

The Social Behavior Competencies of Self-Identified Bullies as Assessed

by Students Themselves Plus Parents and Teachers

by

Shannon Leanna Rupp

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved February 2020 by the
Graduate Supervisory Committee:

Stephen N. Elliott, Chair
Becky Kochenderfer-Ladd
Nancy Eisenberg
Tashia Abry

ARIZONA STATE UNIVERSITY

May 2020

ABSTRACT

This two-study investigation examined the social behavior competencies of a sample of students ages 8 to 18 who identified themselves as either bullies or non-bullies based on ratings of items on a comprehensive behavior rating scale. Specifically, the purpose of Study 1 was to establish criteria using the Social Skills Improvement System – Student Rating Scale (SSIS-S) to identify students from a nationally representative standardization sample who displayed high frequencies of bullying behaviors. The social behavior ratings for these self-identified bullies were then compared with all other students in the national sample and analyzed to determine differences among various domains of social skills and problem behaviors. In Study 2, the same students’ social behaviors were rated by adult informants to determine if there was added value in including parents and teachers in the assessment of the self-identified bullies. Finally, the extent of concurrent agreement was examined for all students among the teachers, parents, and students’ ratings of social skills and problem behavior domains. Study 1 revealed that self-identified bullies are not a homogeneous group. The main findings from Study 2 showed parents and teachers may add to the overall predictive validity of the student self-report assessment, but not the accuracy of classifying the students as bullies. Study 2 showed differences and similarities exist across the ratings provided by each rater. The relative value of including adult reports in the self-assessment likely lies in the reported differences from each rater, as they provide a more complete social behavior profile for each student. These findings are discussed in terms of existing research and theories regarding children and youths’ bullying behavior. Limitations and recommendations for future research conclude the report.

ACKNOWLEDGMENTS

This process would not have taken place without the encouragement of mentors, friends, and family. I would like to acknowledge my faculty mentor and committee chair, Steve Elliott, for his dedication, support, and encouragement. Thank you for providing me excellent mentorship throughout my time in the graduate program. Your continual support and mentorship were central to my success. Furthermore, I would like to thank my committee members: Becky Kochenderfer-Ladd, Nancy Eisenberg, and Tashia Abry. Your comments and feedback throughout the process have tested me intellectually helping me to improve the quality of my work. I would also like to thank my undergraduate advisor Dr. Saralyn DeSmet for her ongoing encouragement and recommendation over the years.

I want to thank my good friends Carla and Samantha, who supported me and gave me the strength to keep going even when I wanted to give up. I also want to thank my friends and extended family, in particular, my sister and best friend, Nichole, who has been with me from the beginning of this journey and many others. Finally, an especially great thank you to Matt Herman for his unwavering support throughout my time at ASU and in the rest of life.

TABLE OF CONTENTS

	Page
LIST OF TABLES	vi
LIST OF FIGURES	viii
CHAPTER	
1 INTRODUCTION	1
Framework for Understanding Bullying	2
Overview of Study 1 and Study 2	6
Summary	8
2 STUDY #1 BULLYING AND RELATED SOCIAL BEHAVIORS IN A NATIONALLY REPRESENTATIVE SAMPLE OF STUDENTS.....	11
Social Skills and Problem Behaviors	12
The Social Skills and Problem Behaviors of Students Who Bully.	16
Assessing Bullying and Related Social Skills.....	24
Normative Trends in Bullying Behavior.....	29
Research Questions and Predictions	32
Research Question 1: What are the Demographics and Social Skills that Characterize Self-Identified Bullies	32
Method	34
Participants.....	34
Procedures.....	35
Measures	35
Plan of Analysis.....	40

CHAPTER	Page
Results.....	41
Discussion.....	51
Strengths, Limitations, and Future Directions	57
Conclusion.....	58
 3 STUDY #2 AN EXAMINATION OF THE ADDED VALUE OF ADULT INFORMANTS' RATINGS FOR THE IDENTIFICATION OF BULLIES	 59
The Value of Using Multiple Informants.....	60
Considerations When Choosing an Assessment.....	64
Purpose of the Study.....	66
Research Questions and Predictions	67
Research Question 1: How Accurate are Parents and Teachers at Identifying Students Who Self-Identify as Bullies or Non-bullies.	67
Research Question 2: What is the Relative Value in Including the Ratings from Teachers and Parents to Those of Students Who Self- Identify as Bullies?	67
Hypothesis 1: Parents and Teachers Were Expected to Add Incrementally to the Overall Predictive Validity of the Student Assessments.....	67
Hypothesis 2: It Was Expected That There Would be Modest Overlap Observed Between the Different Informants' Reports, Resulting in Low to Moderate Agreement Across Ratings and Differences in the Social Behavior Profiles Reported..	68

CHAPTER	Page
Method.....	69
Participants.....	69
Procedures.....	71
Measures.....	73
Plan of Analysis.....	74
Results.....	80
Discussion.....	93
Strengths, Limitations, and Future Directions.....	99
Conclusion.....	100
4 INTEGRATED SUMMARY OF STUDIES 1 AND 2.....	101
Key Findings.....	101
Study 1.....	101
Study 2.....	103
Noted Limitations.....	104
Implications for Future Research.....	105
Implications for Practice.....	106
Specific Conclusions to the Research Questions.....	107
REFERENCES.....	111
APPENDIX	
A SUPPLEMENTAL TABLES.....	131

LIST OF TABLES

Table	Page
1. SSIS Standardization Sample of Valid Self-Report Ratings Age 8 and Above by Sex, Age Group, and Race/Ethnicity (N = 1375)	34
2. SSIS Rating Scale--Bullying Subscale Items on Teacher, Parent, Student Forms	38
3. Variable Descriptions	39
4. SSIS Standardization Sample of Self-Identified Bullies Age 8 and Above by Sex, Age Group, and Race/Ethnicity (N = 167).....	42
5. Logistic Regression Predicting Likelihood of Self-identifying as a Bully based on Sex, Age, Race/Ethnicity, Total Social Skills, and Total Problem Behaviors	45
6. Logistic Regression Predicting Likelihood of Self-identifying as a Bully based on Sex, Age, Race/Ethnicity, Social Skills and Problem Behaviors Subscales.....	46
7. Logistic Regression Predicting Likelihood of Females Self-Identifying as a Bully based on Age, Race/Ethnicity, Total Social Skills, and Total Problem Behaviors	47
8. Logistic Regression Predicting Likelihood of Females Self-identifying as a Bully based on Age, Race/Ethnicity, Social Skills and Problem Behaviors Subscales ..	48
9. Logistic Regression Predicting Likelihood of Males Self-identifying as a Bully based on Age, Race/Ethnicity, Total Social Skills, and Total Problem Behaviors	49
10. Logistic Regression Predicting Likelihood of Males Self-identifying as a Bully based on Age, Race/Ethnicity, Social Skills and Problem Behaviors Subscales ..	50

Table	Page
11. Paired Samples Test of the Differences between Adult and Self-Reports of Bullying Ratings	81
12. Conditional Probability Framework of Student and Parent Identification of Bullies.....	82
13. Conditional Probability Framework of Student and Teacher Identification of Bullies.....	82
14. Coefficients of Differences in Predicting Bullying for Teacher, Parent, and Self-Reports Using Total Social Skills and Total Problem Behaviors Scales	88
15. Coefficients of Differences in Predicting Bullying for Teacher, Parent, and Self-Reports Using Social Skills and Problem Behavior Subscales	91

LIST OF FIGURES

Figure	Page
1. Integrative Theoretical Framework for Bullying.....	3
2. Functional Model for Understanding Bullying.....	5
3. Conditional Probability Framework.....	76

CHAPTER 1

INTRODUCTION

Despite over 30 years of research on school bullying, the field continues to debate issues involving the nature and assessment of students who bully others. Key issues concern: (a) whether bullies are deficient in critical social skills (e.g., Georgiou & Stavrinides, 2008; Haynie et al., 2001; Rodkin, 2012) or are alternatively, socially competent manipulators (e.g., Gasser & Keller, 2009; Perren & Alsaker, 2006; Rigby, 2007; Sharp & Smith, 2002; Sutton, Smith, & Swettenham, 1999; 2001), and (b) how the use of multi-informant assessments can improve understanding of bullies and their bullying behavior (e.g., Crothers & Levinson, 2004; Griffin & Gross, 2004; Demaray, Malecki, Secord, & Lyell, 2013; Groeben, Perren, Stadelmann, & von Klitzing, 2011; Juvonen, Nishina, & Graham, 2001; Nowell, Brewton, & Goin-Kochel, 2014; Phillips & Cornell, 2012).

Clear answers to these issues, especially among American researchers, have been elusive for several reasons. First, few, if any, researchers have used data gathered on nationally representative samples of children and youth (ages 8 to 18 years). Second, most researchers have not gathered assessment data for bullies from multiple informants (e.g., self-, parent- and teacher-reports) who rated a range of social skills (e.g., communication, self-control, empathy) and other problem behaviors (e.g., externalizing, internalizing, hyperactive). The goals of the present research were to (a) address the limitations of prior sampling and methodological designs and (b) provide insights that

advance the use of comprehensive assessment practices for students who self-report high frequencies of bullying behaviors.

Framework for Understanding Bullying

No widely accepted comprehensive framework, let alone theory, exists that adequately explains bullying or the distinctions among children and youth who exhibit high frequencies of bullying behavior. There are, however, documented differences about students who bully others. Specifically, there are differences between those who bully others and individuals who bully who were previously victims of bullies (i.e., bully-victims). Characterizations of both types of bullies exist in the research literature. Bully-victims typically are characterized as showing impulsivity, hyperactivity, and more academic problems (Schwartz, Proctor, & Chien, 2001). Researchers also have identified bullies (not bully-victims) who demonstrate high levels of social intelligence (Sutton et al., 1999) and social stability (Vaillancourt, Hymel, & McDougall, 2003). Such findings suggest a distinction between bullies who appear socially marginalized (a group which may include bully-victims) and those individuals who appear socially integrated (Farmer et al., 2010).

Having a framework acknowledging the differences between bullies and their related social behaviors is especially important when using a multi-informant approach to study bullying because context and circumstances can influence perceptions about a student's behavior. Multi-informant assessments typically broaden the context for understanding bullying. For example, if a student appears socially competent and well-

adjusted in one setting, an adult familiar with the student in that particular setting might not be aware of any bullying behavior the student exhibits in other settings.

Postigo, González, Montoya, and Ordoñez (2013) developed a comprehensive framework that acknowledges the multifaceted nature of bullying. Specifically, as shown in Figure 1, they advanced an integrated framework based broadly on the contextual-ecological model developed by Bronfenbrenner and from Sameroff's transactional model.

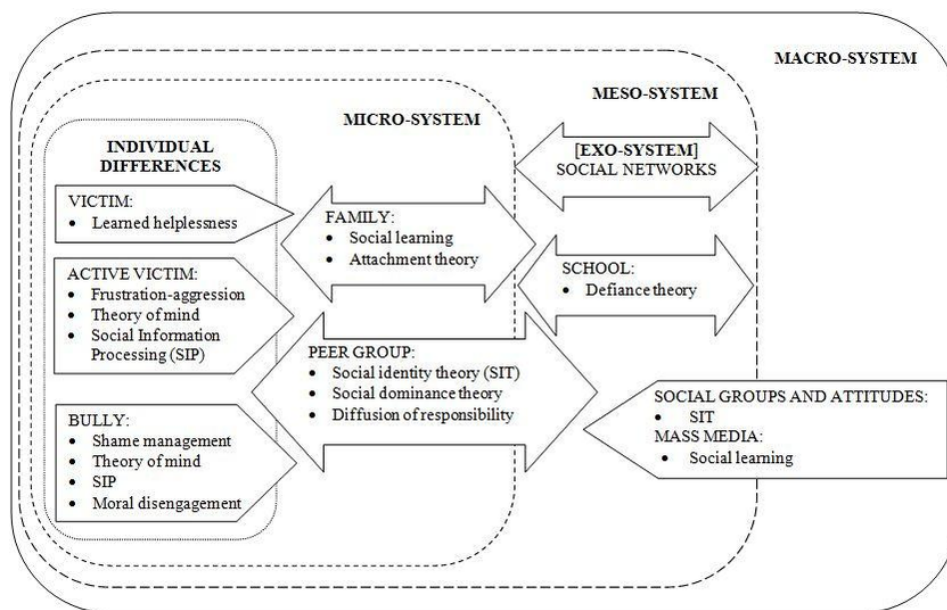


Figure 1. Integrative Theoretical Framework for Bullying (Postigo et al., 2013)

This contextual-ecological model places potential variables affecting bullying into different ecological spheres and highlights differences existing across cultures and contexts (Bronfenbrenner, 1986). The transactional model component emphasizes the two-way influence occurring between personal and contextual factors present during a bullying event (Sameroff, 2009). The integration of these models creates a rather comprehensive framework depicting how a specific model can be used to explain the impact and relations among variables while sorting the variables based upon an

underlying process (e.g., developmental) that influences the relation (Postigo et al., 2013).

The Postigo et al. (2013) integrative theoretical framework for bullying provides a comprehensive view of the numerous variables and contexts that might influence a bullying event. The broader framework addresses different components of bullying such as individual differences, family context in the micro-system, and school in the mesosystem. A narrower, more functional model, however, is needed to frame and situate the context of the current research problem concerning bullying.

Figure 2 on the next page illustrates an example of a more functional behaviorally focused model to supplement the broader framework for bullying. The model was theorized to contextualize bullying assessment research. Note, it is designed to address the situational context (e.g., home, school, community) before a bullying event takes place, specific bullying behaviors observed, the consequences for the bullying behavior, and the outcomes for the bully. This functional model describes the set of behaviors characterizing bullying as observed by the students themselves, teachers, and parents.

Antecedent Situations	Observed Bullying Behaviors Self Teacher Parent	Consequences for Behavior	Outcomes for Bully
Home School Community	Does things to make others feel scared	Social isolation or dominance (Cook et al., 2010)	Lower feelings of school belonging (Swearer, 2011), poor academic engagement (Nansel et al., 2003); changes in brain functioning (Chugani et al., 2001)
Home School Community	Forces others to act against their will	Conduct problems (Cook et al., 2010)	Delinquency (e.g., vandalism, carrying a weapon) (Cook et al., 2010)
Home School Community	Keeps others out of social circles	Popularity (Rodkin, Farmer, Pearl, & Van Acker, 2006)	Higher social standing and development of behavior as part of larger peer group norms (Espelage, Holt, & Henkel, 2003; Salmivalli, Huttunen, & Lagerspetz, 1997)
Home School Community	Is aggressive toward people or objects	Conduct-disordered behavior (Kokkinos & Panayiotou, 2004)	Long-term substance abuse (Brockenbrough, Cornell, & Loper, 2002; Vaughn et al., 2010)

Figure 2. Functional Model for Understanding Bullying (Rupp & Elliott, 2019)

Overview of Study 1 and Study 2

Study 1. In this study, a sample of 1,375 youth, ages 8 to 18 years, was used to address the following research question: *What are the demographics and social skills that characterize self-identified bullies?* This question is crucial to advance assessment practices geared towards the improvement of programs and strategies to reduce children's bullying behavior. In particular, it is important to examine the social skills of individuals who exhibit bullying behaviors. The examination helps to determine if children and youth bully because they lack critical social skills (Farmer et al., 2010) and/or if, conversely, they bully because "they can." That is, they have the social savvy to manipulate others, get what they want, and often avoid detection and punishment (Sutton et al., 1999; 2001).

Answers to these questions about demographics and social skills have important implications for assessment and intervention efforts. If bullies lack skills, then programs that teach children needed skills are appropriate and are more likely to be successful at reducing bullying behavior than if they do not lack such skills. However, if social skill deficiencies are not a salient aspect of the mechanism for bullying, then, appropriate assessment is needed to determine what is the mechanism (e.g., moral disengagement, social norms, lack of consequences, power/dominance). Study 1 contributed data to answer these questions by using a broadband social behavior assessment to examine the social skills (e.g., communication, self-control), and problem behaviors (e.g., hyperactivity, externalizing) of students who self-identify as engaging in bullying.

Study 2. This study expanded upon the findings of Study 1 by including the perspectives of parents and teachers along with the self-reports of students who identify

as bullies. Specifically, Study 2 addressed the issue of determining the value of using a multi-informant approach to studying bullying. The study addressed the following research questions.

Research Question 1: How accurate are parents and teachers at identifying students who self-identify as bullies or non-bullies?

Research Question 2: To what extent do parents', teachers', and students' ratings of social skills and problem behaviors agree with one another for the students who self-identify as bullies and students who do not self-identify as bullies?

Determining the value of including more than one informant in bullying research is part of a continued call from researchers about multi-informants (e.g., self-, parent-, teacher -reports) to assess bullying problems (e.g., Crothers & Levinson, 2004; Griffin & Gross, 2004; Demaray et al., 2013; Groeben et al., 2011; Juvonen et al., 2001; Nowell et al., 2014; Phillips & Cornell, 2012). Yet, there is no precise determination for how these diverse perspectives should be integrated and interpreted to further understanding of bullying. For example, when informants identify different students who have engaged in bullying behavior, are they all “accurate,” but somehow adjusting for potential partiality in other reporters' preconceived perceptions about the students? Alternatively, is it the case that a particular informant is basing their opinion upon preconceived notions or are they perhaps not as well informed about the person in question and thus, an inappropriate informant in some instances?

Study 2 built upon Study 1 findings by comparing adults' perspectives of the social competencies of self-identified bullies to identify commonalities, or discrepancies, of the perceived skillfulness of students who report engaging in such behaviors. The aim of Study 2 was to examine the value of multiple informants in understanding the behavior of self-identified bullies with a comprehensive, multi-rater social behavior measure. The aim also was to provide recommendations for the use of multiple informants by examining the incremental validity of using parent and teacher reports to augment self-ratings of bullying behavior.

Methodology. Both studies addressed several methodological limitations of the existing research by using the Social Skills Improvement System (SSIS) Rating Scales (Gresham & Elliott, 2008). The SSIS assesses bullying behaviors, as well as social skills (e.g., communication, cooperation, assertion, responsibility, engagement, empathy, self-control) and problem behaviors (e.g., bullying, externalizing, hyperactivity/inattention, internalizing). The SSIS is a true multi-informant rating system in which all raters (self, parents, and teachers) complete items with the same root language. These rating scales were normed for boys and girls, ages 3 to 18 years, with a nationally representative sample of students in 2007 (Gresham & Elliott, 2008).

The Student self-rating version of the SSIS is used with children and youth ages 8 to 18. Study 1 relied on student self-reports from 1,375 students, 741 females and 634 males. The racial and ethnic backgrounds were representative of the U.S. population (U.S. Census data, 2017): 10% Black, 21% Hispanic/Latinx, 62% White, 7% Other. Study 2 relied on the same subsample of 1,375 student self-report cases. Within those

1,375 cases, there were 424 cases with all three raters (self, parent, and teacher), 922 with a self-rating and a parent rater, and 29 with a self-report and a teacher rater.

Summary

In summary, there remains a need in the U.S. to estimate the frequency and nature of bullying behavior for a nationally representative sample of students. There also needs to be better documentation of how bullying and related social behaviors vary with age and sex. This documentation is important because while there is some consistent agreement that there are at least two types of bullies—socially deficit and socially savvy—researchers continue to vary in their agreement about specific characterization of bullies and their array of social behaviors. This variation inhibits both assessment and prevention efforts because the most effective interventions appear to benefit from a solid understanding about the complexity of bullies' social behavior strengths and weaknesses (Cantone et al., 2015).

Finally, although strides have been made to address multi-rater assessment issues by conducting studies comparing two or more informants, strong conclusions, along with clear use recommendations, have been elusive. This issue highlights the need for studies specifically exploring the incremental validity of using adult reports in addition to self-reports to assess bullying and social behaviors. This need is in addition to studies examining bullying and related social behaviors using an assessment designed and tested for use with multiple informants, like the SSIS, who observe youths' behavior across multiple settings. More studies like these are expected to contribute to developing a stronger foundation in the literature about the value of the information each informant

contributes to an assessment of bullying. Together, these two studies advanced bullying assessment research and specifically the utility of multi-informant, broadband assessment of the social behavior of students who frequently exhibit behaviors typical of bullies.

CHAPTER 2

STUDY #1 BULLYING AND RELATED SOCIAL BEHAVIORS IN A NATIONALLY REPRESENTATIVE SAMPLE OF STUDENTS

Research over several decades has documented the pervasiveness and harmfulness of bullying for both victims and perpetrators (e.g., Arsenault, Bowes, Shakoor, 2010). In tacit acknowledgment of the harmfulness of bullying on children and youths' well-being, there has been a proliferation of research, programs, and policies to reduce its deleterious effects on youth development, if not prevent its actual occurrence (e.g., Bowllan, 2011; Cantone et al., 2015; Ttofi & Farrington, 2012; Vreeman & Carroll, 2007). Unfortunately, such efforts seem to have had limited success (e.g., Brown, Catalano, Fleming, Haggerty, & Abbott, 2005; Brown, Low, Smith, & Haggerty, 2011; Cantone et al., 2015; Fonagy et al., 2009).

The premise of this study is that bullying prevention efforts are not as successful as hoped partly because the assessment and identification of students who bully others often do not take into account youths' social behavior profiles or characteristics. In other words, by assessing and reliably identifying bullies' existing social competencies – both strengths and weaknesses --interventionists are better able to target skills that students need to improve and use in place of bullying behaviors to accomplish their social goals.

It is important that researchers and policymakers understand the wide range of social behaviors of children who bully given these children do not comprise a homogeneous group. For example, current research suggests at least two distinct subtypes of bullies (though more are possible): (a) socially competent but manipulative bullies

(e.g., Gasser & Keller, 2009; Perren & Alsaker, 2006; Rigby, 2007; Sharp & Smith, 2002) and (b) marginalized/rejected/socially deficient youth who may bully others as a result of being victimized themselves (e.g., bully/victim; see Georgiou & Stavriniades, 2008; Haynie et al., 2001; Rodkin & Roisman, 2010). Taking the perspective that the assessment of bullies should capture the complexity of their social behaviors is important because the most effective interventions appear to benefit from knowledge about the array of bullies' social behavior strengths and weaknesses (Cantone et al., 2015).

The goal of this study was to contribute to advancing assessment practices that have implications for interventions. The first step involved investigating the use of self-report assessment as part of a broadband social skills assessment, the Social Skills Improvement System-Rating Scales (SSIS), to estimate the prevalence of bullies in a nationally representative sample of students ages 8 to 18. The second step was to examine and describe the key demographics and social behaviors associated with a student's status as a self-identified bully (e.g., identifying more boys as bullies, who commit more physical acts of bullying than girls do).

Social Skills and Problem Behaviors

It is important in assessment to understand that the constructs of social skills and social competence are related, but different. Gresham, Elliott, Vance, and Cook (2011) described social skills as the behaviors that individuals use in a given social situation. These skills are developed and carried out with different actions, such as listening to others and making friends in class. The behaviors demonstrate an individual's understanding of the typical behavior acceptable to a specific scenario. Every social

undertaking necessitates the use of various social skills, and an examination of the effects of these skills determines an individual's level of social competence. The presence of acceptable social skills is a requirement for an individual to succeed in most social and academic contexts. Those individuals' lacking in social skills might find poorer social relationships and less acceptance from teachers and peers. Gresham, Elliott, and Kettler (2010) described social competence as the result of comprehending and using an existing set of social skills.

Social skills. The development of an accurate social skills assessment makes it necessary to discuss the potential deficits in a person's social skills revealed by the assessment. Gresham (1981) noted there are two types of deficits in social skills, those regarding performance and acquisition. Distinguishing between acquisition and performance deficits is useful in making decisions about the implementation of interventions (Gresham & Elliott, 2008).

Deficits. When an individual displays a performance deficit, it suggests they do not fully exhibit the desired behavior in a specific social setting or exhibit it at a lower frequency rate than desired (Gresham, 1981). The belief is that the individual understands the desired skills to use, but he/she is reluctant or unable to execute the skills. Having this kind of deficit is thought to be related to motivation or poor understanding of the social situation rather than to the acquisition of behavior. This assumption means an intervention for this type of deficit will be designed to address motivational issues rather than increasing the individual's skill level. A deficit related to an individual's acquisition of skills is when the individual either lacks the skills or the knowledge of the skills

necessary to perform correctly in a specific social setting (Gresham, 1981). In this case, an intervention to address the deficit requires a direct approach that involves teaching the individual the missing skills or how to use the right skills in that specific situation.

Social skills strengths. Besides the identification of social skills deficits, an accurate assessment involves the consideration of an individual's social strengths. Social skills strengths are when the individual possesses the skills and knows when and how to use them consistently when in specific social settings (Gresham & Elliott, 2008). The idea is that these strengths are instrumental in an individual's management of a variety of social settings. These strengths can potentially help to support or cushion the social skills deficits of an individual in some social situations.

Problem behaviors. Problem behaviors are those social behaviors that are considered anti-social, damaging, or interpersonally destructive (Gresham & Elliott, 1990). It is of note that problem behaviors, along with social skills, are usually influenced, if not defined, by one's cultural context. The understanding and influence of each one on an individual's behaviors might have a somewhat different interpretation. Existing studies examining children's problem behaviors and social skills simultaneously typically show a relatively strong negative relationship; however, that is not always the case. The presence of one or more problem behaviors, such as bullying, may not imply the individual lacks in social competency (Gresham, Van, & Cook, 2006). Also, having lower social competency in some areas does not always mean the individual will exhibit problem behavior. With these considerations in mind, social skills and problem behaviors

can be regarded separately under the concept of social competence but measured simultaneously as means for them to complement one another.

A behavioral assessment must also consider the issue of competing problem behaviors with regards to any social skills deficits (Gresham & Elliott, 1990). The behavior of this type can interfere with the performance or acquisition of a social skill causing conflict in an individual's development and/or use of the skill in a social setting. These competing problem behaviors are usually classified as patterns of externalizing behavior (e.g., aggressive behavior) or internalizing behavior (e.g., anxiety). The presence of one of these types of behaviors can inhibit an individual from learning or using a desired behavior. An example of this issue occurs when an individual displays anxious or aggressive behavior. The individual may not learn to share or have the self-control to combat those other behaviors, because they have both a hard time interacting with others and having others wanting to interact with them (Gresham et al., 2006).

In summary, research indicates that children who have a robust set of age-appropriate social skills that promote their competence in social situations have fewer antisocial behaviors and better engagement in those different social settings (Gresham et al., 2004). An accurate assessment and the introduction of early interventions are necessary to prevent and address problem behaviors with the appropriate desired skills development to improve those behaviors. More than a decade ago, a set of meta-analyses from Gresham et al. (2004) concluded social skills include prosocial behavior, social-cognitive functioning, and social interaction. Key correlates of these social skills are the stated problem behaviors and academic performance. Therefore, given this

understanding, it is necessary to have a suitable and applicable assessment connected to a valid intervention that is available to researchers to gauge both social skills and problem behaviors accurately.

The Social Skills and Problem Behaviors of Students Who Bully

The problem behavior of bullying tends to be conceptualized as a subcategory of aggression characterized by hostile intentions, power imbalance, and a behavior repeated over time (Gladden, Vivolo-Kantor, Hamburger, & Lumpkin, 2014; Goldsmid & Howie, 2014; Olweus, 2013; Vlachou, Botsoglou, & Andreou, 2013; Volk, Veenstra, & Espelage, 2017). Consistent with this conceptualization, the present study defines bullying as *repeatedly directing aggressive behavior toward others for the explicit purpose of controlling others (e.g., gaining power over them) and making them feel vulnerable and scared (e.g., intention to harm)* (Olweus, 2013).

Existing studies examining both problem behaviors and social skills show simultaneously that they correlate moderately and negatively based on self-ratings and ratings by parents and teachers (Gresham & Elliott, 2008). The presence of one or more problem behaviors, such as bullying, however, may not imply that the individual lacks social skills (Gresham et al., 2006). Also, having fewer social skills in some areas does not always mean the individual will exhibit problem behavior. Clearly, correlation does not mean causation. In bullying research, however, the debate continues whether or not students who bully lack certain social and emotional skills or simply choose not to use such skills in some interactions. Many researchers argue the students who bully show deficits in social and emotional skills. However, implicit in definitions of bullying, which

include hostile intentions, is a degree of "skillfulness" in manipulating others to their advantage and planning future actions (Sharp & Smith, 2002). This understanding of the definition suggests a contrast to the social skills deficit perspective with some researchers arguing the students who bully demonstrate higher than average levels of social skills, in particular, those who are well-practiced in indirect types of bullying (Rigby, 2007).

Social skills deficit perspective. Crick and Dodge's (1994) social skills deficit model regarding bullying asserted that students identified as bullies demonstrate poor social skills and participate in bullying as a means for improving their self-image and gaining social power and status in their peer group. In reality, however, relatively few researchers have examined the social skills characteristics of children identified as bullies. One exception, Vlachou et al. (2013) described a picture of bullies as arrogant and narcissistic. Other investigators have found that bullies are arrogant, but have fragile egos, mental and personality disorders, anger management, aggressiveness, and poor self-image. Overall, they found those who identify as bullies typically show higher levels of conduct problems, poorer academic outcomes, and lower psychosocial functioning (e.g., Denny et al., 2015; Ttofi, Farrington, Losel, & Loeber, 2011). For example, Wolke, Lereya, Fisher, Lewis, and Zammit (2014) found significant negative relationships between self-reported bullying behaviors (Bullying and Friendship Interview Schedule) and other problem behaviors, such as concurrent internalizing and externalizing behavior and later depression and psychotic problems. Self-reported findings in a study using the Olweus Bullying Questionnaire by Wolke, Woods, Bloomfield, and Karstadt (2000) showed those who demonstrated involvement in direct bullying behaviors had the most

significant increases in overall problem behaviors, hyperactivity, and conduct problems, and they demonstrated lower prosocial behavior scores in comparison to those children who were considered neutral.

Theory of Mind perspective. In contrast to the "pathological" and social skills deficit model, Sutton, Smith, and Swettenham (1999) argued for the Theory of Mind perspective that "is the ability of individuals to attribute mental states to themselves and others in order to explain and predict behavior" (p. 119). They believed the skills and context of bullying depend on how well the bully understands or can manipulate the mind of another individual or group--"a 'theory of mind,' or social cognition" (Sutton et al., 1999, p. 120). That is, bullies use social skills to manipulate people in different social settings and display superior theory of mind. This assumption is because some bullies display strong psychological qualities and achieve higher social standing. As noted earlier, Rodkin and Roisman (2010) suggested bullies represent two types, the socially connected and manipulative and the marginalized who bully in retaliation to being victims (bully-victims).

Existing research supports the perspective that some children who bully others display appropriate social skills and appear socially competent much of the time. For example, Wolke et al. (2000) found children who labeled themselves as relational bullies showed lower behavior problem scores than others labeled as bullies. However, they also were less likely to be prosocial and more likely to demonstrate manipulative tendencies compared to non-bullies. In another example, Gasser and Keller (2009), using an

assessment with ratings developed by Perren and Alsaker (2006)¹, found that bullies scored similarly to prosocial children on moral knowledge (e.g., meaning when moral thoughts are true and justifiable), but lower on moral motivation (e.g., the desire to apply moral knowledge). Perren and Alsaker (2006) found that victims had lower social skills than students in the bullying groups generally disagreeing with the deficit perspective. Bullies identified by Perren and Alasker (2006) seemed to maintain their status by targeting weaker individuals through acts of bullying. Those with stronger leadership skills received support from peers who also engaged in the perpetration of harmful acts.

Current framework/perspective. For this study, I proposed that bullies do not comprise a homogenous group. In fact, I argued that even the most socially unskilled bully might possess social skills that can be assessed and used to guide intervention. Some of the students who bully, likely have a well-developed understanding of social cues (e.g., reading facial expressions) and know exactly how to use the information to their advantage (Gini, Pozzoli, & Hauser, 2011), but may still display some deficits in social skills. For example, as noted by Sutton et al., (1999), the students who use indirect bullying likely need better social skills to carry out their actions. The type of social skills they need to have, or at least to understand, for successfully manipulating the behavior of their peers may include assertion, communication, and empathy, which help to draw in peers who can support the bully's behavior. In contrast, these same bullies may also

¹ Ratings were based on “four victimization and four bullying items (physical, verbal, object-related, exclusion). The 5-point rating-scale consisted of the following categories: *never, seldom, once or several times a month, once a week or several times a week* (e.g., ‘child bullies other children physically’; ‘child is victimized verbally’)” (Perren & Alsaker, 2006, p. 47).

display a lack of social skills in cooperation and self-control due to the aggressive nature of bullying (e.g., Chui & Chan, 2013; Perren & Alasker, 2006). The existing literature about the connections between these social skills and how they characterize the behavior of bullies is limited, and research results are varied. However, the results of these existing studies are discussed below to provide some general expectations about the anticipated results of the current study.

Assertion. Assertion is "initiating behaviors, such as asking others for information, introducing oneself, and responding to the actions of others" (Gresham & Elliott, 2008, p. 1). The existing rationale is that perhaps the more socially savvy bullies are assertive because findings show they use such behavior to skillfully and mindfully control their peers (Boulton & Smith, 1994; Elliott, Hwang, & Wang, 2019; Jenkins, Demaray, & Tennant, 2017; Juvonen, Graham, & Schuster, 2003; Stephenson & Smith, 1989) and to receive frequent leadership nominations (Cairns, Cairns, Neckerman, Gest, & Garipey, 1988). Further evidence suggests bullies display assertive behavior because bullies have higher ratings from peers regarding social desirability and leadership compared to non-bullies (Collins & Bell, 1996). Findings of the characteristics of identified bullies who display assertive behavior, however, are limited.

Communication. Communication skills involve "taking turns and making eye contact during a conversation, using the appropriate tone of voice and gestures, and being polite by saying 'thank you' and 'please'" (Gresham & Elliott, 2008, p. 1). Existing rationale suggest the communication skills displayed by bullies varies depending on characteristics such as age and sex (e.g., Carney, Hazier, & Higgins, 2002). For example,

evidence shows male bullies, middle school (grades to 6 to 8) and younger, tend not to use behaviors such as appropriate tones and gestures, as evidenced by their propensity to participate more often in physical acts of bullying (Carney et al., 2002). Although girls in all age groups might display advanced communication skills given that they participate more often in indirect verbal bullying (name-calling, creating rumors) (e.g., Shahaiean, Razmjooe, Wang, Elliott, & Hughes, 2017; White & Kistner, 2011) and relational bullying, an act of bullying targeting another individual's reputation (Card, Stucky, Sawalani, & Little, 2008; Yubero & Navarro, 2006). These acts often revolve around having conversations with others and using the right tone of voice and gestures to get others to act as they want them to act.

Some researchers, such as Savage and Tokunaga (2017), also suggest that a penchant for acting aggressively, whether impulsively or consciously, indicates access to bullying-related schemas and scripts. So students with access to these knowledge structures, and who display high levels of verbal aggression, likely communicate aggressive behavior no matter the provocation. The behavior suggests advanced communication skills. With that in mind, the literature would indicate that while communication skills may vary across age and sex, some bullies, in particular girls, may show higher ratings of it.

Cooperation. Cooperation involves "helping others, sharing materials, and complying with rules and directions" (Gresham & Elliott, 2008 p. 1). Existing rationale about the behavior of bullies suggests the actions describing cooperation are in contrast to the behaviors typically displayed by bullies. Research supports this rationale showing a

significant negative relationship between cooperation ratings and bullying for girls and boys (Jenkins, Demaray, Fredrick, & Summers, 2016; Perren & Alasker, 2006; Wang, Chen, Xiao, Ma, & Zhang, 2012). Findings from Jenkins et al. (2016) and Perren and Alaskar (2006) suggested bullies were less cooperative than children not involved in bullying. However, it is unknown exactly how cooperation and bullying may vary across age, but these findings point to the results of this study showing differences in cooperation between bullies and non-bullies for both sexes.

Empathy. Empathy involves "showing concern and respect for others' feelings and viewpoints" (Gresham & Elliott, 2008, p. 2). However, the results from existing research about empathy and bullying are varied, including results reported by age and sex. The findings suggest that because some bullies may manipulate others and demonstrate popularity among their peers, this indicates they possess the ability to recognize the emotions and behaviors of others, but they have problems managing their own emotions. These assumptions are supported by existing literature, as students who bully others are shown to have problems with managing their emotional behavior (Espelage, Bosworth, & Simon, 2000; Sutton et al., 1999). The bullies could not understand the emotional harm caused by their actions (Lomas, Stough, Hansen, & Downey, 2012). They recognize emotions, but they do not empathize like their peers do, making them more likely to manipulate others in social situations (Sutton et al., 1999; Warden & Mackinnon, 2003).

Existing studies showed the different results regarding empathy and bullying (van Noorden, Haselager, Cillessen, & Bukowski, 2015). Most results show a negative

association between bullying and any type of empathy (e.g., Caravita, Blasio, & Salmivalli, 2010, Espelage, Mebane, & Adams, 2004; Jolliffe & Farrington, 2006; Jolliffe & Farrington, 2011; Kokkinos & Kipritsi, 2012; Mitsopoulou & Giovazolias, 2015). Others have found no association between bullying and empathy (e.g., Kane, 2015), while others found mixed results depending on the type of empathy, the age, and the sex of the respondent (e.g., Jolliffe & Farrington, 2011). These displays vary by sex, as some studies suggested a negative relationship between bullying and empathy for boys (Jolliffe & Farrington, 2006), but not for girls (Jolliffe & Farrington, 2011). The existing evidence suggested the results of the current study would show similar variations about empathy across age and sex.

Self-control. Self-control involves “responding appropriately in conflict (e.g., disagreeing, teasing) and non-conflict situations (taking turns and compromising)” (Gresham & Elliott, 2008, p. 2). The rationale for the negative relationship between self-control and bullying is that bullies tend to do the opposite of responding appropriately in different situations. They display poor self-control because of their aggressive behavior. This assumption is supported in part by existing research showing identified bullies display poor emotional control, more aggression, and impulsive behavior than non-bullies (Chui & Chan, 2013; Pellegrini, Bartini, & Brooks, 1999; Unnever & Cornell, 2003). However, existing research showed only a general relationship between low self-control and bullying. For example, when controlling for sex and race, Unnever and Cornell (2003) found that lower self-control predicted bullying. These results showed a general negative relationship between bullying and self-control, so it is likely identified bullies of

all types will display a lack of self-control. However, given the limitations of the results from existing studies, it is yet unknown exactly how age and sex influence the relationship between bullying and self-control (Jenkins et al., 2016).

In summary, researchers vary in their agreement about the social skills of bullies. However, there is some consistent agreement that bullies understand aspects of theirs' and others' social-emotional behavior. The bullies, in comparison to non-bullies, can display some prosocial behaviors, understand those emotions and behaviors in others, and use them to their advantage. However, they are less likely to show empathy than non-bullies.

Assessing Bullying and Related Social Skills

Choosing how to assess the social behavior and related social skills of bullies represents another area for debate. Although some researchers stress the need for multi-methods (i.e., observation, surveys, nominations, ratings) or multi-informants (i.e., self-, parent-, peer-, teacher-reports) to assess bullying behavior (Crothers & Levinson, 2004; Griffin & Gross, 2004; Juvonen, Nishina, & Graham, 2001), no consensus has been obtained on the particular combination of methods or informants that is most appropriate. Moreover, many researchers still use single-informant strategies. The reasoning is the assessment they have chosen to use does not provide multiple forms or they want to avoid likely poor to modest agreement between raters and related interpretability challenges (e.g., Achenbach, McConaughy, & Howell, 1987; Branson & Cornell, 2009; Camodeca, Caravita, & Coppola, 2015; Crick & Bigbee, 1998; Hwang, Kim, Koh, Bishop, & Leventhal, 2017; Kuterovic-Jagodic & Velki, 2014; Monks, Palermiti, Ortega,

& Costabile, 2011; Pellegrini, 2001; Pellegrini & Bartini, 2001; Ung et al., 2017; Verlinden et al., 2014).

Self-reports. Rather than using multiple informants, it has been reasonably argued that the best approach is one that has been carefully selected for the research study or the specific age group being studied (e.g., cognitive maturity, ability to reflect on one's own behavior; Camodeca, et al., 2015; Verlinden et al., 2014; Vlachou, Andreou, Botsoglou, & Didaskalou, 2011). Given the current focus of bullying among 8- to 18-year-old students, self-report assessments are used consistently with preadolescent and adolescent samples (Chan, 2006; Cornell & Cole, 2012; Frey, Hirschstein, Edstrom, & Snell, 2009; Juvonen et al., 2001; Yubero & Navarro, 2006). For this age group, self-reports are the most common method of assessing school bullying, especially when a large and diverse sample is required (Branson & Cornell, 2009; Furlong, Sharkey, Felix, Tanigawa, & Green, 2010). Specifically, self-report behavior rating scales are considered the most efficient and cost-effective method for gathering data from large nationally representative samples across multiple developmental periods (Furlong et al., 2010).

Self-reports may be best for estimating the frequency of bullying because they have the advantage of providing information across contexts (e.g., playgrounds, cafeterias, hallways, classrooms, neighborhoods, shopping centers) as well as distinguishing among subtypes of bullying, such as direct, indirect, and relational (Furlong et al., 2010). Moreover, the children themselves are most likely to be aware of their intentions for engaging in aggressive behavior toward others. That is, the students themselves provide information through their first-hand accounts of bullying incidents

that others might not have. This information might be intentionality, power imbalance, and whether aggression was reactive or unprovoked (Juvonen et al., 2001; Pellegrini, 2001; Solberg & Olweus, 2003).

It should be noted, however, that concerns do exist regarding the use of self-reports to assess bullying. These concerns relate to the conceptualization and definition of bullying, the psychometric evaluation of self-report instruments, and the consideration of the developmental age of respondents (e.g., Cornell, Sheras, & Cole, 2006; Crothers & Levinson, 2004; Furlong et al., 2010; Greif & Furlong, 2006; Vlachou et al., 2011). There are variations in the definition of bullying in self-report studies despite researchers' consensus about the definition (Vlachou et al., 2011). The choice to or not to adhere to this researchers' understanding of bullying can influence the definition and measurement of bullying across studies (e.g., Greif & Furlong, 2006). Additionally, the uncertainty in the conceptualization of bullying potentially leads to the different estimations of it seen from reported self-ratings (Finkelhor, Turner, Shattuck, & Hamby, 2015; Iannotti, 2013; Nansel, Craig, Overpeck, Saluja, & Ruan, 2004). There are also only a few empirical articles evaluating the psychometric properties of existing self-reported bullying assessments (Furlong et al., 2010).

Lastly, some self-reports may be biased because existing studies that consider time and repetition (e.g., asking questions about when and how often bullying occurred, such as in the past month or 12 months) of bullying can influence a respondent and raise concerns about how valid responses are. There are difficulties present when asking students to reflect upon a bullying incident if they have to rely on their long-term

memory to respond to the question (Furlong & Sharkey, 2006; Hilton, Harris, & Rice, 1998). Another response issue is that questions about bullying might provoke responses that students deem more socially desirable than an actual accurate response about their behavior (Espelage & Swearer, 2003). Respondents report the perceived desirable response about their bullying experience because they fear potential stigma related to that experience (Furlong et al., 2010).

In summary, there are ongoing concerns about the use of self-report assessment to measure bullying behavior. However, conducting studies using an assessment that addresses these ongoing concerns can help build support for their validity. In addition, self-reports typically have been used as a very efficient, cost-effective way to obtain estimates on bullying in a nationally representative sample of almost 1,500 8-to-18-year-old students. These points are especially important as policymakers and interventionists will most likely rely on self-reports to similarly collect data on large samples of school-age children to estimate the frequency of bullying as an indicator of success (or failure) to reduce school bullying (Solberg & Olweus, 2003).

Choosing an assessment. Besides informant choice, there are many behavior ratings scales to choose from for assessing the social behavior of students (Crowe, Beauchamp, Catroppa, & Anderson, 2011). While multiple measures exist, most only measure one or a few dimensions of social skills. Few use an approach with a focus on multiple domains of prosocial behavior, while at the same time assessing problem behaviors (e.g., bullying) and academic competence across a wide age range. In addition, few consider the concerns about self-report measures. One that does use this approach

and takes these concerns about self-report under consideration is the Social Skills Improvement System (SSIS) Rating Scale, a broadband assessment. That is, it assesses an array of social skills, several classes of problem behaviors including bullying, and academic competencies.

The Social Skills Improvement System (SSIS). The SSIS is a revised version of the Social Skills Rating System (SSRS) (Gresham & Elliott, 1990). The SSRS (Gresham & Elliott, 1990) was developed as a broadband, multi-rater (teacher, parent or caregivers, and student) system to assess a student's social skills, problem behaviors, and level of academic competence (only by teachers). The original goal in creating the SSRS was two-fold: to find out which students were at risk for social behavior problems and to determine which behaviors could be targeted for intervention. Like the SSRS, the SSIS collects ratings from three informants (teachers of children ages 3-18, parents/caregivers of children ages 3-18, and students (ages 8-12 and 13-18) (Gresham & Elliott, 2008). The Social Skills (SS) scales collect data on seven specific skills: (a) Assertion, (b) Communication, (c) Cooperation, (d) Empathy, (e) Engagement, (f) Responsibility, and (g) Self-control. The Problem Behaviors (PB) scale has the same rating for informants to use. The PB Scale consists of five subscales: (a) Autism Spectrum, (b) Bullying, (c) Externalizing, (d) Hyperactivity/Inattention, and (e) Internalizing. Teachers also rate a student's Academic Competence on a short scale.

This assessment was designed to link results directly to intervention through the identification of observable target behaviors with known social importance for functioning in schools and at home (Demaray et al., 1995). The SSIS linked the forms to

the intervention program, so the rating scales are part of a multi-tiered model that allows for the assessment and monitoring of each student assessed using the SSIS. They created a flexible assessment and intervention model by also including an intervention program and guides to the intervention and performance screening (Gresham et al., 2011). The SSIS is frequently used for assessing the social skills and problem behaviors in students ages 3 to 18 both in the United States and internationally. Researchers also use the SSIS and the SSRS widely in research with the results from multiple studies providing evidence of the reliability and validity of both sets of scales (e.g., Crowe et al., 2011; Frey, Elliott, & Gresham, 2011; Gresham & Elliott, 1990, 2008; Gresham et al., 2011; Humphrey et al., 2011; Van der Oord et al., 2005; Walthall, Konold & Pianta, 2005). The evidence and design of both sets of scales addresses some of the concerns about self-reports, such as the inclusion of an index to detect false responses and the inclusion of other informants' ratings.

Normative Trends in Bullying Behavior

Existing differences across the definition of bullying, informant choice, and the assessment used may also contribute to variations in reports about the normative trends in bullying behavior. In the United States, reliable estimates of bullying have been challenging to obtain. Given differences in the ages studied, methodologies used, and other sample characteristics, estimates range from a low of 5% to just over 30% (Finkelhor et al., 2015; Iannotti, 2013; Nansel et al., 2004). For example, in a major report of bullying among youth aged 11 to 15, representing 25 countries, Nansel et al. (2004) reported that self-reports indicated between 5% to 20% reported perpetrating

bullying. Unfortunately, most existing bullying assessments do not have norming studies to allow an examination of behaviors across age, sex, and developmental levels (Griffin & Gross, 2004). Even fewer studies collect data on nationally representative samples of students and with sufficiently large samples to examine behavior over a large age range (Finkelhor et al., 2015; Iannotti, 2013). Moreover, most estimates of bullying either collapse data across males and females, or when they do estimations by sex, collapse across age. The current study addresses these limitations of previous studies by examining the prevalence of bullying both by age and sex for a large nationally representative sample of children in the United States.

Age trends. Researchers argue that accurate bullying assessments are developed with consideration for the different developmental stages and the sex of the potential respondents (Greif & Furlong, 2006). The consideration of developmental levels is necessary because the current research using self-report surveys reports differences in behavior across age groups with the bullying behavior decreasing with age. A sample from Sadinejad et al. (2015) self-reported more bullying among middle-school aged students (ages 12 to 14) than secondary-aged students (age 15-18). Additional reports showed bullying behavior tended to be higher during elementary school, peaking in middle school, and decreased as students aged (e.g., Goldbaum, Craig, Pepler, & Connolly, 2003; Limber, 2014; Sadinejad et al., 2015; Yubero & Navarro, 2006).

Sex differences. The need to also consider sex differences stems from self-reported differences in the estimates for males and females regarding the social skills used and type of involvement in bullying. Self-report results from existing studies using

nationally representative samples, such as Cook, Williams, Guerra, and Kim (2010), Denny et al. (2015), Finkelhor et al., 2015, Limber (2014) Nansel et al. (2001), Sadinejad et al. (2015), and White and Kistner (2011), reported boys as involved in more bullying than girls. Finkelhor et al. (2015) stated that 20% to 23% of boys reported bullying in contrast to about 12 to 23% of girls. However, the estimates also vary by type of bullying. While boys report more overall bullying, self-report research examining both direct (e.g., physical violence, such as hitting or kicking) (Wang & Iannotti, 2012) and indirect types of bullying reports that girls participate more often in indirect verbal bullying (e.g., Shahaieian et al., 2017; White & Kistner, 2011) and relational bullying (Card et al., 2008; Yubero & Navarro, 2006). This act often happens in conjunction with other forms of bullying. This type most often transpires in younger groups with girls perpetuating the acts. In contrast, boys participate more often in direct types of bullying (Sadinejad et al., 2015; White & Kistner, 2011; Yubero & Navarro, 2006). Boys also often use verbal tactics to display domination over another individual while committing a physical act of bullying. Sadinejad et al. (2015) provided one such example finding that 48% of boys versus 31% of girls reported more involvement in physical fighting.

In summary, these normative trends in the existing bullying literature showed that there continue to be reported differences in the results from self-report bullying studies. There are variations in the self-reported data about the social behaviors that characterize bullies based upon age and developmental stages. Lastly, the results of these studies also show the social behaviors of bullies and non-bullies vary widely based upon sex.

Summary of the Study

Overall, this study used a self-report, broadband assessment to (a) estimate the prevalence of bullying for a nationally representative sample of students age 8 to 18, (b) document how patterns of bullying and related social skills vary with age and sex, and (c) provide a critical perspective regarding the assessment of bullying using self-reports collected with the SSIS. With these purposes in mind, Study 1 used the SSIS Student Rating Scale because of (a) the ease in using self-reports to estimate the prevalence of and the specific social behaviors of bullies across a large age range, (b) the widespread use and established reliability and validity of the SSIS, (c) the availability of data from a large nationally representative data set, and (d) the few empirical studies examining the properties of self-reported bullying assessments.

Research Question and Predictions

The research question and predictions guiding this study were:

Research question 1: What are the demographics and social skills that characterize self-identified bullies?

Research predictions. Current reports about the patterns of bullying and related social skills vary with age and sex. However, those reports outlined some general predictions about what to expect from the results of this study. Regarding sex, males were expected to report more often as bullies than females (e.g., Card et al., 2008; Espelage & Holt, 2001; Habashy Hussein, 2013).

For age level differences, those in younger age groups (8 to 14) were expected to report a higher prevalence of bullies than those in higher age groups (15 to 18) (e.g., Cook et al., 2010; Espelage & Holt, 2001; Nansel et al., 2001).

For social skills and problem behaviors, a positive relationship was expected between bullying and assertion at least for the older age levels for both males and females (12 to 18) (e.g., Elliott et al., 2019). A negative relationship was expected for communication skills for males (ages 8-14) (Carney et al., 2002), but a positive relationship was expected for females in all age groups (e.g., Card et al., 2008; Shahaecian et al., 2017; White & Kistner, 2011). A significant negative relationship was expected between bullying and cooperation for males and females across all age levels (Chui & Chan, 2013; Perren & Alasker, 2006; Wang et al., 2012). Existing research suggested the expectation of a negative relationship with empathy and bullying for males (Elliott et al., 2019; Habashy Hussein, 2013; Jolliffee & Farrington, 2006), but not females (Jolliffee & Farrington, 2011). There also was an expected negative relationship between bullying and self-control for both males and females (Unnever & Cornell, 2003). For problem behaviors, there was an expected positive relationship between bullying and externalizing behavior (e.g., physical fighting) and internalizing behavior (e.g., anxiety) for both males and females.

Method

Participants

The sample used was drawn from a larger study conducted on a nationally representative sample of school-age children, age 8 to 18 (Gresham & Elliott, 2008). The subsample uses data on students with self-report data on bullying behavior as well as self-ratings on social skills and problem behaviors. From the sample of Student self-reports ($n = 1,441$), all self-report cases with at least one adult rater were selected for the analysis. (Note that the adult raters were used in Study 2.) This selection yielded 1,375 valid student self-report cases (see Table 1 for the demographic breakdown of this sample).

Table 1

SSIS Standardization Sample of Valid Self-Report Ratings Age 8 and Above by Sex, Age Group, and Race/Ethnicity (N = 1375)

<i>Female (n = 741)</i>										
	Black		Hispanic		Other		White		Total	
	n	%	n	%	n	%	n	%	n	%
age 8-11	39	3%	90	7%	22	2%	235	17%	386	28%
age 12-14	22	2%	43	3%	13	1%	154	11%	232	17%
age 15-18	13	1%	23	2%	14	1%	73	5%	123	9%
<i>Male (n = 634)</i>										
	Black		Hispanic		Other		White		n	%
	n	%	n	%	n	%	n	%		
age 8-11	33	2%	66	5%	16	1%	207	15%	322	23%
age 12-14	17	1%	37	3%	18	1%	126	9%	198	14%
age 15-18	14	1%	26	2%	12	1%	62	5%	114	8%
<i>Combined (n = 1375)</i>										
	Black		Hispanic		Other		White		n	%
	n	%	n	%	n	%	n	%		
age 8-11	72	5%	156	11%	38	3%	442	32%	708	51%
age 12-14	39	3%	80	6%	31	2%	280	20%	430	31%
age 15-18	27	2%	49	4%	27	2%	135	10%	237	17%

Note. US Census data (2017): Combined: Black (13%), Hispanic (16%), Other (7%), White (64%). Age 8-11: Black (15%), Hispanic (23%), Other (7%), White (55%); Age 12-14: Black (15%), Hispanic (22%), Other (7%), White (56%); Age 15-18: Black (16%), Hispanic (21%), Other (7%), White (56%); percentages are identical for male and female groups.

The racial and ethnic backgrounds were diverse and representative of the U.S. population (U.S. Census data, 2017): 10% Black, 21% Hispanic/Latinx, 62% White, 7% Other. All students volunteered to participate after a parent provided consent for them to participate.

Procedures

The SSIS Technical Manual describes the procedure used to collect the rating scale data (Gresham & Elliott, 2008). Briefly, Pearson Assessment field staff recruited school site coordinators in 115 schools in 36 states, who in turn, recruited participants to fit demographic targets based on the 2006 Current Population Survey (U.S. Census Bureau). These site coordinators distributed and collected rating scales from participating teachers. The teachers also distributed consent letters to parents and students, inviting them to participate. Individuals who gave consent received rating forms. However, the final standardization sample was selected from the larger respondent sample to fit current U.S. Census demographics of age, sex, race/ethnicity, and educational status. The teachers and students both completed their rating scales at school, while the parents did so at home and returned the completed forms directly to the coordinator.

Measures

SSIS rating scale. The SSIS manual also provides detailed information about the reliability and validity of the Bullying subscale and the meaning of derived scores, interpretation indexes (standard error of measurement and confidence intervals), validity indexes, and interpretive reports (Gresham & Elliott, 2008). For example, derived scores include standard scores demonstrating where a person's raw score is in relation to the raw scores for a respective normative group (in this case, male, female, and combined). An

equal-interval scale with a mean of 100 and SD of 15 represents these scores. Thus, a score of 115 is one standard deviation above the mean (Gresham & Elliott, 2008, p. 19). The norms combining males and females were used in the study.

The validity indexes provided by the SSIS ensure the results of the assessment are not compromised by a respondent who does something such as intentionally fakes either a bad or good profile or uses a statistically aberrant response pattern to complete the form quickly (Gresham & Elliott, 2008, p., 21). The F index is one such index (also known as the fake bad index) that determines if a respondent negatively rated a student. This index stems from the frequency of rating social skills as *Never* occurring or a problem behavior as *Almost Always* occurring. If there are high scores on the index, this implies the respondent rated deficient levels of prosocial behaviors or very high levels of maladaptive behaviors or both of these. Any of these scores could indicate a respondent rated the behavior more harshly than it should have been rated. An *F*-index score falls into either an Acceptable, Caution, or Extreme Caution range. Scores outside the Acceptable range could skew the results of the assessment and require further inspection. The SSIS Technical manual provides complete descriptions of the *F*-index items numbers and ratings (p. 21).

The self-report ratings *F*-Index raw scores were examined to establish the validity of the self-rating of the sample by determining the number of cases falling into the acceptable range. All cases with scores of *caution* ($n = 39$) and *extreme caution* ($n = 13$) based on the *F*-index were removed from the sample of Student self-reports ($n = 1,441$). The percentage of scores from all the self-report ratings matched the normal

approximation found in the overall norm sample, 1% to 3% for the Caution Range and about 1% for the Extreme Caution Range (Gresham & Elliott, 2008, p. 21).

Bullying behavior. The Student version of the SSIS Rating Scale was used to estimate the frequency of bullying behavior. The Bullying Subscale comprised five behavior descriptors as items: *I make people do what I want them to do; I hurt people when I am angry; I do not let others join my group of friends; I try to make others afraid of me; and I say things to hurt people's feelings* (Gresham & Elliott, 2008). Students rated each item on a 4-point frequency scale using anchor points of *Not True = 0, A Little True = 1, A Lot True = 2, and Very True = 3*. For this study, I used raw scale scores for each subscale. Bullying total raw scores on the SSIS were based on frequency ratings for the five items on the Bullying Subscale. The items described below in Table 2 for the Student version of the Bullying Subscale and the description of bullying in Table 3 operationalizes the definition of bullying in this study.

Table 2

SSIS Rating Scale--Bullying Subscale Items on Teacher, Parent, Student Forms

Teacher Scale
<ul style="list-style-type: none"> • Bullies others • Does things to make others feel scared • Forces others to act against their will • Keeps others out of social circles • Is aggressive toward people or objects
Parent Scale
<ul style="list-style-type: none"> • Bullies others • Does things to make others feel scared • Forces others to act against their will • Keeps others out of social circles • Is aggressive toward people or objects
Student Scale (Ages 8-18)
<ul style="list-style-type: none"> • I make people do what I want them to do • I hurt people when I am angry • I do not let others join my group of friends • I try to make others afraid of me • I say things to hurt people's feelings

For the Bullying Subscale, there are average and above average ranges; no below average range exists because of a floor effect for scores. That is, informants reported the average students rated in the normative sample to very infrequently exhibit many of the bullying behaviors. They assigned the above average level to all scores falling more than one standard deviation above the mean. A student with an above average level on the Bullying subscale typically demonstrated more than the average number of behaviors for students in his or her norm group. Based upon the criteria established by the SSIS, self-identified bullies were those students with a raw score on the Bullying subscale falling 1SD or higher above the mean (score of five or more out of 15 on the Bullying subscale).

Table 3

Variables

Variable	Description
Demographics	<i>Sex</i> : Male and Female; <i>Race/Ethnicity</i> : Black, Hispanic, Other, and White; <i>Age group</i> : 8 to 18 years
Social Skills	Includes "three domains: (a) Social Skills, (b) Problem Behaviors, and (c) Academic Competence. Social skills represent learned behaviors that promote positive interactions while simultaneously discouraging negative interactions when applied to appropriate social situations" (Gresham & Elliott, 2008, p. 1).
Social Skills Subdomains	
Communication	"Taking turns and making eye contact during a conversation, using appropriate tone of voice and gestures, and being polite by saying "thank you" and "please" (Gresham & Elliott, 2008, p. 1).
Cooperation	"Helping others, sharing materials, and complying with rules and directions" (Gresham & Elliott, p. 1).
Assertion	"Initiating behaviors, such as asking others for information, introducing, oneself, and responding to the actions of others" (Gresham & Elliott, 2008, p. 1).
Responsibility	"Showing regard for property or work and demonstrating the ability to communicate with adults" (Gresham & Elliott, 2008, p. 1).
Empathy	"Showing concern and respect for others' feelings and viewpoints" (Gresham & Elliott, 2008, p. 2).
Engagement	"Joining activities in progress and inviting others to join, initiating conversations, making friends, and interacting well with others" (Gresham & Elliott, 2008, p. 2).
Self-Control	"Responding appropriately in conflict (e.g., disagreeing, teasing) and non-conflict situations (taking turns and compromising)" (Gresham & Elliott, 2008, p. 2).
Problem Behaviors Subscales	
Externalizing	"Being verbally or physically aggressive, failing to control temper, and arguing" (Gresham & Elliott, 2008, p. 2).
Bullying	"Forcing others to do something, hurting people physically or emotionally, and not letting others join an activity" (Gresham & Elliott, 2008, p. 2).
Hyperactivity/ Inattention	"Moving about excessively, having impulsive reactions, and becoming easily distracted" (Gresham & Elliott, 2008, p. 2).
Internalizing	"Feeling anxious, sad, and lonely; exhibiting poor self-esteem" (Gresham & Elliott, 2008, p. 2).

Social skills. The Student version of the SSIS Rating Scale was used to collect data on seven specific skills: (a) Assertion, (b) Communication, (c) Cooperation, (d) Empathy, (e) Engagement, (f) Responsibility, and (g) Self-control. For the total Social

Skills Scale, a standard score of 100 is the mean with a standard deviation of 15.

Students with total Social Skills Scale scores between 85 and 115 are considered in the Average range. Students with scores above 115 are in the Above Average range.

Problem behaviors. The Problem Behaviors Scale consists of five subscales: Autism Spectrum (not used in the study), Bullying (used only to determine status as a bully or non-bully), Externalizing, Hyperactivity/Inattention, and Internalizing. A score of more than 115 on the Problem Behaviors scale indicates that an individual displays more problems than the average individual in that population does.

Plan of Analysis

Key variables. For this study, the independent variables were the demographic variables (Sex: Male and Female), Age (8 to 18), Race/Ethnicity (categories of Black, Hispanic, Other compared to White), Social Skills Total Scale and Subscale scores (Assertion, Communication, Cooperation, Empathy, Engagement, Responsibility, and Self-Control), Problem Behavior Total Scale (minus Autism and Bullying Subscale scores) and Subscale (Externalizing, Internalizing, Hyperactivity/Inattention) scores; the dependent variable is Bully Status and consists of two levels (bully or non-bully) (Table 1.3).

Analytic approach. I fit binary logistic regression models to assess if a student's demographics and social behavior profile predict his/her bully status. The binary logistic regression is an appropriate statistical analysis when the purpose of the research is to investigate if a set of independent variables (e.g., demographics, social skills, and problem behaviors) predict a dichotomous dependent variable (bully or not a bully)

(Stevens, 2012). This type of regression is useful when the independent variables are continuous, discrete, or a combination of the two types of variables.

The analysis allowed investigating a student's odds of being placed into either the bully or non-bully group. The combination of predictor variable values determines the group membership. I evaluated the model using the overall model evaluation and a classification table with the percentage of correct predictions. I examined the overall model significance for the regression models using the χ^2 omnibus test of model coefficients. I also checked the Nagelkerke R^2 to evaluate the percent of variance accounted for by the independent variables. I determined the predicted probabilities of an event occurring using $\text{Exp}(\beta)$. Binary logistic regression analysis was used because the model is less restrictive than linear regression, as the model does not assume normality and equal variances. The model also does not assume the normal distribution of the error term variance. In this regression model, it is assumed the student's status as a bully or non-bully is dichotomous. It was determined if there was multi-collinearity between the independent variables, outliers were examined and analyzed, and there were linear relationships between the odds ratio and the independent variables. A large sample and maximum likelihood were used.

Results

The evidence addressing Research Question 1: What are the demographics and social skills that characterize self-identified bullies? The subsample ($n = 1,375$) of students upon which the evidence is based included 741 females and 634 males

(Table 1.1). From this subsample of cases, 12.1% ($n = 167$) of the students self-identified as bullies, 4.9% ($n = 67$) of females, and 7.3% ($n = 100$) of males (Table 4).

Table 4

SSIS Standardization Sample of Self-Identified Bullies Age 8 and Above by Sex, Age Group, and Race/Ethnicity (N = 167)

<i>Female (n = 67)</i>										
	Black		Hispanic		Other		White		Total	
	n	%	n	%	n	%	n	%	n	%
age 8-11	9	5%	7	4%	0	0%	11	7%	27	16%
age 12-14	3	2%	5	3%	2	1%	15	9%	25	15%
age 15-18	1	1%	3	2%	2	1%	9	5%	15	9%
<i>Male (n = 100)</i>										
	Black		Hispanic		Other		White		n	%
	n	%	n	%	n	%	n	%		
age 8-11	11	7%	15	9%	2	1%	30	18%	58	35%
age 12-14	1	1%	3	2%	3	2%	15	9%	22	13%
age 15-18	4	2%	2	1%	4	2%	10	6%	20	12%
<i>Combined (n = 167)</i>										
	Black		Hispanic		Other		White		n	%
	n	%	n	%	n	%	n	%		
age 8-11	20	12%	22	13%	2	1%	41	25%	85	51%
age 12-14	4	2%	8	5%	5	3%	30	18%	47	28%
age 15-18	5	3%	5	3%	6	16%	19	11%	35	34%

Note: US Census data (2017): Combined: Black (13%), Hispanic (16%), Other (7%), White (64%). Age 8-11: Black (15%), Hispanic (23%), Other (7%), White (55%); Age 12-14: Black (15%), Hispanic (22%), Other (7%), White (56%); Age 15-18: Black (16%), Hispanic (21%), Other (7%), White (56%); percentages are identical for male and female groups.

Descriptive results. The majority of analyses that follow are based on this subsample of self-identified bullies. A mean comparison of self-identified bullies versus non-self-identified bullies, as required by definition, showed self-identified bullies had higher average bullying ratings than non-self-identified bullies (Appendix A). Within the non-self-identified bully group, the standard deviations were relatively close. However,

the self-identified bullies' bullying ratings showed a range of at least one to more than three standard deviations above the mean.

Within both the self-identified and non-self-identified bully groups, overall boys showed higher average bullying ratings than females. However, among the self-identified bullies, females age 8 to 11 showed slightly higher average bullying ratings (.52) than males in the same age group. Among the self-identified bullies, White females age 8 to 11 had the highest average bullying ratings at 8.27 and a standard deviation of 3.38. In addition, among the self-identified bullies, Hispanic males ages 12 to 14 had the highest average bullying ratings at 9.33 with a standard deviation of 3.79. Overall, within the group of self-identified bullies, the average bullying ratings decreased (.54) across ages 8 to 18. However, there was a slight increase of average bullying ratings (.51) between the ages of 12 to 14 and 15 to 18. Interestingly, in the group of non-self-identified bullies, the average bullying ratings increased with age for both males and females. (Further information regarding means and SDs available in Appendix.)

Logistic regression. Binary logistic regression models were used to assess if a student's demographics and social behavior profile predict their bully status. The independent variables used were the categorical variables of Sex and Race/Ethnicity and the continuous variable of Age (ages 8 to 18). I also used Social Skills Scale and Subscale raw scores (Assertion, Communication, Cooperation, Empathy, Engagement, Responsibility, and Self-Control), Problem Behaviors Scale (minus Autism and Bullying Subscale scores) and Subscale (Externalizing, Internalizing, Hyperactivity/Inattention)

raw scores; the dependent variable was Bully Status which consists of two levels (bully or non-bully).

To evaluate the degree of collinearity and multi-collinearity in the set of the independent variables, the Variance Inflation Factor (VIF) was calculated in a regression model including all the variables as predictors. Values of VIF were evaluated based upon the assumption that if the model is strong any values exceeding 10 likely indicate multi-collinearity, or if the model is weaker any values above 2.5 would be examined. An initial check of the variables for multi-collinearity revealed VIF values all under 10. However, when considering the type of model and examining the correlations, VIF values over 2.5 were considered.

The Responsibility and Hyperactivity/Inattention Subscales were more highly correlated (i.e., .65 and higher) with the other variables. These variables caused some of the other variables to have VIF values of 2.9 and higher. After removing both of these subscales from the regression analysis, the removal reduced the VIF values of the remaining variables closer to 2 or less. Given this result, the Responsibility and the Hyperactive/Inattentive Subscales raw scores were excluded from both the Subscale analyses and the Social Skills and Problem Behavior Total raw scores resulting in a more parsimonious model.

Males and females. A binomial logistic regression was performed to determine the effects of Sex, Age, Race/Ethnicity, total Social Skills raw scores, and total Problem Behaviors raw scores on the likelihood that students self-identify as a bully. The logistic regression model was statistically significant, $\chi^2(7) = 428.28, p < .001$. The model

explained 51.2% (Nagelkerke r^2) of the variance in the self-reports of students self-identifying as a bully and correctly classified 90.9% (hit rate) of cases.

Table 5

Logistic Regression Predicting Likelihood of Self-identifying as a Bully based on Sex, Age, Race/Ethnicity, Total Social Skills, and Total Problem Behaviors

	<i>B</i>	SE	Wald	<i>df</i>	<i>p</i>	Odds Ratio	95% C.I. for Odds Ratio	
							Lower	Upper
Sex(1)	0.78	0.222	12.345	1	0.000	2.183	1.412	3.373
Age	0.077	0.04	3.646	1	0.056	1.08	0.998	1.17
Race/Ethnicity			14.786	3	0.002			
Race/Ethnicity(1)	0.983	0.305	10.366	1	0.001	2.672	1.469	4.86
Race/Ethnicity(2)	0.789	0.281	7.871	1	0.005	2.202	1.269	3.821
Race/Ethnicity(3)	0.191	0.422	0.205	1	0.650	1.211	0.529	2.771
Social Skills Total Raw Score	-0.025	0.007	14.29	1	0.000	0.975	0.962	0.988
Problem Behavior Total Raw Score	0.192	0.014	187.321	1	0.000	1.212	1.179	1.245
Constant	-5.426	0.894	36.832	1	0.000	0.004		

Note: Sex is for males compared to females. Race/Ethnicity categories are (1) Black, (2) Hispanic, (3) Other compared to White. Social Skills Total Raw Score excludes the Responsibility subscale. Problem Behavior Raw Score includes only scores for the Externalizing and Internalizing subscales.

Table 5 shows four predictor variables were statistically significant in the analysis: Sex, Race/Ethnicity (1&2), Social Skills, and Problem Behaviors total raw scores. Males had 2.183 times higher odds of self-identifying as bullies than females. For Race/Ethnicity, the odds of self-identifying as a bully were 2.672 times greater for Black students and 2.202 times greater for Hispanic students as opposed to White students. A lower Social Skills total raw score and a higher Problem Behaviors total raw score were associated with an increased likelihood of self-identifying as a bully (Table 5).

A second logistic regression was performed to determine the effects of Sex, Age, Race/Ethnicity, Communication, Cooperation, Assertion, Empathy, Engagement, Self-Control, Externalizing, and Internalizing total raw scores on the likelihood that students

self-identify as a bully. The logistic regression model was statistically significant, $\chi^2(13) = 502.59, p < .001$. The model explained 58.6% (Nagelkerke r^2) of the variance in the self-reports of students self-identifying as a bully and correctly classified 91.6% (hit rate) of cases.

Table 6

Logistic Regression Predicting Likelihood of Self-identifying as a Bully based on Sex, Age, Race/Ethnicity, Social Skills and Problem Behaviors Subscales

	B	SE	Wald	df	p	Odds Ratio	95% C.I. for Odds Ratio	
							Lower	Upper
Sex(1)	0.398	0.244	2.652	1	0.103	1.488	0.922	2.402
Age	-0.009	0.047	0.038	1	0.846	0.991	0.903	1.087
Race/Ethnicity			6.016	3	0.111			
Race/Ethnicity(1)	0.74	0.338	4.802	1	0.028	2.096	1.081	4.064
Race/Ethnicity(2)	0.505	0.315	2.569	1	0.109	1.656	0.894	3.07
Race/Ethnicity(3)	0.236	0.443	0.284	1	0.594	1.267	0.531	3.018
Communication Raw Score	-0.089	0.057	2.468	1	0.116	0.915	0.819	1.022
Cooperation Raw Score	0.019	0.047	0.163	1	0.686	1.019	0.93	1.116
Assertion Raw Score	0.063	0.044	2.067	1	0.151	1.065	0.977	1.161
Empathy Raw Score	-0.111	0.054	4.223	1	0.040	0.895	0.804	0.995
Engagement Raw Score	-0.032	0.042	0.568	1	0.451	0.969	0.891	1.052
Self-Control Raw Score	0.101	0.043	5.482	1	0.019	1.106	1.017	1.204
Externalizing Raw Score	0.378	0.032	135.61	1	0.000	1.459	1.369	1.554
Internalizing Raw Score	0.07	0.023	8.996	1	0.003	1.072	1.025	1.123
Constant	-5.751	1.031	31.147	1	0.000	0.003		

Note: Sex is for males compared to females. Race/Ethnicity categories are (1) Black, (2) Hispanic, (3) Other compared to White.

Table 6 shows that of the predictor variables, only five were statistically significant in the analysis: Race/Ethnicity (1), Empathy, Self-Control, Externalizing, and Internalizing raw scores. For Race/Ethnicity, the odds of self-identifying as a bully is 2.096 times greater for Black students as opposed to White students. Lower raw scores for Empathy increased the odds of self-identifying as a bully. Higher raw scores for Self-Control, Externalizing, and Internalizing and lower raw scores for Empathy increased the odds of self-identifying as a bully.

Females only. A third logistic regression was performed to determine the effects of Age, Race/Ethnicity, total Social Skills and total Problem Behaviors raw scores on the likelihood that female students self-identify as a bully. The logistic regression model was statistically significant, $\chi^2(6) = 209.234, p < .001$. The model explained 54.1% (Nagelkerke r^2) of the variance in the self-reports of female students self-identifying as a bully and correctly classified 93.1% (hit rate) of cases.

Table 7

Logistic Regression Predicting Likelihood of Females Self-Identifying as a Bully based on Age, Race/Ethnicity, Total Social Skills, and Total Problem Behaviors

	B	SE	Wald	df	p	Odds Ratio	95% C.I. for Odds Ratio	
							Lower	Upper
Age	0.137	0.066	4.333	1	0.037	1.147	1.008	1.305
Race/Ethnicity			11.526	3	0.009			
Race/Ethnicity(1)	1.512	0.472	10.275	1	0.001	4.537	1.8	11.439
Race/Ethnicity(2)	0.793	0.434	3.347	1	0.067	2.21	0.945	5.169
Race/Ethnicity(3)	-0.018	0.754	0.001	1	0.981	0.982	0.224	4.308
Social Skills Total Raw Score	-0.038	0.011	12.161	1	0.000	0.963	0.942	0.983
Problem Behaviors Total Raw Score	0.218	0.024	84.106	1	0.000	1.244	1.187	1.303
Constant	-5.897	1.428	17.055	1	0.000	0.003		

Note: Race/Ethnicity categories are (1) Black, (2) Hispanic, (3) Other compared to White. Social Skills Total Raw Score excludes the Responsibility subscale. Problem Behaviors Raw Score includes only scores for the Externalizing and Internalizing subscales.

Table 7 shows that of the predictor variables Age, Race/Ethnicity (1), Social Skills and Problem Behaviors raw scores were statistically significant. Increasing age was associated with an increased likelihood of a female student self-identifying as a bully. For Race/Ethnicity, the odds of self-identifying as a bully were 4.537 times greater for Black females as opposed to White females. A lower Social Skills total raw score and a higher Problem Behaviors total raw score were associated with an increased likelihood of self-identifying as a bully.

A fourth logistic regression was performed to determine the effects of Age, Race/Ethnicity, the Communication, Cooperation, Assertion, Empathy, Engagement, Self-Control, Externalizing, and Internalizing total raw scores on the likelihood that female students self-identify as a bully. The logistic regression model was statistically significant, $\chi^2(12) = 258.009, p < .001$. The model explained 64.6% (Nagelkerke r^2) of the variance in the self-reports of female students self-identifying as a bully and correctly classified 94.7% (hit rate) of cases.

Table 8

Logistic Regression Predicting Likelihood of Females Self-identifying as a Bully based on Age, Race/Ethnicity, Social Skills and Problem Behaviors Subscales

	B	SE	Wald	df	p	Odds Ratio	95% C.I. for Odds Ratio	
							Lower	Upper
Age	0.056	0.082	0.471	1	0.493	1.058	0.901	1.242
Race/Ethnicity			6.851	3	0.077			
Race/Ethnicity(1)	1.44	0.552	6.813	1	0.009	4.222	1.432	12.453
Race/Ethnicity(2)	0.482	0.513	0.882	1	0.348	1.619	0.592	4.423
Race/Ethnicity(3)	0.213	0.816	0.068	1	0.794	1.237	0.25	6.119
Communication Raw Score	-0.12	0.099	1.478	1	0.224	0.887	0.731	1.076
Cooperation Raw Score	0.114	0.079	2.074	1	0.150	1.121	0.96	1.31
Assertion Raw Score	-0.024	0.067	0.132	1	0.716	0.976	0.855	1.113
Empathy Raw Score	-0.049	0.093	0.28	1	0.597	0.952	0.792	1.143
Engagement Raw Score	-0.088	0.072	1.507	1	0.220	0.916	0.796	1.054
Self-Control Raw Score	0.067	0.072	0.876	1	0.349	1.07	0.929	1.231
Externalizing Raw Score	0.491	0.061	65.722	1	0.000	1.634	1.451	1.84
Internalizing Raw Score	0.062	0.037	2.784	1	0.095	1.064	0.989	1.146
Constant	-7.43	1.757	17.88	1	0.000	0.001		

Note: Race/Ethnicity categories are (1) Black, (2) Hispanic, (3) Other compared to White.

Table 8 shows that Race/Ethnicity (1) and Externalizing raw scores were statistically significant. For Race/Ethnicity, the odds of self-identifying as a bully were 4.222 times

greater for Black females as opposed to White females. A higher Externalizing raw score was associated with an increased likelihood of self-identifying as a bully.

Males only. A fifth logistic regression was performed to determine the effects of Age, Race/Ethnicity, total Social Skills and total Problem Behaviors raw scores on the likelihood that male students self-identify as a bully. The logistic regression model was statistically significant, $\chi^2(6) = 212.837, p < .001$. The model explained 49% (Nagelkerke r^2) of the variance in the self-reports of male students self-identifying as a bully and correctly classified 87.9% (hit rate) of cases.

Table 9

Logistic Regression Predicting Likelihood of Males Self-identifying as a Bully based on Age, Race/Ethnicity, Total Social Skills, and Total Problem Behaviors

	B	SE	Wald	df	p	Odds Ratio	95% C.I. for Odds Ratio	
							Lower	Upper
Age	0.03	0.052	0.342	1	0.559	1.031	0.931	1.142
Race/Ethnicity			5.326	3	0.149			
Race/Ethnicity(1)	0.596	0.409	2.119	1	0.146	1.814	0.814	4.045
Race/Ethnicity(2)	0.778	0.379	4.214	1	0.04	2.178	1.036	4.578
Race/Ethnicity(3)	0.31	0.515	0.362	1	0.548	1.363	0.497	3.738
Social Skills Total Raw Score	-0.017	0.009	3.808	1	0.051	0.983	0.966	1
Problem Behaviors Total Raw Score	0.18	0.018	98.562	1	0.000	1.198	1.156	1.241
Constant	-4.39	1.106	15.762	1	0.000	0.012		

Note: Race/Ethnicity categories are (1) Black, (2) Hispanic, (3) Other compared to White. Social Skills Total Raw Score excludes the Responsibility subscale. Problem Behaviors Raw Score includes only scores for the Externalizing and Internalizing subscales.

Table 9 shows that of the predictor variables Race/Ethnicity (2) and the Problem Behavior raw score were statistically significant. For Race/Ethnicity, the odds of self-identifying as a bully were 2.178 times greater for Hispanic males as opposed to White males. A higher Problem Behaviors total raw score was associated with an increased likelihood of self-identifying as a bully.

A sixth logistic regression was performed to determine the effects of Age, Race/Ethnicity, the Communication, Cooperation, Assertion, Empathy, Engagement, Self-Control, Externalizing, and Internalizing total raw scores on the likelihood that male students self-identify as a bully. The logistic regression model was statistically significant, $\chi^2(12) = 249.675, p < .001$. The model explained 56% (Nagelkerke r^2) of the variance in the self-reports of male students that self-identify as a bully and correctly classified 87.9% (hit rate) of cases.

Table 10

Logistic Regression Predicting Likelihood of Males Self-identifying as a Bully based on Age, Race/Ethnicity, Social Skills and Problem Behaviors Subscales

	<i>B</i>	SE	Wald	<i>df</i>	<i>p</i>	Odds Ratio	95% C.I. for Odds Ratio	
							Lower	Upper
Age	-0.057	0.061	0.875	1	0.35	0.944	0.838	1.065
Race/Ethnicity			1.855	3	0.603			
Race/Ethnicity(1)	0.287	0.445	0.415	1	0.519	1.332	0.557	3.189
Race/Ethnicity(2)	0.548	0.42	1.702	1	0.192	1.729	0.759	3.938
Race/Ethnicity(3)	0.152	0.556	0.075	1	0.784	1.165	0.392	3.464
Communication Raw Score	-0.111	0.073	2.33	1	0.127	0.895	0.776	1.032
Cooperation Raw Score	-0.037	0.058	0.394	1	0.53	0.964	0.86	1.081
Assertion Raw Score	0.132	0.061	4.747	1	0.029	1.141	1.013	1.285
Empathy Raw Score	-0.143	0.069	4.342	1	0.037	0.867	0.758	0.992
Engagement Raw Score	-0.001	0.054	0	1	0.984	0.999	0.898	1.111
Self-Control Raw Score	0.137	0.058	5.56	1	0.018	1.147	1.023	1.285
Externalizing Raw Score	0.327	0.041	63.814	1	0.000	1.387	1.28	1.503
Internalizing Raw Score	0.082	0.031	6.855	1	0.009	1.085	1.021	1.154
Constant	-4.5	1.279	12.385	1	0.000	0.011		

Note: Race/Ethnicity categories are (1) Black, (2) Hispanic, (3) Other compared to White.

Table 10 shows that of the predictor variables Assertion, Empathy, Self-Control, Externalizing, and Internalizing raw scores were statistically significant. A lower Empathy raw score and higher Assertion, Self-Control, Externalizing, and Internalizing raw scores were associated with an increased likelihood of self-identifying as a bully.

Discussion

The purpose of this study was to use a self-report, broadband assessment to (a) estimate the prevalence of bullying for a nationally representative sample of students age 8 to 18, (b) document how patterns of bullying and related social skills vary with age and sex for a diverse sample, and (c) provide a critical perspective regarding the assessment of bullying using self-reports collected with the SSIS. The research question guiding this study was: *What are the demographics and social skills that characterize self-identified bullies?*

The study reiterated that bullying is a problem among youth in the United States and the assumption that reliable estimates for the percentage of a population identified as displaying bullying behavior falls into the middle (12.1%) of the estimated range, a low of 5% to just over 30% (Finkelhor et al., 2015; Iannotti, 2013; Nansel et al., 2004). As predicted, consistent with previous studies, and as required by definition, self-identified bullies showed higher average bullying ratings than non-self-identified bullies with self-identified bullies having much higher bullying ratings than those students who did not identify as one (Nansel et al. 2015). The results also demonstrated that Sex, Age, Race/Ethnicity, total Social Skills raw scores, and total Problem Behaviors and the Subscale scores for each scale explained more than half of the variance in the self-reports

of students self-identifying as bullies. All of the models, examining males and females combined and separately, correctly classified students as bullies or non-bullies in about 90% of the cases. These results may reflect that examining these combined indicators will help in the early detection of children and youth who might participate in bullying.

The results also supported the framework of the study that bullies do not comprise a homogeneous group, as they corresponded to current reports about the patterns of bullying and related social skills, which vary with age and sex. Regarding sex, as with previous studies, males reported more often as bullies than females (e.g., Card et al., 2008; Espelage & Holt, 2001; Habashy Hussein, 2013; White & Kistner, 2011). However, an interesting exception to this prediction was for female self-identified bullies age 8 to 11 who showed slightly higher average bullying ratings (.52) than males in the same age group. This result may be indicative of higher frequencies of indirect verbal and relational bullying in this group of self-identified female bullies, which are the types of bullying that more often transpire for females in this age group (e.g., Card et al., 2008; Shahaiean et al., 2017; White & Kistner, 2011; Yubero & Navarro, 2006) The overall results about males and females also held when examining the total social skills and problem behavior trends for combined males and females where males were more than twice as likely to self-identify as bullies than females.

The results met general expectations for age level differences, as those in younger age groups (8 to 14) reported a higher frequency of bullies than those in higher age groups (15 to 18) (e.g., Cook et al., 2010; Espelage & Holt, 2001; Nansel et al., 2001). Overall, within the group of self-identified bullies, the average bullying ratings decreased

(.54) across ages 8 to 18. However, there was an exception to this general trend for students between the ages of 12 to 14 and 15 to 18, as there was a slight increase of average bullying ratings (.51). In the group of non-self-identified bullies, the average bullying ratings increased with age for both males and females. Yet, the interesting result was that age was not shown as a significant predictor of identifying as a bully. Age was only a significant predictor when examining the total social skills and problem behaviors scores for females alone. Interestingly, instead of decreasing age, it was increasing age that was associated with a higher likelihood of a female student self-identifying as a bully. These findings may be related to specific group differences. There is research indicating that some groups of secondary aged students, such as older females, are participating more in other kinds of bullying. For example, cyberbullying appears to occur more frequently in the last three years of secondary school than it does during grades six to nine (NCES, 2015).

The results for Race/Ethnicity also revealed some interesting findings. The model examining combined norms for males and females on the Total Social Skills and Problem Behaviors scales scores determined that the odds of self-identifying as a bully were more than twice as high for Black students and Hispanic students as opposed to White students. In the model using the subscale scores, the results showed that the odds of self-identifying as a bully were only twice as high for Black students as opposed to White students.

The results also were different when looking at males and females separately. Among the self-identified bullies, White females age 8 to 11 had the highest average

bullying ratings at 8.27 and a standard deviation of 3.38. However, older Black females (age 12 to 18) had higher average bullying ratings in comparison to White females in the same age group. Of note was that in the model examining the total social skills and problem behaviors scales scores for females overall, the results showed the odds of self-identifying as a bully were more than four times greater for Black females as opposed to White females. For males, the descriptive results corresponded more closely to the model outcome. Hispanic males ages 12 to 14 had the highest average bullying ratings at 9.33 with a standard deviation of 3.79, while the odds of self-identifying as a bully were also more than twice as high for Hispanic males as opposed to White males when examining the model for the total scale scores. These findings may be indicative of internalized feelings about racial identity and other behaviors for Black females and Hispanic males.

The results for the combined norms for males and females corresponded with the expected results that lower Social Skills total raw scores and higher Problem Behaviors total raw score were associated with an increased likelihood of self-identifying as a bully. When examining the subscale scores for both combined males and females and then separately, the results diverged from some of the predicted outcomes. For problem behaviors, there was an expected positive relationship between bullying and externalizing behavior (e.g., physical fighting) and internalizing behavior (e.g., anxiety) for both males and females. The expected relationship held for the combined male and female sample where higher externalizing and internalizing scores increased the odds of self-identifying as a bully. This relationship held for males when examining males and females separately. However, for females, only higher scores for externalizing behavior were a

significant predictor for self-identifying as a bully. The results may be indicative of existing social norms and the stereotyping of bullying behavior that males are typically bullies and mostly only bully other males. However, research suggests females can be strong bullies as well (Chen, 2019). A female's behavior may be overlooked because of existing social norms about the frequency of female bullying, or the female bully's behavior might be passive and quiet in comparison to the behavior associated with male bullying.

For the Social Skills Subscales, a positive relationship was expected between bullying and assertion at least for the older age levels for both males and females (12 to 18) (e.g., Elliott et al., 2019). However, the results showed that higher assertion scores were only predictive of self-identifying as a bully when separately examining males. These results support the limited research about assertive behavior that suggests some bullies are more socially savvy and popular than their peers and can skillfully use their behavior to control their peers (e.g., Boulton & Smith, 1994; Elliott et al., 2019; Jenkins et al., 2017; Juvonen et al., 2003; Stephenson & Smith, 1989).

A negative relationship was expected for communication skills for males (ages 8-14) (Carney et al., 2002), but a positive relationship was expected for females in all age groups (e.g., Card et al., 2008; Shahaieian et al., 2017; White & Kistner, 2011). Surprisingly, communication was not a significant predictor of self-identifying as a bully for any group. However, while the communication subscale score was not a specific predictor of self-identifying as a bully, those who do self-identify as bullies might still display higher ratings for communication than non-bullies.

The expectation of a negative relationship with empathy and bullying for males (Elliott et al., 2019; Habashy Hussein, 2013; Jolliffee & Farrington, 2006), but not females (Jolliffee & Farrington, 2011) corresponded to the results. In both the models, combined males and females and males alone, having lower empathy scores increased the odds of self-identifying as a bully. However, this was not the case for females. Given that the definition of empathy in the SSIS involves "showing concern and respect for others' feelings and viewpoints" (Gresham & Elliott, 2008, p. 2), these results support existing research that self-identified bullies likely possess the ability to recognize the emotions and behaviors of others, but they may have problems managing their own emotions (Espelage et al., 2000; Sutton et al., 1999). The self-identified bullies may be able to identify emotions, but they may not identify with emotions in the same way as their peers do, perhaps making the self-identified bullies more likely to manipulate emotions of others in a social situation (Sutton et al., 1999; Warden & Mackinnon, 2003).

The limitations of the results from existing studies left it unknown exactly how age and sex influence the relationship between bullying and self-control (Jenkins et al., 2016). Given that most existing research shows a general relationship between low self-control and bullying (Chui & Chan, 2013; Pellegrini et al., 1999; Unnever & Cornell, 2003), there was an expected negative relationship between bullying and self-control for both males and females (Unnever & Cornell, 2003). However, the opposite result occurred because males showed that higher self-control ratings predicted a greater likelihood of self-identifying as a bully. Self-control was a not a significant predictor for

female behavior. These results suggest new evidence for how a person's sex may influence the relationship between bullying and self-control.

Strengths, Limitations, and Future Directions

A limitation was the cross-sectional nature of the data, so the direction of the results remains undetermined. While self-reports are a valid measure, they still represent a limitation because of potential variation across individuals. However, the subjectivity of the study was reduced because a specific definition of bullying was included in the assessment. The SSIS Rating Scale is a useful broadband assessment that assesses an array of social skills, several classes of problem behaviors including bullying, and academic competencies that is tied directly to an intervention. However, the SSIS does not capture in-depth details about the specific instances of when a child is participating in bullying. A strength of the study, missing from similar studies (Nansel et al., 2001), is that the sample incorporates a wide age range including elementary--, middle- and high school-aged youth. Another minor limitation was that there are two items overlapping on the Bullying (five total items) and Externalizing (12 total items) Subscales. However, these items account for less than 20% of the total loading items.

The differences emerging from the results might be used to develop future research targeting specific groups. For example, assertive behavior and self-control were predictive indicators of a self-identified male bully. A closer examination of these social skills in groups of males could help to identify situations where males are more likely to participate in bullying. Lastly, while this study provides notable data on the frequency and social behavioral connections of children who self-identify as bullies among youth in

the United States, more research is necessary. Of particular worth would be more studies examining the specific social behaviors connected to students who bully others. In addition, studies with a longitudinal design would help improve knowledge about specific bullying incidents involving those youth who identify as bullies.

Conclusion

This study provided important data about the frequency of and the perceived social behaviors of self-identified bullies. However, the results also introduced more questions that require further research. The higher frequency of self-identified bullies observed in the data indicates the value of increased preventive intervention research examining the social behaviors of children and youth who self-identify as bullies. The differences between groups and behaviors also indicates the necessity of future studies examining specific groups within the larger sample to improve understanding about the differences. Improved prevention efforts will require this improved knowledge about the specific contexts and social behaviors that both enable and prevent the development of bullying. The improved knowledge may be used to advance efforts to develop effective individual and school-based interventions that address bullying and the related consequences.

CHAPTER 3

STUDY #2

AN EXAMINATION OF THE ADDED VALUE OF ADULT INFORMANTS' RATINGS FOR THE IDENTIFICATION OF BULLIES

The purpose of this study was to assess how well behavior ratings by self-identified bullies on the Social Skills Improvement System (SSIS) Rating Scales corresponded with the ratings of these students by their parents and teachers. Thus, the study addressed the frequent recommendation by researchers to use multiple informants to assess bullying behaviors of students (e.g., Crothers & Levinson, 2004; Griffin & Gross, 2004; Demaray, Malecki, Secord, & Lyell, 2013; Groeben, Perren, Stadelmann, & von Klitzing, 2011; Juvonen, Nishina, & Graham, 2001; Nowell, Brewton, & Goin-Kochel, 2014; Phillips & Cornell, 2012). These referenced studies using multiple informants have resulted in important information to answer questions that self-report data alone could not have answered. However, they are limited in scope in that they typically have focused on comparing the behavior of bullies and non-bullies and behavior outcomes (e.g., Sourander et al., 2011; Veenstra et al., 2005; Winsper, Lereya, Zanarini, & Wolke, 2012; Wright & Lee, 2013). They also have focused on the prevalence of bullying (Lee, Smith, & Monks, 2016; Vlachou, Botsoglou, & Andreou, 2013). These studies generally have not focused on how the addition of more informants might improve the identification of bullies.

Additionally, most assessments (e.g., parent, teacher, and self) used have varied in their reliability and validity for use in assessing bullying and any related social behaviors (Bowes et al., 2013; Fisher et al., 2012). Specifically, the use of these different assessments, many of which were not designed for multiple informants, and influenced by varying definitions of bullying, make it difficult to compare findings across studies and to replicate results. The current study contributed knowledge to the existing bullying literature by using multiple informant data collected with the SSIS Rating Scales, which is one of the few multi-rater assessment systems where all participants complete items with the same root language. Thus, this study has the potential to fill a gap in the research by investigating the potential additive value of using multiple informants to assess bullying and related social behaviors.

The Value of Using Multiple Informants

The use of multiple informants to assess the construct and specific behaviors of bullying remains limited (e.g., Demaray et al., 2013; Nowell et al., 2014; Phillips & Cornell, 2012). The expectation when using multiple informants is that the recent memories of two or more people who have observed a child's behavior in different situations – home, school, community – will result in a more complete and more representative picture of the child's behavior repertoire. The challenges of using multiple informants are the meaningful integration and interpretation of the results of all informants. Because of these challenges and the fact that few true multi-rater systems of bullying behavior have been validated, the use of self-reports remains the most common method for gathering data about the behavior of children who bully others (Ledwell &

King, 2015). While the use of these self-reports from students is important in the assessment of children's bullying behaviors, it is argued that the ratings from other informants like parents and teachers are also important (e.g., Demaray et al., 2013; Groeben et al., 2011; Nowell et al., 2014; Phillips & Cornell, 2012). They are especially important with children under the age of 8 who often cannot adequately comprehend items and respond to common behavior rating scale formats. The reasoning for including multiple informants, in particular adults, is that with informant-based assessments it is unlikely a single measure can be both completely free of bias and capture all relevant information about a subject (Carrasco, Holgado, Del Barrio, & Barbero, 2008; De Los Reyes & Kazdin, 2004; Flake & Petway, 2019).

An important aspect of using a multi-informant approach concerns the differences in what one of these sources of information (e.g., teacher, parent, self-report) contributes to the assessment in relation to the other sources of information. Conceptually, adding an informant to an assessment reliably expands the predictive power and decision-making of it, thus improving its validity evidence incrementally (Hunsley & Mash, 2005). Evidence of incremental validity when using multiple assessment informants, however, is not often empirically tested in the assessment of students or adults (Hunsley, 2002), especially not in bullying assessment.

Therefore, one could argue that even if a measure includes strong psychometric properties, it does not necessarily mean it is sufficient for ensuring the incremental validity of using additional measures (e.g., another informant reporting form) in the assessment of a particular construct like bullying. Besides, not only is the research limited that deals directly with incremental validity in assessment, the incremental validity of parent vs. self-report and teacher vs. self-report in bullying is also limited.

Some existing bullying research supports the inclusion of ratings from parents and teachers into assessment because their addition provides valuable information about the assessed construct (Shiner & Allen, 2013). Some multiple informant studies examining children's problem behaviors also provide support for the value of including an adult's perspective over that of the child (Carrasco et al., 2008). The value in the adult perspective becomes apparent across childhood and adolescence. For example, during adolescence, children typically spend more time at school than with family (Shiner & Allen, 2013). It can be argued that teachers, because of their professional training and opportunities to observe many children in many situations, are preferable for identifying students who have bullying tendencies. Teachers likely have experiences distinguishing developmentally normative aggression from bullying behaviors and would appear to have little motivation to be biased toward under or over-reporting of such behaviors (Camodeca, Caravita, & Coppola, 2015; Monks, Smith, & Swettenham, 2003; Verlinden et al., 2014; Vlachou, Andreou, Botsoglou, & Didaskalou, 2011). Without the addition of information from an adult informant, such as a parent or teacher, during these different developmental stages, it might be hard to know when a behavior might only be reflective

of a single occurrence (Hebron & Humphrey, 2014). Therefore, single informant reports by themselves may not offer a complete narrative or picture of bullying behavior.

Another consideration of the value of including multiple informants when assessing bullying behavior is to minimize errors in reporting (Shiner & Allen, 2013). Errors happen when an individual's report about his/her behavior is less accurate because she/he cannot gauge personal actions as well as another person (e.g., a parent) (Gromann, Goossens, Olthof, Pronk, & Krabbendam, 2013). A youth may underreport specific actions due to social desirability, while an adult report could be subject to personal bias regarding that individual or limited because the informant only interacts with an individual in a single context (Henrich & Shahar, 2014). In addition, with older youth, some evidence exists that there is incremental value in using self-reports over adult reports for internalizing and covert behaviors because the adult reports often miss these behaviors. Any of these are possible concerns with the potential for negatively affecting score reliability and validity. Despite the potential for low to moderate agreement between raters, there is still value in examining alternative ratings. This value is especially present when all raters are responding to the same behaviors.

When considering the value of the information provided by each informant, there is always a question about the agreement between informants. A consistent note about child assessment is that informant agreement is operationalized with correlations and they typically are in the low to moderate range (e.g., .20 to .40) (Achenbach, McConaughy, & Howell, 1987; De Los Reyes et al., 2015; Flake & Petway, 2019). Results from relevant studies typically show informants who observe students in the same context (e.g., school)

show more agreement than informants' who observe the students in different contexts (e.g., teacher and parent). There also is usually more agreement regarding externalizing behaviors (e.g., aggression, bullying, hyperactivity/inattention) than for internalizing behaviors (e.g., anxiety) (Achenbach et al., 1987). However, these differences in agreement vary across studies. The agreement may reflect both informant effects and how each source of information contributes to the assessment, which contributes to a need for properly analyzing these differences (Flake & Petway, 2019).

In summary, although no score from an assessment is without error, including multiple sources of data to categorize bullying behavior, in theory, could reduce errors and thus increase the reliability of the data. Even if informants' reports and ratings show differences, those differences might identify and potentially explain the variance and actions that vary by situation. Overall, the inclusion of different informants has the potential to improve the identification, understanding, and analysis of assessed behavior. It is likely that it does so by allowing documentation of a higher percentage of the behaviors in a more significant number of situations.

Considerations When Choosing an Assessment

While bullying researchers note the importance of multiple informants in assessing bullying, they also report finding only general levels of consistency and interpretability across studies with many of the existing assessments. Within the multiple informant studies focused on bullying, researchers have chosen from several of the same existing assessments, and similar adaptations, to measure bullying. The most commonly chosen assessments included the Olweus Bully/Victim Questionnaire (BVQ) (Olweus,

1996); the Berkeley Puppet Interview (BPI; Ablow & Measelle, 1993); the Achenbach System of Empirically Based Assessment (ASEBA; Achenbach, & Rescorla, 2001): the Child Behavior Checklist (CBCL), which includes the Teacher Report Form (TRF), and the Youth Self-Report (YSR); The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997); the Bullying and Friendship Interview Schedule (Wolke, Woods, Bloomfield & Karstadt, 2000); the Social Behavior Questionnaire (SBQ; Tremblay et al., 1991); and an interview protocol developed by Monks et al. (2003).

Even though most of these assessments used in multi-rater studies are based on developmentally appropriate methods and surveys, there remains no established criteria or set of standards about the psychometric properties of relevant assessments specifically measuring bullying (Blake, Banks, Patience & Lund, 2014; Olweus, 2013; Volk, Veenstra, & Espelage, 2017). Each set of rating scales has different evidence to make conclusions about reliability and validity, areas of focus, age, and sometimes sex norms, and the type of informant used to collect the data. The empirical evaluation of many of these assessment instruments is limited. While several have substantial psychometric evidence regarding certain types of score validity (e.g., concurrent validity) and reliability for their score inferences, the results from those studies vary in establishing strong reliability and validity for using the assessment to study bullying and related social behaviors (Ezpeleta et al., 2013; Lee & Cornell, 2010; Waller, Hyde, Baskin-Sommers, & Olson, 2017). Many of these multiple measures only measure one or a few dimensions of social skills. Few use an approach with a focus on multiple domains of prosocial behavior, while at the same time assessing problem behaviors (e.g., bullying) and

academic competence across a wide age range. Also, many of the assessments facilitate the use of multiple informants; however, few were standardized and normed for use with multiple informants (e.g., teacher, parent, student).

The Social Skills Improvement System (SSIS) is the best example of a truly multi-informant assessment of children's social behavior that includes specific bullying behaviors (Gresham & Elliott, 2008). The SSIS Rating Scale is a broadband assessment. That is, it assesses an array of social skills, several response classes of problem behaviors including bullying, and academic competencies. This assessment was designed to link results directly to intervention through the identification of observable target behaviors with known social importance for functioning in schools and at home (Demaray et al., 1995).

In summary, there is a need for a study examining bullying and related social behaviors using an assessment designed and tested for use with multiple informants. Such a study would contribute to developing a stronger foundation in the literature about the kind and value of the information each informant contributes to an assessment. The present study provides such support by using the SSIS as a multi-rater assessment to examine the added value of including more informants in the study of the social behaviors characterizing bullying.

Purpose of the Study

This study addressed the lack of information about the incremental validity of adding adults' behavior ratings of students to the students' self-perceptions about their status as a bully and the related social behaviors. The agreement in ratings also was

examined across these multiple informants. Thus, the aim of the study was to (a) explore the incremental validity of including the ratings of teachers and parents to the social behavior assessment of students who self-identify as bullies and (b) determine the extent of agreement between the adult and student ratings of social skills and problem behaviors.

Research Questions and Predictions

Specifically, I addressed two questions:

Research Question 1: How accurate are parents and teachers at identifying students who self-identify as bullies?

Research Question 2: What is the relative value in including the ratings from teachers and parents to those of students who self-identify as bullies for predicting the social skills and problem behaviors of the identified bullies?

There were overlapping hypotheses for both research questions.

Hypothesis 1: Parents and teachers were expected to add to the overall predictive validity of the student assessments.

Limited research exists specifically showing how uniquely useful ratings from individual informants are in providing additional information in predicting behavior outcomes, in particular when accounting for other informants' ratings. There have been recommendations to build on the findings from studies using self-report assessment to determine if the outcomes from those studies would be replicated by incorporating the ratings from more informants (Martino, Ellickson, Klein, McCaffrey, & Edelen, 2008). Existing research has focused more broadly on identifying maladaptive behaviors like

aggression and externalizing behaviors and the prediction of adult behaviors rather than on bullying (e.g., Di Giunta et al., 2010; Hill, Coie, Lochman, & Greenberg, 2004; Lochman, 1995). However, those studies did show utility in including multiple informants for predicting both behavior outcomes and improving the classification of participants. Specifically, Lochman (1995) reported improved classification and prediction of externalizing behavior outcomes for first graders when including the parent ratings to the teachers' ratings in the assessment. In addition, Di Giunta et al. (2010) found value in including both the children's and mothers' reports of aggression in predicting and classifying overt and covert antisocial behaviors for young adults. Given the existing research, I expected that adding the adult ratings to the self-reports of the identified bullies would improve the accuracy of classifying the students as bullies and in predicting their social behavior characteristics.

Hypothesis 2: It was expected that there would be modest overlap observed between informants' reports, resulting in low to moderate agreement across ratings and differences in the social behavior characteristics reported.

This expectation stemmed from existing research theorizing that adults and students see behaviors in various situations using different interpretive lenses. Research using multiple informants in general and in bullying assessment consistently has shown the type of informant influences the results of an assessment (e.g., Achenbach et al., 1987; Rønning et al., 2009; Rupp, Elliott, & Gresham, 2018; Wegge, Vandebosch, Eggermont, Van Rossem, & Walrave, 2016). These results often lead to differences observed between the reports of different informants, resulting in low to moderate

agreement and differences in the behavior characteristics reported. However, some recent research has found value in both the agreement and disagreement between the ratings found, as they can be predictive of a variety of behavior outcomes (De Los Reyes, 2011). While research focused specifically on individual social skills and problem behaviors is limited, a few studies have highlighted the potential value in examining agreement and its predictive value. For example, Ferdinand, Van der Ende and Verhulst (2004; 2006) examined informant agreement/disagreement from the ratings of children and parents (they used the CBCL and YSR). The results revealed parents rated more problem behaviors than students' self-reports of problem behaviors (e.g., attention problems) and were predictive of a greater likelihood of certain behavior outcomes such as delinquency. Another study of Ferdinand, Van der Ende, and Verhulst (2007) examined agreement in the ratings between teachers and parents found differences in the ratings of aggression, with parents providing higher ratings than teachers and predicting greater risk for suicidal ideation in the children. The authors proposed predictive utility in examining the rating agreement because the similarities or differences in the ratings would provide useful information. The results highlighted similarities and differences in behavior across contexts and the communication between the dyads of parents and students and teachers and students, which in turn influenced related outcomes.

Method

Participants

The sample used was drawn from a larger study conducted on a nationally representative sample of school-age children, age 8 to 18 (Gresham & Elliott, 2008). The

subsample uses data on students with self-report data on bullying behavior as well as self-ratings on social skills and problem behaviors. The self-report ratings on items comprising the Fake or *F*-Index were examined to establish the validity of the sample by determining the number of cases falling into the acceptable range. All cases with scores of *caution* ($n = 39$) and *extreme caution* ($n = 13$) based on the *F*-index were removed from the sample of Student self-reports ($n = 1,441$). The percentage of scores from all the self-report ratings matched the normal approximation found in the overall norm sample, 1% to 3% for the Caution Range, and about 1% for the Extreme Caution Range (Gresham & Elliott, 2008, p. 21). From the sample of Student self-reports ($n = 1441$), all self-report cases with at least one adult rater were selected for the analysis. This selection yielded 1,375 valid student self-report cases. Within those 1,375 cases, there were 424 cases with all three raters (self, parent, and teacher), 922 with a self-rater and a parent rater, and 29 with a self-rater and a teacher rater. The subsample included 741 females and 634 males, and the demographics for those cases are provided in Table 1.1 (see page 34).

The subsample of self-identified bullies was also derived from those 1,375 cases, and it includes those scoring one standard deviation or more above the mean on the Bullying Subscale (a score of five or more out of 15). The Bullying Subscale raw scores on the SSIS used to determine a participant's status as a bully or non-bully were based on frequency ratings for the five items on the Bullying Subscale. From this sample of cases, 12.1% ($n = 167$) of the students self-identified as bullies, with 4.9% ($n = 67$) females and 7.3% ($n = 100$) males (See Table 1.4 on page 42). The racial and ethnic backgrounds were representative of the U.S. population (U.S. Census data, 2017): 10% Black, 21%

Hispanic/Latinx, 62% White, 7% Other. All students volunteered to participate after a parent provided consent for them to participate.

Due to the smaller sample size, a post hoc statistical power analysis using G*Power 3.1 was also performed for computing the achieved power of the Race/Ethnicity variable included in the analysis. The effect size in this study was 0.37, considered to be medium using Cohen's (1988) criteria. With an alpha = .05, the sample size of $n = 167$, and the number of groups ($n = 4$), the achieved power is computed as 0.985 ($F = 2.66$) based upon ANOVA: fixed effects, omnibus, one-way testing. Therefore, the results suggest the sample of self-identified bullies appears adequate for the main objective of the study in examining agreement/disagreement across raters.

Procedures

The SSIS Technical Manual describes the procedure used to collect the rating scale data (Gresham & Elliott, 2008). Briefly, Pearson Assessment field staff recruited school site coordinators in 115 schools in 36 states, who in turn, recruited participants to fit demographic targets based on the 2006 Current Population Survey (U.S. Census Bureau). These site coordinators distributed to and then collected rating scales from teachers. The teachers also distributed consent letters to parents and students, inviting them to participate. Individuals who gave consent received rating forms. However, the final standardization sample was selected from the larger respondent sample to fit U.S. Census demographics of age, sex, race/ethnicity, and educational status. The teachers and students both completed their rating scales at school, while the parents did so at home and returned the completed forms directly to the coordinator.

Measures

SSIS rating scale. The SSIS Technical Manual (Gresham & Elliott, 2008) provides detailed information about the reliability and validity of the Scales and Subscales and the meaning of derived scores, interpretation indexes (standard error of measurement and confidence intervals), validity indexes, and interpretive reports. For example, derived scores include standard scores (Mean = 115, SD = 15) demonstrating where a person's raw score is in relation to the raw scores for a respective normative group (in this case, male, female, and combined). Thus, a standard score of 115 is one standard deviation above the mean and a score of 85 one standard deviation below the mean (Gresham & Elliott, 2008, p. 19).

The validity indexes provided by the SSIS maximizes the likelihood that the results of the assessment are not compromised by a respondent who does something such as intentionally fakes either a bad or good profile or uses just one response to complete the form quickly (Gresham & Elliott, 2008, p., 21). The F-index is one such index (also known as the fake bad index) that determines if a respondent negatively rated a student. This index stems from the frequency of rating five social skills items as *Never* occurring or five problem behavior items as *Almost Always* occurring. If there are high scores on the index, this implies the respondent rated deficient levels of prosocial behaviors or very high levels of maladaptive behaviors or both of these. Any of these scores could indicate a respondent rated the behavior more harshly than it should have been rated. An *F*-index score falls into either an Acceptable, Caution, or Extreme Caution range. Scores outside the Acceptable range could skew the results of the assessment and require further

inspection. The SSIS Technical manual provides complete descriptions of the *F*-index item numbers and ratings (p. 21). Scores indicative of caution or extreme caution were removed and not included in the sample used for analysis.

Bullying behavior. The Teacher, Parent, and Student versions of the SSIS Rating Scale were used to examine bullying. (Note: the items of the student scale are written as “I” statements, but they share the same root language as the adult versions). The adult scales comprised 5-behavior descriptors as items: *Bullies others; Does things to make others feel scared; Forces others to act against their will; Keeps others out of social circles; Is aggressive toward people or objects* (Gresham & Elliott, 2008) (see Table 1.3, page 39). The items and scoring of the Social Skills and Problem Behaviors scales for the teacher and parent versions use a 4-point frequency scale with *Never = 0, Seldom = 1, Often = 2, and Almost Always = 3* as scale anchor points. The Student version of the scale also uses a 4-point frequency scale for the SS and PB Scales with anchor points of *Not True = 0, A Little True = 1, A Lot True = 2, and Very True = 3*. For this study, I used raw scale scores for each subscale. The description of bullying in Table 1.3 (see page 39) and the items described in Table 1.4 that show the Bullying subscale items included in each version of the SSIS operationalize the definition of bullying. Bullying total raw scores on the SSIS were based on frequency ratings for five items on the Bullying Subscale (Table 1.2 on page 38).

For the Bullying Subscale, there are average and above average ranges; no below average range exists because of a floor effect for scores. That is, informants reported the average students rated in the normative sample to very infrequently exhibit many of the

bullying behaviors. They assigned the above average level to all scores falling more than one standard deviation above the mean. A student with an above average level on the Bullying subscale typically demonstrates more than the average number of behaviors for students in his or her norm group. Based upon the criteria established by the SSIS, self-identified bullies were those with a raw score on the Bullying Subscale falling one standard deviation or higher above the mean.

Social skills. The Teacher, Parent, and Student version of the SSIS Rating Scale was used to collect data on seven specific skills: (a) Assertion, (b) Communication, (c) Cooperation, (d) Empathy, (e) Engagement, (f) Responsibility, and (g) Self-control. For the total Social Skills Scale, a standard score of 100 is the mean with a standard deviation of 15. Students with total Social Skills Scale scores between 85 and 115 are considered in the Average range. Students with scores above 115 are in the Above Average range.

Problem behaviors. The Problem Behaviors Scale consists of five subscales: Autism Spectrum (not used in the study), Bullying, Externalizing, Hyperactivity/Inattention, and Internalizing. A standard score of more than 115 on the Problem Behaviors scale indicates that an individual displays more problems than the average individual in that population does.

Plan of Analysis

Key variables. In this study, the level-1 variables were the ratings from each informant, and students comprised the level-2 units. The outcome variable was the Bullying Subscale raw score. The level-2 student variables included Sex (1 = male, 2 = female), Age, and Race/Ethnicity. The level-1 variable of rater included the individual

raters labeled 1 to 3 (teacher, parent, and student). The ID was the student id variable, the variable that identified the specific student. Table 3 provides detailed descriptions of the variables, including the social skills subdomains and the problem behaviors subscales (see page 39).

Descriptive analyses. The descriptive data (total N, mean, and standard deviation) was reported and organized by Sex, Age, and Race/Ethnicity. The results of paired samples t-tests of students and adult ratings were also provided.

Analytic approach for research question 1. A conditional probability analysis provides evidence for the relative accuracy or inaccuracy of a screening system (Kettler, Glover, Albers, & Feeney-Kettler, 2014). Specifically, such an analysis shows how well various indicators predict if a student is at risk of bullying or not. In the present study, the indicators are represented by parents' and teachers' SSIS behavior ratings of students' bullying behaviors and the risk outcome are the result of students' self-ratings of bullying (i.e., Bully or non-Bully) on the SSIS completed within a month of those by a parent and teacher. Figure 3 below shows the four potential categories in a conditional probability framework: a true positive, a false positive, a false negative, or a true negative (Kettler et al., 2014).

Conditional Probability Framework

Screening result	Reality (defined by established measure)		Total
	At risk	Not at risk	
At risk	a (true positives)	b (false positives)	a + b
Not at risk	c (false negatives)	d (true negatives)	c + d
Total	a + c	b + d	a + b + c + d

Note. Sensitivity = $[a / (a + c)]$, specificity = $[d / (b + d)]$, positive predictive value = $[a / (a + b)]$, negative predictive value = $[d / (c + d)]$, base rate = $[(a + c) / (a + b + c + d)]$, and hit rate = $[(a + d) / (a + b + c + d)]$. From "Screening Systems and Decision-Making at the Preschool Level: Application of a Comprehensive Validity Framework," by R. J. Kettler and K. A. Feeney-Kettler, 2011, *Psychology in the Schools*, 48, p. 436. Copyright 2011 by Wiley-Blackwell. Adapted with permission.

Figure 3. Conditional probability framework (Kettler et al., 2014)

Each student case falls into only one of these categories, and once cases are assigned to a category, a set of indices are developed regarding the accuracy of the screening system.

These indices include the following: (a) sensitivity, the likelihood the screening accurately identifies a student at risk for bullying who is truly at risk; (b) specificity, the likelihood the screening accurately identifies a student as not at risk for bullying who is actually not at risk; (c) the positive predictive value (PPV), the likelihood a student identified as at risk for bullying actually is at risk; (d) the negative predictive value (NPV), the likelihood a student identified as not at risk for bullying really is not at risk.

This is a conditional probability framework between the teacher and parent assessments and reality. It is also possible to calculate hit rate data from these numbers by adding the true positives and true negatives together and dividing these by the total sample size. The hit rate acts as a single indicator of the accuracy and utility of a risk indicator.

An evaluation of an assessment used for screening or identification of students in need of help considers these indices together because developmental outcomes and features of a sample could maximize one and minimize another (Kettler et al., 2014). The values for each of the indices range from 0.0 to 1.0, where a higher value indicates better accuracy in the results. Of note, when considering the values for the indices is Kettler et al.'s (2014) assertion that the evaluation of an assessment should be developed based upon the outcome decision's specific criteria for what are the acceptable values of the conditional probability indices. These are founded on the relative cost of a false positive versus a false negative as well as based upon the base rate of the necessity for preventive intervention in the population of interest. Essentially, determining a general rule for interpretation of these indices depends upon one's specific situation.

In the case of bullying, incorrectly identifying a student (false positive) as a bully who then receives an unnecessary intervention is perhaps less problematic than not identifying a student who is a bully (false negative) who then does not receive the needed intervention (Kettler et al., 2014). Given this scenario, a potentially acceptable interpretation of values, as developed by Kettler et al. (2014) for the indices of PPV and specificity, is that indices are moderate and considered acceptable if the indices are greater than or equal to .6 but are less than or equal to .8. A result is low and considered acceptable under some circumstances if the indices are equal to or greater than .4 but not greater than .6. Any indices less than .4 are low and could indicate a screening system not suited to the population and purpose being tested. Given this interpretation, Receiver Operating Characteristic (ROC) curves were used to assess how accurately parents and

teachers identified self-rated bullies and non-bullies. The predictive validity was assessed using the area under the curve. The hit rate of the data was also identified.

Analytic approach for research question 2. A multilevel model (Cohen, Cohen, West, & Aiken, 2014) was used to examine the agreement across the social skills and problem behaviors (minus the bullying and autism subscale scores) ratings from self-reports, parents, and teachers. The choice of a multilevel model came from the ability of the model to handle nesting in the data, data missing at random, and the improvement in estimation because the model leverages all available data (Little & Rubin, 1989). This method increases the robustness of results by minimizing the impact of any missing information. The multilevel model also handles the nesting of participants within higher-order settings (Singer, 1998). Lastly, a multilevel model provides flexibility as a method by allowing any data (binary, ordinal) to be provided by each informant. A well-integrated assessment allows for a proper evaluation of the individual and common factors presented by the different informants' ratings.

There were multiple measurements of perceived bullying behavior across three different informants as the level-1 unit. These raters' scores were nested within the individual students at level-2 in the model. The SAS statistical package with PROC MIXED syntax and output were used to conduct the analysis (Singer, 1998). The data was kept in the long format showing the scores for each student. A model building process was used to establish the best fit for the data. The equations for the model followed this structure:

$$\text{Equation 1: } Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + e_{ij}$$

The above equation signifies a basic model with a level-1 predictor, and Y_{ij} represents an individual bullying behavior score for a single student. The first coefficient β_{0j} signifies the average bullying behavior score for student j , X_{ij} represents a level-1 predictor for a bullying behavior score i for student j . β_{1j} denotes the slope connected to X_{ij} . The value shows the relationship between the level-1 variable and bullying behavior. The final term e_{ij} represents the level-1 error term that is understood to be distributed normally with the covariance R.

$$\text{Equation 2: } \beta_{0j} = \gamma_{00} + \gamma_{01}W_j + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

Equation 2 represents the random intercept model, as it demonstrates the basic level 2 model that includes a single student (level-2) predictor, so γ_{00} represents the intercept. The intercept was the grand mean of bullying behavior scores across raters and students, and W_j depicts a student-level predictor for student j , and γ_{01} becomes the regression coefficient connected with W_j , and μ_{0j} was the error term demonstrating a single effect associated with rating j . γ_{10} provides an estimate of the average effect of the level-1 predictor. There was no error term in the β_{1j} equation because the effect of the level-1 predictor remains fixed across students. The student level errors were distributed normally with covariance G.

$$\text{Equation 3: } \beta_{0j} = \gamma_{00} + \gamma_{01}W_j + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \mu_{0j}$$

Equation 3 depicts a random intercept and random slope model, which demonstrates the effect of the level-1 predictor by modeling it to vary across the student-level units. This

equation has a random effect for the level-1 predictor, modeled using, X_{ij} . The single change in this model was the addition of μ_{0j} in the second level equation. The addition of this term allows the relationship between the level-1 predictor X_{ij} and outcome of Y_{ij} to vary across students.

The next stage involved placing the values β_{0j} and β_{1j} into the level-1 equation. The combined equation became the fourth model representing the regression aspect underlying the model. The complete two-level model:

$$Y_{ij} = \gamma_{00} + \gamma_{01}W_j + \gamma_{10}X_j + \mu_{0j} + \mu_{1j}X_{ij} + e_{ij}$$

Full maximum likelihood (ML) was used to estimate each of the models. The use of ML allowed for the comparison of the models using Akaike's Information Criterion (AIC), which changes when a model changes with fixed or random effects. Smaller values for AIC indicate a better fit (Raudenbush & Bryk, 2002).

Results

Descriptive analyses. The means for the total Bullying Subscale ratings provided by each rater were examined. The possible range of scores for students on this subscale are a minimum of 0 to a maximum of 15. The standard and raw score scale ranges on each rater's version of the scale are the same. I examined the different ratings from students and parents and students and teachers.

Table 11

Paired Samples Test of the Differences between Adult and Self-Reports of Bullying Ratings

	Mean	SD	SE	95% Confidence Interval of the Difference		<i>t</i>	<i>df</i>	Sig. (2- tailed)
				Lower	Upper			
Student-Teacher								
Female	4.833	2.496	0.51	3.779	5.887	9.485	23	0.000
Male	3.545	3.202	0.557	2.41	4.681	6.36	32	0.000
Overall	4.088	2.972	0.394	3.299	4.876	10.385	56	0.000
Student-Parent								
Female	5.047	2.572	0.322	4.404	5.689	15.696	63	0.000
Male	5.083	2.736	0.279	4.529	5.638	18.202	95	0.000
Overall	5.069	2.664	0.211	4.653	5.485	24.069	159	0.000

Table 11 shows the results of the paired samples testing for both groups. The mean correlations of $-.14$ for parents and students and $-.05$ for teachers and students were not significant. The full correlation matrices were not included in the results because the multilevel model used in the study provides a summary of significant correlations for each rater and if/how the raters correlate with each other across variables.

Research question 1. How accurate are parents and teachers at identifying students who self-identify as bullies or non-bullies?

Using students' self-reported bullying status as a criterion, a conditional probability analysis revealed that parent reports of bullying correctly classified 85.9% of cases (hit rate) (Table 12). The sensitivity (True Positives) was 8.8%, specificity (True Negatives) was 96.3%, positive predictive value was 24.1%, and negative predictive value was 88.7%. The area under the curve (AUC) was $.525$ and not significant.

Table 12

Conditional Probability Framework of Student and Parent Identification of Bullies

Parent Response	Reality (Student Response)		Total
	Bully	Not Bully	
Bully	14	44	58
Not Bully	146	1142	1288
Total	160	1186	1346

By comparison, the teacher reports of bullying correctly classified 84.3% of cases (hit rate) (Table 13). The sensitivity was 17.5%, specificity was 93.9%, positive predictive value was 29.4%, and negative predictive value was 88.8%. The AUC was .557 and not significant.

Table 13

Conditional Probability Framework of Student and Teacher Identification of Bullies

Teacher Response	Reality (Student Response)		Total
	Bully	Not Bully	
Bully	10	24	34
Not Bully	47	372	419
Total	57	396	453

Research question 2. What is the relative value in including the ratings from teachers and parents to those of students who self-identify as bullies for predicting the social skills and problem behaviors of the identified bullies?

Hypothesis 1: Parents and teachers were expected to add incrementally to the overall predictive validity of the student assessments.

Hypothesis 2: It was expected that there would be modest overlap observed between the different informants' reports, resulting in low to moderate agreement across ratings and differences in the social behavior profiles reported.

The multilevel model represented the rater's scores nested within students, so the student level predictors were at level-2 with the raters at level-1. The main focus of the research question regarded the value of and the potential differences between raters for the students who predicted themselves as bullies. Since raters were the focus, and it was unknown if the effects differed by rater, rater interactions needed to be included to account for all the effects of interest. However, before the creation of the full models used to answer the research question, initial model building was done to test out the correct fit and estimated parameters for the full model. Variance components were the best fit for the covariance structure, because they allowed the model to converge successfully. This structure was added to the SAS code for each of the tested models.

Unconditional model. The first fitted model was an unconditional model with no predictors. The results for the random effects in the model showed the estimated intercept as 0 and the estimated residual as 10.11. Hypothesis testing of these estimates revealed that only the residual component, representing the variance within the students' individual ratings, significantly differed from zero ($p < .001$). Therefore, the results suggested that each student differed in his/her individual bullying ratings from each informant, but there was not a significant difference in the variance between each student.

The estimated student ratings effect of 3.74 in the model refers to the average student Bullying Subscale self-ratings within the sampled pool of student ratings.

Conditional model. The next model was a conditional model that included Rater as a level-1 predictor. The results of the random effects in this model showed the estimated intercept was 0.09 and the estimated residual was 3.99. Hypothesis testing of these estimates again showed that only the residual component, the variance within students' individual ratings, significantly differed from zero ($p < .001$). However, the residual component decreased from the initial model indicating some of the variance within students' bullying ratings was accounted for by the addition of the predictor for Rater. The AIC also was smaller in this model having been reduced from 1982.3 to 1640.2, indicating a better fit for the data.

The results showed the estimated Intercept as 6.53. The estimates for teachers (-4.59) and parents (-5.06) implied that, on average, teachers and parents rated students lower on bullying behavior than the students rate themselves. The standard errors for teachers ($SE = .31$) and parents ($SE = .22$) demonstrated a significant effect of teachers' and parents' ratings on the average of the bullying ratings given to a student from each rater. The hypothesis testing supported the significant effect of both adult raters on the mean of a student's three different bullying ratings.

Random intercepts and slopes model. The next model included random intercepts and slopes. It included Rater as predicting the outcome as a function of the students bullying ratings, but it also specified the relationship between the outcome and that the raters' scores for each student may vary across students. The Rater variable acted as both

a fixed and as a random effect. The model included intercepts and slopes that varied across students. The overall results were similar to the previous model. Specifically, the results for the random effects showed that the variance component for Rater was significantly different from zero, but the residual component was not significantly different from zero. This result indicated that a reduced model, like the previous model, not including slopes varying across students was likely a better fit. This point also was indicated by a small increase in the AIC from 1640.2 to 1642.2.

Full model. After the initial testing, a full model, with student self-report ratings as the overall reference group, was fitted that included a random intercept and the addition of all student level predictors including Sex, Age, Race/Ethnicity, and Total Social Skills and Total Problem Behaviors raw scores. Initially, interactions were also included for each student level predictor and the Rater variable to uniquely estimate the effects of the other predictors. However, if the effect of the student level predictors did not vary across the raters, the interaction term could be removed to constrain the effect to be constant across all raters. The determination for keeping or removing the interaction was based upon the multivariate hypothesis testing for the interaction. If the interaction was not significant, it conveyed that the effect of, say, Sex was constant for student, parent, and teacher ratings and did not need to be estimated separately for each rater.

If after removing the interaction, the multivariate hypothesis testing of again, say, Sex was statistically significant it suggested there were statistically significant differences between the coefficients for students' self-report ratings and the adult ratings for the student level variable. However, the effect was constant. Please note, if also comparing

the effects of student level predictors for two non-reference group ratings (meaning comparing teacher to parent ratings), this step would require additional testing. However, that additional comparison was not included because the focus of this study was regarding agreement between student raters and the adult raters.

Full model 1. The random effects for the first full model showed the estimated intercept was 0.10 and the estimated residual was 1.82. Hypothesis testing of these estimates again showed that only the residual component, the variance within students' scores, remained significantly different from zero ($p < .001$). However, the residual component decreased significantly from the initial model indicating more of the variance within the students' individual bullying ratings was accounted for by the added predictors. The AIC was much smaller from the previous model going from 1640.2 to 1380.7, indicating a better fit for the data.

As part of the next step in fitting the first full model, all of the predictors were examined to determine if they varied across raters or if the effect of the predictor was constant. After the initial run, the hypothesis tests and the AIC were examined to determine model fit. The first run showed that the interaction for Sex and Rater was not significant, so this did not need to be estimated separately for each rater. In addition, the AIC decreased slightly to 1377. In the next run, the interaction was removed for Race/Ethnicity because it was also not significant, which provided a further reduction of the AIC to 1375.6. The non-significant interaction between Rater and Total Social Skills was then removed, which again decreased the AIC to 1373.5. Lastly, the interaction between Age and Rater was initially removed because it was not significant. However,

the removal of the interaction resulted in a slight increase of the AIC to 1374.6, so the interaction for Age and Rater remained in the model even though it was not significant. The best model fit appeared to be the one that included all of the predictors constrained across Raters except for the interactions for Age and Total Problem Behaviors.

Table 14 below shows the outcome that the predictors of Sex, Rater, Age for parents and self-reports, Race/Ethnicity, and Problem Behaviors ratings between teachers and students had fixed effects that differed significantly from zero. The table also shows if the effect of a student level predictor did not vary across the raters. If the effect was constant, the interaction term was removed and replaced with a pooled estimate. In Table 14, the pooled estimate was reported for Sex, Race/Ethnicity, and Total Social Skills. The separate ratings for teacher, parent, and self-report ratings were reported for the interactions of Age and Total Problem Behaviors. The multivariate hypothesis testing for each pooled estimate and/or interaction showed if the coefficients for the predictor were different between parent reports and self-reports and teacher reports and self-reports, but they did not examine differences between parent and teacher reports.

Table 14

Coefficients of Differences in Predicting Bullying for Teacher, Parent, and Self-Reports Using Total Social Skills and Total Problem Behaviors Scales

	Coefficient (SE) [p-value]			Pooled Estimates(c)	Multivariate Hypothesis Testing <i>F</i>
	Teacher	Parent	Self(a)		
Rater	- 3.46(1.09)* [.002]	-4.74(.81)* [<.0001]	4.81(.98)* (b) [<.0001]		17.31**
Sex Female (reference group)		-			
Male				-.36(.15)* [.02]	5.51**
Age	.03(.08) [.66]	.13(.06)* [.03]	-.09(.04)* [<.03]		2.56
Race/Ethnicity White (reference group)		-			
Black				-.18(.06)* [.005]	8.25**
Hispanic					
Other					
Total Social Skills(c)		-		.001(.004) [.77]	0.09
Total Problem Behaviors	.08(.02)* [.002]	.01(.02) [.56]	.13(.01)* [<.0001]		5.47**

Note: (a) Self-reports from students are the overall reference group. (b) The intercept is the mean of the outcome for the reference group (the self-reports ratings from students). (c) The coefficient was constrained to a pooled estimate because the hypothesis testing conveyed that the effect of the variable is constant for all three raters and does not need separate estimation. *The outcome of the test for the individual rater on the predictor variable is significant at the alpha .05 level. **Multivariate hypothesis test is significant at the alpha .05 level.

The first line of the results in Table 14 showed the Intercept was 4.81 for student self-ratings, so the estimates for teachers (-3.46) and parents (-4.74) again implied that, on average, teachers and parents rated students lower on bullying behavior than the students rated themselves. The multivariate hypothesis testing supported the statistically significant effect of both adult raters on the mean of the students' individual bullying ratings from each rater.

The results for the individual student level predictors are reported beginning on the second line of the table, starting with the variable of Sex. Sex was significant across all raters with boys having higher ratings than girls. The multivariate hypothesis testing showed that the coefficients for Sex were statistically significantly different between parent ratings and self-ratings and teacher ratings and self-ratings. The results for the student level predictors showed that bullying ratings changed with age separately for parent reports and self-reports. The bullying rating was higher by .13 for parent reports, but it was lower by -.09 for self-reports. However, the multivariate hypothesis testing showed that the coefficients for Age were not statistically significantly different between parent reports and self-reports or teacher reports and self-reports. Race/Ethnicity was found to be significant across all raters. The multivariate hypothesis testing showed that the coefficients for Race/Ethnicity were statistically significantly different between parent ratings and self-ratings and teacher ratings and self-ratings. Lastly, the results showed that bullying ratings changed with Total Problem Behaviors ratings for teacher and self-reports. The bullying rating was higher by .08 for teacher reports and by .13 for

self-reports. Multivariate hypothesis testing showed the coefficients for reports of Total Problem Behaviors were statistically significantly different between teacher reports and self-reports but not parent reports and self-reports.

Full model 2. In this version of the full model, the same predictors were examined and followed the same steps as in the previous full model. However, the various Social Skills and Problem Behaviors Subscale scores were included as predictors in place of the Total Social Skills and Problem Behaviors Scale Scores. The Responsibility and Hyperactivity Inattention Subscales were not included due to multicollinearity. The model included the student level predictors for Sex, Age, Race/Ethnicity, Communication, Cooperation, Assertion, Empathy, Engagement, Self-Control, Externalizing, and Internalizing subscales. Interactions with the Rater variable were again tested for each student level predictor to estimate the effects of the other predictors. Only the interactions for Rater by Age, Rater by Communication, and Rater by Externalizing were retained in the model. In the final version of full model 2, the random effects showed the estimated intercept was 0 and the estimated residual was 1.45. Hypothesis testing showed the residual component, the variance within students' individual ratings, significantly differed from zero ($p < .001$). Therefore, the results again suggested that the students differed in their bullying ratings from each informant, but that there was not a statistically significant difference in the variance between the different students. The final full model 2 showed a lower AIC of 1277.2 than full model 1's AIC of 1373.5, indicating a more parsimonious model than the model with the Total Social Skills and Problem Behaviors scores.

Table 15

Coefficients of Differences in Predicting Bullying for Teacher, Parent, and Self-Reports Using Social Skills and Problem Behavior Subscales

	Coefficient (SE) [p-value]			Pooled Estimates(c)	Multivariate Hypothesis Testing <i>F</i>
	Teacher	Parent	Self(a)		
Rater	6.42(1.33)* [<.0001]	5.33(1.06)* [<.0001]	4.21(.73)*(b) [<.0001]		18.21**
Sex-					
Female (reference group)		-			
Male				-.17(.14) [.21]	1.59
Age	.10(.07) [.13]	.18(.05)* [.0006]	-.11(.04)* [<.0013]		2.56**
Race/Ethnicity					
White (reference group)		-			
Black				-.05(.05) [.37]	0.8
Hispanic					
Other					
Communication	.17(.07)* [.02]	-.09(.06) [.13]	-.03(.04) [.57]		6.42
Cooperation				.01(.03) [.75]	0.1
Assertion				-.01(.04) [.69]	3.29
Empathy				-.04(.03) [.15]	2.14
Engagement				.03(.02) [.26]	1.29
Self-Control				.07(.03)** [.002]	9.79**
Externalizing	.18(.04)* [<.0001]	.003(.02) [.88]	.24(.02)* [<.0001]		12.46**
Internalizing				.02(.02) [.18]	1.83

Note: (a) Self-reports from students are the overall reference group. (b) The intercept is the mean of the outcome for the reference group (the self-reports ratings from students). (c) The coefficient was constrained to a pooled estimate because the hypothesis testing conveyed that the effect of the variable is constant for all three raters and does not need separate estimation. *The outcome of the test for the individual rater on the predictor variable is significant at the alpha .05 level. **Multivariate hypothesis test is significant at the alpha .05.

The results in Table 15 shows the outcome of the analyses where student self-ratings are again the overall reference group. This model was the most parsimonious. The ratings from each rater had fixed effects that differed significantly from zero for Age for parents and students, the Communication subscale for teachers, the Self-Control subscale across raters, and the Externalizing subscale between teachers and students. The table also again shows if the effect of a student level predictor did not vary across the raters. If the effect was constant, the interaction term was removed and replaced with a pooled estimate. In Table 15, the pooled estimate was reported for Sex, Race/Ethnicity, Cooperation, Assertion, Empathy, Engagement, Self-Control, and Internalizing. The separate ratings for teacher, parent, and self-report ratings were reported for the interactions of Age, Communication, and Externalizing. The multivariate hypothesis testing for each pooled estimate and/or interaction showed if the coefficients for the predictor were statistically significantly different between parent reports and self-reports and teacher reports and self-reports, but they did not examine differences between parent and teacher reports.

The first line in Table 15 shows the Intercept for student self-ratings was 4.21, so the estimates for teachers (-6.42) and parents (-5.33) again implied that, on average, teachers and parents rated students lower on bullying behavior than the students rated themselves. The multivariate hypothesis testing showed a statistically significant effect of both adult raters on the mean of the students' three bullying ratings.

As with Table 14, the results for the individual student level predictors are again reported beginning on the second line of the table, starting with the variable of Sex. The

results showed that bullying ratings changed with age separately for parent reports and self-reports. The bullying ratings were higher by .18 for parent reports, but lower by -.11 for self-reports. However, unlike full model 1, the multivariate hypothesis testing showed the coefficients for age were statistically significantly different between parent reports and self-reports, but not teacher reports and self-reports. Communication scores had a statistically significant effect but only for teacher reports. In this sample, teachers who gave the students the higher bullying ratings also gave them higher communication ratings. Students with higher self-control ratings also reported higher bullying ratings across raters. The multivariate hypothesis testing showed that the coefficients for self-control were statistically significantly different between parent reports and self-reports and teacher reports and self-reports. Lastly, the results showed that bullying ratings changed with externalizing ratings for teacher and self-reports. The bullying ratings were higher by .18 for teacher reports and by .24 for self-reports. Multivariate hypothesis testing showed the coefficients for externalizing were statistically significantly different between teacher reports and self-reports but not parent reports and self-reports.

Discussion

This study addressed the incremental validity of adding adults' information to students' self-ratings of bullying behaviors and status. In addition, the agreement of ratings was examined across these informants regarding the students' social skills and problem behaviors. Thus, the aim of the study was to (a) explore the incremental validity of including the ratings of teachers and parents to the social behavior assessment of

students who self-identify as bullies, and (b) determine the extent of agreement between the adult and student ratings of social skills and problem behaviors.

Parents and teachers added to the overall predictive validity of students' self-report assessment. The first research question explored how accurate parents and teachers were at identifying students who self-identified as bullies or non-bullies. A conditional probability framework between the teacher and parent assessments and reality (based on students' self-identification) helped to answer this question. The results for the four potential categories used in a conditional probability framework include true positive, false positive, false negative, and true negative rates (Kettler et al., 2014). Each student case fell into only one of these categories, and once the case was assigned to a category, the set of indices were developed regarding the accuracy of the screening system used. An evaluation of the assessment used for adult identification of self-identified bullies considered these indices together, but with a focus on the PPV (true positive), specificity (true negative), and hit rate (overall accuracy), because developmental outcomes and features of a sample could maximize one and minimize another (Kettler et al., 2014). The values for each of the indices ranged from 0.0 to 1.0, where a higher value indicated better accuracy in the results.

In the case of bullying, the general rule for interpretation of these indices was based on the understanding that incorrectly identifying a student (false positive) as a bully who then receives an unnecessary intervention was considered less problematic than not identifying a student who is a bully (false negative) who then does not receive the needed intervention (Kettler et al., 2014). Given this scenario, a potentially acceptable

interpretation of values, as developed by Kettler et al. (2014) for the indices of PPV and specificity, was that indices are moderate and considered acceptable if the indices were greater than or equal to .6 but are less than or equal to .8. A result was low and considered acceptable under some circumstances if the indices were equal to or greater than .4 but not greater than .6. Any indices less than .4 were low and could indicate a screening system not suited to the population and purpose being tested.

Given the stated interpretation, the Specificity provided an acceptable index for showing how parent reports accurately identified a student as not at risk for bullying who was actually not at risk in 96.3% of cases. In contrast, the PPV index for parent reports was low and not considered acceptable, because parent reports identified a student as at risk for bullying who actually was at risk in only 24.1% of cases. However, the overall accuracy of and utility of parent reports in classifying cases (85.9%) appeared useful. There were similar results for the teacher reports, although the Specificity (93.9%) was slightly lower than for parent reports, the results indicated the screening was accurate in identifying a student as not at risk for bullying who was actually not at risk. For PPV, the index was also low because teacher reports identified a student as at risk for bullying who actually was at risk in only 29.4% of the cases. Overall, teacher reports were moderately accurate (84.3%) in correctly classifying cases.

I expected that adding the adult ratings to the self-reports of the identified bullies would improve the accuracy of classifying the students as bullies. However, the results of the assessment of the predictive validity of parent reports and teacher reports indicated that neither model alone was better than chance at distinguishing between those who

were identified as bullies and those who were not. Parents (85.9%) and teachers (84.3%) were also less accurate than self-reports alone (91.6%) were when correctly classifying cases as a bully or non-bully. Based upon the stated interpretation, the results suggest that overall reports from teachers and parents do show some accuracy in identifying students who self-identify as bullies and non-bullies. However, when used alone, the adult reports might be an unsuitable screening assessment for ensuring a student at risk of bullying is identified and receives the necessary intervention.

There is relative value in including the ratings from teachers and parents to those of students who self-identify as bullies for improving predictions about their social behavior profiles. The second question focused on determining the relative value of including behavior ratings from teachers and parents with those of students for predicting the social skills and problem behaviors of the self-identified bullies. The conditional probability analysis suggested that while adult reports are not as accurate in identifying a student as at risk for bullying who actually is at risk, there is still value provided by the adult reports. In particular, they are accurate in identifying a student as not at risk for bullying who is actually not at risk. These results may indicate that the relative value of including adult reports of bullying to existing self-reports lies in the reported differences in the information provided by each rater rather than on their similarity to the self-reported data. Existing research using multiple informants in general and in bullying assessment, supports this interpretation of the results, as it consistently, shows the type of rater influences the results of an assessment (e.g., Achenbach et al., 1987; Rønning et al., 2009; Rupp et al., 2018; Wegge et al., 2016). There is also value

found in both the agreement and disagreement between the ratings, as they can be predictive of a variety of behavior outcomes (De Los Reyes, 2011). This information can be informative for prevention and intervention activities.

Using a multilevel model also helped alleviate the concern of reporting errors when assessing bullying (Shiner & Allen, 2013) because the model provided for the use of all the available data from each rater within a single model. The model allowed for formal testing to determine if any associations between the predictive variables and bullying ratings differed between raters. When a rater effect occurred, the model could get a distinct group of estimates for each rater.

The results were similar with findings from other studies (e.g., Ferdinand et al., 2004, 2006, 2007) showing overall differences in the ratings provided by different informants. Therefore, these results could suggest there is importance in examining the differences in rater agreement, as the students did differ significantly in their bullying ratings from each informant. Teachers and parents rated students lower on bullying behavior than the students rated themselves (Nansel et al., 2015). The non-significant difference in the variance between the different students across raters suggested the stability of these results.

Higher age was associated with higher bullying ratings from parents, whereas it is associated with lower bullying ratings for self-reports. The association with age for self-reports is consistent with existing research that reports of bullying typically decrease with age. The higher age being associated with increased reports of bullying from parents may be indicative of behaviors observed by parents that may be underreported by adolescents

because of issues like the social desirability of the behavior (Henrich & Shahar, 2014) or simply that parents and students interpret the intent or effect of the same behaviors differently.

The analysis provided valuable statistical results demonstrating the ways in which the association between predictor variables and bullying ratings differed by rater. In addition, the analysis showed how when no rater-predictor interaction existed, one could get a pooled parameter estimate from each rater. In these cases, the standard error is likely smaller than it would be if individual regression analyses were carried out for each one. An example of this pooling of coefficients was done with the variable of Self-Control, because the multivariate hypothesis testing for the interaction showed no significant difference for the coefficients, as all three raters reported higher bullying ratings for students with higher frequency of Self-Control behaviors. These results supported the relative value in separately examining the information from the different raters.

Using the average of the ratings from each rater without determining the significance of any differences in the ratings could be deceiving. For example, only teachers gave the children with higher bullying ratings higher communication scores. A potential significance of this finding was seen in understanding that teachers are perhaps reporting a behavior in self-identified bullies that is not self-reported or seen by a parent (Henrich & Shahar, 2014). Similarly, this point also is demonstrated by the finding that higher Externalizing ratings were associated with increased bullying ratings from teachers and from self-reports but not parents. The agreement on externalizing behavior

between teachers and self-reports also is consistent with existing research showing that raters who observe behavior in the same context typically have more agreement, in particular regarding externalizing behaviors (Achenbach et al., 1987). Averaging the ratings from each rater without determining the significance, or using only one rater, might mean losing or overlooking this valuable contextual information. Knowing the statistical significance of each set of ratings can help to determine both rater effects and the way in which each source of information contributes to the assessment (Flake & Petway, 2019).

Strengths, Limitations, and Future Directions

A strength of the study was that the multiple rater data were collected with the SSIS Rating Scales, one of the few multi-rater assessment systems where all participants complete items with the same root language. A limitation of the study was the cross-sectional nature of the dataset. However, while the data was cross-sectional, a strength of the data was the large representative national sample of students. Future studies would benefit from a longitudinal design and the examination of different groups within the sample. Some of the limitations in using self-reports alone also were addressed by examining the addition of adult raters to the self-reported data. However, a limitation in the addition of the adult raters is for those cases not having a rating from both a parent and a teacher. Another minor limitation was that there are two items overlapping on the Bullying (five total items) and Externalizing (12 total items) Subscales. However, these items account for less than 20% of the total loading items.

The missing data were also a potential limitation, but the type of analysis helped to reduce this limitation. The multilevel model approach provided for the use of all existing data in the sample. Using such a model also allowed for the testing of all the statistical differences between raters. Essentially, the strength of this type of analysis is that when multiple raters provide similar ratings, the statistical precision of the analysis is improved. In addition, when raters provide differing results the model can adjust for the contrasting information. Future research should include similar studies that could provide comparable results for examining the utility of including multiple raters as part of an overall bullying assessment. More results could also be obtained from including more predictive variables, such as parental education, special needs status, and location.

Conclusion

The current study provided a means for investigating if there was any relative value in including data from adult raters to that of self-identified bullies. Specifically, the analyses investigated if there were differences and similarities between raters. A conditional probability analysis showed the inaccuracy of using parent and teacher reports alone as a method to identify students who are likely to participate in relatively high levels of bullying. The overall findings showed that differences and similarities do exist across the ratings provided by each rater. However, the results demonstrated that there is relative value in including the parent and teacher reports with a self-assessment of bullying as means for providing a more complete social behavior profile for each student.

CHAPTER 4

INTEGRATED SUMMARY OF STUDIES 1 AND 2

Key Findings

The subjectivity of both studies was reduced because a specific definition of bullying is included in the assessment. The definition of bullying used in this set of studies was based upon the items from the SSIS Bullying Subscale. It also was based upon the following definition used in the SSIS Technical Manual, "Forcing others to do something, hurting people physically or emotionally, and not letting others join an activity" (Gresham & Elliott, 2008, p. 2).

Study 1

For Study 1, it was proposed that bullies do not comprise a homogenous group. In fact, it was argued that even the most socially unskilled bully might possess social skills that can be assessed and used to guide intervention. Some of the students who bully, likely have a well-developed understanding of social cues (e.g., reading facial expressions) and know exactly how to use the information to their advantage (Gini et al., 2011), but may still display some deficits in social skills.

The purpose of Study 1 was to establish criteria using the SSIS (Gresham & Elliott, 2008) to identify students from a nationally representative standardization sample who displayed high frequencies of bullying behaviors. The social behavior ratings for these self-identified bullies were then compared with all other students in the national sample and analyzed to determine differences among various domains of social skills and problem behaviors.

To review, there were seven key findings reported for Study 1 regarding students' self-ratings of bullying and broader social behavior skills. These were:

1. As expected, in the self-identified bully group, overall, males showed higher average bullying ratings than females did, except for females ages 8 to 11 who showed slightly higher ratings than males did.
2. Overall, bullying appeared to decrease with age.
3. Each of the logistic regression models explained more than 50% of the variance in the self-reports of bullying and correctly classified 91.6% of the combined self-report cases as a bully or non-bully with females showing the highest classification rate at 94.6%.
4. Race/Ethnicity was significant for females where the odds of self-identifying as a bully were more than 4 times greater for Black females as opposed to White females.
5. Also as expected, a decrease in empathy was associated with an increased likelihood of self-identifying as a bully for males, but not females.
6. Interestingly, increases in assertion and self-control ratings were associated with an increased likelihood of self-identifying as a bully for males, but not females.
7. As expected, increases in externalizing behavior (e.g., physical fighting) and internalizing behavior (e.g., anxiety) were associated with an increased likelihood of self-identifying as a bully for males. However, only increases in externalizing behavior were associated for females.

Study 2

In Study 2, an important feature was a focus of the multilevel analysis on the raters and not on the bullying ratings alone. The same students' social behaviors from Study 1 were rated by adult informants to determine if there was added value in including parents and teachers in the assessment of the self-identified bullies. Finally, the extent of concurrent agreement was examined for all students among the teachers, parents, and students' ratings of social skills and problem behavior domains. To review, there were six key findings reported for Study 2. These were:

1. Parents (85.9%) and teachers (84.3%) had lower hit rates than the self-reports (91.6%) did when using the same predictors (e.g., Sex, Age, Race/Ethnicity, Total Social Skills and Problem Behaviors Scales and Subscales) to correctly classify cases as a bully or non-bully.
2. Students did differ significantly in their bullying ratings from each rater with teachers and parents rating the students lower on bullying behavior than the students rate themselves.
3. Higher age was associated with higher bullying ratings by parents, whereas it was associated with lower bullying ratings for self-reports.
4. All raters reported higher bullying ratings for students with higher frequency of self-control behaviors.
5. Interestingly, teachers rated children with higher communication scores, higher bullying scores.

6. Higher externalizing ratings were associated with higher bullying ratings by both teachers and students, but not parents.

Noted Limitations

A limitation was the cross-sectional nature of the data, so the direction of the results for both studies remains undetermined. Although cross-sectional in nature, the data provided a strength to both studies, missing from similar studies (Nansel et al., 2001). The sample incorporated a wide age range including elementary-, middle- and high school-aged youth. In addition, while self-reports are a valid measure, they still represent a limitation because of potential variation across individuals. Some of the limitations in using self-reports alone were addressed in Study 2 by examining the addition of adult raters to the self-reported data.

Another overall strength of the data used was that the multiple rater data was collected with the SSIS Rating Scales, one of the few multi-rater assessment systems where all participants complete items with the same root language. The SSIS Rating Scale is a useful broadband assessment that assesses an array of social skills, several classes of problem behaviors including bullying, and academic competencies that is tied directly to an intervention. However, a noted limitation is that the SSIS does not capture in-depth details about the specific instances of when a child is participating in bullying; no behavior rating scales accomplishes this, thus the need for supplemental direct observational assessments. Another minor limitation was that there are two items overlapping on the Bullying (five total items) and Externalizing (12 total items) Subscales. However, these items account for less than 20% of the total loading items.

The missing adult data were also a potential limitation, but the type of analysis in the Study 2 aided in helping to reduce the limitation by using all available data in the sample. Using such a model made it possible to test the statistical differences between raters. The strength of the analysis was the ability to make use of any similar ratings from multiple raters, so the statistical precision of the analysis is improved. In contrast, when raters have differing ratings the model can also adjust for the differences in information.

Implications for Future Research

The differences emerging from the results of both studies suggested a need for the development of future research targeting specific sex and age groups. For example, further examination of differences based upon sex might reveal valuable information about specific group behaviors. Assertive behavior and self-control were predictive indicators of a self-identified male bully, but not a female one. Whereas, increasing age was associated with female self-identified bullies, but not males. A closer examination of these variables, and the addition of more predictive variables, within specific groups could aid in the early identification of situations or circumstances where children and youth are more likely to participate in bullying.

While both studies provided notable data on the frequency and social behavior connections of children who self-identify as bullies among youth in the United States, more research is necessary. Future research should include similar self-report studies that provide more data on the social behaviors of self-identified bullies.

Future studies using multilevel models and multiple rater data could also provide results that might be useful for comparing results across studies. Of particular worth would be future studies using a longitudinal design examining the specific social behaviors connected to students who bully others.

Implications for Practice

Both studies provided important information that can be used to advance assessment practices designed to improve programs and strategies aimed at reducing children's bullying behavior. An examination of the demographics and social behaviors of individuals who display bullying behavior helped to demonstrate that bullies are not a homogeneous group. Some children self-identified as bullies lacked or demonstrated limitations in important social skills (Farmer et al., 2010), such as empathy. However, there were other children who self-identified as bullies who displayed higher social skills, such as assertion and self-control. These skills were actually predictive of identifying with bullying behavior. These children perhaps have the ability to avoid detection or can get others to do what they want (Sutton et al., 1999; 2001).

A subsequent look at complementary data provided by an adult rater, for each of the self-identified bullies, revealed similar results and suggested that informants see a range of behaviors that may be predictive of bullying behavior. Given the range of demographics and social behaviors displayed by so-called bullies, programs will need to adjust any existing understanding of bullying to be successful at reducing bullying behavior. A way to make the necessary adjustments and to identify all of the potential cases is by using a multi-rater broadband social behavior assessment, like the SSIS.

Such an assessment allows for the examination of a range of social skills and problem behaviors that can be used to identify students who engage in bullying.

The results also supported the importance of including multiple perspectives in bullying assessment research in order to identify all at-risk children, as suggested by existing researchers (e.g., Crothers & Levinson, 2004; Griffin & Gross, 2004; Demaray, et al., 2013; Groeben et al., 2011; Juvonen et al., 2001; Nowell et al., 2014; Phillips & Cornell, 2012). No precise determination exists for how these diverse perspectives should be integrated and interpreted to further understanding of bullying. However, the findings from Study 2 supported the recommendation to include data from multiple raters, and the use of a statistical model, such as a multilevel model, that incorporates all available information from multiple raters. It is potentially problematic for bullying assessment research if studies do not include multiple perspectives, or they just average the data from each rater without determining the statistical significance of it. Doing either of these things may mean missing out on identifying a child at risk of bullying, confirming that a behavior is indicative of bullying, or not adjusting for any potential partiality in a rater's preconceived perceptions about the child at risk.

Specific Conclusions to the Research Questions

In closing, I return to the research questions that motivated this two-study dissertation with the goals of succinctly answering them and pointing to key research implications.

Study 1. The research question addressed was: *What are the demographics and social skills that characterize self-identified bullies?* The study revealed the self-

identified bullies are not a homogeneous group. They represent a range of demographics and social behavior profiles. Although, the results did support some of the general findings from previous studies, such as males typically participate more in bullying than females. The results, however, also introduced more important questions requiring more research. The observed frequency of self-identified bullies (12.1%), which falls into the middle of the reported ranges, suggested the necessity for increasing and improving preventive intervention and bullying assessment research. The differences between groups and behaviors also suggested the need for more detailed studies that examine specific groups, such as by sex and race/ethnicity.

Through such studies, researchers and practitioners may improve their understanding about what potential bullies actually look like or why some social behaviors are more predictive of self-identification in that specific group. For example, the study revealed that Black females were much more likely to self-identify as bullies than White females. To improve intervention efforts, this improved knowledge about such specific contexts and social behaviors is necessary for both enabling and preventing the development of bullying. The improved knowledge may be used to advance efforts to develop effective individual and school-based interventions that address bullying and the related consequences.

Study 2. This study built upon Study 1 findings by comparing adults' perspectives of the social competencies of self-identified bullies to identify commonalities, or discrepancies, of the perceived skillfulness of students who report engaging in bullying. The study answered two research questions. *How accurate are*

parents and teachers at identifying students who self-identify as bullies or non-bullies?

To what extent do parents', teachers', and students' ratings of social skills and problem behaviors agree with one another for the students who self-identify as bullies and students who do not self-identify as bullies?

With regard to the first question, a conditional probability analysis showed that parents and teachers adds to the overall predictive validity of the student self-report assessment. However, adding the adult ratings to the self-reports of the identified bullies did not specifically improve the accuracy of classifying the students as bullies. Using either set of adult ratings alone was no better than chance at distinguishing between those who are identified as bullies and those who are not. However, the results pointed out that reports from teachers and parents still showed some accuracy in correctly identifying students who self-identify as bullies and non-bullies. However, when used alone, the adult reports might be an unsuitable screening assessment for ensuring a student at risk of bullying is identified and receives the necessary intervention. The results suggested that, if multiple raters are not available, using self-reports may render the most accurate classification of cases as a bully or non-bully than using teacher or parent reports alone would.

With regard to the second question, the overall findings showed that differences and similarities do exist across the ratings provided by each rater. The conditional probability analysis suggested that while adult reports are not as accurate in identifying a student as at risk for bullying who actually is at risk, there is still value provided by the adult reports. The relative value of including adult reports of bullying to existing self-

reports likely lies in the reported differences in the information provided by each rater rather than on their similarity to the self-reported data. There is relative value in including the parent and teacher reports in a self-assessment of bullying as means for providing a more complete social behavior profile for each student. In addition, the results showed how using a multilevel model alleviates some of the concerns about how to interpret multi-rater data. It helps to handle potential reporting errors, uses all the available data, and tests all of the potential associations between and across raters.

In summary, there remain some concerns about the use of self-report ratings to measure bullying behavior and the validity of any obtained results. The results of this set of studies helps to build additional support for the use of self-report assessments. This support is based on the valid findings from students who self-identify as displaying bullying behavior. The sample of students came from a large national sample who reported bullying using a psychometrically sound measure with a built-in measure to detect false responses. This sample was powerful and diverse, thus providing a rich source for exemplifying a persistent and important problem with 12 to 15% of youth and supports a clearer definition of bullying. Having such supportive evidence is important, as self-reports will likely continue to be a popular choice because self-reports typically offer a very efficient, cost-effective way to obtain estimates on bullying. Policymakers, researchers, and interventionists will most likely rely on this self-reported data to collect information on large samples of school-age children to estimate the frequency of bullying as an indicator of success (or failure) to prevent and reduce school bullying (Solberg & Olweus, 2003).

REFERENCES

- Ablow, J. C., & Measelle, J. R. (1993). Berkeley Puppet Interview: Administration and scoring system manuals. *Berkeley, CA: University of California.*
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin, 101*(2), 213.
- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the Child Behavior Checklist/6-18*. Burlington, VT: University of Vermont Department of Psychiatry.
- Arseneault, L., Bowes, L., & Shakoor, S. (2010). Bullying victimization in youths and mental health problems: ‘much ado about nothing’?. *Psychological Medicine, 40*(5), 717-729.
- Bauer, D. J., Howard, A. L., Baldasaro, R. E., Curran, P. J., Hussong, A. M., Chassin, L., & Zucker, R. A. (2013). A tri-factor model for integrating ratings across multiple informants. *Psychological Methods, 18*(4), 475.
- Blake, J. J., Banks, C. S., Patience, B. A., & Lund, E. M. (2014). School-based mental health professionals' bullying assessment practices: A call for evidence-based bullying assessment guidelines. *Professional School Counseling, 18*(1), 136-147.
- Boulton, M.J., & Smith, P.K. (1994). Bully/victim problems in middle-school children: Stability, self-perceived competence, peer perceptions and peer acceptance. *British Journal of Developmental Psychology, 12*, 315–329.
- Bowen, R., & Holtom, D. (2010). A survey into the prevalence and incidence of school bullying in Wales. Main report. Retrieved from: <https://dera.ioe.ac.uk//563/>
- Bowes, L., Maughan, B., Ball, H., Shakoor, S., Ouellet-Morin, I., Caspi, A., ... & Arseneault, L. (2013). Chronic bullying victimization across school transitions: The role of genetic and environmental influences. *Development and Psychopathology, 25*(2), 333-346.
- Bowllan, N. M. (2011). Implementation and evaluation of a comprehensive, school-wide bullying prevention program in an urban/suburban middle school. *Journal of School Health, 81*(4), 167–173, <http://dx.doi.org/10.1111/j.1746-1561.2010.00576.x>.
- Branson, C. E., & Cornell, D. G. (2009). A comparison of self and peer reports in the assessment of middle school bullying. *Journal of Applied School Psychology, 25*(1), 5-27.

- Brockenbrough, K.K., Cornell, D.G., & Loper, A.B. (2002). Aggressive attitudes among victims of violence at school. *Education and Treatment of Children, 25*, 273-287.
- Bronfenbrenner, U. (1986). Ecology of the family as a context for human development: Research perspectives. *Developmental psychology, 22*(6), 723.
- Brown, E. C., Catalano, R. F., Fleming, C. B., Haggerty, K. P., & Abbott, R. D. (2005). Adolescent substance use outcomes in the Raising Healthy Children project: A two-part latent growth curve analysis. *Journal of Consulting and Clinical Psychology, 73*(4), 699.
- Brown, E. C., Low, S., Smith, B. H., & Haggerty, K. P. (2011). Outcomes from a school-randomized controlled trial of STEPS to RESPECT: A bullying prevention program. *School Psychology Review, 40*(3).
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. *Sage focus editions, 154*, 136-136.
- Cairns, R.B., Cairns, B.D., Neckerman, H.J., Gest, S.D., & Gariépy, J. (1988). Social networks and aggressive behavior: Peer support or peer rejection? *Developmental Psychology, 24*, 815–823.
- Camodeca, M., Caravita, S. C. S., & Coppola, G. (2015). Bullying in preschool: The associations between participant roles, social competence, and social preference. *Aggressive Behavior, 41*(4), 310-321. doi:10.1002/ab.21541
- Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.
- Cantone, E., Piras, A. P., Vellante, M., Preti, A., Daniélsdóttir, S., D'Aloja, E., ... & Bhugra, D. (2015). Interventions on bullying and cyberbullying in schools: A systematic review. *Clinical Practice and Epidemiology in Mental Health: CP & EMH, 11*(Suppl 1 M4), 58.
- Caravita, S. C., Di Blasio, P., & Salmivalli, C. (2010). Early adolescents' participation in bullying: Is ToM involved?. *The Journal of Early Adolescence, 30*(1), 138-170.
- Card, N.A., Stucky, B.D., Sawalani, G.M., and Little, T.D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development, 79*(5), 1185-1229.

- Carney, J. V., Hazier, R. J., & Higgins, J. (2002). Characteristics of school bullies and victims as perceived by public school professionals. *Journal of School Violence, 1*(3), 91-106.
- Carrasco, M. A., Holgado, F. P., Del Barrio, M. V., & Barbero, M. I. (2008). Incremental validity: An applied study using different informants and measurements. *Acción Psicológica, 5*, 65-76.
- Chan, J. H. (2006). Systemic patterns in bullying and victimization. *School Psychology International, 27*(3), 352-369.
- Chen, G. (2019, February 15). *Bullied to death: Should public schools be more responsible for mean girls?* Retrieved from <https://www.publicschoolreview.com/blog/bullied-to-death-should-public-schools-be-more-responsible-for-mean-girls>
- Chugani, H.T., Behen, M.E., Muzik, O., Juhasz, C., Nagy, F., & Chugani, D.C. (2001). Local brain functional activity following early deprivation: A study of postinstitutionalized Romanian orphans. *NeuroImage, 14*, 1290-1301.
- Chui, W. H., & Chan, H. C. (2013). Association between self-control and school bullying behaviors among Macanese adolescents. *Child Abuse and Neglect, 37*, 237–242. doi:10.1016/j.chiabu.2012.12.003
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284.
- Cohen, J. (1988). The concepts of power analysis. *Statistical Power Analysis for the Behavioral Sciences, 2*, 1-17.
- Cohen, P., West, S. G., & Aiken, L. S. (2014). *Applied multiple regression/correlation analysis for the behavioral sciences*. Psychology Press.
- Collins, K., & Bell, R. (1996). Peer perceptions of aggression and bullying behavior in primary schools in Northern Ireland. *Annals of the New York Academy of Sciences, 794*(1), 336-338.
- Cook, C.R., Williams, K.R., Guerra, N.G., and Kim, T. (2010). Variability in the prevalence of bullying and victimization. In S.R. Jimerson, S.M. Swearer, and D.L. Espelage (Eds.), *Handbook of Bullying in Schools: An International Perspective* (pp. 347-362). New York: Routledge.

- Cornell, D. G., & Cole, J. C. M. (2012). "Assessment of bullying," in S. R. Jimerson, A. B. Nickerson, M. J. Mayer, & M. J. Furlong (eds.), *Handbook of school violence and school safety* (pp. 289–304). New York, NY: Routledge.
- Cornell, D. G., Sheras, P. L., & Cole, J. C. (2006). "Assessment of bullying," in S. Jimerson & M. Furlong (eds.), *Handbook of school violence and school safety: From research to practice* (pp. 191–210). Mahwah, NJ: Erlbaum.
- Crick, N. R., & Bigbee, M. A. (1998). Relational and overt forms of peer victimization: a multiinformant approach. *Journal of Consulting and Clinical Psychology, 66*(2), 337.
- Crick, N. R., & Dodge, K. A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*, 74–101. doi:10.1037/0033-2909.115.1.74
- Crothers, L. M., & Levinson, E. M. (2004). Assessment of bullying: A review of methods and instruments. *Journal of Counseling & Development, 82*(4), 496-503.
- Crowe, L. M., Beauchamp, M. H., Catroppa, C., & Anderson, V. (2011). Social function assessment tools for children and adolescents: A systematic review from 1988 to 2010. *Clinical Psychology Review, 31*(5), 767-785.
- De Bolle, M., & Tackett, J. L. (2013). Anchoring bullying and victimization in children within a Five-Factor Model-based person-centred framework. *European Journal of Personality, 27*(3), 280-289.
- De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin, 141*(4), 858.
- De Los Reyes, A., & Kazdin, A. E. (2004). Measuring informant discrepancies in clinical child research. *Psychological Assessment, 16*(3), 330.
- De Los Reyes, A., Youngstrom, E. A., Pabón, S. C., Youngstrom, J. K., Feeney, N.C., & Findling, R. L., (2011). Internal consistency and associated characteristics of informant discrepancies in clinic referred youths age 11 to 17 years. *Journal of Clinical Child & Adolescent Psychology, 40*, 1-9. doi:10.1080/15374416.2011.533402
- Di Giunta, L., Pastorelli, C., Eisenberg, N., Gerbino, M., Castellani, V., & Bombi, A. S. (2010). Developmental trajectories of physical aggression: prediction of overt and covert antisocial behaviors from self-and mothers' reports. *European Child & Adolescent Psychiatry, 19*(12), 873-882.

- Dirks, M. A., De Los Reyes, A., Brigg-Gowan, M., Cella, D., & Wakschlag, L. S. (2012). Annual research review: Embracing not erasing contextual variability in children's behavior-theory and utility in the selection and use of methods and informants in developmental psychopathology. *Journal of Child Psychology and Psychiatry*, 53, 558- 574. doi:10.1111/j.1469-7610.2012.02537.x
- Demaray, M. K., Malecki, C. K., Secord, S. M., & Lyell, K. M. (2013). Agreement among students', teachers', and parents' perceptions of victimization by bullying. *Children and Youth Services Review*, 35(12), 2091-2100.
- Demaray, M. K., Ruffalo, S. L., Carlson, J., Busse, B. T., Olson, A. E., McManus, S. M., & Leventhai, A. (1995). Social skills assessment: A comparative evaluation of six published rating scales. *School Psychology Review*, 24, 648-671.
- Demaray, M. K., Summers, K. H., Jenkins, L. N., & Becker, L. D. (2014). The Bullying Participant Behaviors Questionnaire (BPBQ): Establishing a reliable and valid measure. *Journal of School Violence*. doi:10.1080/15388220.2014.964801
- Denny, S., Peterson, E. R., Stuart, J., Utter, J., Bullen, P., Fleming, T., ... & Milfont, T. (2015). Bystander intervention, bullying, and victimization: A multilevel analysis of New Zealand high schools. *Journal of School Violence*, 14(3), 245-272.
- Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika*, 65(2), 241-261.
- Elliott, S. N., Hwang, Y. S., & Wang, J. (2019). Teachers' ratings of social skills and problem behaviors as concurrent predictors of students' bullying behavior. *Journal of Applied Developmental Psychology*, 60, 119-126.
- Espelage, D. L., Bosworth, K., & Simon, T. R. (2000). Examining the social context of bullying behaviors in early adolescence. *Journal of Counseling & Development*, 78(3), 326-333.
- Espelage, D. L., & Holt, M. K. (2001). Bullying and victimization during early adolescence: Peer influences and psychosocial correlates. *Journal of Emotional Abuse*, 2(2-3), 123-142.
- Espelage, D.L., Holt, M.K., & Henkel, R.R. (2003). Examination of peer-group contextual effects on aggression during early adolescence. *Child Development*, 74, 205- 220.
- Espelage, D. L., Mebane, S. E., & Adams, R. S. (2004). Empathy, caring and bullying: Toward an understanding of complex associations. In D. L. Espelage & S.

- M. Swearer (Eds.), *Bullying in American schools: A social-ecological perspective on prevention and intervention* (pp. 37–61). Mahwah, NJ: Erlbaum.
- Espelage, D. L., & Swearer, S. M. (2003). Research on school bullying and victimization: What have we learned and where do we go from here? *School Psychology Review, 32*, 365–383.
- Espelage, E., & Swearer, S. (2010). A social-ecological model for bullying prevention and intervention: Understanding the impact of adults in the social ecology of youngsters. In S. R. Jimerson, S. M. Swearer, & D. L. Espelage (Eds.), *Handbook of bullying in schools: An international perspective* (pp. 347–362). New York, NY: Taylor & Francis.
- Ezpeleta, L., Granero, R., la Osa, N. d., Penelo, E., & Domènech, J. M. (2013). Psychometric properties of the strengths and difficulties Questionnaire3–4 in 3-year-old preschoolers. *Comprehensive Psychiatry, 54*(3), 282-291. doi:10.1016/j.comppsy.2012.07.009
- Farmer, T. W., Petrin, R., Robertson, D., Fraser, M., Hall, C., Day, S., & Dadisman, K. (2010). Peer relations of bullies, bully-victims, and victims: The two social worlds of bullying in second-grade classrooms. *The Elementary School Journal, 110*, 364–392. <http://dx.doi.org/10.1086/648983>
- Ferdinand, R. F., van der Ende, J., & Verhulst, F. C. (2004). Parent-adolescent disagreement regarding psychopathology in adolescents from the general population as a risk factor for adverse outcome. *Journal of Abnormal Psychology, 113*(2), 198.
- Ferdinand, R. F., van der Ende, J., & Verhulst, F. C. (2006). Prognostic value of parent–adolescent disagreement in a referred sample. *European Child & Adolescent Psychiatry, 15*(3), 156-162.
- Ferdinand, R. F., Van der Ende, J., & Verhulst, F. C. (2007). Parent–teacher disagreement regarding psychopathology in children: A risk factor for adverse outcome? *Acta Psychiatrica Scandinavica, 115*, 48-55. doi: 10.1111/j.1600-0447.2006.00843.x
- Finkelhor, D., Turner, H., Ormrod, R., Hamby, S., and Kracke, K. (2012). *Children’s Exposure to Violence: A Comprehensive Survey*. U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention. Available: <https://www.ncjrs.gov/pdffiles1/ojjdp/grants/248444.pdf> [June 2016].

- Finkelhor, D., Turner, H.A., Shattuck, A., and Hamby, S.L. (2015). *Violence, Crime, and Abuse Exposure in a National Sample of Children and Youth: An Update*. U.S. Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention. Available: <http://www.ojjdp.gov/pubs/248547.pdf> [June 2016].
- Fisher, H. L., Moffitt, T. E., Houts, R. M., Belsky, D. W., Arseneault, L., & Caspi, A. (2012). Bullying victimisation and risk of self-harm in early adolescence: Longitudinal cohort study. *BMJ*, *344*, e2683.
- Flake, J. K., & Petway, K. T. (2019). Methodologies for Investigating and Interpreting Student–Teacher Rating Incongruence in Noncognitive Assessment. *Educational Measurement: Issues and Practice*, *38*(1), 63-77.
- Fonagy, P., Twemlow, S. W., Vernberg, E. M., Nelson, J. M., Dill, E. J., Little, T. D., & Sargent, J. A. (2009). A cluster randomized controlled trial of child-focused psychiatric consultation and a school systems-focused intervention to reduce aggression. *Journal of Child Psychology and Psychiatry*, *50*(5), 607-616.
- Frey, J. R., Elliott, S. N., & Gresham, F. M. (2011). Preschoolers’ social skills: Advances in assessment for intervention using social behavior ratings. *School Mental Health*, *3*(4), 179-190.
- Frey, K. S., Hirschstein, M. K., Edstrom, L. V., & Snell, J. L. (2009). Observed reductions in school bullying, nonbullying aggression, and destructive bystander behavior: A longitudinal evaluation. *Journal of Educational Psychology*, *101*(2), 466.
- Furlong, M. J., Sharkey, J. D., Felix, E. D., Tanigawa, D., & Green, J. G. (2010). “Bullying assessment: A call for increased precision of self-reporting procedures,” in S. R. Jimerson, S. M. Swearer, & D. L. Espelage (eds.), *Handbook of bullying in schools: An international perspective* (pp. 329-345). New York, NY: Routledge.
- Gasser, L., & Keller, M. (2009). Are the competent the morally good? Perspective taking and moral motivation of children involved in bullying. *Social Development*, *18*, 798–816. doi:10.1111/j.1467-9507.2008.00516.x
- Georgiou, S. N., & Stavrinides, P. (2008). Bullies, victims and bully-victims: Psychosocial profiles and attribution styles. *School Psychology International*, *29*(5), 574-589.

- Gini, G., Pozzoli, T., & Hauser, M. (2011). Bullies have enhanced moral competence to judge relative to victims but lack moral compassion. *Personality and Individual Differences, 50*(5), 603-608.
- Gladden, R.M., Vivolo-Kantor, A.M., Hamburger, M.E., and Lumpkin, C.D. (2014). *Bullying Surveillance among Youths: Uniform Definitions for Public Health and Recommended Data Elements, Version 1.0*. Atlanta, GA: Centers for Disease Control and Prevention and U.S. Department of Education.
- Goldbaum, S., Craig, W. M., Pepler, D., & Connolly, J. (2003). Developmental trajectories of victimization: Identifying risk and protective factors. *Journal of Applied School Psychology, 19*(2), 139-156.
- Goldsmid, S., & Howie, P. (2014). Bullying by definition: An examination of definitional components of bullying. *Emotional and Behavioural Difficulties, 19*(2): 210–225. doi:10.1080/13632752.2013.844414
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586.
- Greif, J. L., & Furlong, M. J. (2006). The assessment of school bullying: Using theory to inform practice. *Journal of School Violence, 5*(3), 33-50.
- Gresham, F. M. (1981). Social skills training with handicapped children: A review. *Review of Educational Research, 51*(1), 139-176.
- Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System manual*. Circle Pines, MN: American Guidance Service.
- Gresham, F.M, & Elliott, S.N. (2008). *Social skills improvement system: Rating scales manual*. Minneapolis, MN: NCS Pearson, Inc.
- Gresham, F. M., Elliott, S. N., Cook, C. R., Vance, M. J., & Kettler, R. (2010). Cross-informant agreement for ratings for social skill and problem behavior ratings: An investigation of the Social Skills Improvement System—Rating Scales. *Psychological Assessment, 22*(1), 157.
- Gresham, F. M., Elliott, S. N., Vance, M. J., & Cook, C. R. (2011). Comparability of the Social Skills Rating System to the Social Skills Improvement System: Content and psychometric comparisons across elementary and secondary age levels. *School Psychology Quarterly, 26*, 27–44. doi:10.1037/a0022662
- Gresham, F. M., Reschly, D., Tilly, W. D., III., Fletcher, J., Burns, M. K., Christ, T. Prasse, D., Vanderwood, M., & Shinn, M. (2004). Comprehensive evaluation of

- learning disabilities: A response-to-intervention perspective. *Communiqué*, 33(4), 34-35.
- Gresham, F. M., Van, M. B., & Cook, C. R. (2006). Social skills training for teaching replacement behaviors: Remediating acquisition deficits in at-risk students. *Behavioral Disorders*, 31(4), 363-377.
- Griffin, R. S., & Gross, A. M. (2004). Childhood bullying: Current empirical findings and future directions for research. *Aggression and Violent Behavior*, 9(4), 379-400.
- Goldbaum, S., Craig, W. M., Pepler, D., & Connolly, J. (2003). Developmental trajectories of victimization: Identifying risk and protective factors. *Journal of Applied School Psychology*, 19(2), 139-156.
- Groeben, M., Perren, S., Stadelmann, S., & von Klitzing, K. (2011). Emotional symptoms from kindergarten to middle childhood: Associations with self-and other-oriented social skills. *European Child & Adolescent Psychiatry*, 20(1), 3-15.
- Gromann, P. M., Goossens, F. A., Olthof, T., Pronk, J., & Krabbendam, L. (2013). Self-perception but not peer reputation of bullying victimization is associated with non-clinical psychotic experiences in adolescents. *Psychological Medicine*, 43(4), 781-787.
- Guerra, N. G., Williams, K. R., & Sadek, S. (2011). Understanding bullying and victimization during childhood and adolescence: A mixed methods study. *Child development*, 82(1), 295-310.
- Habashy Hussein, M. (2013). The social and emotional skills of bullies, victims, and bully-victims of Egyptian primary school children. *International Journal of Psychology*, 48(5), 910-921.
- Haynie, D. L., Nansel, T., Eitel, P., Crump, A. D., Saylor, K., Yu, K., & Simons-Morton, B. (2001). Bullies, victims, and bully/victims: Distinct groups of at-risk youth. *The Journal of Early Adolescence*, 21(1), 29-49.
- Hebron, J., & Humphrey, N. (2014). Exposure to bullying among students with autism spectrum conditions: A multi-informant analysis of risk and protective factors. *Autism*, 18(6), 618-630.
- Henrich, C. C., & Shahar, G. (2014). Moderators of the effect of peer victimization during fifth grade on subsequent symptoms of (anxious) depression: The roles of engagement in bullying and baseline symptomatology. *Prevention Science*, 15(6), 888-896.

- Hill, L. G., Coie, J. D., Lochman, J. E., & Greenberg, M. T. (2004). Effectiveness of early screening for externalizing problems: issues of screening accuracy and utility. *Journal of Consulting and Clinical Psychology, 72*(5), 809.
- Hilton, N. Z., Harris, G. T., & Rice, M. E. (1998). On the validity of self-reported rates of interpersonal violence. *Journal of Interpersonal Violence, 13*(1), 58-72.
- Holt, T. J., Chee, G., Ng, E. A. H., & Bossler, A. M. (2013). Exploring the consequences of bullying victimization in a sample of Singapore youth. *International Criminal Justice Review, 23*(1), 25-40.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1-51.
- Humphrey, N., Kalambouka, A., Wigelsworth, M., Lendrum, A., Deighton, J., & Wolpert, M. (2011). Measures of social and emotional skills for children and young people: A systematic review. *Educational and Psychological Measurement, 71*(4), 617-637.
- Hunsley, J. (2002). Psychological testing and psychological assessment: A closer examination. *American Psychology, 57*, 139-140.
- Hunsley, J., & Mash, E. J. (2005). Introduction to the special section on developing guidelines for the evidence-based assessment (EBA) of adult disorders. *Psychological assessment, 17*(3), 251.
- Hwang, S., Kim, Y. S., Koh, Y., Bishop, S., & Leventhal, B. L. (2017). Discrepancy in perception of bullying experiences and later internalizing and externalizing behavior: A prospective study. *Aggressive Behavior, 43*(5), 493-502.
doi:10.1002/ab.21707
- Hymel, S., & Swearer, S. M. (2015). Four decades of research on school bullying: An introduction. *American Psychologist, 70*(4), 293.
- Iannotti, R.J. (2012). *Health Behavior in School-Aged Children (HBSC), 2005-2006*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.
- Iannotti, R.J. (2013). *Health Behavior in School-Aged Children (HBSC), 2009-2010*. Ann Arbor, MI: Inter-university Consortium for Political and Social Research.
- Jansen, D.E., Veenstra, R., Ormel, J., Verhulst, F.C., and Reijneveld, S.A. (2011). Early risk factors for being a bully, victim, or bully/victim in late elementary and early

- secondary education. The Longitudinal Trails Study. *BMC Public Health*, *11*(1), 1-7.
- Jansen, P.W., Verlinden, M., Dommisse-van Berkel, A., Mieloo, C., van der Ende, J., Veenstra, R., Verhulst, F.C., Jansen, W., and Tiemeier, H. (2012). Prevalence of bullying and victimization among children in early elementary school: Do family and school neighbourhood socioeconomic status matter? *BMC Public Health*, *12*(1), 1-10.
- Jenkins, L. N., Demaray, M. K., Fredrick, S. S., & Summers, K. H. (2016). Associations among middle school students' bullying roles and social skills. *Journal of School Violence*, *15*(3), 259-278.
- Jenkins, L. N., Demaray, M. K., & Tennant, J. (2017). Social, emotional, and cognitive factors associated with bullying. *School psychology review*, *46*(1), 42-64.
- Jimerson, S. R., & Furlong, M. J. (2006). The handbook of school violence and school safety. *From research to practice*.
- Jolliffe, D., & Farrington, D. P. (2006). Examining the relationship between low empathy and bullying. *Aggressive Behavior*, *32*, 540–550. doi:10.1002/ab.20154
- Jolliffe, D., & Farrington, D. P. (2011). Is low empathy related to bullying after controlling for individual and social background variables? *Journal of Adolescence*, *34*, 59–71. doi:10.1016/j.adolescence.2010.02.001
- Juvonen, J., Graham, S., & Schuster, M. A. (2003). Bullying among young adolescents: The strong, the weak, and the troubled. *Pediatrics-English Edition*, *112*(6), 1231-1237.
- Juvonen, J., Nishina, A., & Graham, S. (2001). Self-view versus peer perceptions for victim status among early adolescents. In J. Juvonen & S. Graham (Eds.), *Peer harassment in school: The plight of the vulnerable and victimized* (pp. 105–124). New York, NY: Guilford Press.
- Kane, K. (2015). THINK. In *Students, Teachers, and Leaders Addressing Bullying in Schools* (pp. 77-78). SensePublishers, Rotterdam.
- Kettler, R. J., Glover, T. A., Albers, C. A., & Feeney-Kettler, K. A. (2014). *Universal screening in educational settings: Evidence-based decision making for schools*. American Psychological Association.
- Kokkinos, C. M., & Kipritsi, E. (2012). The relationship between bullying, victimization, trait emotional intelligence, self-efficacy and empathy among

- preadolescents. *Social Psychology of Education*, 15, 41–58. doi:10.1007/s11218-011-9168-9
- Kokkinos, C., & Panayiotou, G. (2004). Predicting bullying and victimization among early adolescents: Associations with disruptive behavior disorders. *Aggressive Behavior*, 30, 520-533.
- Kollerová, L., & Smolík, F. (2016). Victimization and its associations with peer rejection and fear of victimization: Moderating effects of individual-level and classroom-level characteristics. *British Journal of Educational Psychology*, 86(4), 640-656.
- Kuterovac-Jagodić, G., & Velki, T. (2014). Different measurement approaches as sources of differences in data on the prevalence of peer violence. *Društvena Istraživanja - Časopis Za Opća Društvena Pitanja*, 23(2), 259-281.
- Ledwell, M., & King, V. (2015). Bullying and internalizing problems: Gender differences and the buffering role of parental communication. *Journal of Family Issues*, 36(5), 543-566.
- Lee, T., & Cornell, D. (2009). Concurrent validity of the Olweus bully/victim questionnaire. *Journal of School Violence*, 9(1), 56-73.
<https://doi.org/10.1080/15388220903185613>
- Lee, S., Smith, P. K., & Monks, C. P. (2016). Participant roles in peer-victimization among young children in South Korea: Peer-, self-, and teacher-nominations. *Aggressive Behavior*, 42(3), 287-298. doi:10.1002/ab.21623
- Limber, S.B. (2014). *Bullying Among Children and Youth*. Unpublished manuscript of a study commissioned by the Planning Committee on Increasing Capacity for Reducing Bullying and Its Impact on the Lifecourse of Youth Involved. Department of Psychology and Institute on Family and Neighborhood Life, Clemson University, SC.
- Limber, S.B., Olweus, D., and Luxenberg, H. (2013). *Bullying in U.S. Schools, 2012 Status Report*. Center City, MN: Hazelden Foundation.
- Little, J. A., & Rubin, D., B. (1987). Statistical analysis with missing data. *Hoboken, NJ: Wiley*.
- Lochman, J. E. (1995). Screening of child behavior problems for prevention programs at school entry. *Journal of Consulting and Clinical Psychology*, 63(4), 549.

- Lomas, J., Stough, C., Hansen, K., & Downey, L. A. (2012). Brief report: Emotional intelligence, victimisation and bullying in adolescents. *Journal of adolescence, 35*(1), 207-211.
- Marsh, H. W. (1989). Confirmatory factor analyses of multitrait-multimethod data: Many problems and a few solutions. *Applied Psychological Measurement, 13*(4), 335-361.
- Marsh, H. W., & Grayson, D. (1995). Latent variable models of multitrait-multimethod data. In: Hoyle, R. H., editor. *Structural Equation Modeling: Concepts, Issues, and Applications*. Thousand Oaks, CA: Sage; 1995. p. 177-198.
- Martino, S. C., Ellickson, P. L., Klein, D. J., McCaffrey, D., & Edelen, M. O. (2008). Multiple trajectories of physical aggression among adolescent boys and girls. *Aggressive Behavior, 34*(1), 61-75.
- Maxwell, J. A. (2012). *Qualitative research design: An interactive approach* (Vol. 41). Thousand Oaks, CA: Sage.
- Mitsopoulou, E., & Giovazolias, T. (2015). Personality traits, empathy and bullying behavior: A meta-analytic approach. *Aggression and Violent Behavior, 21*, 61-72.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine, 6*(7), e1000097.
- Monks, C. P., Palermiti, A., Ortega, R., & Costabile, A. (2011). A cross-national comparison of aggressors, victims and defenders in preschools in England, Spain and Italy. *The Spanish Journal of Psychology, 14*(1), 133-144.
Doi:10.5209/revSJOP.2011.v14.n1.11
- Monks, C.P., and Smith, P.K. (2006). Definitions of bullying: Age differences in understanding of the term and the role of experience. *British Journal of Developmental Psychology, 24*(4), 801-821.
- Monks, C. P., Smith, P. K., & Swettenham, J. (2003). Aggressors, victims, and defenders in preschool: Peer, self-, and teacher reports. *Merrill-Palmer Quarterly, 49*(4), 453-469. Doi:10.1353/mpq.2003.0024
- Nansel, T. R., Craig, W., Overpeck, M. D., Saluja, G., & Ruan, W. J. (2004). Cross-national consistency in the relationship between bullying behaviors and psychosocial adjustment. *Archives of Pediatrics & Adolescent Medicine, 158*(8), 730-736.

- Nansel, T. R., Overpeck, M., Pilla, R. S., Ruan, W. J., Simons-Morton, B., & Scheidt, P. (2001). Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. *Jama*, 285(16), 2094-2100.
- NCES. (2015, April). *Student reports of bullying and cyber-bullying: Results from the 2013 school crime supplement to the National Crime Victimization Survey*. Retrieved from <https://nces.ed.gov/pubs2015/2015056.pdf>
- Neal, J. W., Cappella, E., Wagner, C., & Atkins, M. S. (2011). Seeing eye to eye: Predicting teacher-student agreement on classroom social networks. *Social Development*, 20(2), 376-393.
- Nowell, K. P., Brewton, C. M., & Goin-Kochel, R. P. (2014). A multi-rater study on being teased among children/adolescents with autism spectrum disorder (ASD) and their typically developing siblings: associations with ASD symptoms. *Focus on Autism and Other Developmental Disabilities*, 29(4), 195-205.
- Olino, T. M., & Klein, D. N. (2015). Psychometric comparison of self- and informant-reports of personality. *Assessment*, 22(6), 655–664. <http://doi.org/10.1177/1073191114567942>
- Olweus, D. (1996). Revised Olweus bully/victim questionnaire. *British Journal of Educational Psychology*.
- Olweus, D. (2010). Understanding and researching bullying. In S. R. Jimerson, S. M. Swearer, & D. L. Espelage (Eds.), *Handbook of bullying in schools* (pp. 9–33). New York, NY: Routledge.
- Olweus, D. (2013). School bullying: Development and some important challenges. *Annual Review of Clinical Psychology*, 9, 751-780.
- Pellegrini, A. D. (2001). “Sampling instances of victimization in mile school: A methodological comparison,” in J. Juvonen & S. Graham (eds.), *Peer harassment in school: The plight of the vulnerable and victimized* (pp. 125–144). New York, NY: Guilford Press.
- Pellegrini, A. D., & Bartini, M. (2001). Dominance in early adolescent boys: Affiliative and aggressive dimensions and possible functions. *Merrill-Palmer Quarterly*, 142-163.
- Pellegrini, A. D., Bartini, M., & Brooks, F. (1999). School bullies, victims, and aggressive victims: Factors relating to group affiliation and victimization in early adolescence. *Journal of Educational Psychology*, 91(2), 216.

- Pepler, D., Craig, W., & O'Connell, P. (2010). "Peer processes in bullying: Informing prevention and intervention strategies," in S. R. Jimerson, S. M. Swearer, & D. L. Espelage (eds.), *Hand- book of bullying in schools: An international perspective* (pp. 469– 479). New York, NY: Routledge.
- Perren, S., & Alsaker, F. D. (2006). Social behavior and peer relationships of victims, bully-victims, and bullies in kindergarten. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 47, 45–57.
doi:10.1111/j.1469-7610.2005.01445.x
- Perren, S., Forrester-Knauss, C., & Alsaker, F. D. (2012). Self- and other-oriented social skills: Differential associations with children's mental health and bullying roles. *Journal for Educational Research Online*, 4, 99–123.
- Phillips, V., & Cornell, D. (2012). Identifying victims of bullying: Use of counselor interviews to confirm peer nominations. *Professional School Counseling*, 15(3), 123-131.
- Pohl, S., Steyer, R., & Kraus, K. (2008). Modeling method effects as individual causal effects. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 171(1), 41-63.
- Postigo, S., González, R., Montoya, I., & Ordoñez, A. (2013). Theoretical proposals in bullying research: a review. *Anales de psicología*, 29(2).
- Power-Elliott, M., & Harris, G. (2012). Guidance counselor strategies for handling bullying. *British Journal of Guidance & Counselling*, 40(1), 83-98.
- Rigby, K. (2007). *Bullying in schools: And what to do about it*. Aust Council for Ed Research.
- Rodkin, P. C. (2012). Bullying and children's peer relationships. *Colleagues*, 8(2), 4.
- Rodkin, P.C., Farmer, T.W., Pearl, R., & Van Acker, R. (2006). They're cool: Social status and peer group supports for aggressive boys and girls. *Social Development*, 15, 175-204.
- Rodkin, P. C., & Roisman, G. I. (2010). Antecedents and correlates of the popular-aggressive phenomenon in elementary school. *Child Development*, 81, 837–850.
<http://dx.doi.org/10.1111/j.1467-8624.2010.01437.x>
- Rønning, J. A., Sourander, A., Kumpulainen, K., Tamminen, T., Niemelä, S., Moilanen, I., ... & Almqvist, F. (2009). Cross-informant agreement about bullying and victimization among eight-year-olds: Whose information best predicts psychiatric

- caseness 10–15 years later?. *Social Psychiatry and Psychiatric Epidemiology*, 44(1), 15-22.
- Rupp, S., Elliott, S. N., & Gresham, F. M. (2018). Assessing elementary students' bullying and related social behaviors: Cross-informant consistency across school and home environments. *Children and Youth Services Review*, 93, 458-466.
- Sadinejad, M., Bahreynian, M., Motlagh, M. E., Qorbani, M., Movahhed, M., Ardalan, G., ... & Kelishadi, R. (2015). Frequency of aggressive behaviors in a nationally representative sample of Iranian children and adolescents: The CASPIAN-IV study. *International Journal of Preventive Medicine*, 6.
- Salmivalli, C., Huttunen, A., & Lagerspetz, K.M.J., (1997). Peer networks and bullying in schools. *Scandinavian Journal of Psychology*, 38, 205-312.
- Sameroff, A. (2009). *The transactional model*. American Psychological Association.
- Savage, M. W., & Tokunaga, R. S. (2017). Moving toward a theory: Testing an integrated model of cyberbullying perpetration, aggression, social skills, and Internet self-efficacy. *Computers in human Behavior*, 71, 353-361.
- Sawyer, A.L., Bradshaw, C.P., & O'Brennan, L.M. (2008). Examining ethnic, gender, and developmental differences in the way children report being a victim of “bullying” on self-report measures. *Journal of Adolescent Health*, 43, 106-114. doi:10.1016/j.jadohealth.2007.12.011.
- Schwartz, D., Proctor, L. J., & Chien, D. H. (2001). The aggressive victim of bullying. *Peer harassment in school: The plight of the vulnerable and victimized*, 147-174.
- Shahaeian, A., Razmjooe, M., Wang, C., Elliott, S.N., & Hughes, C. (2017). Understanding relational aggression during early childhood: An examination of the association with language and other social and cognitive skills. *Early Childhood Research Quarterly*, 40(3), 204-214.
- Sharp, S., & Smith, P. (2002). *School bullying: Insights and perspectives*. Routledge.
- Shiner, R. L., & Allen, T. A. (2013). Assessing personality disorders in adolescents: Seven guiding principles. *Clinical Psychology: Science and Practice*, 20(4), 361-377.
- Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, 23(4), 323-355.

- Snijders, T. A. B., & Bosker, R. J. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London, UK: Sage.
- Solberg, M.E. (2003). Prevalence estimation of school bullying with the Olweus bully/victim questionnaire. *Aggressive Behavior, 29*(3), 239-268.
- Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus Bully/Victim Questionnaire. *Aggressive Behavior, 29*, 239–268.
doi:10.1002/ab.10047
- Sourander, A., Brunstein Klomek, A., Kumpulainen, K., Puustjärvi, A., Elonheimo, H., Ristkari, T., . . . Ronning, J. A. (2011). Bullying at age eight and criminality in adulthood: Findings from the Finnish nationwide 1981 birth cohort study. *Social Psychiatry and Psychiatric Epidemiology, 46*(12), 1211-1219.
doi:10.1007/s00127-010-0292-1
- Stephenson, P., & Smith, D. (1989). Bullying in the junior schools. In D.P. Tattum & D.A. Lane (Eds.), *Bullying in schools* (pp. 45–57). Stoke-on-Trent: Trentham Books.
- Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. Routledge.
- Sutton, J., Smith, P. K., & Swettenham, J. (1999). Bullying and ‘theory of mind’: A critique of the ‘social skills deficit’ view of anti-social behaviour. *Social Development, 8*(1), 117-127.
- Swearer, S. M. (2011). Risk factors for and outcomes of bullying and victimization.
- Thomas, H. J., Connor, J. P., Lawrence, D. M., Hafekost, J. M., Zubrick, S. R., & Scott, J. G. (2017). Prevalence and correlates of bullying victimisation and perpetration in a nationally representative sample of Australian youth. *Australian & New Zealand Journal of Psychiatry, 51*(9), 909-920.
- Tremblay, R. E., Loeber, R., Gagnon, C., Charlebois, P., Larivee, S., & LeBlanc, M. (1991). Disruptive boys with stable and unstable high fighting behavior patterns during junior elementary school. *Journal of Abnormal Child Psychology, 19*(3), 285-300.
- Ttofi, M. M., & Farrington, D. P. (2012). Risk and protective factors, longitudinal research, and bullying prevention. *New Directions for Youth Development, 2012*(133), 85-98.

- Ttofi, M. M., Farrington, D. P., Lösel, F., & Loeber, R. (2011). The predictive efficiency of school bullying versus later offending: A systematic/meta-analytic review of longitudinal studies. *Criminal Behavior and Mental Health, 21*, 80–89. doi:10.1002/cbm.808
- Ung, D., Boone, D. M., McBride, N., Howie, F., Scalli, L., & Storch, E. A. (2017). Parent and teacher agreement of behavior problems in youth diagnosed with and without autism spectrum disorders. *Journal of Child and Family Studies, 26*(2), 370-380. doi:10.1007/s10826-016-0566-7
- Unnever, J. D., & Cornell, D. G. (2003). Bullying, self-control, and ADHD. *Journal of Interpersonal Violence, 18*, 129–147. doi:10.1177/0886260502238731
- U.S. Department of Education. (2015). *Student Reports of Bullying and Cyber-Bullying: Results from the 2013 School Crime Supplement to the National Crime Victimization Survey*. Washington DC: Author. Available: <http://nces.ed.gov/pubs2015/2015056.pdf> [April 2015].
- Vaillancourt, T., Hymel, S., & McDougall, P. (2003). Bullying is power: Implications for school-based intervention strategies. *Journal of Applied School Psychology, 19*, 157–176. http://dx.doi.org/10.1300/J008v19n02_10
- Vaillancourt, T., McDougall, P., Hymel, S., Krygsman, A., Miller, J., Stiver, K., & Davis, C. (2008). Bullying: Are researchers and children/youth talking about the same thing? *International Journal of Behavioral Development, 32*(6), 486-495.
- Van der Oord, S., Van der Meulen, E. M., Prins, P. J., Oosterlaan, J., Buitelaar, J. K., & Emmelkamp, P. M. (2005). A psychometric evaluation of the social skills rating system in children with attention deficit hyperactivity disorder. *Behaviour research and therapy, 43*(6), 733-746.
- Van Noorden, T. H., Haselager, G. J., Cillessen, A. H., & Bukowski, W. M. (2015). Empathy and involvement in bullying in children and adolescents: A systematic review. *Journal of Youth and Adolescence, 44*(3), 637-657.
- Vaughn, M. G., Fu, Q., Bender, K., DeLisi, M., Beaver, K. M., Perron, B. E., & Howard, M. O. (2010). Psychiatric correlates of bullying in the United States: Findings from a national sample. *Psychiatric Quarterly, 81*(3), 183-195.
- Veenstra, R., Lindenberg, S., Oldehinkel, A. J., De Winter, A. F., Verhulst, F. C., & Ormel, J. (2005). Bullying and victimization in elementary schools: a comparison of bullies, victims, bully/victims, and uninvolved preadolescents. *Developmental Psychology, 41*(4), 672.

- Verlinden, V., Tiemeier, H., Veenstra, R., Mieloo, C., Jansen, W., Jaddoe, V., . . . Jansen, P. (2014). Television viewing through ages 2-5 years and bullying involvement in early elementary school. *BMC Public Health, 14*(1), 157. doi:10.1186/1471-2458-14-157
- Vlachou, M., Andreou, E., Botsoglou, K., & Didaskalou, E. (2011). Bully/victim problems among preschool children: A review of current research evidence. *Educational Psychology Review, 23*(3), 329.
- Vlachou, M., Botsoglou, K., & Andreou, E. (2013). Assessing bully/victim problems in preschool children: A multimethod approach. *Journal of Criminology, 2013*. <http://dx.doi.org/10.1155/2013/301658>
- Volk, A. A., Dane, A. V., & Marini, Z. A. (2014). What is bullying? A theoretical redefinition. *Developmental Review, 34*, 327–343.
- Volk, A. A., Veenstra, R., & Espelage, D. L. (2017). So you want to study bullying? Recommendations to enhance the validity, transparency, and compatibility of bullying research. *Aggression and Violent Behavior, 36*, 34-43.
- Vreeman, R. C., & Carroll, A. E. (2007). A systematic review of school-based interventions to prevent bullying. *Archives of Pediatrics & Adolescent Medicine, 161*(1), 78-88.
- Waller, R., Hyde, L. W., Baskin-Sommers, A. R., & Olson, S. L. (2017). Interactions between callous unemotional behaviors and executive function in early childhood predict later aggression and lower peer-liking in late-childhood. *Journal of Abnormal Child Psychology, 45*(3), 597-609.
- Walthall, J. C., Konold, T. R., & Pianta, R. C. (2005). Factor structure of the social skills rating system across child gender and ethnicity. *Journal of Psychoeducational Assessment, 23*(3), 201-215.
- Wang, F. M., Chen, J. Q., Xiao, W. Q., Ma, Y. T., & Zhang, M. (2012). Peer physical aggression and its association with aggressive beliefs, empathy, self-control, and cooperation skills among students in a rural town of China. *Journal of Interpersonal Violence, 27*, 3252–3267. doi:10.1177/0886260512441256
- Wang, J., & Iannotti, R. J. (2012). Bullying among US adolescents. *The Prevention Researcher, 19*(3), 3-6.
- Wegge, D., Vandebosch, H., Eggermont, S., Van Rossem, R., & Walrave, M. (2016). Divergent perspectives: Exploring a multiple informant approach to cyberbullying

- victimization and perpetration. *European Journal on Criminal Policy and Research*, 22(2), 235-251.
- White, B. A., & Kistner, J. A. (2011). Biased self-perceptions, peer rejection, and aggression in children. *Journal of Abnormal Child Psychology*, 39(5), 645-656. doi:10.1007/s10802-011-9506-6
- Winsper, C., Lereya, T., Zanarini, M., & Wolke, D. (2012). Involvement in bullying and suicide-related behavior at 11 years: A prospective birth cohort study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(3), 282.e3. doi:10.1016/j.jaac.2012.01.001
- Wolke, D., Lereya, S. T., Fisher, H. L., Lewis, G., & Zammit, S. (2014). Bullying in elementary school and psychotic experiences at 18 years: A longitudinal, population-based cohort study. *Psychological Medicine*, 44(10), 2199-2211. doi:10.1017/S0033291713002912
- Wolke, D., Woods, S., Bloomfield, L., & Karstadt, L. (2000). The association between direct and relational bullying and behaviour problems among primary school children. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 41(8), 989-1002.
- Yanagida, T., Gradinger, P., Strohmeier, D., Solomontos-Kountouri, O., Trip, S., & Bora, C. (2016). Cross-national prevalence of traditional bullying, traditional victimization, cyberbullying and cyber-victimization: Comparing single-item and multiple-item approaches of measurement. *International Journal of Developmental Science*, 10(1-2), 21-32.
- Yubero, S., & Navarro, R. (2006). Students' and teachers' views of gender-related aspects of aggression. *School Psychology International*, 27(4), 488-512.
- Zimmerman, F.J., Glew, G.M., Christakis, D.A., and Katon, W. (2005). Early cognitive stimulation, emotional support, and television watching as predictors of subsequent bullying among grade-school children. *Archives of Pediatrics & Adolescent Medicine*, 159(4), 384-388.

APPENDIX A
SUPPLEMENTAL TABLES

*Raw Score Mean Comparison of Self-Identified Bullies and Non-Self-Identified Bullies
(n = 1375)*

Bully Status	Sex	Age Group	Race/Ethnicity	Mean	N	SD	% of Total Sum	% of Total N
No	Female	8 to 11	Black	0.70	30	1.18	1%	2%
			Hispanic	1.01	83	1.32	4%	6%
			Other	1.00	22	1.54	1%	2%
			White	0.78	224	1.15	8%	16%
			Total	0.84	359	1.22	13%	26%
		12 to 14	Black	0.79	19	0.86	1%	1%
			Hispanic	0.58	38	0.92	1%	3%
			Other	0.64	11	0.81	0%	1%
			White	0.97	139	1.20	6%	10%
			Total	0.86	207	1.12	8%	15%
		15 to 18	Black	1.17	12	1.03	1%	1%
			Hispanic	0.65	20	1.04	1%	2%
			Other	0.83	12	0.84	0%	1%
			White	1.22	64	1.41	3%	5%
			Total	1.06	108	1.26	5%	8%
		Total	Black	0.82	61	1.06	2%	4%
			Hispanic	0.84	141	1.20	5%	10%
			Other	0.87	45	1.22	2%	3%
			White	0.91	427	1.22	17%	31%
			Total	0.88	674	1.20	26%	49%
			Male	8 to 11	Black	1.73	22	1.42
Hispanic	0.73				51	1.10	2%	4%
Other	1.07				14	1.39	1%	1%
White	1.08				177	1.43	8%	13%
Total	1.07				264	1.38	12%	19%
12 to 14	Black			1.38	16	1.41	1%	1%
	Hispanic			0.79	34	1.20	1%	3%
	Other			1.87	15	1.36	1%	1%
	White			1.23	111	1.42	6%	8%
	Total			1.22	176	1.39	9%	13%
15 to 18	Black			1.30	10	1.70	1%	1%
	Hispanic			1.17	24	1.31	1%	2%
	Other			1.13	8	1.64	0%	1%

*Raw Score Mean Comparison of Self-Identified Bullies and Non-Self-Identified Bullies
(n = 1375)*

Bully Status	Sex	Age Group	Race/Ethnicity	Mean	N	SD	% of Total Sum	% of Total N
			White	1.50	52	1.39	3%	4%
			Total	1.36	94	1.41	6%	7%
		Total	Black	1.52	48	1.46	3%	4%
			Hispanic	0.84	109	1.18	4%	8%
			Other	1.41	37	1.44	2%	3%
			White	1.20	340	1.42	18%	25%
			Total	1.17	534	1.39	27%	39%
	Total	8 to 11	Black	1.13	52	1.37	3%	4%
			Hispanic	0.90	134	1.24	5%	10%
			Other	1.03	36	1.46	2%	3%
			White	0.92	401	1.29	16%	29%
			Total	0.94	623	1.30	25%	45%
		12 to 14	Black	1.06	35	1.16	2%	3%
			Hispanic	0.68	72	1.06	2%	5%
			Other	1.35	26	1.29	2%	2%
			White	1.09	250	1.31	12%	18%
			Total	1.03	383	1.26	17%	28%
		15 to 18	Black	1.23	22	1.34	1%	2%
			Hispanic	0.93	44	1.21	2%	3%
			Other	0.95	20	1.19	1%	2%
			White	1.34	116	1.40	7%	8%
			Total	1.20	202	1.34	11%	15%
		Total	Black	1.13	109	1.29	5%	8%
			Hispanic	0.84	250	1.19	9%	18%
			Other	1.11	82	1.34	4%	6%
			White	1.04	767	1.32	34%	56%
			Total	1.01	1208	1.29	53%	88%
Yes	Female	8 to 11	Black	5.89	9	1.05	2%	1%
			Hispanic	7.57	7	1.62	2%	1%
			Other	0.00	0	0.00	0%	0%
			White	8.27	11	3.38	4%	1%
			Total	7.30	27	2.54	9%	2%
		12 to 14	Black	5.67	3	1.16	1%	0%
			Hispanic	6.20	5	1.30	1%	0%

*Raw Score Mean Comparison of Self-Identified Bullies and Non-Self-Identified Bullies
(n = 1375)*

Bully Status	Sex	Age Group	Race/Ethnicity	Mean	N	SD	% of Total Sum	% of Total N
			Other	6.00	2	1.41	1%	0%
			White	5.47	15	0.74	4%	1%
			Total	5.68	25	0.95	6%	2%
		15 to 18	Black	7.00	1	.	0%	0%
			Hispanic	6.33	3	1.16	1%	0%
			Other	5.50	2	0.71	1%	0%
			White	5.56	9	1.01	2%	1%
			Total	5.80	15	1.01	4%	1%
		Total	Black	5.92	13	1.04	3%	1%
			Hispanic	6.87	15	1.51	5%	1%
			Other	5.75	4	0.96	1%	0%
			White	6.37	35	2.35	10%	3%
			Total	6.36	67	1.92	18%	5%
	Male	8 to 11	Black	6.91	11	2.77	3%	1%
			Hispanic	7.33	15	2.44	5%	1%
			Other	5.00	2	0.00	0%	0%
			White	6.57	30	2.18	9%	2%
			Total	6.78	58	2.33	17%	4%
		12 to 14	Black	5.00	1	.	0%	0%
			Hispanic	9.33	3	3.79	1%	0%
			Other	5.00	3	0.00	1%	0%
			White	5.80	15	1.27	4%	1%
			Total	6.14	22	2.05	6%	2%
		15 to 18	Black	5.75	4	0.96	1%	0%
			Hispanic	6.00	2	1.41	1%	0%
			Other	8.00	4	0.00	1%	0%
			White	7.00	10	2.00	3%	1%
			Total	6.85	20	1.66	6%	2%
		Total	Black	6.50	16	2.39	5%	1%
			Hispanic	7.50	20	2.61	7%	2%
			Other	6.33	9	1.58	3%	1%
			White	6.44	55	1.95	15%	4%
			Total	6.65	100	2.15	29%	7%
	Total	8 to 11	Black	6.45	20	2.19	6%	2%

*Raw Score Mean Comparison of Self-Identified Bullies and Non-Self-Identified Bullies
(n = 1375)*

Bully Status	Sex	Age Group	Race/Ethnicity	Mean	N	SD	% of Total Sum	% of Total N
			Hispanic	7.41	22	2.18	7%	2%
			Other	5.00	2	0.00	0%	0%
			White	7.02	41	2.62	13%	3%
			Total	6.94	85	2.40	26%	6%
		12 to 14	Black	5.50	4	1.00	1%	0%
			Hispanic	7.38	8	2.77	3%	1%
			Other	5.40	5	0.89	1%	0%
			White	5.63	30	1.03	7%	2%
			Total	5.89	47	1.56	12%	3%
		15 to 18	Black	6.00	5	1.00	1%	0%
			Hispanic	6.20	5	1.10	1%	0%
			Other	7.17	6	1.33	2%	0%
			White	6.32	19	1.73	5%	1%
			Total	6.40	35	1.50	10%	3%
		Total	Black	6.24	29	1.90	8%	2%
			Hispanic	7.23	35	2.20	11%	3%
			Other	6.15	13	1.41	4%	1%
			White	6.41	90	2.10	25%	7%
			Total	6.53	167	2.06	47%	12%