td>
<td>10.275</td>
<td>3</td>
<td>.016</td>
<td>-.037</td>
<td>-.073</td>
<td></td>
</tr>
</tbody>
</table>
5.4 Summary of Factorial Invariance Tests

In summary, among the five measurement models examined for factorial invariance, two were found to have some level of variance when examined across different groups of individuals based on reading and listening skills and language mastery. Specifically, the measurement models for sex beliefs and parental sex beliefs were found to vary by Language/US index.
Chapter 8: Results for Specific Aim 3

Testing for Moderating Effects of Cognitive Burden on Outcomes

Chapter Overview

Specific Aim 3: To determine whether and how individual characteristics modify the relationships between mode and outcomes (Path C)

The third aim of the study focused on the moderating effects of the cognitive burden variables (i.e. sentence comprehension, listening comprehension, Language/US index, reading enjoyment, reading difficulty, inattentiveness) on the relationship between mode and outcomes. That is, do the observed mode effects on data outcomes vary by individual characteristics such as sentence and listening abilities and preferences, language mastery or attentiveness? To test this, interaction terms were created by crossing the individual characteristics (centered) with each mode, with SAQ serving as the index variable. Thus, each moderating effect of an individual characteristic is represented by two variables, variable*pda and variable*apda. Moderation was only tested for models where both mode and the individual characteristic variable were found to be significant in previous analyses (Cohen, Cohen, West and Aiken, 2003). As mode was not significantly related to overall honesty, discomfort, or enjoyment, moderation of individual characteristics was not assessed for these outcomes. However, moderation of significant individual characteristics was assessed for number of questions answered and missing data and results are provided below.
1. Moderation within Number of Questions Answered

In Specific Aim 1, mode was found to be significantly related to the number of questions answered. In Specific Aim 2, sentence comprehension stanine, listening comprehension, reading difficulty, and inattentiveness were found to be significantly related to the number of questions answered. That is, the moderation analysis should answer the following questions:

1. Does the impact of mode on the number of questions answered vary by students’ sentence comprehension?
2. Does the impact of mode on the number of questions answered vary by students’ listening comprehension?
3. Does the impact of mode on the number of questions answered vary by students’ self-reported reading difficulty?
4. Does the impact of mode on the number of questions answered vary by student inattentiveness?

1.1 Mode and Sentence Comprehension

In this analysis, the direct effects of demographic variables (age, gender, race, two parents in the home), mode, and sentence comprehension on the number of questions answered (left side of Table 8.1) were examined, followed by the introduction of the cross-product terms (e.g. SC*PDA, SC*APDA) in a second step (right side of Table 8.1). To create the cross-product terms, sentence comprehension was centered by subtracting the mean value from each score. The resulting centered variable was then multiplied by both the PDA dummy variable and the APDA dummy variable, resulting in two cross-
product terms. As shown below in Table 8.1, the interaction terms were not significant, indicating lack of moderation. Thus, students taking the survey on APDA answered significantly more questions than students taking the survey on SAQ or PDA, regardless of their reading ability, as measured by sentence comprehension.

| Table 8.1: Regression Parameters for Questions Answered and Sentence Comprehension |
|-----|-----|-----|-----|-----|-----|
| Variable | B   | Se  | P   | B   | Se  | P   |
| Age     | 2.357 | 2.424 | 0.332 | 2.498 | 2.420 | 0.303 |
| Gender: Female (1) vs. Male (0) | 4.899 | 3.457 | 0.158 | 4.571 | 3.453 | 0.187 |
| Race: Black (1) vs. White (0) | -2.057 | 4.930 | 0.677 | -1.987 | 4.920 | 0.687 |
| Race: Hispanic (1) vs. White (0) | -5.975 | 4.713 | 0.206 | -5.447 | 4.711 | 0.249 |
| Race: Other (1) vs. White (0) | -7.525 | 6.596 | 0.255 | -6.147 | 6.628 | 0.355 |
| Two parents in home (1) vs. Other (0) | 2.068 | 3.555 | 0.561 | 1.645 | 3.555 | 0.644 |
| Mode: PDA (1) vs. SAQ (0) | 2.392 | 4.274 | 0.576 | 2.734 | 4.270 | 0.523 |
| Mode: APDA (1) vs. SAQ (0) | 32.507 | 4.286 | 0.000 | 32.712 | 4.276 | 0.000 |
| Sentence Comprehension | 6.660 | 1.293 | 0.000 | 7.084 | 2.168 | 0.001 |
| SC * PDA | 3.101 | 3.304 | 0.349 |
| SC * APDA | -2.688 | 2.882 | 0.352 |

1.2 Mode and Listening Comprehension

In this analysis, the direct effects of demographic variables (age, gender, race, two parents in the home), mode, and listening comprehension on the number of questions answered (left side of Table 8.2) were examined, followed by the introduction of the cross-product terms (e.g. LC*PDA, LC*APDA) in a second step (right side of Table 8.2). To create the cross-product terms, listening comprehension was centered by subtracting the mean value from each score. The resulting centered variable was then multiplied by both the PDA dummy variable and the APDA dummy variable, resulting in two cross-product terms. As shown below in Table 8.2, the interaction terms were not significant,
indicating lack of moderation. Thus, students taking the survey on APDA answered significantly more questions than students taking the survey on SAQ or PDA, regardless of their listening ability, as measured by their listening comprehension score.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>se</th>
<th>p</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.042</td>
<td>2.463</td>
<td>.673</td>
<td>1.228</td>
<td>2.466</td>
<td>.619</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>6.106</td>
<td>3.547</td>
<td>.086</td>
<td>6.203</td>
<td>3.547</td>
<td>.081</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-4.526</td>
<td>5.034</td>
<td>.369</td>
<td>-4.659</td>
<td>5.037</td>
<td>.356</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-7.545</td>
<td>4.829</td>
<td>.119</td>
<td>-7.336</td>
<td>4.852</td>
<td>.132</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-10.155</td>
<td>6.731</td>
<td>.133</td>
<td>-9.396</td>
<td>6.808</td>
<td>.169</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>2.208</td>
<td>3.665</td>
<td>.547</td>
<td>2.322</td>
<td>3.671</td>
<td>.528</td>
</tr>
<tr>
<td>Mode: PDA (1) vs. SAQ (0)</td>
<td>4.212</td>
<td>4.422</td>
<td>.342</td>
<td>4.849</td>
<td>4.446</td>
<td>.276</td>
</tr>
<tr>
<td>Mode: APDA (1) vs. SAQ (0)</td>
<td>33.400</td>
<td>4.391</td>
<td>.000</td>
<td>33.355</td>
<td>4.407</td>
<td>.000</td>
</tr>
<tr>
<td>Listening Comprehension</td>
<td>3.206</td>
<td>.931</td>
<td>.001</td>
<td>2.051</td>
<td>1.608</td>
<td>.203</td>
</tr>
<tr>
<td>LC * PDA</td>
<td></td>
<td></td>
<td></td>
<td>3.029</td>
<td>2.194</td>
<td>.169</td>
</tr>
<tr>
<td>LC * APDA</td>
<td></td>
<td></td>
<td></td>
<td>.387</td>
<td>2.241</td>
<td>.863</td>
</tr>
</tbody>
</table>

1.3 Mode and Reading Difficulty

In this analysis, the direct effects of demographic variables (age, gender, race, two parents in the home), mode, and reading difficulty on the number of questions answered (left side of Table 8.3) were examined, followed by the introduction of the cross-product terms (e.g. RD*PDA, RD*APDA) in a second step (right side of Table 8.3). To create the cross-product terms, reading difficulty was centered by subtracting the mean value from each score. The resulting centered variable was then multiplied by both the PDA dummy variable and the APDA dummy variable, resulting in two cross-product terms.
The reading difficulty interaction term by PDA (index is SAQ) was found to be significant (p=0.019), and models were probed for effect of the moderation.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>se</th>
<th>p</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.440</td>
<td>2.494</td>
<td>0.564</td>
<td>1.627</td>
<td>2.470</td>
<td>0.511</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>6.575</td>
<td>3.578</td>
<td>0.067</td>
<td>6.052</td>
<td>3.542</td>
<td>0.089</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-5.839</td>
<td>5.080</td>
<td>0.251</td>
<td>-6.056</td>
<td>5.043</td>
<td>0.231</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-10.080</td>
<td>4.778</td>
<td>0.036</td>
<td>-10.015</td>
<td>4.724</td>
<td>0.035</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-12.341</td>
<td>6.749</td>
<td>0.069</td>
<td>-11.311</td>
<td>6.696</td>
<td>0.092</td>
</tr>
<tr>
<td>Two parents in home (1) vs.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other (0)</td>
<td>3.009</td>
<td>3.663</td>
<td>0.412</td>
<td>2.372</td>
<td>3.668</td>
<td>0.518</td>
</tr>
<tr>
<td>Mode: PDA (1) vs. SAQ (0)</td>
<td>3.264</td>
<td>4.422</td>
<td>0.461</td>
<td>3.733</td>
<td>4.384</td>
<td>0.395</td>
</tr>
<tr>
<td>Mode: APDA (1) vs. SAQ (0)</td>
<td>34.267</td>
<td>4.412</td>
<td>0.000</td>
<td>34.211</td>
<td>4.368</td>
<td>0.000</td>
</tr>
<tr>
<td>Reading difficulty</td>
<td>-6.140</td>
<td>2.061</td>
<td>0.003</td>
<td>-2.120</td>
<td>3.832</td>
<td>0.581</td>
</tr>
<tr>
<td>RD * PDA</td>
<td></td>
<td></td>
<td></td>
<td>-11.807</td>
<td>5.022</td>
<td>0.019</td>
</tr>
<tr>
<td>RD * APDA</td>
<td></td>
<td></td>
<td></td>
<td>0.281</td>
<td>5.097</td>
<td>0.956</td>
</tr>
</tbody>
</table>

Using the mean +/-1 standard deviation approach to probe (Cohen, Cohen, West and Aiken, 2003; Holmbeck 2002), multivariate linear regression models were run to assess the modifying effect of reading difficulty on the relationship between mode and the number of questions answered. The results of this probing are provided in Figure 8.1 below. Results at one standard deviation below the mean are plotted as “low” on the x-axis, and results at one standard deviation above the mean are plotted as “high” on the x-axis; number of questions answered on the survey is plotted on the y-axis. SAQ respondents are represented by the blue diamonds, PDA respondents are represented by the pink squares, and APDA respondents are represented by the yellow triangles. The plot illustrates the number of questions answered above or below the mean value for the reference category, SAQ respondents. As shown in the plot, the slope for SAQ and
APDA respondents similarly decreases, but the slope for PDA respondents is much more pronounced, with an extreme downward slope as reading difficulty increases. This plot indicates that among PDA respondents, students with low self-assessed reading difficulty answer many more questions than students with high self-assessed reading difficulty. Thus, it appears that the technology of the PDA, in the absence of the audio-enhancement, may hinder student progress in answering survey questions for students with self-assessed reading difficulty.

Figure 8.1: Plot of Moderating Effect of Reading Difficulty on Questions Answered

1.4 Mode and Inattentiveness

In this analysis, the direct effects of demographic variables (age, gender, race, two parents in the home), mode, and inattentiveness on the number of questions answered
(left side of Table 8.4) were examined, followed by the introduction of the cross-product terms (e.g. Inattent*PDA, Inattent*APDA) in a second step (right side of Table 8.4). To create the cross-product terms, inattentiveness was centered by subtracting the mean value from each score. The resulting centered variable was then multiplied by both the PDA dummy variable and the APDA dummy variable, resulting in two cross-product terms. As shown below in Table 8.4, the interaction terms were not significant, indicating lack of moderation.

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>Se</th>
<th>P</th>
<th>B</th>
<th>Se</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.847</td>
<td>2.437</td>
<td>0.728</td>
<td>0.901</td>
<td>2.442</td>
<td>0.713</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>1.780</td>
<td>3.619</td>
<td>0.623</td>
<td>1.899</td>
<td>3.638</td>
<td>0.602</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-3.469</td>
<td>4.992</td>
<td>0.488</td>
<td>-4.138</td>
<td>5.047</td>
<td>0.413</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-10.606</td>
<td>4.698</td>
<td>0.025</td>
<td>-11.221</td>
<td>4.741</td>
<td>0.019</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-13.927</td>
<td>6.656</td>
<td>0.037</td>
<td>-14.520</td>
<td>6.710</td>
<td>0.031</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>2.381</td>
<td>3.607</td>
<td>0.510</td>
<td>2.421</td>
<td>3.619</td>
<td>0.504</td>
</tr>
<tr>
<td>Mode: PDA (1) vs. SAQ (0)</td>
<td>2.176</td>
<td>4.338</td>
<td>0.616</td>
<td>1.659</td>
<td>4.376</td>
<td>0.705</td>
</tr>
<tr>
<td>Mode: APDA (1) vs. SAQ (0)</td>
<td>32.011</td>
<td>4.409</td>
<td>0.000</td>
<td>32.042</td>
<td>4.447</td>
<td>0.000</td>
</tr>
<tr>
<td>Inattentiveness</td>
<td>-5.784</td>
<td>1.500</td>
<td>0.000</td>
<td>-7.840</td>
<td>2.455</td>
<td>0.002</td>
</tr>
<tr>
<td>Inattent * PDA</td>
<td>3.597</td>
<td>3.660</td>
<td>0.327</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inattent * APDA</td>
<td>2.909</td>
<td>3.427</td>
<td>0.397</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Moderation within Missing Data

In Specific Aim 1, mode was found to be significantly related to the missing data. In Specific Aim 2, inattentiveness was found to be significantly related to missing data. The moderation analysis should be able to inform us as to the impact of inattentiveness on the relationship between mode and missing data.
2.1 Mode and Inattentiveness

In this analysis, the direct effects of demographic variables (age, gender, race, two parents in the home), mode, and inattentiveness on missing data (left side of Table 8.5) were examined, followed by the introduction of the cross-product terms (e.g. Inattent*PDA, Inattent*APDA) in a second step (right side of Table 8.5). To create the cross-product terms, inattentiveness was centered by subtracting the mean value from each score. The resulting centered variable was then multiplied by both the PDA dummy variable and the APDA dummy variable, resulting in two cross-product terms. As shown below in Table 8.5, the interaction terms were not significant, indicating lack of moderation.

<table>
<thead>
<tr>
<th>Table 8.5: Regression Parameters for Missing Data and Inattentiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
</tr>
<tr>
<td>Mode: PDA (1) vs. SAQ (0)</td>
</tr>
<tr>
<td>Mode: APDA (1) vs. SAQ (0)</td>
</tr>
<tr>
<td>Inattentiveness</td>
</tr>
<tr>
<td>Inattent * PDA</td>
</tr>
<tr>
<td>Inattent * APDA</td>
</tr>
</tbody>
</table>
3. Summary of Results

In summary, perceived reading difficulty was the only variable found to act as a moderating effect on the relationship between data collection mode and the number of questions answered. Students with lower perceived reading difficulty answered more questions on the PDA compared to the SAQ respondents, but students with higher perceived reading difficulty answered fewer questions on PDA when compared to SAQ respondents.
Chapter 9: Results for Specific Aim 4

Testing Perceived Confidentiality and Privacy for Mediation and Moderation

Chapter Overview

Specific Aim 4: To clarify the role of perceived confidentiality and perceived privacy on the number of questions answered, data quality and student evaluation as either a moderator (Path D1) or a mediator (Paths D2).

The fourth and final aim of the study focused on clarifying the roles of perceived confidentiality and perceived privacy as either moderating effects or mediating effects on the relationship between mode and outcomes. Both sets of analyses are described in detail below.

1. Testing for Moderation

Interaction terms were created by crossing perceived confidentiality and perceived privacy (centered) with each mode, with SAQ serving as the index variable. Thus, each moderating effect of an individual characteristic is represented by two variables, variable*pda and variable*apda. Moderation was only tested for models where both mode and the individual characteristic variable were found to be significant in previous analyses (Cohen, Cohen, West and Aiken, 2003). As mode was not significantly related to overall honesty, discomfort, or enjoyment, moderation was not assessed in these models. Additionally, neither perceived confidentiality nor perceived
privacy was related to the number of questions answered. Thus, moderation was assessed
missing data; results are provided below.

1.1 Moderating Effect of Perceived Privacy on the Number of Questions Answered

Although the direct effect of perceived privacy was non-significant at a
significance level of 0.05 in the multivariate analysis conducted in Aim 2, the p-value
was only slightly outside of this bound. As this study was exploratory in nature, it was
decided that privacy should be examined for potential moderating effects. In this
analysis, the direct effects of demographic variables (age, gender, race, two parents in the
home), mode, and perceived privacy on presence of missing data (left side of Table 9.1)
were examined, followed by the introduction of the cross-product terms (e.g.
Privacy*PDA, Privacy*APDA) in a second step (right side of Table 9.1). As shown
below in Table 9.1, the interaction terms were not significant, indicating that perceived
privacy was not a moderator.

<table>
<thead>
<tr>
<th>Variable</th>
<th>b</th>
<th>se</th>
<th>p</th>
<th>b</th>
<th>se</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>1.327</td>
<td>2.565</td>
<td>.605</td>
<td>1.162</td>
<td>2.577</td>
<td>.652</td>
</tr>
<tr>
<td>Gender: Female (1) vs. Male (0)</td>
<td>6.151</td>
<td>3.649</td>
<td>.093</td>
<td>5.960</td>
<td>3.662</td>
<td>.105</td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>-4.604</td>
<td>5.135</td>
<td>.371</td>
<td>-4.666</td>
<td>5.147</td>
<td>.365</td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>-10.968</td>
<td>4.872</td>
<td>.025</td>
<td>-10.975</td>
<td>4.886</td>
<td>.026</td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>-13.399</td>
<td>6.938</td>
<td>.055</td>
<td>-13.422</td>
<td>6.952</td>
<td>.055</td>
</tr>
<tr>
<td>Two parents in home (1) vs. Other (0)</td>
<td>4.043</td>
<td>3.736</td>
<td>.280</td>
<td>3.530</td>
<td>3.783</td>
<td>.352</td>
</tr>
<tr>
<td>Mode: PDA (1) vs. SAQ (0)</td>
<td>.697</td>
<td>4.608</td>
<td>.880</td>
<td>.966</td>
<td>4.714</td>
<td>.838</td>
</tr>
<tr>
<td>Mode: APDA (1) vs. SAQ (0)</td>
<td>32.394</td>
<td>4.604</td>
<td>.000</td>
<td>33.231</td>
<td>4.703</td>
<td>.000</td>
</tr>
<tr>
<td>Privacy</td>
<td>2.480</td>
<td>1.799</td>
<td>.169</td>
<td>.685</td>
<td>3.058</td>
<td>.823</td>
</tr>
<tr>
<td>Privacy * PDA</td>
<td>4.192</td>
<td>4.458</td>
<td>.348</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Privacy * APDA</td>
<td>1.445</td>
<td>4.090</td>
<td>.724</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
1.2 Moderating Effect of Perceived Privacy on Missing Data

In this analysis, the direct effects of demographic variables (age, gender, race, two parents in the home), mode, and perceived privacy on presence of missing data (left side of Table 9.2) were examined, followed by the introduction of the cross-product terms (e.g. Privacy*PDA, Privacy*APDA) in a second step (right side of Table 9.2). As shown below in Table 9.2, the interaction terms were not significant, indicating that perceived privacy was not a moderator.

<table>
<thead>
<tr>
<th></th>
<th>95% CI</th>
<th>Exp(B)</th>
<th>95% CI</th>
<th>Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower</td>
<td>Upper</td>
<td>p</td>
<td>Lower</td>
</tr>
<tr>
<td>Gender: Female (1)</td>
<td>1.256 .643 2.455 .505</td>
<td>1.283 .655 2.513 .467</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>.917 .582 1.445 .710</td>
<td>.921 .585 1.449 .721</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race: Black (1) vs. White (0)</td>
<td>1.501 .610 3.695 .376</td>
<td>1.512 .614 3.726 .368</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race: Hispanic (1) vs. White (0)</td>
<td>.958 .385 2.380 .926</td>
<td>.950 .383 2.361 .913</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race: Other (1) vs. White (0)</td>
<td>1.298 .369 4.571 .685</td>
<td>1.305 .372 4.584 .678</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Two parents in home</td>
<td>.679 .340 1.353 .271</td>
<td>.719 .356 1.450 .356</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mode: PDA vs. SAQ</td>
<td>.382 .177 .824 .014</td>
<td>.358 .165 .779 .010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mode: APDA vs. SAQ</td>
<td>.130 .047 .365 .000</td>
<td>.122 .043 .344 .000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Privacy</td>
<td>.839 .605 1.163 .291</td>
<td>.962 .619 1.494 .863</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Privacy * PDA</td>
<td>.696 .338 1.434 .696</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Privacy * APDA</td>
<td>.859 .351 2.102 .859</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. Testing for Mediation

Following the criteria for testing mediation as outlined by Baron and Kenny (1986), perceived confidentiality and perceived privacy were assessed as mediators. The simple model outlined below in Figure 9.1 shows A (independent variable), B (the mediating variable), and C (the outcome variable). Using this model, the criteria to assess mediation are as follows: (1) A must be associated with C; (2) B must be
associated with C; (3) A must be associated with B; and (4) in the presence of B, the relationship of A and C becomes non-significant (i.e. complete mediation) or is diminished (i.e. partial mediation).

Figure 9.1: Model for Testing for Mediation

Drawing from the results of Aim 1 indicating an absence of the A-C relationship between mode and honesty, mode and overall discomfort, and mode and enjoyment completing the survey, these three outcomes were not assessed for mediation effects of perceived confidentiality and perceived privacy. Drawing from the results of Aim 2 indicating an absence of the B-C relationship between perceived confidentiality and the number of questions answered and perceived confidentiality and missing data, perceived confidentiality was effectively ruled out as a potential mediator. The results for perceived privacy as a mediator for the number of questions answered and missing data are provided below.

2.1 Mediating Effect of Perceived Privacy on Number of Questions Asked

As described above, earlier analyses confirmed two of the four criteria necessary for mediation: (1) A-C: mode was associated with the number of questions answered; and (2) B-C: perceived privacy was marginally (p=.052) associated with the number of
questions answered. To satisfy the third criterion of the A-B relationship, a Person’s correlation coefficient was calculated in SPSS, yielding a significant coefficient of 0.226, \( p<.001 \). The final criterion was examined utilizing the multivariate analysis run above in Section 1.1 which included demographic variables, mode, and perceived privacy. In examining the left-hand side of Table 9.1, if evidence of mediation were illustrated, the relationship between mode and the number of questions answered would diminish or disappear. Instead, the relationship between perceived privacy and the number of questions answered diminished in the presence of mode, indicating that privacy may be somewhat confounded with mode. Thus, based on the failure of the fourth criterion, perceived privacy was not found to be a mediator of the relationship between mode and the number of questions answered.

2.2 Mediating Effect of Perceived Privacy on Missing Data

As described above, earlier analyses confirmed two of the four criteria necessary for mediation: (1) A-C: mode was associated with missing data; and (2) B-C: perceived privacy was associated with missing data. To satisfy the third criterion of the A-B relationship, a Person’s correlation coefficient was calculated in SPSS, yielding a significant coefficient of 0.226, \( p<.001 \). The final criterion was examined utilizing the multivariate analysis run above in Section 1.2 which included demographic variables, mode, and perceived privacy. In examining the left-hand side of Table 9.2, if evidence of mediation were illustrated, the relationship between mode and missing data would diminish or disappear. Instead, the relationship between perceived privacy and missing data diminished in the presence of mode, indicating that privacy may be somewhat
confounded with mode. Thus, based on the failure of the fourth criterion, perceived privacy was not found to be a mediator of the relationship between mode and missing data.

2.3 Summary of Results

In summary, perceived confidentiality and perceived privacy of the survey environment were found to be neither moderators nor mediators of the relationships between mode and outcomes.
Chapter 10: Discussion

Chapter Overview

This chapter provides a discussion of the impact of mode of data collection and individual respondent characteristics on data outcomes and provides a context in which the study findings contribute to the literature regarding data collection with adolescents, and the role of reading ability and preferences, language mastery and inattentiveness on data outcomes. Implications of the results are provided. Limitations of the study and future studies are also discussed.

1. Impact of Data Collection Mode on Data Outcomes

Self-administered data collection methods have been shown to be preferred when data are potentially impacted by self-presentation bias (Aquilino, 1994, Aquilino and LoSciuto, 1990; Boekeloo et al, 1994; Epstein, Barker and Kroutil, 2001 Hewett, 2002; Jobe et al, 1997; London and Williams, 1990; Metzger et al, 2000; Mosher and Duffer, 1994; Robinson and West, 1992; Tourangeau et al, 1997; Schober et al, 1992; Tourangeau and Smith, 1996; and Turner, Lessler and DeVore, 1992; Waterton and Duffy, 1984; Mott, 1985). Further, evidence has shown that computer-assisted self-interviewing (CASI) systems provide researchers with cleaner data (e.g. fewer missing responses) and higher rates of reporting of sensitive behaviors in adult populations and in some youth populations (Le et al, 2006; Romer et al, 1997; Ellen et al, 2002; Supple, Aquilino, and Wright, 1999; Turner, Ku, et al, 1998; Turner et al, 1992; Wright, Aquilino, and Supple, 1998). The addition of audio-enhancement has also been shown to contribute to increased data quality, yet few comparative studies have examined the effects of data
collection mode on data quality and student experience in schools. This research intended to extend the current knowledge not only with regard to the use of technology and data collection, but also to the understanding of how adolescents in schools respond to survey questions and the mode in which they are delivered.

Our analyses revealed a number of significant findings regarding the impact of the technology and audio enhancement on the data outcomes. First, students completing the survey on the PDA and especially the APDA were able to complete significantly more questions than those completing the SAQ. Second, the technology of the PDA and the programmed survey greatly reduced missing data among the PDA and APDA respondents compared to the SAQ respondents. Third, the consistency of responses to at least three unique constructs appeared to vary by mode (i.e., factor variant), suggesting the method in which data is collected does in fact influence the way adolescents answer questions. Fourth, mode did not appear to influence the reporting of non-sensitive behaviors such as fast food consumption, but had a marginal impact on the reporting of sexual intercourse. Finally, student evaluation of the data collection experience (i.e. overall honesty, discomfort, and enjoyment) did not vary by mode. These findings are each discussed in greater detail below.

First, students completing the survey on the APDA answered significantly more questions than both PDA and SAQ respondents. These results require a small caveat with regard to how the data were collected. To assess the impact of the technology and audio-enhancement on the impact of survey administration, we had two options: allow students to work through the full survey and record completion time relative to start time; or limit the time allowed and count the number of questions answered. In order to be
respectful of the disruption we were already providing in the school day, we chose the latter based on past experiences where we’ve found that students have a wide range in the amount of time needed to complete surveys. Results indicated that the number of questions answered by students during the 25-minute survey period was impacted by addition of audio, but not by the PDA technology per se as the number of questions answered by the SAQ respondents was comparable to that of PDA respondents. Relieving the burden of reading survey questions allowed students to move through the survey more quickly, with APDA respondents answering on average over one question per minute more than PDA and SAQ respondents.

Next, the ability to program the survey via Surveyor and execute it on the PDA greatly reduced the prevalence of missing data. This finding was consistent with findings in other studies implementing computerized survey technology (Beebe et al, 1997; Couper, Singer, and Tourangeau, 2003; Couper, et al, 1998; Hallfors et al, 2000; Ramos, Sedivi and Sweet 1998), with both PDA and APDA respondents having significantly less missing data than SAQ respondents. Issues of failed skip patterns and logically inconsistent data were also moot for the PDA and APDA groups due to the programmable skip patterns inherent in the software (Trapl, 2004; Trapl et al, 2005). Among the SAQ respondents, up to one third of students failed to follow a skip pattern. More interestingly, logically inconsistent data among responses to simple skip patterns was found to be less problematic than failed skip patterns with a maximum of 9% of students who should have followed the skip (5.6% of the entire SAQ sample) providing a logically inconsistent response. Thus, a majority of students who failed to follow a skip pattern still provided logically consistent responses to subsequent questions when a
logically consistent option was available. While technology can solve the resulting symptom of failed skip patterns and logically inconsistent data, we still have very little understanding as to what contributes to a failed skip pattern or a logically inconsistent response given by a respondent.

In order to assess data consistency across mode, we next examined the factor structures of five different constructs (i.e. reading enjoyment, reading difficulty, sex beliefs, consequences of early sexual behavior, and parental sex views) within each of the three modes. Results from the tests of factorial invariance revealed that the measurement models for three of the five constructs were not consistent across the three groups. However, due to a lack of post-hoc probing procedures with this technique (i.e. factorial invariance tests), our results cannot indicate that one mode provided “better” data than another. Nonetheless, researchers using multiple modes of data collection within a single study, particularly in longitudinal studies, should carefully examine factor structure of latent constructs to ensure that measurement models are similar across modes, and thus, not contributing additional error to the measurement.

Next, this study supported the findings of the school-based studies of Beebe and colleagues (1998) who found increased reporting of sensitive behaviors by SAQ and Hallfors and colleagues (2000), who found no difference in reporting of drug and alcohol behaviors by mode, as compared to a large body of research indicating increased reports of sensitive behaviors among CASI or ACASI respondents. Reporting of sexual intercourse was the only behavior found to vary significantly by mode of data collection. Contrary to what was initially hypothesized based on CASI and ACASI literature, reporting of sexual intercourse was significantly higher among SAQ respondents when
compared to both PDA and APDA respondents. In fact, the reporting of sexual intercourse among SAQ respondents was almost twice as much as that reported by APDA respondents (30% vs. 15% respectively). In an attempt to validate at least one of these values, prevalence rates of sexual intercourse among this population were sought from two alternate sources. The SAQ prevalence rate of sexual intercourse was similar to the rate reported for 7th grade students (31.7%) from the 2005 administration of the Youth Risk Behavior Survey in Cleveland Municipal School District (Case Western Reserve University, 2005). Similarly, the APDA prevalence rate of sexual intercourse was similar to the rate reported for two cohorts of 7th grade students (n=1331) participating in the Healthy Teens Building Healthy Schools baseline survey in Cleveland Municipal School District in 2004 and 2005 (Trapl et al, 2006). Thus, both the SAQ sexual intercourse prevalence and the APDA sexual intercourse prevalence were supported by local data.

The findings related to the unexpected reporting of sexual intercourse by mode led the authors to think more about sensitive behaviors generally, and ask the question, what is a sensitive behavior among this particular population? Illegal behaviors typically thought to be considered sensitive, such as marijuana use or showed no impact of mode. Perhaps these findings are simply an indication of a change in the trend of what is sensitive, or more, what students consider to be “cool” or desirable among their peers, an illustration of social norms at work. Students’ responses on SAQ were much more visible to the other students seated at their table. This was supported in the data by PDA and APDA respondents’ reporting lower scores of survey visibility (i.e. do you think the people sitting next to you could see your answers?) compared to the SAQ respondents. Further, because a “no” response to sexual experience instructed the respondent to skip
ahead in the survey, student sexual experience may have been even more obvious to students at the table. Thus, to maintain an appearance of “cool”, it is possible that some students among the SAQ group may have indicated sexual experience when, in fact, this was untrue.

As has been suggested in other research (Nass et al, 1997), there was also the possibility that the voice used for the audio-enhancement contributed to the “humanization” of the APDA approach, thus contributing to the feeling of an interviewer-administered survey and subsequent social desirability, and a subsequent downward reporting trend. However, the fact that the PDA report of sexual experience was not significantly different from the APDA report, and in fact, was much more similar to the reporting of the APDA respondents than the SAQ respondents, indicates that this was not likely.

Finally, there were no mode differences among student evaluation of the survey experience. While this was the finding for a one-time, 25-minute survey, it would be interesting to see if students might have a different opinion if they were asked on the second, third, or even fourth time completing a similar survey, as often is the case with longitudinal studies. Further, it is possible that the twenty-five minute survey attenuated some findings related to student experience that may have become more apparent during a much longer survey. Also, by only exposing students to only one mode of data collection, they may not have had the exposure and subsequent understanding of other modes to inform their mode preference.

2 Impact of Individual Characteristics on Data Outcomes
The addition of audio-enhancement to computerized survey approaches has allowed for the inclusion of survey research participants who have limited literacy levels or English language mastery perhaps by relieving the burden of reading among those who reported high levels of reading difficulty. While the impact of adult literacy on data outcomes has been examined, the literacy literature is much sparser for adolescents as well as studies examining the effects of other types of cognitive burden (e.g., low reading enjoyment, attentiveness) on data outcomes.

We were pleased to find that we were able to enroll students with a wide range of cognitive and language abilities. Several students were second-language speakers of English, and, as indicated by the lower mean sentence comprehension stanine, the overall study sample exhibited lower reading comprehension than the standardization population. Further, there was variability in the sample’s reading enjoyment, reading difficulty, and inattentiveness.

Many individual characteristics were found to be related to students answering fewer questions on the survey, including lower sentence comprehension, lower listening comprehension, higher self-assessed reading difficulty, and higher teacher-rated inattentiveness. While this finding is highly intuitive and expected, this study, when published, will be the first documenting the impact of these respondent characteristics on a measure of survey administration, specifically the number of survey questions answered. Of the eight characteristics examined, inattentiveness was the only respondent characteristic related to missing data. This finding is also somewhat intuitive in that we might expect students who have difficulty focusing on a task would also find it difficult to focus on a survey, especially when the survey requires attention to detail and
instructions on how to proceed. It is also reassuring to find an absence of relationship between other cognitive burden measures and missing data. For instance, while students with lower sentence comprehension answer fewer questions than those with higher sentence comprehension, they are no more likely to have missing data; thus, it is likely that these students spent more time carefully reading through the survey to ensure accuracy in their responses, thus reducing the number of questions they could answer in the limited time frame.

Interestingly, students’ perceptions of confidentiality had no relationship to the number of questions answered, yet there was a relationship between students’ perceptions of privacy (i.e. of other students seeing their survey responses) and the number of questions answered. This would seem to indicate that SAQ students may have been concerned with keeping their responses covered or even may have been distracted by the responses of other students to the point where this distraction limited students’ ability to stay on task.

In examination of factorial invariance for five constructs (i.e. reading enjoyment, reading difficulty, sex beliefs, consequences of early sexual behavior, parental sex beliefs), all five constructs were found to be invariant by listening comprehension. That is, students with poorer listening comprehension skills answered questions related to these five construct similarly to those with higher listening comprehension skills. In contrast, when the constructs were examined by sentence comprehension, the measurement model assessing sex beliefs yielded a significant chi-square test at p=.049. However, as discussed in the Chapter 7, Section 5.1.3, the structure was determined to be invariant across sentence comprehension group based on other fit statistics. Thus, as with
listening comprehension, students with poorer sentence comprehension skills answered questions related to these five constructs similarly to students with higher listening comprehension skills.

While the results were fairly consistent with regard to sentence and listening comprehension, there were some inconsistencies found in the measurement models when groups were examined by English mastery as measured by the Language/US index. The measurement models for both sex beliefs and parental sex beliefs were found to vary by Language/US index group (no indicators compared to at least one indicator). Unfortunately, a limit to this relatively new statistical technique (e.g., testing factorial invariance) is that it does not indicate where the consistency lies, just that it exists. Thus, we do not know which group has “better” data. Three possible interpretations are offered. The first two involve the qualities of the constructs and questions used to tap the latent constructs. For example, the questions contained within the sex beliefs and parental sex beliefs scales ask students to identify beliefs around sexual behavior, for example “I believe it is okay for people my age to have sex with a steady boyfriend or girlfriend” (from sex beliefs scale), or “Do you think your parents would approve or disapprove of teens your age having sex?”. First, it is possible that these questions may not be as easy to comprehend and readily identify a response among non-native English speakers and US-newcomers as questions about reading behavior or even the potential negative consequences of early sexual activity, thus, the questions are answered differently. Alternately, it is possible that students without any indicators (i.e. US born or US-dwelling for six+ years, speak English with family and friends at least half of the time) who have been raised in the current culture are responding differently to questions
within these scales due to the inconsistent sexual messages that pervade society today. Each of the groups may have different experiences from which to draw to answer these questions. A third interpretation may be that these groups may have inherent differences that lead to both groups answering the questions equally “well”, but just answering differently. Factorial invariance does not necessarily inform us as to “bad” data, it is simply a method to alert researchers to the varying underlying structures which represent how individuals answer questions.

In assessing student evaluation across individual characteristics, it does appear that individual characteristics have an impact on student evaluation of the data collection experience. Nearly all of the indicators of cognitive burden (e.g. sentence comprehension, Language/US index, reading difficulty, and inattentiveness) were associated with decreased overall honesty, while inattentiveness was associated with increased reported discomfort with the data collection process and low reading enjoyment was associated with decreased enjoyment completing the survey. Relationships between increased discomfort and inattentiveness and between reading enjoyment and enjoyment completing the survey are somewhat intuitive, yet still important contributions to the literature in their confirmation. Similarly important is the absence of associations with the other cognitive burden measures and discomfort. Thus, it is reassuring that while a student may have lower reading comprehension, the student’s perception of discomfort is not significantly different from students with higher reading comprehension.

However, the lower reporting of overall honesty among students with lower sentence comprehension, higher Language/US index, higher reading difficulty, and
higher inattentiveness is alarming. With the exception of students with higher inattentiveness, these students did not report differences in discomfort or enjoyment completing the survey, so there may have been other factors impacting their self-reported honesty. Conversely, we could also interpret this finding as students with lower cognitive burden reporting higher levels of overall honesty. Perhaps the cognitive burden variables are confounded by other, unobserved variables that are driving the association between overall honesty and the cognitive burden variables. Also important to note was the fact that these questions were collected via the Debriefing Survey, which was only designed as a paper-based questionnaire.

Similarly, lower perceptions of confidentiality were related to decreased overall honesty, increased overall discomfort, and decreased enjoyment completing the survey. This finding is also intuitive and in line with the body of literature discussing the importance of confidentiality to adolescent study participants and study participants in general. During the data collection sessions of this study, students would often ask if their parents or their teachers would know how they answered questions, and students seemed reassured when survey administration staff explained the mechanisms put in place to ensure the confidentiality of their data. Perceptions of response privacy were not related to students’ evaluation of the data collection experience. It is interesting to note the dichotomy between concern for confidentiality and concern for survey environment privacy. In assessing concern for confidentiality, students were responding to the idea of others knowing how they answered questions, while survey environment privacy assessed the extent to which students believed that other students could see their survey responses. This finding supports the idea that students may self-present to their peers in a
fashion that would not necessarily be as readily put forth to a parent, teacher, or unknown other.

3. Modifying Impact of Cognitive Burden on Relationship Between Mode and Outcomes

Based on the significant direct effects of mode and individuals characteristics on data outcomes, there were only four potential modifying relationships examined: (1) modifying effect of sentence comprehension on mode and the number of questions answered; (2) modifying effect of listening comprehension on mode and the number of questions answered; (3) modifying effect of self-reported reading difficulty on mode and the number of questions answered; (4) modifying effect of inattentiveness on mode and missing data. For the most part, the individual characteristics had little impact on the relationship between mode and outcomes. That is, most relationships between mode and outcomes did not vary based on different values of individual characteristics. The exception to this was the modifying effect of reading difficulty on the number of questions answered. Results of post-hoc probing indicated that students with lower perceived reading difficulty (as measured by one standard deviation below the mean score) answered more questions using the PDA when compared to those completing the survey on SAQ, similar to the overall trend by mode. In contrast, students with higher reading difficulty (as measured by one standard deviation above the mean score) answered fewer questions using the PDA when compared to students completing the survey using SAQ. Thus, students who do not necessarily acknowledge much reading burden still benefit from the PDA technology, but more importantly, it appears that
students with self-assessed reading difficulty are hindered in the survey process by the PDA technology. Moreover, this finding implies that when considering a population with self-assessed reading difficulty, if researchers do not have the resources to employ audio-enhancement of the PDA, use of the PDA alone may not be a significant benefit over SAQ data collection.

4. Consideration of Perceived Confidentiality and Privacy as Moderating or Mediating Effects

Aim 4 of the study explored the role of confidentiality and privacy in relationship between mode and data outcomes. As described in Chapter 2, we could provide a rationale for exploring these variables as both moderators and mediators. As discussed in our methods chapter, pre-existing direct effects were first examined to determine which relationships were possible (i.e., both mode and the variable of interest must be associated with the outcome for moderating effects to be examined and mode-outcome, mode-confidentiality/privacy and confidentiality/privacy-outcomes must be first established for examining mediation).

As a result, only moderating effects associated with perceived privacy on the mode to number of questions answered and missing data were examined. Mode and perceived confidentiality were not significantly associated with similar outcomes. Interestingly, upon adding the perceived privacy variable to multivariate model along with mode, the association between perceived privacy and both outcomes was greatly diminished to the point of (removing) statistical significance, indicating a correlation between mode and perceived privacy. Regardless, these findings indicated that the
varying levels of perceptions of privacy did not impact the relationship between mode and outcomes, specifically the number of questions asked and missing data.

Perceived privacy was next assessed for mediating effects for the same relationships. That is, we examined whether mode impacted perceived privacy, which in turn influenced data outcomes. As expected from the analyses described above, Pearson’s correlation coefficient indicated a significant relationship between mode and perceived privacy of 0.226 (p<.001). However, as indicated from the multivariate analysis mentioned above, the association between perceived privacy and outcomes was diminished in the presence of mode. Thus, perceived privacy does not mediate the relationship between mode and outcomes, and in fact, mode may actually confound the relationship between perceived privacy and outcomes.

These combined findings could be related to poor measurement of privacy or a lack of power, both described more fully below. However, perhaps our modeling of perceived privacy as the mediator is incorrect, and other variables not included in the model are impacting these relationships. More research in this area is needed to more fully understand the roles of perceived confidentiality and perceived privacy of survey responses on adolescent, school-based research.

5. Implications

This research provides strong support for the use of audio-enhanced personal digital assistants for school-based data collection. The findings provide clear indication that the audio-enhancement greatly reduced the burden of reading each of the survey questions, allowing students to move through the survey more quickly, in addition to
reducing missing data. Further, data collected via APDA appear to have greater internal consistency based on Cronbach’s alpha for four of five unique latent constructs. Beyond the support provided by the data-driven outcomes, the ease in which APDA data collection systems can be implemented in the school setting should not be overlooked. Moreover, findings of varying factor structures across mode indicate that researchers should be cognizant of the impact of multi-mode data collection strategies, especially when designing longitudinal studies that intend to use a variety of modes to engage a respondent (i.e., using paper/pencil as a backup for computerized data collection). Our data would suggest that similar modes may be acceptable (i.e., collection baseline with PDA and APDA and then for respondents who have moved away, collection of follow up with web-based methods), but different methods (paper vs. computerized) could introduce unnecessary measurement error into data analyses.

In addition to understanding the impact of mode, a goal of this study was to better understand the role individual student characteristics play on survey data quality and experience. Having found through past research that students with a range of cognitive and language abilities could be included in the data collection process, this study found that while these students may take more time to complete the survey, data quality as measured by prevalence of missing data was not impacted by student sentence comprehension, listening comprehension, reading enjoyment nor reading difficulty. Further, factor structure for five unique constructs was found to be invariant across sentence comprehension, listening comprehension and attentiveness groups, indicating that students of varying reading/listening abilities and attentiveness levels responded similarly to these questions. This finding is a significant contribution to a body of
literature that does not directly address issues of survey item comprehension by adolescents with a range of literacy abilities in survey research. However, there were some differences found among two constructs by our measure of English mastery (i.e. Language/US index), indicating that researchers may need to carefully examine data collected among less acculturated students.

Moreover, lack of effect modification among almost all individual characteristics on the relationships between mode and data outcomes provide further evidence that students of all abilities should be encouraged to participate. More importantly, given the moderating effect of reading difficulty on the relationship between mode and the number of questions answered, it is clear that among populations with potentially high levels of self-assessed reading burden, if researchers do not have the resources to incorporate audio-enhancement with a PDA-based survey, use of PDA alone may not be a significant improvement over the use of SAQ.

6. Limitations

There are several limitations to the current study. First, the participating students self-selected into the study and were likely very different than those students who chose not to participate or those with consent but not attending school on the day of the data collection sessions. Students not attending school, typically those with lower academic achievement, were also less likely to be engaged in the study. Thus, we do not know the abilities of the students who declined participation in the study, nor do we know why these students did not want to participate. For instance, it is plausible to believe that students with increased cognitive burden would not find the opportunity to take a survey
to be interesting or beneficial. While students did receive a t-shirt as a token of appreciation for their participation and were removed from class (unfortunately seen as a benefit to many students), this was not incentive enough for those students who were completely disinterested.

Furthermore, while the students recruited into the study were diverse with regard to cognitive burden, special education students and students speaking English as a second language are special populations who may not have been adequately represented in the current study due to the very same reasons identified by McGrew (1993). If students were not on a 7th grade roster, we would not have known to recruit the youth into the study. Further, if a student’s teacher indicated that the student could not participate in the survey for physical, mental, or cognitive purposes, we did not try to recruit the youth into the study. Due to the K-8 structure of this school district, smaller schools often do not have the resources to accommodate a stand-alone special education or ESL class, and it was unclear if these students were simply included in mainstream classrooms or reassigned to larger schools with the capacity to accommodate students with special needs.

Second, students participating in the study completed a survey using only one mode of data collection, so inter-mode correspondence of respondent reports could not be calculated. This approach could have provided additional insight into some of the discrepancies found across mode, such as reporting of sexual experience or lack of factorial invariance. Also, by not exposing students to other modes of data collection aside from the mode to which they were randomized, the student’s ability to make a comparison judgment on future mode preference may have been reduced. Additionally,
our comparison did not include CASI or ACASI survey administration, allowing us to draw conclusions solely from the comparisons of SAQ, PDA, and APDA.

Third, because youth are not typically the focus of mode comparison studies, reliable and validated tools for collecting “debriefing” information regarding the student’s experience while taking the survey do not exist. For example, the word “private” may have many different meanings to a young population with limited literacy. Anecdotally, in conversations with students immediately following the data collection, some students indicated to data collection staff that they thought the survey was very “private” and asked too many personal questions, while others indicated that they didn’t think the survey was very “private” because other students sitting at their table could see their responses. Further, while all attempts were made to ensure that survey and debriefing questions were grade reading-level appropriate, data collection administrators were consistently asked what certain words meant, such as embarrassing or sympathetic.

Fourth, several variables important in predicting data outcomes may have been missing from the models. While we were most interested in variables related to reading abilities and attention, there may have been other measures of cognitive burden that were overlooked by the authors that could have enhanced the current study. Additionally, asking other questions may have helped explain some findings. For example, student responses to a question assessing student perceptions of social norms for sexual intercourse might have contributed to a better understanding of the differential reporting of sexual intercourse by mode.

Moreover, some of the most interesting patterns of response and student engagement in survey data collection would seem to benefit from some level of face-to-
face interviewing. When examining the data, it was not uncommon to discuss with other colleagues what may have been going through an adolescent respondent’s head when responding to certain questions. However, the answers to these questions cannot be obtained through close-ended responses to simple questions.

7. Future Research

Future studies should be designed to implement the APDA with older adolescents and younger adolescents to further explore feasibility and superiority, inferiority, or equivalence of APDA when compared to alternative data collection modes. Similarly, it would be helpful to better understand the limitations of the APDA system, for example, to know if the APDA system could be effectively implemented for larger data collection sessions, such as data collection sessions sometimes held in cafeterias or gymnasiums in order to reduce classroom interruption.

Given our findings related to the higher reporting of sexual intercourse among SAQ respondents compared to APDA respondents in addition to our findings related to relationships between perceived confidentiality and outcomes and (lack of) relationships between perceived privacy of survey responses and outcomes, there is clearly a need to better understand the cognitive processes experienced by an adolescent engaged in the survey process. Sexual intercourse and other sex-related measures are often the primary outcomes for adolescent prevention programs aimed at reducing sexual initiation or risky sexual behavior (e.g. not using condoms). If student responses to these types of questions can be so seemingly easily swayed by such factors as who is sitting next to them in the survey environment and what data collection mode is used to respond to the questions,
then researchers in the area of adolescent sexual behavior should be very concerned about
the validity of their data. While much of the population is hesitant to engage young
adolescents in conversations around sexual behavior, most would agree that it is
appropriate to engage youth who are developmentally ready to engage in such
conversations (e.g. those already having sexual intercourse), and researchers must find
ways to ethically identify these youth and engage them in processes that draw from them
the social and psychological components driving responses to such sensitive questions.

Finally, future research should further explore the factorial invariance (or lack of)
among latent constructs important in behavioral change and evaluation of behavioral
interventions when collected across different modes and in populations with varying
levels of English mastery. Latent constructs are often the only way to measure beliefs and
attitudes of individuals targeted by behavioral intervention programs. Thus, if these
constructs are not answered similarly across the study sample, then researchers could be
introducing error into their models, potentially leading to incorrect results. The results of
this study indicate that factor structures may not be consistent across mode nor across
levels of English mastery. Studies implementing a test-retest design utilizing multiple
mode pairs may inform this area of research. Health behavior research, in addition to
other related fields, would greatly benefit from better understanding the impact of data
collection mode, individual characteristics, and the impact on the measurement models of
latent constructs in order to build a stronger, more well-developed and understood cadre
of reliable measures.
8. Conclusions

Self-report questionnaires are oftentimes the mechanism by which data is collected to assess the efficacy and effectiveness of interventions designed to minimize high-risk behavior and prevent chronic diseases. Thus, researchers must acknowledge the important role of data collection and data quality on informing these programs.

The findings of this study indicate strong benefits to be gained by the use of audio-enhanced personal digital assistants for school-based data collection with adolescents. Further, while students higher levels of cognitive burden yielded data of somewhat lower quality (e.g. fewer questions answered, more missing data), there was observed consistency of the students’ responses across a variety of constructs, regardless of their level of cognitive burden. This provides strong evidence that if audio-supported data collection methods are used, researchers can and should include students with a wide range of cognitive abilities (reading levels and preferences, language mastery, attention issues) in school-based intervention research.
APPENDIX

Data Collection Tools:

(1) Teen Health Survey

(2) Debriefing Survey

(3) Teacher Rating of Student Academic & Behavior Performance
Instructions

Welcome to the Teen Health Survey Study. With your help, we hope to learn more about how students like you answer questions on different types of surveys. In this survey, we’re going to ask you about what you know, what you believe, and what you do.

Please DO NOT write your name on this survey. Remember, your answers are completely confidential. The ONLY way anyone will know what you say is if YOU tell them. We appreciate your honesty!

When you are ready to begin, please turn to the next page and begin answering the questions.

Please use a pencil or a black/blue ink pen to complete the survey. Be sure to make heavy, dark marks that fill the circle completely.

Example:  ○ ○ ○ ○

FOR OFFICE USE ONLY

V1  □ □ □ □
1. How old are you?
   - 12
   - 13
   - 14
   - 15

2. Are you male or female?
   - Male
   - Female

3. What is your race or ethnicity?
   - White or Caucasian
   - Black or African-American
   - Hispanic
   - Asian
   - Other, please specify

4. Please fill the circle next to all the people who live in your home.
   - Mother or step-mother
   - Father or step-father
   - Brothers or sisters
   - Grandparent
   - Aunts, uncles, or cousins
   - Other adults I am not related to

5. What kind of grades do you usually get in school?
   - Mostly A's
   - Mostly C's
   - Mostly A's and B's
   - Mostly C's and D's
   - Mostly B's
   - Mostly D's
   - Mostly B's and C's
   - Mostly D's and F's

6. Are your parents married?
   - Yes, they're married to each other
   - Yes, they're both married, but to different people
   - Only my mom is married
   - Only my dad is married
   - Neither of my parents are married
   - I don't know

7. How important is religion in your life?
   - Not at all important
   - Somewhat important
   - Important
   - Very important

8. What languages do you usually speak with your friends?
   - Only English → GO TO #10
   - Mostly English → GO TO #10
   - About half English and half my other language
   - Mostly my other language
   - Only my other language

9. Which language do you speak with your friends?
   - Spanish
   - Chinese
   - Russian
   - Arabic
   - Other, please specify

10. What languages do you usually speak with your family?
    - Only English → GO TO #12
    - Mostly English → GO TO #12
    - About half English and half my other language
    - Mostly my other language
    - Only my other language
11. What languages are spoken the most in your home?
   ○ Spanish
   ○ Chinese
   ○ Russian
   ○ Arabic
   ○ Other, please specify

12. How long have you lived in the United States?
   ○ All my life
   ○ Less than 1 year
   ○ 1 to 3 years
   ○ 4 to 6 years
   ○ More than 6 years, but not my whole life

13. I love to read.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

14. When I am at home I read a lot.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

15. I need a lot of help in reading.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

16. I can read but I don’t understand what I’ve read.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

17. I’m the kind of person who enjoys a good book.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

18. I spend a lot of my spare time reading.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree
19. I try very hard, but I just can't read very well.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

20. I quickly forget what I have read even if I have just read it.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

21. Reading is one of my favorite activities.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

22. I like to read a book whenever I have free time.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

23. I remember the things people tell me better than the things I read.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

24. I have trouble understanding what I read.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

25. I believe people my age should wait until they are older to have sex.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

26. I believe it is okay for people my age to have sex with a steady boyfriend or girlfriend.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

27. I believe it is okay for people my age to have sex as long as they use a condom.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree
28. I believe it is okay for people my age to have sex with a casual friend.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

29. I believe it is okay for people my age to have sex, even if they are not in love.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

30. Having sexual intercourse as a teenager makes it harder for someone to get a good job or be successful in a career.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

31. Having sexual intercourse as a teenager makes it harder for someone to study and stay in school.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

32. Having sexual intercourse as a teenager makes it harder for someone to have a good family life in the future.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

33. Having sexual intercourse as a teenager makes it harder for someone to develop a loving relationship in the future.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

34. When used properly, a condom prevents girls from getting pregnant.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree

35. When used properly, a condom prevents HIV.
   ○ Strongly agree
   ○ Agree
   ○ In the middle
   ○ Disagree
   ○ Strongly disagree
36. When used properly, a condom prevents sexually transmitted diseases such as Chlamydia and herpes.

- Strongly agree
- Agree
- In the middle
- Disagree
- Strongly disagree

37. A man can't have a good sex life if he's not in love with his sexual partner.

- Strongly agree
- Agree
- In the middle
- Disagree
- Strongly disagree

38. A woman can't have a good sex life if she's not in love with her sexual partner.

- Strongly agree
- Agree
- In the middle
- Disagree
- Strongly disagree

39. Sexual intercourse is best if you just do it because it feels good, not because you are in love.

- Strongly agree
- Agree
- In the middle
- Disagree
- Strongly disagree

40. Sexual intercourse is better—more enjoyable, intense and satisfying—when two people are married to each other.

- Strongly agree
- Agree
- In the middle
- Disagree
- Strongly disagree

41. Having sex with many different people who are just friends can be just as satisfying as having sex with one person you love.

- Strongly agree
- Agree
- In the middle
- Disagree
- Strongly disagree

42. In the past two weeks, how many times did you eat at a fast food restaurant?

- None
- Once
- 2-3 times
- 4-6 times
- 7 times or more

43. In the past seven days, how many times did you eat breakfast?

- 0 days
- 1 day
- 2 days
- 3 days
- 4 days
- 5 days
- 6 days
- 7 days
44. Please describe your experience with smoking. This includes cigarettes and little cigars, such as Black & Milds.
   ○ I have never tried smoking
   ○ I have only experimented with smoking
   ○ I smoke occasionally
   ○ I smoke a few times a week
   ○ I smoke everyday
   ○ I used to smoke regularly, but I quit

45. Have you ever had a drink of alcohol, even just a sip?
   ○ No → GO TO #48
   ○ Yes

46. Have you ever gotten drunk?
   ○ No ○ Yes

47. Have you had a drink of alcohol in the past month?
   ○ No ○ Yes

48. Have you ever smoked marijuana?
   ○ No → GO TO #50
   ○ Yes

49. Have you smoked marijuana in the last month?
   ○ No ○ Yes

50. Have you ever cheated on a test in school?
   ○ No ○ Yes

51. Have you ever copied from someone else's homework?
   ○ No ○ Yes

52. Have you ever shoplifted, or stolen something from a store?
   ○ No ○ Yes

53. Have you ever beaten up another person?
   ○ No ○ Yes

54. Have you ever kissed someone on the lips, other than a parent or relative?
   ○ No → GO TO #56
   ○ Yes

55. Now, how about in the past three months. Have you kissed someone on the lips, other than a parent or relative?
   ○ No ○ Yes

56. Have you ever French kissed, put your tongue in someone's mouth while kissing?
   ○ No → GO TO #58
   ○ Yes

57. Now, how about in the past three months. Have you French kissed?
   ○ No ○ Yes
58. Have you ever touched a girl’s breasts or, if you are a girl, has anyone ever touched your breasts?
   ○ No → GO TO #60
   ○ Yes

59. How about in the past 3 months. Have you ever touched a girl’s breasts or, if you are a girl, has anyone ever touched your breasts?
   ○ No   ○ Yes

60. Have you ever touched someone else’s private parts below the waist under their clothes?
   ○ No → GO TO #62
   ○ Yes

61. Now how about in the past 3 months. Have you ever touched someone else’s private parts below the waist under their clothes?
   ○ No   ○ Yes

62. Has someone touched your private parts below the waist under your clothes?
   ○ No → GO TO #64
   ○ Yes

63. Now how about in the past 3 months. Has someone touched your private parts below the waist under your clothes?
   ○ No   ○ Yes

64. Have you ever had sexual intercourse?
   ○ No → GO TO #70
   ○ Yes

65. How old were you the first time you had sexual intercourse?
   ○ 9 years old or younger
   ○ 10 years old
   ○ 11 years old
   ○ 12 years old
   ○ 13 years old
   ○ 14 years old

66. How many people have you had sexual intercourse with in your life?
   ○ One
   ○ 2
   ○ 3
   ○ 4
   ○ 5 or more

67. In your lifetime, who have you had sex with?
   ○ Boys only
   ○ Girls only
   ○ Boys and girls

68. The last time you had sex, did you or your partner use a condom?
   ○ No   ○ Yes

69. In the past three months, have you had sexual intercourse?
   ○ No   ○ Yes
70. Do you think your parents would approve or disapprove of teens your age having sex?
   - Strongly disapprove
   - Disapprove
   - Agree
   - Strongly approve
   - I don’t know what my parents think

71. Do you think your parents would approve or disapprove of *you* having sex right now?
   - Strongly disapprove
   - Disapprove
   - Agree
   - Strongly approve
   - I don’t know what my parents think

72. Do you think your parents want you to wait until you are older to have sex?
   - Definitely want me to wait
   - Probably want me to wait
   - Sort of want me to wait
   - Don’t want me to wait
   - I don’t know what my parents think

73. Do you think your parents want you to wait until you are married to have sex?
   - Definitely want me to wait
   - Probably want me to wait
   - Sort of want me to wait
   - Don’t want me to wait
   - I don’t know what my parents think

74. If I had a question about sexual intercourse, I know I could talk to my parents about it.
   - Strongly agree
   - Agree
   - In the middle
   - Disagree
   - Strongly disagree

75. Have your parents told you that they do not want you to smoke cigarettes?
   - No
   - Yes

76. Have your parents told you that they do not want you to drink alcohol?
   - No
   - Yes

77. Have your parents told you that you cannot date until you are older?
   - No
   - Yes

78. Have your parents told you that they expect you to wait until you are older to have sexual intercourse?
   - No
   - Yes

79. Have your parents told you that they expect you to wait until you are married to have sexual intercourse?
   - No
   - Yes
How often have your parents talked to you about:

80. Advice in your personal problems, such as problems with your friends.
   - Often
   - Sometimes
   - Hardly ever
   - Never

81. What your parents think about teenagers having sex
   - Often
   - Sometimes
   - Hardly ever
   - Never

82. Getting HIV/AIDS
   - Often
   - Sometimes
   - Hardly ever
   - Never

83. Sexually transmitted diseases.
   - Often
   - Sometimes
   - Hardly ever
   - Never

84. How to prevent pregnancy.
   - Often
   - Sometimes
   - Hardly ever
   - Never

85. When you should have sexual intercourse
   - Often
   - Sometimes
   - Hardly ever
   - Never

86. Using a condom during sex
   - Often
   - Sometimes
   - Hardly ever
   - Never

87. Sexual development of boys
   - Often
   - Sometimes
   - Hardly ever
   - Never

88. Sexual development of girls
   - Often
   - Sometimes
   - Hardly ever
   - Never
89. I take my schoolwork seriously.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

90. I like school because I learn new thing all the time.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

91. I feel a sense of accomplishment when I learn something new.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

92. I take part in classroom discussion often.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

93. Before a quiz or exam, I plan out how I will study.
   ○ Never
   ○ Rarely
   ○ Sometimes
   ○ Usually
   ○ Always

94. When I finish doing practice problems or homework, I check my work for errors.
   ○ Never
   ○ Rarely
   ○ Sometimes
   ○ Usually
   ○ Always

95. If I have trouble understanding something I go over it again until I understand it.
   ○ Never
   ○ Rarely
   ○ Sometimes
   ○ Usually
   ○ Always

96. I try to plan an approach in my mind before I actually start homework or studying.
   ○ Never
   ○ Rarely
   ○ Sometimes
   ○ Usually
   ○ Always
97. When learning new information I try to put the idea in my own words.
   ○ Never
   ○ Rarely
   ○ Sometimes
   ○ Usually
   ○ Always

98. When doing an assignment I make sure I know what I am being asked to do before I begin.
   ○ Never
   ○ Rarely
   ○ Sometimes
   ○ Usually
   ○ Always

99. When I study I'm usually aware of what I understand and what I don't understand.
   ○ Never
   ○ Rarely
   ○ Sometimes
   ○ Usually
   ○ Always

100. I answer practice problems to check my understanding.
    ○ Never
    ○ Rarely
    ○ Sometimes
    ○ Usually
    ○ Always

101. I make sure I understand the ideas that I study.
    ○ Never
    ○ Rarely
    ○ Sometimes
    ○ Usually
    ○ Always

102. Education is the key to success in the future.
    ○ Strongly disagree
    ○ Disagree
    ○ In the middle
    ○ Agree
    ○ Strongly agree

103. If everyone in America gets a good education, we can end poverty.
    ○ Strongly disagree
    ○ Disagree
    ○ In the middle
    ○ Agree
    ○ Strongly agree

104. Doing well in school now leads to getting a good job later.
    ○ Strongly disagree
    ○ Disagree
    ○ In the middle
    ○ Agree
    ○ Strongly agree

105. Young people like me have a real chance of making it if we do well in school.
    ○ Strongly disagree
    ○ Disagree
    ○ In the middle
    ○ Agree
    ○ Strongly agree
106. All I need to learn for my future is to read, write, and how to make change.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

107. Studying in school rarely pays off later with good jobs.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

108. Based on their experiences, my parents say people like us are not always paid or promoted according to our education.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

109. Although my parents tell me to get a good education in order to get a good job, they have problems achieving success in their jobs.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

110. People in my family haven't been treated fairly at work no matter how much education they have.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

111. I don't mind speaking in front of a group of people.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

112. I feel comfortable raising my hand in class.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

113. I feel comfortable stating my opinion even when I don't know what other people think.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree
114. I have good conversations with other people.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

115. I don’t mind writing.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

116. For me, answering essay questions is not a problem.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

117. Having to write is one of the worst parts of school.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

118. I enjoy putting my thoughts on paper.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

119. When with a group, I can get people working well together.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

120. I am a good leader when things need to get done.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

121. I enjoy working with a group on a common goal or project, particularly if I’m the leader.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree

122. I’m usually the one who steps forward when someone needs something done.
   ○ Strongly disagree
   ○ Disagree
   ○ In the middle
   ○ Agree
   ○ Strongly agree
123. As you look to the future, how important is it to you that you go to a high school that challenges you?
   ○ Not at all important
   ○ Not very important
   ○ Somewhat important
   ○ Pretty important
   ○ Extremely important

124. How important is it to you to graduate from high school?
   ○ Not at all important
   ○ Not very important
   ○ Somewhat important
   ○ Pretty important
   ○ Extremely important

125. How important is it to you to go to college?
   ○ Not at all important
   ○ Not very important
   ○ Somewhat important
   ○ Pretty important
   ○ Extremely important

126. How important is it to you to graduate from college?
   ○ Not at all important
   ○ Not very important
   ○ Somewhat important
   ○ Pretty important
   ○ Extremely important

127. How important is it to you to develop a lifelong career (like being a teacher, doctor, or lawyer) as compared to finding a job that you like a lot?
   ○ Not at all important
   ○ Not very important
   ○ Somewhat important
   ○ Pretty important
   ○ Extremely important

128. I help around the house.
   ○ Not at all like me
   ○ Somewhat like me
   ○ Mostly like me
   ○ Very much like me

129. I keep my room and belongings tidy.
   ○ Not at all like me
   ○ Somewhat like me
   ○ Mostly like me
   ○ Very much like me

130. I keep my clothes clean and tidy.
   ○ Not at all like me
   ○ Somewhat like me
   ○ Mostly like me
   ○ Very much like me
131. I shower and keep myself clean.
  - Not at all like me
  - Somewhat like me
  - Mostly like me
  - Very much like me

132. I help with cooking at home.
  - Not at all like me
  - Somewhat like me
  - Mostly like me
  - Very much like me

133. I help with the cleaning after meals.
  - Not at all like me
  - Somewhat like me
  - Mostly like me
  - Very much like me

134. I volunteer on a regular basis to help others in my community.
  - Not at all like me
  - Somewhat like me
  - Mostly like me
  - Very much like me

135. I work to make my community a better place.
  - Not at all like me
  - Somewhat like me
  - Mostly like me
  - Very much like me

136. I know where to volunteer in my community.
  - Not at all like me
  - Somewhat like me
  - Mostly like me
  - Very much like me

137. I am a person who tells others about my community.
  - Not at all like me
  - Somewhat like me
  - Mostly like me
  - Very much like me

138. Most of my friends stay out of trouble.
  - Strongly disagree
  - Disagree
  - In the middle
  - Agree
  - Strongly agree
139. Most of my friends follow the rules that their parents make for them.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

140. Most of my friends are responsible.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

141. Most of my friends do well in school.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

142. I prefer to do things that I can do well rather than things that I do poorly.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

143. I'm happiest when I can do things where I won't make any mistakes.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

144. I try hard to improve on my past performance.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

145. The opinions others have about how well I can do certain things are important to me.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

146. I like to work on tasks that I have done well in the past.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

147. The opportunity to learn new things is important to me.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree
148. The opportunity to do challenging work is important to me.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

149. When I fail to complete a task, I plan to try harder next time I work on it.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

150. I prefer to work on tasks that force me to learn new things.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

149. I feel smart when I can do something better than most people.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

150. I do my best when I’m working on a fairly difficult task.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

153. The things I enjoy the most are the things that I do the best.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

154. When I have difficulty solving a problem, I enjoy trying different approaches to see which one will work.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

155. I feel smart when I do something without making any mistakes.
   - Strongly disagree
   - Disagree
   - In the middle
   - Agree
   - Strongly agree

156. Good at schoolwork
   - Sounds not at all like me
   - Sounds a little like me
   - Sounds mostly like me
   - Sounds very much like me
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>157. Popular with kids.</strong></td>
<td><strong>162. Remembers things easily</strong></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>○ Sounds not at all like me</td>
<td>○ Sounds not at all like me</td>
</tr>
<tr>
<td>○ Sounds a little like me</td>
<td>○ Sounds a little like me</td>
</tr>
<tr>
<td>○ Sounds mostly like me</td>
<td>○ Sounds mostly like me</td>
</tr>
<tr>
<td>○ Sounds very much like me</td>
<td>○ Sounds very much like me</td>
</tr>
<tr>
<td><strong>158. Likes school and does well</strong></td>
<td><strong>163. Understands what he or she reads</strong></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>○ Sounds not at all like me</td>
<td>○ Sounds not at all like me</td>
</tr>
<tr>
<td>○ Sounds a little like me</td>
<td>○ Sounds a little like me</td>
</tr>
<tr>
<td>○ Sounds mostly like me</td>
<td>○ Sounds mostly like me</td>
</tr>
<tr>
<td>○ Sounds very much like me</td>
<td>○ Sounds very much like me</td>
</tr>
<tr>
<td><strong>159. Can figure out answers</strong></td>
<td><strong>164. Has a lot of friends</strong></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>○ Sounds not at all like me</td>
<td>○ Sounds not at all like me</td>
</tr>
<tr>
<td>○ Sounds a little like me</td>
<td>○ Sounds a little like me</td>
</tr>
<tr>
<td>○ Sounds mostly like me</td>
<td>○ Sounds mostly like me</td>
</tr>
<tr>
<td>○ Sounds very much like me</td>
<td>○ Sounds very much like me</td>
</tr>
<tr>
<td><strong>160. Important to classmates</strong></td>
<td><strong>165. Easy to like</strong></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>○ Sounds not at all like me</td>
<td>○ Sounds not at all like me</td>
</tr>
<tr>
<td>○ Sounds a little like me</td>
<td>○ Sounds a little like me</td>
</tr>
<tr>
<td>○ Sounds mostly like me</td>
<td>○ Sounds mostly like me</td>
</tr>
<tr>
<td>○ Sounds very much like me</td>
<td>○ Sounds very much like me</td>
</tr>
<tr>
<td><strong>161. Finishes schoolwork quickly</strong></td>
<td><strong>166. Does things with other kids</strong></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>○ Sounds not at all like me</td>
<td>○ Sounds not at all like me</td>
</tr>
<tr>
<td>○ Sounds a little like me</td>
<td>○ Sounds a little like me</td>
</tr>
<tr>
<td>○ Sounds mostly like me</td>
<td>○ Sounds mostly like me</td>
</tr>
<tr>
<td>○ Sounds very much like me</td>
<td>○ Sounds very much like me</td>
</tr>
</tbody>
</table>
167. Just as smart as others
   - Sounds not at all like me
   - Sounds a little like me
   - Sounds mostly like me
   - Sounds very much like me

168. Easy to make friends
   - Sounds not at all like me
   - Sounds a little like me
   - Sounds mostly like me
   - Sounds very much like me

169. Most kids like him or her
   - Sounds not at all like me
   - Sounds a little like me
   - Sounds mostly like me
   - Sounds very much like me

170. I think about the past a lot.
   - Not at all like me
   - Somewhat like me
   - Mostly like me
   - Very much like me

171. Generally, I am more focused on what is going on now than on what will happen in the future.
   - Not at all like me
   - Somewhat like me
   - Mostly like me
   - Very much like me

172. I think a lot about what life was like when I was younger.
   - Not at all like me
   - Somewhat like me
   - Mostly like me
   - Very much like me

173. I take care of what needs done before having fun.
   - Not at all like me
   - Somewhat like me
   - Mostly like me
   - Very much like me

174. The joy in my life comes from what I am doing now, not from what I am going to do later.
   - Not at all like me
   - Somewhat like me
   - Mostly like me
   - Very much like me

175. I am able to resist temptation when there is work to be done.
   - Not at all like me
   - Somewhat like me
   - Mostly like me
   - Very much like me
176. I try to live one day at a time.

☐ Not at all like me
☐ Somewhat like me
☐ Mostly like me
☐ Very much like me

177. Thinking about the past makes me very emotional.

☐ Not at all like me
☐ Somewhat like me
☐ Mostly like me
☐ Very much like me

178. I get things done by working at a steady pace.

☐ Not at all like me
☐ Somewhat like me
☐ Mostly like me
☐ Very much like me

Thank you for completing the Teen Health Survey. You are now finished. Please raise your hand and a staff member will collect this survey.

Remember, all of your answers are completely confidential. None of the questions you answered will be shared with your teachers, your parents, or your friends.
Debriefing Questions

These questions concern your reactions to the survey you just completed. We are always trying to improve our studies, and it would be helpful to learn how you felt about this survey. Please place a check mark next to the best answer for each question.

A1. What did the staff say about whether answering the questions was voluntary or not?

___ The survey was completely voluntary and I didn’t have to answer any questions I didn’t want to.
___ The survey was voluntary but if I agreed to complete the survey, I had to answer all the questions.
___ The staff didn’t say anything about this.
___ I don’t remember

A2. What did the survey say about who would be able to find out the answers you gave in the survey?

___ My answers would be confidential unless I shared my answers with someone else
___ My answers would be shared with my parents and my school
___ The survey didn’t say anything about this.
___ I don’t remember.

A3. How likely do you think it is that your answers will remain confidential?

Do you think your answers are...

___ Certain to remain confidential
___ Likely to remain confidential
___ As likely as not to remain confidential
___ Unlikely to remain confidential
___ Certain NOT to remain confidential

A4. How likely do you think it is that your answers will fall into the wrong hands?

Do you think your answers are ...

___ Certain to fall into the wrong hands
___ Likely to fall into the wrong hands
___ As likely as not to fall into the wrong hands
___ Unlikely to fall into the wrong hands
___ Certain NOT to fall into the wrong hands
A5. Suppose the survey staff found out your answers to the questions in the interview. Do you think they would be sympathetic, judgmental, or couldn’t you tell how they’d react?

___ Sympathetic
___ Judgmental
___ Don’t think they would care
___ Didn’t notice
___ Couldn’t tell

A6. Overall, how honest were you when completing the survey?

___ Completely honest
___ Mostly honest
___ Somewhat honest
___ Not very honest
___ Not honest at all

A7. How honest were you when answering questions about reading?

___ Completely honest
___ Mostly honest
___ Somewhat honest
___ Not very honest
___ Not honest at all

A8. How honest were you when answering questions about drugs and alcohol?

___ Completely honest
___ Mostly honest
___ Somewhat honest
___ Not very honest
___ Not honest at all

A9. How honest were you when answering questions about sexual behavior?

___ Completely honest
___ Mostly honest
___ Somewhat honest
___ Not very honest
___ Not honest at all
B. People find different kinds of questions upsetting. How upsetting or embarrassing did you find the various topics in the interview? Please rate each topic on a ten-point scale, where a rating of 1 indicates that the topic was not upsetting at all and a rating of 10 indicates that the topic was very upsetting.

Circle the number that corresponds to your rating.

<table>
<thead>
<tr>
<th>Not upsetting at all</th>
<th>Very</th>
</tr>
</thead>
</table>

Upsetting

B1. Eating fast food? ........ 1 2 3 4 5 6 7 8 9 10
B2. Shoplifting? .......... 1 2 3 4 5 6 7 8 9 10
B3. Cheating at school? ....... 1 2 3 4 5 6 7 8 9 10
B4. Drinking alcohol? ......... 1 2 3 4 5 6 7 8 9 10
B5. Smoking marijuana? ....... 1 2 3 4 5 6 7 8 9 10
B6. Sexual behaviors? .......... 1 2 3 4 5 6 7 8 9 10

B7. Think about the topic that you found most upsetting to answer questions about. How embarrassed do you think you’d be if the survey staff found out your answers to those questions?

___ Very embarrassed
___ Embarrassed
___ Somewhat embarrassed
___ A little embarrassed
___ Not embarrassed at all

B8. Suppose your parents found out your answers to those questions. How embarrassed do you think you’d be?

___ Very embarrassed
___ Embarrassed
___ Somewhat embarrassed
___ A little embarrassed
___ Not embarrassed at all
B9. Suppose someone you didn’t know found out your answers to those questions. How embarrassed do you think you’d be?

___ Very embarrassed
___ Embarrassed
___ Somewhat embarrassed
___ A little embarrassed
___ Not embarrassed at all

B10. On the whole, how private did you feel the interview was? Would you say...

___ Completely private
___ Private
___ Somewhat private
___ Not very private
___ Not private at all

B11. Do you think the people sitting next to you could see your answers?

___ Definitely could see my answers
___ Probably could see my answers
___ Probably could not see my answer
___ Definitely could not see my answers

B12. How comfortable did you feel answering the questions? Would you say...

___ Completely comfortable
___ Comfortable
___ Somewhat comfortable
___ Not very comfortable
___ Not comfortable at all

B13. If you had the choice, would you have preferred to answer the survey?

___ In person with an interviewer
___ By PDA (handheld computer)
___ By PDA (handheld computer), with the questions read to me through earphones
___ By a written questionnaire that I could fill out myself
___ On the phone with an interviewer
___ No preference, doesn’t matter
B14. How much did you like completing the survey?

___ Liked it a lot
___ Liked it a little bit
___ Didn’t like it or dislike it
___ Didn’t like it too much
___ Didn’t like it at all

IF YOU USED A PDA TO ANSWER YOUR SURVEY, PLEASE ANSWER THE NEXT SET OF QUESTIONS.

C1. Before today, had you ever used a handheld computer, or PDA?

___ Yes
___ No

C2. Please tell us what you liked about using the PDA to complete your survey:

____________________________________________________________________
____________________________________________________________________

C3. Please tell us what you did not like about using the PDA to complete your survey:

____________________________________________________________________
____________________________________________________________________

Thanks for taking part in our study!
We appreciate your time and effort.
STUDYID:__

RATING OF STUDENT ACADEMIC PERFORMANCE: TEACHER

PART 1: STUDENT/TEACHER BACKGROUND

Your Gender: M  F (please circle)

Today’s Date: ____________

Your Role: I teach this child _______ period(s) out of the total _______ periods of the day.

I teach this child the following subjects: __________________________________________

There are _______ students in this child’s class.

1. For how many months have you known this pupil? _______ (months)

2. How well do you know this student? __ Not Well __ Moderately Well __ Very Well

3. How much time does this student spend in your class per week? ____________

4. During a typical school day, does this student speak a language other than English? Yes  No (please circle)

PART 2: STUDENT’S CURRENT ACADEMIC PERFORMANCE

Directions: Please check the box that indicates the student’s current performance for each item compared to the established academic standards for 7th grade.

<table>
<thead>
<tr>
<th>1. READING COMPREHENSION</th>
<th>Far Below Grade</th>
<th>Somewhat Below Grade</th>
<th>At Grade Level</th>
<th>Somewhat Above Grade</th>
<th>Far Above Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. accurately follows written instructions</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>b. recalls facts after reading out loud</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>c. can compare, make inferences, and/or draw conclusions about what is read</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>d. can find a main idea statement accurately</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. LISTENING COMPREHENSION</th>
<th>Far Below Grade</th>
<th>Somewhat Below Grade</th>
<th>At Grade Level</th>
<th>Somewhat Above Grade</th>
<th>Far Above Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. attends to peers when they are taking</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>b. can follow simple oral instructions</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>c. detects humor or sarcasm in the verbal expression of others</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>d. understands the context while listening to others read orally</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>
PART 3: COMPARISON OF STUDENT TO OTHERS

Directions: Please check the box that indicates the student’s performance for each item in comparison to typical students of the same age in Cleveland Municipal School District.

<table>
<thead>
<tr>
<th></th>
<th>Far Below Others</th>
<th>Somewhat Below Others</th>
<th>At Same Level as Others</th>
<th>Somewhat Above Others</th>
<th>Far Above Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. READING COMPREHENSION</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. accurately follows written instructions</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>b. recalls facts after reading out loud</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>c. can compare, make inferences, and/or draw conclusions about what is read</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>d. can find a main idea statement accurately</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>4. LISTENING COMPREHENSION</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. attends to peers when they are taking</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>b. can follow simple oral instructions</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>c. detects humor or sarcasm in the verbal expression of others</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>d. understands the context while listening to others read orally</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

Directions: Please check the box that indicates the student’s performance for each item in comparison to typical students of the same age in Cleveland Municipal School District.

<table>
<thead>
<tr>
<th></th>
<th>Far Below Others</th>
<th>Somewhat Below Others</th>
<th>At Same Level as Others</th>
<th>Somewhat Above Others</th>
<th>Far Above Others</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. ENGLISH LANGUAGE MASTERY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a. makes few grammatical errors when speaking English</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>b. easily initiates speech, asks questions and speaks spontaneously in English</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>c. almost always responds appropriately to questions/statements</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>
RATING OF STUDENT BEHAVIOR PERFORMANCE: TEACHER

Children differ in their abilities to focus attention, control activity, and inhibit impulses. For each item listed below, how does this child compare to other children of the same age? Please select the best rating based on your observations over the past month. Place an “X” under the column that best describes this student. Compared to other children, how does this child do the following:

<table>
<thead>
<tr>
<th>Compared to other children of the same age, how does this child compare?</th>
<th>Far below</th>
<th>Below</th>
<th>Slightly below</th>
<th>Average</th>
<th>Slightly above</th>
<th>Above</th>
<th>Far above</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Give close attention to detail and avoid careless mistakes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Sustain attention on tasks or play activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Listen when spoken to directly</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Follow through on instructions &amp; finish school work/chores</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Organize tasks and activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Engage in tasks that require sustained mental effort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Keep track of things necessary for activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Ignore extraneous stimuli</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Remember daily activities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Sit still (control movement of hands/feet or control squirming)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Stay seated (when required by class rules/social conventions)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Modulate motor activity (inhibit inappropriate running/climbing)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Play quietly (keep noise level reasonable)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Settle down and rest (control constant activity)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Modulate verbal activity (control excess talking)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Task</td>
<td>Far below</td>
<td>Below</td>
<td>Slightly below</td>
<td>Average</td>
<td>Slightly above</td>
<td>Above</td>
<td>Far above</td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>-----------</td>
<td>-------</td>
<td>----------------</td>
<td>---------</td>
<td>----------------</td>
<td>-------</td>
<td>-----------</td>
</tr>
<tr>
<td>16. Reflect on questions (control blunting out answers)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Await turn (stand in line and take turns)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Enter into conversations &amp; games (control interrupting/intruding)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Control temper</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Avoid arguing with adults</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Follow adult requests or rules</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Avoid deliberately doing things that annoy others</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Assume responsibility for mistakes or misbehavior</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. Ignore annoyances of others</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. Control anger and resentment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. Control spitefulness or vindictiveness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Avoid quarreling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. Remain focused on task (does not stare into space/daydream)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. Maintain appropriate energy level (is not sluggish or drowsy)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. Engage in goal directed activity (is not apathetic or unmotivated)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Thank you for completing this assessment.
REFERENCES


Boone, WJ; Rogg, SR; Kahle, JB; Damnjanovic A. *Race, gender, test length, and missing data: Why estimates of performance may be clouded.* Paper presented at the Annual National Association for Research in Science Teaching.


Case Western Reserve University, Department of Family Medicine, The Division of Adolescent Health. Middle School Youth Risk Behavior Survey: CMSD 2005.


DigSee SURE. Accessed March 9, 2004; See www.mobilentry.com/eng/.


Mott, F. Evaluation of fertility data and preliminary analytic results from the 1983 survey of the national longitudinal surveys of work experience of youth. A report to the National Institute of Child Health and Human Development by the Center for Human Resources Research. Columbus, OH: Center for Human Resources Research, 1995.


PocketPC PocketSurvey. Accessed March 9, 2004; See www.pocketsurvey.co.uk/.


