BULLYING MEASUREMENT STUDY

By

CAROLYN S. LITTLE

Bachelor of Arts in Psychology

Auburn University

Auburn, Alabama

2004

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 2007

BULLYING MEASUREMENT STUDY

Thesis Approved:

Dr. Cynthia Hartung
_____
Thesis Adviser

Dr. Melanie Page
_____

Dr. Douglas Scambler
_____

Dr. A. Gordon Emslie
_____
Dean of the Graduate College

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Measurement of Bullying and Victimization in Schools

Bullying is a pervasive problem in the United States (US) and internationally (Nansel,

Overpeck, Pilla, Ruan, Simons-Morton & Scheidt, 2001). In recent years bullying has been at the

forefront of news in the US due to incidents of school violence (e.g., Littleton, Colorado and

Jonesboro, Arkansas). Bullying is repeated exposure to a negative action, such as physical or

verbal aggression, by one or more individuals (Olweus, 1993). Olweus defines a negative action

as a purposeful attempt to injure or inflict discomfort on someone, either directly or indirectly.

Bullying can be inflicted by a group or an individual, but a key to determining whether the

behavior in question meets the definition of bullying is whether there is an imbalance of power

among the individuals involved. This imbalance of power could be physical or social, such as

differences in physical stature or popularity (Olweus, 1993). Direct bullying includes physical

aggression (e.g., hitting or shoving) and verbal aggression (e.g., name-calling, shouting, or

accusing). Indirect bullying includes inflicting emotional pain through social isolation, group

exclusion, and/or manipulation of relationships or friendships (Olweus, 1993).

In the literature aggression has received much more attention than bullying. A child who

meets the definition of bullying does so because they use aggressive means towards another

person. Aggression can be physical, verbal, or relational. However, just because a child acts

aggressively does not mean that they are bullying. The behavior becomes bullying if the

aggressive acts are repeated, directed to someone of unequal status, and meant to cause some sort

of intentional harm to the recipient (Olweus, 1993). Victims of aggression are subject to some of the same psychosocial problems as victims of bullying (Pellegrini and Bartini, 2000). Due to the greater quantity of information on aggression and the strong relation of these two constructs, it is appropriate to include literature on aggression within the bullying literature.

Estimates of Bullying Prevalence

Research has shown that in the US, approximately 10-20% of students have experienced bullying (Pellegrini, Bartini, & Brooks, 1999; Perry, Kusel & Perry, 1988). A nationwide survey of students in grades six through ten (n = 15,686) indicated that approximately 29.9% of the sample reported being involved in bullying as either a victim or a bully. Of this group, 13.0% were considered bullies, 10.6% were considered victims and 6.3% were considered both bullies and victims. These authors found that boys were more likely to be involved in bullying than girls, as both victims and bullies, and there was an increase in prevalence in grades six through eight (Nansel et al., 2001). This increase in middle school is consistent with prior research showing that bullying peaks in the middle school years and then decreases with age (Nansel et al., 2001; Olweus, 1993; Swearer & Carey, 2003). Indirect bullying is as prevalent for boys as it is for girls. As age increases physical bullying may decline, but indirect bullying remains high (Perry, Kusel, & Perry, 1988). However, name-calling seems to decrease with maturity (Dennis & Satcher, 1999; Embry & Luzzo, 1996). Research has consistently shown that boys display higher rates of direct physical aggression than girls (Lindeman, Harakka, Keltikangas-Jarvinen, 1997; Nansel et al., 2001; Olweus, 1993). In contrast, girls display higher rates of relational aggression than direct physical aggression (Casey-Cannon, Hayward, & Gowen, 2001; Perry, Kusel, & Perry, 1988). Although boys display more direct physical aggression than girls, rates of relational aggression are similar for boys and girls (Olweus, 1994).

Where Bullying Occurs

Bullying occurs in places where teachers are less likely to be directly observing students' behaviors, such as the playground, school bus, and lunchroom (Olweus, 1993; Pellegrini &

Bartini, 2000; Pepler & Craig, 1995). Even though teachers are present in these settings, they cannot always see what each student is doing and they are often too far away to hear every conversation that is occurring. When teachers pay attention to, and reinforce, appropriate behaviors the number of aggressive incidents is reduced. For example, Roderick, Pitchford, and Miller (1997) found that giving students (aged five to seven) a raffle ticket when they displayed appropriate playground behaviors, resulted in a reduction in physical aggression by 47-75%. A study by Colvin, Sugai, Good, and Lee (1997) looked at differences in aggressive behaviors during transition periods, such as walking to the cafeteria, when teachers were instructed to use active supervision and precorrection. For the purpose of this study active supervision consisted of the teachers moving around, visually scanning areas, and interacting with students. Precorrection consisted of teachers reminding students about appropriate behaviors immediately before entering into the transition area. When the teachers used precorrection and active supervision, bullying behaviors were about five times less frequent than at baseline (Colvin et al., 1997).

Correlates of Bullying and Victimization

Involvement in bullying, as either a bully or a victim, has serious short-term and long-term correlates (Bernstein & Watson, 1997; Dulmus, Theriot, & Sowers, 2004; Gilmartin, 1987; Olweus, 1993). For example, Dulmus et al. (2004) found that victimized students were more likely to dislike or feel neutral about school than non-victimized students (32.9% vs. 8.8%). Victims have also been found to display school phobia, concentration difficulties and/or psychological distress (Bernstein & Watson, 1997). In addition, victims have been shown to be anxious, insecure, and/or suffer from low self-esteem, compared to other students (Olweus, 1993). There is also a correlation with relationship problems later in life; for example, Gilmartin (1987) found that men who suffered victimization in childhood also had difficulty initiating and maintaining intimate relationships in adulthood.

Olweus (1994) suggested that although bullies act aggressively, they may have anxious personality styles. Children who bully throughout childhood and adolescence are more aggressive

and more disruptive than children who only bully in childhood or do not bully at all (Scholte, Engels, Overbeek, Kemp, & Haselager, 2007). Olweus (1994) suggested that bullies may be more likely to hold pro-violence attitudes and display behaviors that are viewed as more antisocial than other children (Olweus, 1994). In terms of long-term correlates, bullies are at risk for such problems as domestic violence, criminality, and substance abuse (Olweus, 1993; Olweus, 1994). Furthermore, bullies are also more likely to drop out of high school (Olweus, 1994).

CHAPTER II

REVIEW OF LITERATURE

Measurement of Bullying and Victimization

In the US, federal and state governments are beginning to take notice of the problems that bullying is causing in schools, and some states are even legislating that schools adopt anti-bullying policies and/or bullying prevention programs. However, research is needed to examine the effectiveness of these policies and programs. In order to determine if these programs are effective, levels of bullying and victimization following program implementation must be compared to levels of bullying and victimization prior to program implementation. Although multiple measures have been created to assess bullying and victimization, many of these measurement tools have not been subjected to research evaluation, have not shown adequate psychometric properties, or have not been cross-validated in multiple laboratories. Whereas some of the measures have demonstrated poor reliability or validity, the majority of measures have demonstrated adequate reliability, but inadequate validity. Until sufficient validity is demonstrated, it will not be clear whether these measures are accurately estimating the frequency of bullying and victimization in schools. As a result, it will be difficult to determine whether policies and/or prevention programs are effective. Methods such as unstructured observations, structured observations, structured interviews, sociometric procedures, and questionnaires offer different techniques for assessing the problem of bullying, while offering different benefits and disadvantages. Each of these procedures will be described and the psychometric properties of existing measures will be reviewed.

Unstructured Observations

During unstructured observations, an observer chooses a time when bullying is likely to occur and then monitors and records factors such as social status, social isolation, social withdrawal, and bullying. A study by Evans and Eder (1993) used unstructured observations in a high school cafeteria. During this three-year study, observers attempted to join different groups on the school social hierarchy (e.g., athletes or band members) during lunchtime. The researchers sat with their assigned group in the cafeteria and took notice of the bullying/victimization that occurred. They attempted to fit in as much as possible by dressing casually, not taking notes, and not conversing with teachers or school staff during lunch time. They were able to observe who was bullying and who was being victimized. After three years of observations, the researchers concluded that appearance, unattractiveness, gender-inconsistent behavior, lack of certain social skills and mental immaturity were predictors of victimization.

Although unstructured observations are beneficial since researchers can observe behaviors as they occur in the real world, there are also problems with this method (Crothers & Levinson, 2004). One such problem is reactivity, meaning that the mere presence of researchers may change behavior. Additionally, it may also be difficult to classify, categorize, and measure specific behavior without using a structured coding system and therefore may be difficult to obtain sufficient inter-rater reliability. Unstructured observations are also financially expensive and time-consuming. Finally, unstructured interviews require consent and assent from all parents and students. This may influence the sample, due to the characteristics of who returns the forms.

Structured Observations

In a structured observation the researcher chooses a time where bullying is likely to occur and then observes and tracks specific variables. The variables are operationally defined for the behaviors that are being measured. Boulton (1993) investigated aggressive fighting in middle school students using consecutive studies that built upon the limitations of the first study. In the first study, the students (n = 110) were interviewed individually about their attitudes toward

school and teachers, aspects of peer relationships, and playground activities including fighting and aggression. The students reported that physical altercations were usually caused by name calling, teasing, response to aggression, and disputes over the game currently being played. A kappa of .91 was found for inter-rater reliability when an independent rater compared the responses of the observers.

In the second study (Boulton, 1993) used a structured observation to assess 86 students on the playground. For each recess period, one child was picked as a target child and a portable audio tape-recorder was used to document this child's verbalizations. Inter-rater reliability was determined by one observer watching the child in close proximity, listening to the audio and rating behaviors while the second observer videotaped from further away. Cohen's kappa for inter-rater reliability was found to be .93. Later two different raters coded whether or not the two other researchers verbal descriptions matched the videotape. The researchers observed 114 instances of aggressive fighting. One limitation of this study is that since the researchers were attempting to be as discreet as possible, they did not ask the participants' about their perception of the situations were not included. Another possible limitation of this study, is that it was conducted in the United Kingdom and may not be generalizable to the United States.

One very effective method of conducting structured observations is to take advantage of technology. Pepler and Craig (1995) observed peer interactions of aggressive and non-aggressive students using video cameras and microphones. Inter-rater reliability was established through coding and comparing behaviors. To establish external validity they compared the ratings of observations to other measures of aggression. Children who were considered aggressive based on other measures were observed to be aggressive (physically and verbally) every three to eight minutes. Children who were not considered aggressive based on other measures were observed to be aggressive every five to eleven minutes. Although girls admitted to lower levels of bullying than boys on questionnaires, Pepler and Craig (1995) observed girls to bully as often as boys when direct and indirect bullying were both measured. Although using technology seems to be a

very effective method for assessing the rates of bullying, it is not always feasible due to the high cost and time commitment. Another issue with feasibility is that when assessing the whole school, the researcher will have to get consent and assent from all children and all parents in the school. As mentioned previously, this is a potential problem for both unstructured and structured observations.

Structured Interviews

A structured interview includes a standard set of questions and an individual's responses to the questions determine subsequent questions. Casey-Cannon, Hayward, and Gowen (2001) used a structured interview to assess relational aggression in a sample of adolescent girls. This study was part of a larger study about body image and a subset of the original sample was selected to return and answer questions about victimization and bullying. In addition to the qualitative interview, they also administered the Social Experience Questionnaire (SEQ). The qualitative interview included open-ended questions about bullying and victimization. Results corroborated prior findings that suggested that girls experienced more relational and verbal aggression than physical aggression. The authors did not provide any information about the psychometric properties of these measures.

A structured interview allows the researcher to obtain a better understanding of the context and interpretation of experiences (Crothers & Levinson, 2004). Due to the retrospective and subjective nature of the interview, it is important to not rely exclusively on interviews and to also administer quantitative measures. One threat to the internal validity of studies that use only self-report measures such as an interview is that the participants may respond in socially desirable ways (Crothers & Levinson, 2004).

Sociometric Procedures

The sociometric procedures or the nomination method identifies children involved in the bullying process, based upon information from multiple informants (Ortega et al., 2001). This method can be as simple as having children identify which children in their class best fit a

behavioral descriptor or it can be more complicated, such as having questionnaires within a self-perception scale (Crothers and Levinson, 2004). This method is a time-efficient way to gather information and approximately 60% of the class needs to participate to ensure accuracy (Hoza et al., 2005). However, there are problems with this method such as limited information that does not include feelings, causes, or contexts (Ortega et al., 2001).

Dodge, Coie, Petit, and Price (1990) assessed social status among boys in 25 first grade and 28 third grade classrooms in 11 schools. Students were asked to identify the three students they liked best and the three they liked least in their classroom. Researchers used the procedures from Coie, Dodge, and Coppotelli (1982) to determine sociometric status. A social preference score was calculated based on the standardized difference of the liking and disliking ratings. *Popular* students were those who were nominated more in the 'like best' category than in the 'like least' category. *Rejected* children were those who were nominated more in the 'like least' category than in the 'like most' category. A social impact score was calculated as the standardized sum of 'like most' and 'like least' categories. *Neglected* children fell somewhere in the middle of "like most" and "like least" and are ignored. *Average* children were those who some students reported liking the most and some reported liking the least.

After assessing the preference ratings, the researchers created 24 playgroups, each containing one popular, one neglected, two average and two rejected boys (Dodge et al., 1990). The researchers coded the boys' behaviors in the playgroups along 16 categories (e.g., solitary focused behavior, cooperative play, rough play, bullying). Researchers were trained on the coding system and the inter-rater reliability ranged from .41 to .91. After only a few sessions, social preference within the playgroups was correlated with social preference in the classroom. Rejected boys showed a greater frequency of angry reactive aggression (i.e., they became visibly upset based on the actions of another student and would respond with equal or greater intensity) and instrumental aggression (i.e., they used aggressive means to get to an end goal), than Average

boys (Dodge et al., 1990). Popular first grade boys displayed more bullying than Average first grade boys. However, this difference was not found for third grade boys (Dodge et al., 1990).

Pakaslahti and Keltikangas-Jarvinen (2000) assessed aggression in 2,002 students ages 11, 14, and 17. The researchers used peer nomination to assess peers' views of bullies and victims, a teacher questionnaire to rate aggression, and a self-report questionnaire of aggressive behavior. When all ages were combined, significant correlations were found between peer and teacher ratings of direct aggression ($r = .40$), peer and self-report ratings of direct aggression ($r = .15$), teacher and self-report ratings of direct aggression ($r = .20$), peer and teacher ratings of indirect aggression ($r = .17$), peer and self-report ratings of indirect aggression ($r = .10$), and teacher and self-report ratings of indirect aggression ($r = .15$).

For the 11-year-olds correlations between peer and teacher ratings of direct and indirect aggression were higher than correlations between peer and self-report ratings or peer and teacher ratings on these two measures. For the 14-year-olds correlations between peer and teacher ratings of direct aggression were stronger than correlations between self-report ratings and peer or teacher ratings. However, correlations between teacher ratings and self-report ratings correlated more highly than correlations between peer ratings and self-report or teacher ratings. For the 17-year-olds there were significant differences related to direct aggression, for correlations between peer ratings and teacher ratings. Teacher and self-report ratings correlated more strongly than peer and self report ratings.

The authors concluded that peer and teacher ratings were probably more similar because they both involve outside observers rating an individual. The self ratings were probably less similar because the individual understands personal intentions and reasons for behavior, or because they attempt to answer in a socially desirable manner (Pakaslahti & Keltikangas-Jarvinen, 2000). These differences among raters may be due to amount of interaction with the individual, while teachers only observe students in the classroom, peers observe each other in multiple domains, such as the playground, lunchroom, and classroom, and the individual in

question has information about his or her behavior in all situations. This study showed good correspondence between peer nominations and teacher ratings because peer and teacher ratings resulted in similar categorization of students as bullies and/or victims (Pakaslahti & Keltikangas-Jarvinen, 2000).

Peer Nomination Inventory (PNI; Wiggins & Winder, 1961). The Peer Nomination Inventory (PNI) is a measure that asks children to identify which classmates match a particular behavior. The PNI has 26 items, including seven that measure victimization and seven that measure aggression (Wiggins & Winder, 1961). This measure has been used in various studies (Perry & Bussey, 1977; Perry, Perry, & Rasmussen, 1986), has high reliability for the victimization subscale and has demonstrated validity (Perry, Kusel, & Perry, 1988). Next, details of the psychometric properties of this instrument will be reported.

Researchers used self-report, teacher report, and peer nomination to assess victimization for students (n = 165) in third through sixth grades (Perry, Kusel, & Perry, 1988). The PNI was used to assess aggression and victimization. To assess peer acceptance and rejection researchers gave the children names of same sex classmates and asked them who they "would like to play or work with the most (or least)." The teachers were given a modified PNI and matched the behavior descriptor with the child who best fit. A victimization score was given to each child by summing the number of victim items the teacher endorsed for each child. The children also provided self-report of victimization by responding to a questionnaire asking them to describe how often each of the events happened to them. There was much variability in the way students viewed their peers as either victims or aggressors, with children's scores on the PNI victimization ranging from 0 to 612. However, about 10% of the sample fell far outside of the median (n = 37). These children were categorized as severely victimized.

The high internal consistency of the victimization scale was evident by the alpha value of .96 (Perry, Kusel, & Perry, 1988). There was a correlation between students' individual victimization items when they were compared to their sum of scores for all victimization items.

When the scores on the physical victimization items were averaged and compared to the average of the verbal victimization items the resulting correlation was high ($r = .76$). The test-retest reliability at 3-month follow-up was also high ($r = .93$). There were also similarities between the peer group report and the self-report in terms of identifying victims (grade by sex alphas ranged from .78 to .98). The similarity between the teacher ratings, r (163) = .62, $p < .001$, and self ratings, r (163) = .42, $p < .001$ with the PNI also helped to establish the validity of the instrument.

Children's Social Behavior Scale-Peer Report (CSBS-P; Crick, 1996) This is a group administered peer nomination instrument that measures physical, relational, and prosocial behavior. The physical aggression and relational aggression subscale each consist of 5 items, and the prosocial behavior subscale consists of 4 items. Test-retest reliability for the relational aggression subscale has been shown to be .82 and for the physical aggression subscale to be .90. Research has shown Cronbach's alphas for the relational aggression subscale to range from .82 to .89 and .94 to .97 for the physical aggression subscales (Crick, 1996; Crick, 1997, Crick, Ostrov, & Werner, 2006).

Crick (1997) administered the CSBS-P and a teacher measure of psychosocial adjustment to 6th grade children (N = 1,166). A factor analysis was conducted for each item and as the authors predicted this provided evidence for three separate factors. Overt aggression accounted for 50.6 of the variation, relational aggression accounted for 17.6% of the variation, and prosocial behavior accounted for 10.6% of the variation. Factor loadings for overt aggression ranged from .79 to .90, factor loadings for relational aggression ranged from .70 to .83, and factor loadings for prosocial behavior ranged from .80 to .90. Overtly aggressive children were found to exhibit more symptoms of maladjustment than nonaggressive peers (Crick, 1997).

Questionnaires

Bullying questionnaires include items designed to measure multiple aspects of the bullying experience. Questionnaires may target aspects of bullying such as frequency,

perceptions, locations and knowledge of what constitutes bullying (Ortega et al., 2001).

Questionnaires may also include items to assess attitudes towards bullying, bullies, and victims.

Questionnaires may be designed for students, parents, teachers, and/or school officials.

Questionnaires offer some benefits over other forms of measurement. They are a low-cost way to evaluate a large sample and obtain first-hand information based on self-report (Ortega et al., 2001). Questionnaires can provide information about bullying, social relationships, intervention strategies, and demographics. This can be done in a way that is simple enough for non-experts, such as classroom teachers who simply want to assess the level of peer victimization in their classroom (Ortega et al., 2001). If the questionnaires are completed anonymously, the validity of the responses will likely improve. However, researchers will be unable to examine within subject differences or changes over time if questionnaires are completed anonymously (Ortega et al., 2001).

Despite the benefits offered by questionnaires, there are some potential drawbacks. Questionnaires may be too superficial to assess the problem or may not include all relevant topics. For example, in an interview the answer to one question may lead to another question, whereas questionnaires cannot easily be structured this way (Ortega et al., 2001). Another limitation of questionnaires is that it may be difficult to maintain consistency across responses, the respondent may answer differently on different days or throughout the questionnaire. Also, the questions asked may not take into account the socio-cultural context, so it is important to validate the questionnaires across ethnic groups and cultures (Ortega et al., 2001).

Another issue with questionnaires is the issue of what should constitute "good" psychometric properties. In the area of reliability Nunnally (as cited in Mash & Hundsley, 2005) recommended that for questionnaires past the stages of piloting, the reliability should be at least .90. However, there has been much criticism that this criteria is too strict. After reviewing clinical, personality, neuropsychological, educational, and intelligence measures, Charter (as cited in Mash & Hundsley, 2005) determined the median of most measures to be .85; this would mean

if abiding by the .90 criteria, 75% of all studies would be rejected. However, Caplan, Naidu, and Tripati (as cited in Pedhauzer & Schmelkin, 1991) stated that an alpha of .50 and higher should be judged as adequate for research purposes. For the purpose of the upcoming review of measures, reliability of .50 to .84 will be considered adequate and .85 and above will be considered good.

When assessing convergent validity Campbell and Fiske (1959) state that the best way to assess convergent validity is to use different methods of assessment. For example, two self-report measures of victimization will probably result in strong convergent validity, since the questions are assessing the same construct. In order to accurately determine convergent validity, the methods should be different; this can be done by using self-report in combination with either report by others, self-monitoring, or direct observation. They also state that it is important to establish discriminant validity, tests can be invalidated if they are correlated too highly with tests from which they should differ. In order to establish discriminant validity more than at least two traits and two methods should be used. On survey studies, different raters, such as a child and a peer can represent different methods. Convergent and discriminant validity are related to construct validity, before it can be said that something demonstrates construct validity, it must be known that the employed measures accurately assess the construct of interest.

The Revised Olweus Bully/Victim Questionnaire (OBVQ) (Olweus, 1993). The OBVQ is one of the most frequently used measures to assess bullying. The OBVQ begins with a definition of bullying and examples of different behaviors that fit this definition. This questionnaire is designed to measure how often bullying occurs, what form it takes, where it occurs, who the perpetrators are, and students' perceptions of teachers' responses to bullying. A recent study by Solberg and Olweus (2003) examined the OBVQ with 5,171 students in grades five through nine. The purpose of the study was to assess the functionality of bully and victim variables, as well as to determine appropriate cutoff scores for categorizing children as bullies, victims, and bully/victims. To assess the functionality of the bully and victim constructs, the researchers also

gave measures of social disintegration, global negative self-evaluations, depressive tendencies, general aggression, and antisocial behavior. Victims reported substantially less social acceptance, more negative self-evaluations, and more depressive tendencies than non-victims. Students who reported being bullied more frequently had worse psychosocial adjustment than students who reported being bullied less frequently. Students who reported bullying others reported more aggressive and antisocial behaviors than students who did not report bullying. The authors used the Spearman-Brown formula to give a reliability coefficient, the reliability estimate approached .90, .88 for victims, and .87 for bullies. The prevalence estimates from the global measures and bullying and victimization variables were strongly correlated ($r = .80$).

In order to determine the cutoff scores for categorizing a student as a bully or a victim frequency was used. If a student bullied or was victimized "2 or 3 times a month" he/she was categorized as a bully or a victim, respectively (Solberg & Olweus, 2003). Bullies reported more aggression and antisocial tendencies than non-bullies. The authors chose this frequency to classify as either a bully or victim because the students at this frequency showed marked differenced in their psychosocial adjustment (Solberg & Olweus, 2003).

Early validity studies (Olweus, 2004) showed that bullying or victimization items correlated significantly ($r = .40 - .60$) with other measures (Olweus, 2004). Construct validity for bullying and victimization was also shown. Specifically, victimization items were significantly correlated with self-report of depression, low self-esteem, and peer rejection. In addition, bullying items were significantly correlated with antisocial behavior (Olweus, 2004). The internal consistency reliability was adequate to good (alpha = .80).

Bullying-Behaviour Scale and Peer Victimization Scale (Austin & Joseph, 1996; Neary & Joseph, 1994). Austin and Joseph (1996) created the Bullying Behaviour Scale (BBS). In order to test the psychometric properties of this scale and their earlier scale, the Peer Victimization Scale (PVS; Neary & Joseph, 1994) they mixed items from the BBS and PVS with items from Harter's (1985) Self-Perception Profile for Children. The items were mixed with Harter's items in order to

disguise them from being recognized as a bullying/victimization assessment. Both the BBS and the PVS included six forced-choice items. On the BBS three items referred to being the perpetrator of physical bullying and three items referred to verbal bullying. On the PVS, three items referred to being the victim of physical bullying and three items referred to verbal bullying. In the first study by Neary and Joseph (1994) the PVS was administered to elementary school girls (n = 60). Higher scores on the PVS were associated with less self-worth and more depression. In the second study, Austin and Joseph (1996) found good internal consistency reliability for the PVS (Cronbach's alpha = 0.83) and the BBS (Cronbach's alpha = 0.82). Although reliability has been established and can be considered adequate, there has not been a study investigating the validity. This measure also only offers a limited number of items, which may not get at all aspects of the bullying problem. Used together the PVS and the BBS can give researchers both a bullying and a victimization subscale. Another limitation of this measure is that it these studies were completed in Ireland, which may limit the generalizability of the measures.

The Colorado School Climate Survey (CSCS; Garrity, Jens, Porter, Sager, & Short-Camilli, 2004). The Colorado School Climate Survey (CSCS) is a student self-report questionnaire designed to measure bullying behaviors, perceptions of school safety, and overall school climate. Frequency of victimization is assessed by asking about specific behaviors that qualify as physical bullying, indirect bullying, or verbal bullying. These behaviors are assessed using quantity specific responses (i.e., 0 = never, 1 = less than once a week, 2 = once a week, 3 = 2 to 4 times a week, and 4 = 5 or more times a week over the past month). Next, students are asked how they responded to the incident (i.e., "I got help from an adult," "I did nothing," "I hit, kicked, or pushed the kid.") Students are then asked, "Who was it done by?" "Where did it happen?" and "Who did you tell?" The next section of the questionnaire follows the same format but involves questions about observations of other students being bullied. School safety is assessed by asking students to indicate how safe they feel in different locations (e.g., classroom,

playground, hallways) using quantity specific responses (i.e., 0 = very unsafe/scared, 1 = unsafe/scared, 2 = kind of unsafe, 3 = kind of safe, 4 = safe, 5 = very safe).

For elementary school students (n = 8677) internal consistency reliability for victimization (alpha = .79), observations of bullying (alpha = .83), feelings of school safety (alpha = .80), and school climate (alpha = .79) were all adequate (Plog, Epstein, Jens, & Porter, 2006). For secondary school students (n = 6586), internal consistency reliability for victimization (alpha = .81), observations of bullying (alpha = .90), school safety (alpha = .90), and school climate (alpha = .81) were good to excellent. For both the parent and teacher forms the internal consistencies ranged from adequate to excellent (alphas = .75 to .93). Convergent validity was established through significant correlations of the comparisons of the experienced bullying and observed bullying questions (r = .58, p < .001, n = 7,235 and r = .51, p < .001, n = 4,910 for elementary and secondary schools). Although these authors did attempt to establish convergent validity, they attempted to do this by comparing observing bullying to experiencing bullying. Depending on whether or not the student has been victimized, this could or could not be asking essentially the same question. Since you can observe bullying without being a part of it and vice versa, this measure should have validity established by comparing it to another instrument. This was also taken from an unpublished manuscript and has not been replicated in multiple laboratories. This scale also only produced a victimization scale but no bullying scale.

Name Calling Survey (NCS; Embry, 1995). The NCS was developed as a questionnaire for children in grades one to six to measure the extent of name-calling (Embry & Luzzo, 1996). It includes 35 broad categories aimed at determining the specific nature of the name-calling (e.g., "A name that made fun of your weight"). Students respond by circling "yes" or "no." Split-half reliability for this questionnaire was good (r = .88; Embry & Luzzo, 1996). Test-retest reliability was also good (r = .87) for third, fourth, and fifth graders (n = 80) who completed the measure twice a week for two weeks (Embry & Luzzo, 1996).

Dennis and Satcher (1999) assessed the psychometric properties of this instrument, Cronbach's alpha was used to determine internal consistency ($r = .89$) the content validity of this measure was established by counselor and counselor educators reviewing the instrument (Dennis & Satcher, 1999). Although the psychometric properties of this measure are good, this measure is too specific. Name calling in only a small part of bullying, in order to be a truly adequate measure of bullying the measure must be more comprehensive. The validity information for this instrument was not provided, other than stating that it was adequate. There was also no description of how the counselors reviewed the instruments.

Life in School Booklet (LSB; Arora & Thompson, 1987). The authors conducted a two part study that investigated children's definitions of bullying. The first study began by investigating the occurrence of certain types of events, but without using the word "bullying," so as not to influence the children. The second study used the events reported during the first study to uncover which of these events children considered as bullying. Researchers gave children a list of events and asked them to classify them with regard to bullying by answering "yes," "no," or "sometimes." There was also a space at the bottom of the questionnaire that allowed for the students to add other incidents that they considered to be bullying. After the researchers agreed upon the events they considered bullying, they calculated the percentage of events that reportedly occurred for each student. For example, 65% of 12-year-old boys reported that someone tried to hurt them during the previous week. One major problem with this measure is that there has not been any reliability or validity established.

Peer Relations Questionnaire (PRQ; Rigby & Slee, 1993). The Peer Relations Questionnaire (PRQ) is a 20-item measure which includes six items about bullying others, six items about being victimized, four prosocial items, and four "filler" items (Rigby & Slee, 1993). Students are asked how often the statements were true of them on a four-point scale (not at all to very). Factor analysis showed that the bullying and victimization items loaded on separate factors. For students ages 12 to 18 years old (n = 1,162) from two different schools, the PRQ has

also shown adequate to good internal consistency reliability for the items, the Bully Scale (alpha = .75 and .78) the Victim Scale (alpha = .78 and .86), and the Prosocial Scale (alpha = .71 and .74). The validity of this measure was shown through the correlation of this measure to related constructs. Self-esteem was significantly positively correlated with prosocial tendencies (r = .11, p = .01) and significantly negatively correlated with tendency to be victimized (r = -.22, p = .001; Rigby & Slee, 1993). Measures of happiness were significantly positively correlated with prosocial tendencies (r = .36, p = .001), and significantly negatively correlated with age (r = -.18, p = .001), victimization (r = -.10, p = .01), and bullying (r = -.26, p = .001). Liking for school was significantly positively correlated with prosocial tendencies (r = .28, p = .001) and significantly negatively correlated with age (r = -.23, p = .001) and bullying (r = -.26, p = .001). For boys at one school there was a negative correlation with bullying and prosocial behaviors (r = .29, p < .001; Rigby & Slee, 1993).

Although this measure provided adequate reliability, the measure is not comprehensive enough to assess all aspects of the bullying experience, as indicated by only having 12 items about bullying and victimization. In terms of validity, the authors stated that the measure demonstrated adequate validity, but all the measures were based on a student's self-report. Due to the content of these measures it was already expected for them to show convergent and discriminant validity, in order to get the best estimate of validity, it would be necessary to compare the self-report measures with other forms of measurement (teacher report or direct observation).

Reynolds Bully-Victimization Scale (BVS). The BVS is designed for students in grades 3-12 or ages 7-20. The BVS contains 46 items, 23 on the bullying scale and 23 on the victimization scale (Reynolds, 2003). Responses for each item are based on a four point scale (i.e., 0 = never, 1 = sometimes, 2 = a lot of the time, and 3 = five or more times). Factor analyses were conducted on the BVS items and a two factor solution was found (Reynolds, 1993). Items from the bullying

scale were correlated with items from the victimization scale ($r = .36$). The factor loadings for the bullying scale were .52 to .70 and for the victimization scale were .41 to .71 (Reynolds, 2003).

The bullying items were designed to measure the frequency of being a perpetrator of overt aggression, relational aggression, and/or harassment. The victimization items are designed to measure the frequency of being a victim of overt aggression, relational aggression, and harassment. The coefficients for internal consistency reliability, for both the bullying and the victimization scales, were excellent (alpha = .93). These coefficients were high across gender and grade level. The test-retest reliability was assessed by administering the scale 1-2 weeks after the first administration (n = 207). Analyses resulted in good reliability for the bullying ($r = .81$) and victimization items ($r = .80$) (Reynolds, 2003).

Criterion-related validity was assessed by comparing the student self-report BVS with the teacher-report BVS (Reynolds, 2003). Teachers completed a rating scale on each individual student. There was a moderate correlation found for the total sample ($r = .46, p < .001$) and for students in grades 3-6 ($r = .54, p < .001$); however, the correlation for the students in grades 7-8 was only lower (r = .24, NS). In addition, there was a stronger correlation for older students ($r = .47, p < .001$). The BVS bullying items had moderate correlations with the Beck Youth Inventory (BYI) Disruptive Behavior Scale ($r = .54, p < .001$) and with the BYI Anger Scale ($r = .38, p < .001$). In addition, the BVS bullying items showed a moderate correlation of ($r = .48, p < .001$) with the Reynolds Adolescent Adjustment Screening Inventory (RAASI), Antisocial Behavior Scale and a moderate high correlation ($r = .60, p < .001$) with the RAASI anger control scale. The BVS victimization scale had moderate correlations with anxiety ($r = .58, p < .001$) and depression ($r = .50, p < .001$) and with disruptive behavior ($r = .32, p < .001$; Reynolds, 2003).

Discriminant validity occurs when measures of two different constructs are not highly correlated with one another (Reynolds, 2003). For BVS bullying items there was a low correlation ($r = .08$) with the RAASI Emotional Disturbance scale. The BVS bullying items also showed low correlations with the BYI scales of anxiety and depression (r's = .11 and .12). Thus,

as expected bullying was not related to these measures of internalizing disorders. Also consistent with expectations, the BVS victimization items showed convergent validity with Emotional Disturbance ($r = .34$, $p < .001$), with Anger Control ($r = .44$, $p < .001$), and also showed discriminant validity with a low correlation with the RAASI Antisocial Behavior Scale ($r = .17$). Although this measure has demonstrated good reliability and adequate validity the results have not been replicated across laboratories.

Social Experience Questionnaire (SEQ; Crick & Grotpeter, 1996). The SEQ measures victimization and prosocial behavior, it consists of three subscales: relational victimization, overt victimization, and receipt of prosocial behavior (Crick & Grotpeter, 1996). Each subscale includes five items. Responses are on a five-point scaling ranging from "never" to "all the time." Students in third through sixth grade (n = 474) completed the study. The reliability for all the scales were adequate for relational victimization (*alpha* = .80), overt victimization (*alpha* = .78), and prosocial recipient scale (*alpha* = .77; Crick & Grotpeter, 1996). Factor analysis was used to assess validity information, the factor loadings for the relational victimization questions ranged from .66 to .80, the factor loadings for the overt victimization questions ranged from .69 to .81, and the factor loadings for the prosocial questions ranged from .67 to .75. One benefit of this scale is that it measures relational aggression separately. This is beneficial since girls often display this kind of aggression rather than physical aggression. This measure shows adequate reliability and validity. However, this scale is limited to aggression and does not assess bullying or victimization.

Self-Rating Questionnaire on Aggressive Behavior (SQAB; Lindeman, Harakka, & Keltikangas-Jarvinen, 1997). The Self-Rating Questionnaire on Aggressive Behavior (SQAB) assesses adolescents' reactions to conflicts based on aggression, prosociality, and withdrawal (Lindeman, Harakka, & Keltikangas-Jarvinsen, 1997). On the SQAB two interpersonal conflict situations are presented. The first situation involves overt aggression and the second situation involves indirect aggression. After reading each situation the respondent chooses among

21

responses that represent direct aggression, indirect aggression, a prosocial strategy, or withdrawal.

Internal consistency reliability for the three subscales found moderate alphas. Specifically, internal consistency for the aggression items was adequate (alpha = .81), internal consistency for the prosocial items was adequate (alpha = .70), and internal consistency for the withdrawal items was adequate (alpha = .73; Lindeman et al., 1997). A more recent study has found stronger internal consistency for the aggressive items (alpha = .88) and slightly weaker internal consistency for the prosocial items (alpha = .65; Keltikangas-Jarvinen, 2002). Factor analysis was used to assess the validity of the instrument. The factor loadings for the three variables differed based on the questions. The eigenvalue for aggression was 6.12, for prosociality was 2.25, and for withdrawal was 1.39. The SQAB has demonstrated adequate reliability and validity (Lindeman et al., 1997). Although this measure assesses how the student would react in the hypothetical situation, it does not assess the rates of bullying and victimization that the student has experienced.

Teacher Measures

Teacher ratings are useful when assessing bullying and victimization in the classroom. The use of teacher ratings can aid in establishing external validity to self-report measures and are often correlated with peer report (Pellegrini and Bartini, 2000). There are some issues with teacher ratings, one of which is that teachers may not observe students in all contexts and may underestimate the amount of bullying and victimization that occurs (Crothers and Levinson, 2004). However, if a teacher spends long amounts of time with students and sees them in a variety of contexts, this bias can be minimized (Pellegrini and Bartini, 2000). Teacher report should not be used alone in research on bullying and victimization but is useful when combined with another form of measurement. Another issue with teacher report is that there needs to be a quick and easy method for obtaining a bullying and victimization subscale from a teacher. It is difficult for a teacher to find the time to complete a multiple question survey for each child in his

or her class. In order to increase the likelihood that teachers will complete the measure it is necessary to make it as convenient and take the least amount of time possible.

Teacher ratings of bullying and victimization are one area where the research is particularly lacking. There are psychometrically sound teacher report measures (Edelbrock &Achenbach, 1984), but these measures are not specific to bullying and victimization. Although there are a few instruments designed for this purpose, there does not yet seem to be a gold standard. This leaves some researchers to adapt a measure or create their own. In the Pakaslahtu and Keltikangas-Jarvinen (2000) study the authors asked teachers to rate individual students on behavioral descriptions on a scale of one to five. They did this by asking two items on direct aggression, one item on indirect aggression, and four items about prosocial behaviors. They reported the reliability for direct aggression to be alpha = .72 and correlations between the scales to be at the $p < .001$ level.

*Colorado School Climate Survey-Teacher (Garrity, Jens, Porter, Sager, & Short-Camilli, 2004).*The CSCS offers a teacher survey in conjunction with a self-report and parent-report measure. This survey assesses the teachers view of victimization that occurs at their school and in their classroom, it also assesses how they perceive their school to intervene with and prevent bullying and victimization. This measure does not assess individual children but rather the classroom climate. As previously reviewed, the CSCS was given to parents, students, and school staff (N = 27, 597). Of the staff participants, they represented 45% of elementary school staff, 20% of middle school staff, and 35% of high school staff. After exclusion due to missing items there were 983 elementary school staff included and 1495 secondary school staff included. The items on the CSCS staff version are aimed at student victimization, school safety, and school climate. Internal consistency reliabilities for the elementary school staff ranged from adequate to good (bullying = .88; safety = .88, climate-student = .76, climate-adult = .76). For secondary school staff the alphas ranged from adequate to near excellent (bullying = .93, safety = .91, climate-student = .75, climate-adult = .81). Correlations showed evidence of convergent validity

of the staff bullying, safety, and climate scales. All correlations were significant at the p < .01 level. One problem with this information is that it is from an unpublished manuscript. Another problem is that it does not ask about individual students but rather the classroom or school as a whole. This scale also only assesses victimization but not rates of bullying.

Preschool Peer Victimization Measure-Teacher Report (Crick, Casas, and Hu 1999) This scale was adapted for preschoolers from another teacher report measure. The authors began by asking teachers who were not participating in the study about examples of victimization they observed in their classroom. They compiled these answers and developed a measure with two items to assess relational victimization, two items to assess physical victimization, and three items to assess receiving prosocial treatment. The behaviors were rated on a scale of one (never or almost never) to five (always or almost always). The participants were children (N = 129) ages three to five. In order to assess the psychometric properties of the instrument the authors conducted a factor analysis, physical victimization accounted for 55% of the variation, receiving prosocial treatment accounted for 15% of the variation, and relational victimization accounted for 9% of the variation. Cronbach's alpha was .77 for relational victimization, .88 for physical victimization, and .84 for receiving prosocial behavior. A recent study by Garner and Lemerise (2007) found the alpha for the physical victimization questions to be .87 and the alpha for the relational victimization to be .89. Using this scale, the authors stated that 11.62% of children were relationally victimized and 16.30% were physically victimized (Crick, Casas, and Hu, 1999). This scale has demonstrated adequate to good psychometric properties. However, the psychometrics of this scale are for preschool students which may limit its generalizability to older children. This scale also only assesses receiving victimization and not about perpetrating victimization.

Social Experience Questionnaire-Teacher Report (SEQ-T; Cullerton-Sen., Crick, 2005) The SEQ-T assesses teacher perceptions of children's physical and relational victimization experiences. This measure consists of six items. Internal consistency reliability for the relational victimization subscale was .82 and .93 for the physical victimization subscale. The correlation

between the two subscales suggests was $p < .01$, suggesting evidence for construct validity. The teacher report was compared to peer and self- report on victimization and adjustment problems. The teacher reported physical and relational aggression was similar but slightly less than peer and self report. Teacher report about physical victimization was predictive of externalizing disorders and teacher report of relational victimization was predictive of peer rejection, externalizing and internalizing behaviors. Although this scale has demonstrated adequate to good psychometric properties, it asks multiple questions about each child and does not assess bullying.

Teacher Bully-Victimization Rating Scale (TBVRS; Reynolds 2003.) is an unpublished measure that was used in conjunction with the reliability and validity studies of the BVS. This consists of a 12-item Bullying subscale and a 10-item Victimization subscale. Teachers report on each student's behavior over the past six months. One major limitation of this instrument is that it is unpublished and psychometrics are unknown.

Teachers Checklist (Dodge & Coie, 1987). This is a 45 item measure where children are rated on a scale of one to five. In order to validate this measure the authors had 12 statements in reference to proactive aggression, retroactive aggression, and nonspecific aggression, these statements were embedded within 12 other statements in order to disguise the focus on aggression. This instrument was administered to 3rd through 6th grade teachers ($N = 259$). Teachers completed this survey about every child in their classroom. A factor analysis was conducted and the three items that loaded the most strongly on each factor were used to create two scales. The Proactive/Aggression scale had an alpha of .91 and the Reactive/Aggression scale had an alpha of .90. The two scales are correlated between eachother, but the correlations between the items in each scale are higher than that, supporting the construct validity of the measure. Modified versions of this checklist have been used in other studies (Pellegrini and Bartini, 2000; Pellegrini, Bartini, & Brooks, 1999). Dodge and Coie (1987) reported the factor of aggression to have an alpha of .92. Pellegrini and Bartini (2000) found this 14-item factor to have an alpha = .88. Pellegrini, Bartini, and Brooks (1999) found the Aggression/Reactive scale to

have an alpha of .93, the Aggression/Proactive scale to have an alpha of .94, and the Dominance Scale to have an alpha of .88. This scale asks multiple questions about each child, which may be time consuming for the teacher. However, one benefit of this scale is that it has a multitude of research that has demonstrated adequate to extremely good psychometrics.

<center>*Problems with Current Measurement Tools*</center>

Although some of the measurement tools discussed so far have demonstrated adequate reliability and validity, the majority have not. The OBVQ, SQAB, and BVS have demonstrated good reliability and adequate validity. However, the psychometric properties for the BVS have not been replicated across laboratories. The psychometrics for the OBVQ have only been assessed beginning with fifth grade, despite the instrument being marketed to children as young as third grade. The SQAB asks only about reactions to conflicts, but does not assess the rates of bullying and victimization. The majority of the other instruments (BBS, PVS, SEQ) have demonstrated adequate reliability but not validity. The remaining instruments (i.e., NCS, PRQ, and CSCS) that have demonstrated both reliability and validity are not optimal for other reasons, such as not being specific and comprehensive in regards to bullying. While the PRQ does have validity information, validity was established through the use of all self-report information. The LSB did not provide any information on psychometric properties. The most valid methods of measurement (video and audio taping observations) are time consuming, expensive, and create ethical problems in terms of consent. Other means of assessing bullying and victimization are needed. Due to their low cost and easy administration, questionnaires seem to be the best method of assessing the behavior in an economical and convenient manner (Crothers & Levinson, 2004).

Observational assessment has weaknesses, including that the measures do not necessarily correlate over time (Crothers & Levinson, 2004). This may be because the behavior is observed in only a few situations and for short periods of time. Another problem with observations is that if sensitive instruments are not used, the researcher may not be able to capture the covert aspects of victimization. The presence of the researchers may also change the behavior of some of the

<center>26</center>

students who are regularly involved in bullying. Unless the researcher has the time and money to ensure that there is a strict level of control for the experiment, then they should consider using a different method. However, even if they do have the time and money to conduct these observations, they may want to still want to consider using this method with some form of self-report, so they can obtain detailed information from the student.

Interviews or other qualitative measures may yield different problems for the researcher. The validity of the interview may be compromised due to various factors, such as social desirability causing the participant to self-select information (Crothers & Levinson, 2004). Different interviewers may also elicit different information. Interviews and observations can also be biased due to the preconceptions or stereotypes of the interviewers and observers (Crothers & Levinson, 2004). The perceptions of teachers may be skewed and this could elicit a problem with interviews and questionnaires (Crothers & Levinson, 2004). Teachers may not be present when the majority of bullying is occurring and in the classroom they may only notice direct bullying and may be unaware of more indirect forms of bullying. They may also view some aggressive behaviors as play instead of bullying.

Self-reports of aggression are usually underestimates because the perpetrator of bullying may not report all of his or her bullying behaviors due to issues of social desirability (Pellegrini & Bartini, 2000). A study by Pellegrini and Bartini (2000) compared methods of identifying aggressive/bullying and victimized students in a middle school. Peer nominations and self-report measures required the students to identify three aggressive students and three victimized students (Schwartz, Dodge, & Coie, 1993); the other peer nomination method was by Perry et al., (1988). The self report measure was the OBVQ. They began by comparing researchers and teachers responses to the Teacher Checklist (Dodge & Coie, 1987). The Teacher Checklist was validated in the original study and the authors reported excellent internal consistency (alpha = .92). They found a moderate correlation (r = .50) between researchers and teachers reports. Researcher direct observations of student aggression and victimization were compared with diary measures

kept by students. For the victimization items diary measures were correlated to both the peer nomination and the peer report (r's = .21 and .26) and self-report measures (r = .34). The direct observations were also correlated with the peer nomination and the peer report measures for the victimization items (r's = .22 and .34). For the bullying items direct observations were correlated with both the peer nomination and the peer report (.41 and .52), self-report (.35), and teacher report (.50). The self report of bullying was also correlated with peer nomination and peer report (.18 and .47) and direct observations (.20; Pellegrini & Bartini, 2000).

Achenbach, McConaughy, and Howell (1987) assessed ratings of children's emotional and behavioral problems by conducting a meta-analysis of studies using the Pearson r to compare ratings among parents, teachers, mental health workers, observers, peers, and student self-reports. People in the same/similar roles, such as a mother and a father, had significant correlations (r approximately equal to .60). The correlation was small across different informants, such as the correlation between mental health worker and parent (r = .24) (Achenbach, McConaughy, & Howell, 1987). There were also no significant correlations between self-reports and those of other informants. Due to the lack of significant correlations across informants, this study reinforced the idea of assessments being conducted across multiple situations, domains, and informants (Achenbach, McConaughy, & Howell, 1987).

Hartung, McCarthy, Milich, and Martin (2005) examined adolescent and parent ratings of inattention, hyperactivity, and conduct problems through the use of self-report, parent report, and interviews. Parents seemed to be more accurate reporters of behaviors consistent with inattention (z = 5.50 and 5.02) and hyperactivity (z = 4.83 and 4.28) than the adolescents. However, adolescents reported engaging in more antisocial behaviors than the parents reported (z = 4.80 and 3.55). This indicates that while parents may notice behaviors consistent with inattention and hyperactivity, the adolescent may not notice him or her self engaging in these behaviors. Therefore, adolescent self-report may not be the best indicator of these behaviors. However, the adolescent may be a better reporter for antisocial behaviors because the behaviors may occur at a

low frequency and when adults are not present, therefore the parents may have no knowledge of these behaviors.

In short, there is a need for better tools to assess bullying. If there are going to be expensive, time-consuming programs implemented in the school systems, then measures that have been validated and cross-validated are needed to reliably determine if these programs are working. As of today, the majority of the available tools offer mostly reliability information but little or no information on validity. While some of the studies assessing these tools have shown good reliability and/or validity, the sample sizes were small or the techniques were prohibitively expensive, time-consuming, or invasive. In order to assess the validity of instruments, research should be conducted that focuses exclusively on the measurement instruments. Most of the studies use these measures as outcome measures with other goals for the study including the effectiveness of interventions.

The Current Study

The current study has been designed to examine the psychometric properties of different bullying/victimization questionnaires with elementary school children. Students will complete two measures: the Revised Olweus Bully/Victim Questionnaire and the Reynolds Bully Victimization Questionnaire. Teachers will complete a measure designed for this study that will demonstrate which students they view as bullies and/or victims. Although there are other teacher measures specific to bullying and victimization, the measures are time consuming and ask multiple questions about each student. We chose to design a teacher checklist that would obtain the information needed but still be as convenient as possible for the teachers. Internal consistency will be calculated to assess reliability of the measures. Student reports will be correlated with one another and with teacher reports of bullying and victimization. These reports will be correlated to assess convergent and discriminant validity. Since the BVS may not have some of the previously mentioned limitations of the OBVQ, the current study is designed to assess the utility and psychometric properties of the measures by comparing the reliability and validity of the two

29

instruments with each other and with the teacher report. We hypothesize that 3rd and 4th grade students will have better validity and reliability for the BVS scale compared to the OBVQ. We also hypothesize that 5th grade students will not exhibit a marked difference in terms of the psychometric properties of the scales.

CHAPTER III

METHODOLOGY

Participants

According to a power analyses that was conducted with *g power* (Erdfelder, Faul, & Buchner, 1996) our power was 1.00 with the total sample of 627 students, alpha at .05, and a medium effect size of .30 is desired (Cohen, 1988).

Participants were 627 elementary school children in third (n = 220), fourth (n = 183), and fifth (224) grade at six public elementary schools in a small Southwestern college town. Students were recruited by either their teacher asking their parents to sign an informed consent at a parent teacher conference or through returning an informed consent that was sent home in a weekly folder. The students whose parents signed an informed consent were given the opportunity to sign an assent form. The students who did not participate were either due to their parents not consenting, their parents not returning the consent, the student not assenting, or their teacher not distributing the consent form. The total sample of participants represented 60% of the students, with 62% of 3rd graders, 56% of 4th graders , and 62% of 5th graders participating. Of the participants 53.7% were boys and 46.3% were girls. The ethnic diversity of the sample was 76.6% Caucasian, 3.9% African American, 3.7% American Indian, 3.2% Asian American, 2.7% Hispanic/Latino, and 9.8% other. Most students in the "other" category are those who's parent checked two or more categories. There were 42 teachers who participated out of a possible 55. Some of the teachers did not participate because too few students in their class participated while others did not participate because they failed to return their rating scale.

Measures

Two questionnaires containing the bullying and victimization measures designed to assess the frequency and severity of bullying, locations in which bullying and victimization occur, attitudes toward bullying and victimization, and responses to bullying and victimization were used. The measures that were used in the current study are described below.

Teacher's Checklist. A teacher's checklist was created for this study (see Appendix A). The checklist consisted of a roster for the teachers to fill in the name of each student in their class who's parents consented to the study. The directions included a definition of bullying based on Olweus (1993). The teacher was asked to rate each child on a Likert scale based on the frequency with which the teacher believed the student was victimized and the frequency with which the teacher believed the student bullied other children. The Likert scale had five categories, which vary from "never" to "very often." This scale was created as a very brief measure of the frequency of bullying and victimization for each student.

The Revised Olweus Bully/Victim Questionnaire. This English version of this 39-item questionnaire was used in the current study (Olweus, 2004). This student self-report questionnaire included items about aspects of bully/victim problems such as physical, verbal, indirect, racial, or sexual forms of bullying, location of bullying, attitudes about bullying, and the extent to which others are informed about and react to bullying. The question about sexual forms of bullying was removed for this study because it was not approved by the school board.

This questionnaire has been used to assess bullying in schools (Austin & Joseph, 1996) and has been shown to correlate with peer nomination (Ross, 1996). The psychometric properties, as reviewed previously, suggest adequate reliability and validity. Pellegrini, Bartini, and Brooks (1999) used the OBVQ in combination with a peer nomination measure, and teacher ratings to

assess bullying and victimization. Cronbach's alpha for bullying was .76, exposure to direct and

indirect bullying was .78, and negative attitudes towards bullying was .66. Early validity studies

demonstrated construct validity and were significantly correlated with related measures (Olweus,

2004). Despite the good psychometric properties of the OBVQ, there are some limitations. The

questionnaire is 13 pages long and includes approximately the same number of questions as the

BVS questionnaire which is only 2 pages long. The formatting of the OBVQ may be confusing

for younger children or children with attention disorders because the answer choices change

throughout the questionnaire. The OBVQ begins with a general definition of bullying and

includes questions in relation to this definition, this may not be as beneficial as asking about the

frequency of specific events. The cultural validity of this measure is also a limitation; although

this measure has been used throughout Europe, there has been limited use in the United States. A

final limitation of this measure is that while this questionnaire is supposed to be valid for 3rd

grade students and older, there is limited information on the psychometric properties of this

measures use with 3rd and 4th graders.

   Reynolds Bully-Victimization Scale (BVS). The BVS was also used in the current study

(Reynolds, 2003). This scale was designed for students in grades 3-12 or ages 7-20. The BVS

contains 46 items, 23 on the bullying scale and 23 on the victimization scale (Reynolds, 2003).

Responses for each item are based on a four point scale (i.e., 0 = never, 1 = sometimes, 2 = a lot

of the time, and 3 = five or more times). The BVS bullying items were designed to measure the

frequency with which the student is a perpetrator of overt aggression, relational aggression,

and/or harassment. The BVS victimization items were designed to measure the degree to which

the student is a victim of overt aggression, relational aggression, and harassment. The

psychometric properties, as reviewed previously, suggest adequate reliability and validity. The

internal consistency reliability for both subscales was .93 and test re-test reliability was .81 for bullying subscale and .80 for victimization subscale. Criterion related validity ranged from moderate to high. This scale also demonstrated convergent and discriminant validity with measures of psychopathology. These properties have not been cross-validated in another laboratory.

Procedure

The university IRB and the school board approved all the procedures for the study. Active consent was obtained from parents. The consent was originally going to be sent home in a weekly folder with the students, but after discussions with school administrators, they stated that they believed they would have more students participate if the informed consent was passed out during the parent-teacher conferences. Teachers were instructed to pass out the consents during the conferences, however, some teachers left the consents out in the hallway for the parents to find for themselves, other teachers did not pass them out at all, some went home in the weekly folder, and some were signed during the conference.

Questionnaire packets contained child assent forms and the bullying/victimization questionnaires. Students completed the questionnaires during two regularly scheduled guidance class meetings approximately one week apart. The questionnaires were administered in counterbalanced order across grade and school. Students were identified with a number and did not write their names on either questionnaire. The assent forms and the questionnaires were read aloud by a research assistant. While being read aloud a graduate student or psychologist and the school guidance counselor walked around the room to answer questions and to help if any students became emotionally distressed due to the content of the questionnaires. The BVS

34

questionnaire took approximately 20 minutes to complete and the OBVQ took approximately 40 minutes to complete.

The teacher consent forms and rating forms were either given to the teachers by the researchers or the guidance counselor. The teacher rating form included only the students with parental consent and assent for participation in the study. The teachers were asked to return the informed consent and the questionnaire to the guidance counselor within a week.

CHAPTER IV

FINDINGS

Results

Bullying and Victimization Rates

For the BVS survey, each subscale included 23 items, rated from 0 to 3. Therefore, the possible range of scores was 0 to 69. Rates of bullying and victimization resulting from student self-report on the BVS are shown in Table 1. Rates of bullying on the BVS Bullying Subscale were at the 56th percentile for the total sample. Third graders reported rates at the 67th percentile, 4th graders reported rates at the 69th percentile, and 5th graders reported rates at the 65th percentile. A one-way analysis of variance (ANOVA) with grade as the independent variable (3rd, 4th, and 5th grade) did not show a main effect. Rates of victimization on the BVS Victimization Subscale were at the 75th percentile for the total sample. Third graders reported rates at the 66th percentile, 4th graders reported rates at the 53rd percentile, and 5th graders reported rates at the 50th percentile. A one-way ANOVA revealed a statistically significant main effect of grade ($F = 22.06$, $p = .000$). Results are displayed in Table 1.

For the OBVQ survey, each subscale included 9 items out of the original 10. As mentioned in the methods, one item was removed due to it not being approved by the school board. Rates of bullying and victimization resulting from student self-report on the OBVQ are shown in Table 2. For the OBVQ, percentiles are not provided because the normative data is not available in the same format. For the OBVQ bullying subscale a one-way ANOVA did not show a main effect of grade. For the OBVQ Victimization Subscale, a one-way ANOVA revealed a significant main effect of grade ($F = 7.19$, $p = .001$).

The teacher survey included one item about bullying and one about victimization, each rated from 0 to 4. Rates of bullying and victimization resulting from teacher report are shown in Table 3. One-way ANOVAs did not show significant main effects of grade for either bullying or victimization.

Internal Consistency Reliability

Internal consistency reliability is the degree of consistency among the items within a measure (Kazdin, 2003). To establish internal consistency Cronbach's alpha was calculated for the bullying and victimization subscales within each student self-report measure. These values are shown on the diagonal in Table 4. For the total sample, BVS internal consistency was calculated for both the victimization and bullying subscales. Internal consistency reliability for BVS Bullying was .932 and for BVS Victimization was .941. Internal consistency reliability for OBVQ Bullying was .846 and for OBVQ Victimization was .864. Thus, internal consistency was excellent for the BVS subscales and good for the OBVQ subscales when grade was not considered.

Cronbach's alpha was also calculated separately by grade. Results for 3rd, 4th, and 5th graders are shown on the diagonals in Tables 5, 6, and 7, respectively. For OBVQ Bullying, 4th graders had adequate reliability (.785) whereas 3rd and 5th graders had good reliability (.856 and .866, respectively). For BVS Bullying, 4th and 5th graders had excellent reliability (.951 and .944, respectively) whereas 3rd graders had very good reliability (.886). For OBVQ Victimization, 3rd graders had very good reliability (.883) whereas 4th and 5th graders had moderately good reliability (.831 and .853, respectively). For BVS Victimization, reliability was excellent for 3rd, 4th, and 5th graders (.941, .944, and .930, respectively).

Concurrent Validity

Concurrent validity is when two measures are correlated with performance during the same time period (Kazdin, 2003). In order to establish concurrent validity student-report was correlated with teacher report of bullying and victimization (see Table 8). For the total sample, Teacher Bullying was significantly correlated with OBVQ Bullying (p = .000) and BVS Bullying (p = .000). In addition, Teacher Victimization was significantly correlated with OBVQ Victimization (p = .000) and BVS Victimization (p = .000). Pairwise comparisons between correlations were conducted, using Steiger's formula (1980), to determine whether the BVS subscales were significantly more strongly correlated with Teacher ratings. A Bonferroni correction was used since multiple statistical comparisons were conducted. For the total sample and each grade, two comparisons were conducted. Therefore, .05 was divided by 2 and the alpha value was set at p < .025. One comparison for BVS Bullying with Teacher Bullying was significant, all other comparisons were nonsignificant, however the one comparison for BVS Victimization with Teacher Victimization and one comparison for OBVQ Bullying with Teacher Bullying did approach significance. For the total sample, the pairwise comparisons showed that BVS Bullying was more strongly correlated with Teacher Bullying than was OBVQ Bullying (p < .0005 compared to p < .025 respectively).

For 3rd graders, Teacher Bullying was significantly correlated with BVS Bullying (p = .002) but only showed a trend towards significance with OBVQ Bullying (p = .052). Similarly, Teacher Victimization was significantly correlated with BVS Victimization (p = .009) but not OBVQ Victimization (p = .095). Pairwise comparisons between correlations were conducted to determine whether the BVS subscales were significantly more strongly correlated with Teacher ratings. All comparisons were non-significant.

For 4th graders, Teacher Bullying was significantly correlated with OBVQ Bullying ($p =$ .000) and BVS Bullying ($p =$ .000). However, Teacher Victimization was significantly correlated with BVS Victimization ($p =$ .000) but only showed a trend with OBVQ Victimization (p = .055) for 4th graders. Pairwise comparisons between correlations were conducted to determine whether the BVS subscales were significantly more strongly correlated with Teacher ratings. Both comparisons between the BVS Bullying and Teacher Bullying were significant, however all other comparisons were non-significant. For 4th graders, the pairwise comparisons showed that BVS Bullying was more strongly correlated with Teacher Bullying than was OBVQ Bullying.

For 5th graders, Teacher Bullying was significantly correlated with OBVQ Bullying ($p =$ .004) and BVS Bullying ($p =$ .000). In addition, Teacher Victimization was significantly correlated with OBVQ Victimization ($p =$ .000) and BVS Victimization ($p =$ .000). Pairwise comparisons between correlations were conducted to determine whether the BVS subscales were significantly more strongly correlated with Teacher ratings. For both BVS Bullying with Teacher Bullying and OBVQ Bullying with Teacher Bullying one comparison was significant at $p <$ .025. For OBVQ Victimization with Teacher Victimization one comparison was significant at $p <$ .0005. For BVS Victimization with Teacher Victimization, both comparisons were significant one at $p <$ .0005 and the other at $p <$ .025. For 5th graders, the pairwise comparisons showed that BVS Bullying and OBVQ Bullying were equally correlated with Teacher Bullying. BVS Victimization was more strongly correlated with Teacher Victimization than was OBVQ Victimization.

Convergent and Discriminant Validity

For the current study, multitrait-multimethod (MTMM) matrixes were created to allow examination of convergent and discriminant validity (Campbell & Fiske, 1959). Convergent validity is obtained when different measures of the same construct are significantly correlated

39

(Kazdin, 2003). It was expected that the two within-trait, cross-method correlations would reach statistical significance. For the total sample (see Table 4), OBVQ and BVS Bullying were significantly correlated (p = .000) as were OBVQ and BVS Victimization (p = .000). For 3rd graders (see Table 5), OBVQ and BVS Bullying were significantly correlated (p = .000) as were OBVQ and BVS Victimization (p = .000). For 4th graders (see Table 6), OBVQ and BVS Bullying were significantly correlated (p = .001) as were OBVQ and BVS Victimization (p = .000). For 5th graders (see Table 7), OBVQ and BVS Bullying were significantly correlated (p = .000) as were OBVQ and BVS Victimization (p = .000). Since convergent validity was excellent for all grades, pairwise comparisons were not conducted.

Discriminant validity is obtained when different measures of the same construct correlate more strongly than similar or different measures of different constructs (Kazdin, 2003). Correlations between different measures of the same construct should be larger than correlations between similar or different measures of different constructs. In terms of discriminant validity, the two cross-trait cross-method correlations between bullying and victimization subscales were expected to be significantly lower than the two within-trait cross-method correlations. Pair-wise comparisons between correlations were conducted, using Steiger's formula (1980), to determine whether correlations were significantly different from one another. A Bonferroni correction was used since multiple statistical comparisons were conducted. For the total sample and each grade, eight comparisons were conducted. Therefore, .05 was divided by 8 and the alpha value was set at p < .006.

For the total sample (see Table 9), each of the two within-trait, cross method correlations was compared to the two cross-trait, within-method correlations and the two cross-method, cross-trait correlations. All eight of these comparisons were statistically significant at p < .006

suggesting good discriminant validity for the total sample. For 3rd graders (see Table 10), only three of the eight comparisons were statistically significant at $p < .006$ suggesting less than adequate discriminant validity. For 4th graders (see Table 11), all eight comparisons were statistically significant at $p < .006$ suggesting good discriminant validity for 4th graders. For 5th graders (see Table 12), all eight comparisons were statistically significant at $p < .006$ suggesting good discriminant validity for 5th graders.

CHAPTER V

CONCLUSION

Bullying and Victimization Rates

For both the OBVQ and BVS Victimization Subscales, self-reports of victimization decreased as grade increased. This finding suggests differing views about the frequency of victimization with age. Prior research has found that bullying increases throughout the elementary school years, peaks in middle school, and then declines in high school (Nansel et al., 2001; Olweus, 1993; Swearer & Carey, 2003). However, there is also research indicating that there may be more victimization occurring in lower grades but the intensity of victimization increases with age (Dennis & Satcher, 1999; Embry & Luzzo, 1996). Dennis and Satcher (1999) found that 3rd graders participated in more name calling than 5th graders. Embry and Luzzo (1996) found that 2nd graders reported more name calling than 6th graders. Thus, certain forms of bullying, such as name calling may decrease with age whereas other forms of bullying such as sexual harassment may increase with age.

On both the BVS and OBVQ students reported more victimization than bullying at all grade levels. Stated differently, students were likely to report having been a victim of bullying but much less likely to report having been a perpetrator of bullying. This could mean that a few students are doing all of the bullying or it could mean that students are willing to admit being victimized but are not being forthcoming about their own bullying behaviors. This finding is also consistent with prior research (e.g., Pellegrini et al., 1999; Solberg & Olweus, 2003). Another explanation is that the students who bully may attempt to answer questions in a socially desirable

manner. In order to control for this tendency, the confidentiality of the questionnaire was strongly emphasized in the instructions as was the anonymity of the responses.

One advantage of the BVS is that normative data for grades 3-12 is provided. According to the normative data, the bullying scores from the current study ranged from the 65th percentile for 5th graders to the 69th percentile for 4th graders, with the total sample falling at the 56th percentile. This indicates that our sample reported slightly higher than average levels of bullying. Victimization scores ranged from the 50th percentile for 5th graders to the 66th percentile for the 3rd graders, with the total sample falling at the 75th percentile. Thus, our sample also reported higher than average levels of victimization.

Although Olweus (2004) states that the OBVQ is appropriate for students as young as 3rd grade, we did not find any studies evaluating this measure in grades below 5th. Therefore, we compared 5th graders in our study to 5th graders studied by Solberg and Olweus (2003). Solberg and Olweus (2003) classified 12.3% of 5th graders as victims and 4.6% as bullies. In our sample we classified 10.9% of 5th graders as victims and 3.4% of 5th graders as bullies. The 5th graders in our study are reporting slightly less than the 5th graders in the Olweus (2004) study.

Although students self-report more victimization than bullying, teachers reported similar levels of bullying and victimization. There are several possible explanations for this difference between teacher and student reports. One possibility is that the teachers are not observing much of the bullying and victimization (Olweus, 1994). Students may not tell the teacher every time that they are victimized, so they only know about a small number of incidents. Another possibility is that teachers are more aware of bullying behaviors because a small number of students tend to display these behaviors repeatedly; whereas, more students may be impacted by

victimization but to a lesser degree. Therefore, the teachers eventually become aware of whom the bullies are because of their repeated bullying behaviors. However, they may be less likely to become of aware of victimization if the bullies have multiple victims and the victims do not always tell the teacher about these incidents.

In summary, based on the BVS, self-reported rates of victimization in our study were above average. However, self-reported rates of bullying were average. In addition, self-reported levels of victimization decreased with age. Based on the OBVQ, more students in our sample were classified as victims than as bullies. This may be due to a few bullies victimizing multiple students or students not admitting to socially undesirable behaviors. In contrast to these findings, teachers reported similar levels of bullying and victimization. Thus, there is a discrepancy between student and teacher reports of bullying and victimization. In order to understand these findings it must be determined whether students are over-reporting victimization or whether teachers are under-reporting victimization (Olweus, 1994; Pellegrini & Bartini, 2000).

Internal Consistency Reliability

Since an alpha of .85 and above is considered good (Nunnally; as cited in Mash & Hundsley, 2005), both of the BVS Subscales showed good to excellent reliability in this sample. Although the alphas for the OBVQ scales were lower than the BVS scales, internal consistency was still good for the OBVQ. Internal consistency varied by grade for certain subscales. For instance, on the BVS Bullying Subscale, 4th graders had the highest internal consistency. In addition, on the BVS Victimization Subscale, 3rd graders had the highest internal consistency. Finally, on the OBVQ Bullying Subscale, 5th graders had the highest internal consistency. Internal consistency for the OBVQ Bullying Subscale for 4th graders was .785 which is only adequate.

Achenbach et al. (1987) has previously stated that self-report from children under the age of 11 should interpreted with caution due to a lack of reliability. In the current study evidence of good internal consistency reliability was found and did not differ depending on grade.

In summary, the BVS Bullying Subscale had stronger alphas across grade than did the OBVQ Bullying Subscale, indicating that BVS Bullying was slightly more reliable than OBVQ Bullying. This pattern was also found for BVS Victimization being slightly more reliable than the OBVQ Victimization.

Concurrent Validity

For 3rd graders, Teacher Bullying was significantly correlated with BVS Bullying but not OBVQ Bullying. Similarly, Teacher Victimization was significantly correlated with BVS Victimization but not OBVQ Victimization. Thus, for 3rd graders the BVS demonstrated better concurrent validity than the OBVQ. For 4th graders, Teacher Bullying was significantly correlated with both of the self-report bullying scales. Teacher Victimization was significantly correlated with both of the self-report victimization scales. This indicates that the BVS and OBVQ showed good concurrent validity for 4th graders. Finally, for 5th graders, Teacher Victimization was significantly correlated with BVS and OBVQ Victimization. In addition, Teacher Bullying was significantly correlated with BVS and OBVQ Bullying. Thus, the BVS and OBVQ showed good concurrent validity for 4th and 5th graders. This may be because the format and vocabulary of the OBVQ is not as appropriate for the 3rd graders as it is for 4th and 5th graders.

Convergent and Discriminant Validity

Although for the total sample the within-trait cross-method correlations reached statistical significance, the correlations were more variable when examined across grade. All of the within-

trait cross-method correlations met statistical significance across all grades providing evidence of convergent validity. All grades reached approximately the same level of statistical significance. In terms of the cross-trait within-method correlations, the BVS and OBVQ correlations were the strongest for the 3rd grade students compared to the other grades. This could indicate that 3rd graders are more likely to admit to being victimized and are also more likely to admit to bullying other students. This could also indicate that 3rd graders were also more open to responding honestly rather than withholding information. This interpretation is supported by 3rd graders having the highest percentile score on BVS Victimization.

Discriminant validity for the total sample was significant for comparisons of the correlations between the OBVQ and BVS scales. This indicates that these subscales are related through the content of the items and not just because they are from the same scale. In other words, the constructs of bullying and victimization do exist on these measures and the results obtained are not only from the fact that the questions were developed similarly but due to them assessing the same construct. For the total sample, the BVS subscales showed stronger discriminant validity than the OBVQ subscales.

For the discriminant validity comparisons by grade, the pattern is once again that the significance increased as grade level increased. For 3rd graders, comparisons for the OBVQ and BVS Victimization correlations had the best discriminant validity, followed by BVS and OBVQ Bullying correlations. This indicates that for 3rd graders, the content of the victimization subscales is better discriminated against. Fourth graders, had better discriminant validity than 3rd graders. The comparisons between the correlations of the BVS and OBVQ constructs were both significant, indicating that the individual measures were related based on the construct, not

because they were created developed in the same measure. Fifth graders also had better discriminant validity than the 3rd graders. Indicating that as grade level increased there is a better discrimination of what their role is in the bully-victim process. It seems that students in lower grades identify themselves as being involved in this process more frequently but may not discriminate as readily about the specifics of their involvement.

In summary, the within-trait cross-method correlations for the entire sample reached significance. Cross-trait within-method correlations for the OBVQ and BVS were strongest for 3rd graders. For discriminant validity of the total sample the BVS showed stronger significance. Across grades discriminant validity was stronger as grade level increased, so that 5th and 4th graders showed better discrimination than 3rd graders.

Conclusions and Significance

As previously stated, being victimized is correlated with many negative consequences. Before schools implement expensive and time-consuming programs aimed at decreasing the bullying and victimization occurring in the schools, researchers need to know how to accurately measure the change that is occurring to be able to state that the programs are working. While there are many measurement tools available that can estimate the bullying and victimization that is occurring, these tools also have problems associated with them, such as a lack of information about the psychometric properties, lack of evidence for cross-cultural validity, or a lack of replication by laboratories other than the one where they were created. The current study aimed to close this gap by providing evidence for the psychometric properties of the measures. The OBVQ created in Norway, is considered by some to be the gold standard of the bullying and victimization measurement tools. The BVS is a more recently developed instrument with little

information about the psychometric properties other than what is provided in the manual created by the authors.

In order to understand what this current study means for future research it is best to examine this by grade. For 3rd graders the BVS was found to be more reliable and valid than the OBVQ. This is important to note for many reasons. As already stated the Olweus measure states that it is appropriate for children of this age but based on the psychometric properties, it may not be the best measure available. Although the psychometric properties of the OBVQ improved as grade increased, there was some variability in these properties, whereas this was not present for the BVS. For 5th graders the psychometric properties of both measures were good. The decision over which measure to use should be based on what the school personnel is trying to achieve. If an early intervention program is the task at hand then the BVS is probably the better of the two instruments. The measure is reliable and valid for children as young as 3rd grade and can be used at follow-up for years later. However, for older students there may be some advantages to using the OBVQ. The OBVQ includes items related to the bullying experience and offers more than just bully and victimization subscales. The use of the OBVQ can aid in gathering information about perceptions related to bullying, locations of bullying, who the child informs about the bullying, and attitudes towards bullying. For older students, the OBVQ may be a more desirable instrument because of the wealth of information that can be obtained from it.

Limitations and Future Directions

One limitation of this study is related to the teacher measure. This measure was created based on convenience for the teachers. As previously stated other measures include multiple items about individual students and an increase in the amount of time a teacher takes to complete

it. We created this measure in an effort to provide a quick measure of bullying and victimization from the teachers. One reason for this was the desire to increase in the likelihood that the teachers would actually complete and return the measure. Although this measure was deemed appropriate for this study, if resources allow, future studies should include a lengthier teacher measure. Future studies could also use another form of other report, such as peer or parent reports. This would be beneficial to answer the question about if the bullying report is accurate. Although the teacher measure asked whom they viewed as a bully it would be beneficial to see how students (who would possibly be the recipient of the bullying) view other students. This could help determine whether bullies are underreporting their behaviors or if there is a small number of children victimizing multiple students.

All students who participated in this study were from a small southwestern college town. The generalizability of the results of this study could be limited for other geographic regions. Future studies should also attempt to have a population of students that are more ethnically diverse. A final limitation with our sample is that although we had a large number of students participate; this was only 60% of the possible students. This leaves open the possibility of bias in the sample based on which parents returned the informed consent. The school administrators originally suggested that teachers pass out the forms during the parent teacher conferences. However, not all teachers passed out the forms and some of the teachers sent the consents home with their students. This meant that only students whose parents either attended the conference or returned the informed consent to the teacher participated in the study.

A possible limitation of this study is that the schools were already attempting interventions with some students. These interventions were conducted through the guidance

counselor, classroom teacher, or principal and involved tactics such as education about what constitutes bullying, how to handle bullying, and some peer mediation with student conflict. These interventions were not part of a comprehensive school-wide bullying and victimization intervention and the results should be generalizable to other schools that are beginning to make an effort to rectify this problem.

Another limitation of the present study was removing an item from the Olweus measure. The current school system where this measurement study took place did not approve of the question about sexual forms of bullying. Thus, only 9 out of the 10 items were used. This could have an effect on the psychometrics of the OBVQ, however, since we were comparing the OBVQ to the BVS it may be a fair comparison since the BVS does not as questions about sexual forms of bullying. It is important to note that on qualitative items, some children's descriptions of other forms of bullying would have qualified for sexual bullying. However, this could be seen as a limitation of the OBVQ. In the United States this question could limit the number of parents willing to participate, thus if the item is left in it could also raise problems. If this study is replicated in the future the researchers many want to consider including the item for older students but omitting it for the 3rd graders.

Conclusions

In summary, the current study found that the Reynolds BVS was able to assess bullying and victimization with better psychometric properties than the OBVQ for 3rd and 4th graders. Specifically it seemed while internal consistency was good for all students, the BVS had slightly higher internal consistency than the OBVQ. Both measures had good psychometric properties for 5th graders. Convergent and discriminant validity comparisons had stronger significance for 5th

graders compared to 3rd graders. The BVS had stronger concurrent validity for 3rd graders than the OBVQ. For 5th graders both measures had good concurrent validity. Future research that plans to include younger students (such as 3rd graders) should consider using the BVS measure to ensure greater accuracy in the results.

REFERENCES

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, 213-232.

Arora, C. M. J., & Thompson, D. A. (1987). Defining bullying for a secondary school. *British*

Psychological Society, V, 110-120.

Austin, S., & Joseph, S. (1996). Assessment of bully/victim problems in 8 to 11 year-olds. *British Journal*

of Educational Psychology, 66, 447-456.

Bernstein, J. Y., & Watson, M. W. (1997). Children who are targets of bullying: A victim pattern. *Journal*

of Interpersonal Violence, 12, 483-498.

Boivin, M., & Hymel, S. (1997). Peer experiences and social self-perceptions: A sequential model.

Developmental Psychology, 33, 135-145.

Boulton, M. J. (1993). Proximate causes of aggressive fighting in middle school children. *British Journal*

of Educational Psychology, 63, 231-244.

Campbell, D. T., & Fiske, D. (1959). Convergent and discriminant validation by the multitrait-

multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Casey-Cannon, S., Hayward, C., & Gowen, K. (2001). Middle-school girls' reports of peer victimization:

Concerns, consequences, and implications. *Professional School Counseling, 5,* 138-147.

Coie, J. D., Dodge, K. A., & Coppotelli, H. (1982). Dimensions and types of social status: A cross-age

perspective. *Developmental Psychology, 12,* 557-570.

Colvin, G., Sugai, G., Good, R. H., & Lee, Y. (1997). Using active supervision and precorrection to

improve transition behaviors in an elementary school. *School Psychology Quarterly, 12,* 344-363.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Crick, N. R., & Grotpeter, J. K. (1996). Children's treatment by peers: Victims of relational and overt aggression. *Development and Psychopathology, 8,* 367-380.

Crik, N.R. (1996). The role of overt aggression, relational aggression, and prosocial behavior in children's social adjustment. *Child Development, 67,* 2317-2327.

Crick, N.R. (1997). Engagement in gender normative versus non-normative forms of aggression: Links to social-psychological adjustment. *Developmental Psychology, 33*, 610-617.

Crick, N.R., Casas, J.F., Ku, H.C. (1999).Relational and physical forms of peer victimization in preschool. *Developmental Psychology, 35,* 376-385.

Crick, N.R., Ostrov, J.M., & Werner, N.E. (2006). A longitudinal study of relational aggression, physical aggression, and children's social-psychological adjustment. *Journal of Abnormal Child Psychology*, 34, 131-142.

Crothers, L. M., & Levinson, E. M. (2004). Assessment of bullying: A review of methods and instruments. *Journal of Counseling & Development, 82*, 496-503.

Cullerton-Sen, C., & Crick, N.R.(2005). Understanding the effects of physical and relational victimization: The utility of multiple perspectives in predicting social-emotional adjustment. *School Psychology Review, 34*, 147-160.

Dennis, M. J. B., & Satcher, J. (1999). Name calling and peer beliefs of elementary school children. *Professional School Counseling, 3, 76-*80.

Dodge, K. A., & Coie, J. D. (1987). Social-information-processing factors in reactive and proactive aggression in children's peer groups. *Journal of Personality and Social Psychology, 53,* 1146-1158.

Dodge, K. A., Coie, J. D., Pettit, G. S., & Price, J. M. (1990). Peer status and aggression in boys' groups: Developmental and contextual analyses. *Child Development, 61,* 1289-1309.

Edelbrock, C.S. & Achenbach, T.M. (1984).The teacher vesion of the child behavior profile:In boys aged

6-11. *Journal of Consulting and Clinical Psychology, 52,* 207-217.

Embry, S. L. & Luzzo, D. A. (1996). The relationship between name-calling and peer beliefs among

elementary school children: Implications for school counselors. *Elementary School Guidance &*

Counseling, 31, 122-130.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior*

Research Methods, Instruments, & Computers, 28, 1-11.

Evans, C., & Eder, D. (1993). "No Exit": Processes of social isolation in the middle school. *Journal of*

Contemporary Ethnography, 22, 139-170.

Garrity, C., Jens, K., Porter, W., Sager, N., & Short-Camilli, C. (2004). *Bully-proofing your school:*

Working with victims and bullies. Manual. Longmont, Colorado: Sopris West Educational

Services.

Gilmartin, B. G. (1987). Peer group antecedents of severe love-shyness in males. *Journal of Personality,*

55, 467-488.

Harter, S. (1985). *The self-perception profile for children: Revision of the perceived competence scale for*

children. Manual. Denver, CO: University of Denver.

Hartung, McCarthy, Milich, & Martin (2005). Parent-adolescent agreement on disruptive behavior

symptoms: A multitrait-multimethod model. *Journal of Psychopathology and Behavioral*

*Assessment, 27,*159-168.

Hoza, B., Mrug, S., Gerdes, A.C., Bukowski, W.M., Kraemer, H.C., Wigal, T., Hinshaw, S.P., Gold, J.A.,

Pelham, W.E., Arnold, L.E. (2005). What aspects of peer relationships are impaired in children

with attention-deficit/hyperactivity disorder? *Journal of Consulting and Clinical Psychology,*

2005, 411-423.

Kazdin, A. E. (2003). *Research design in clinical psychology*. Boston, MA: Bacon & Allyn.

Lindeman, M., Harakka, T., & Keltikangas-Jarvinen, L. (1997). Age and gender differences in

    adolescents' reactions to conflict situations: Aggression, prosociality, and withdrawal. *Journal of*

    Youth and Adolescence, 26, 339-351.

Mash E.J. & Hunsley J. (2005), Evidence-based assessment of child and adolescent disorders: Issues and

    challenges. *Journal of Clinical Child and Adolescent Psychology, 34*, 362-379.

Murray-Close, D., Ostrov, J.M., Crick, N.R. (2007). A short-term longitudinal study of growth of

    relational aggression during middle childhood: Associations with gender, friendship intimacy,

    and internalizing problems. *Development and Psychopathology, 19,* 187-203.

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. *Journal of*

    American Medical Association, 285, 2094-2100.

Neary, A. & Joseph, S. (1994). Peer victimization and its relationship to self-concept and depression

    among schoolgirls. *Personality and Individual Differences, 16,* 183-186.

Olweus, D. (1993). *Bullying at school: What we know and what we can do.* Cambridge, MA: Blackwell

    Publishers.

Olweus, D. (1994). Annotation: Bullying at school: Basic facts and effects of a school based intervention

    program. *Association for Child Psychology and Psychiatry, 35*, 1171-1190.

Olweus, D. (2004). Brief Psychometric Information. Unpublished Manuscript.

Ortega, R, Mora-Merchan, J. A., Singer, M., Smith, P. K., Pereira, B., & Menesini, E. (2001). Final report

    of the working group on general survey questionnaires and nomination methods concerning

    bullying

Pakaslahti, L. & Keltikangas-Jarvinen, L. (2000). Comparison of peer, teacher, and self-assessments on

    adolescents on adolescent direct and indirect aggression. *Educational Psychology, 20,* 177-190.

Pedhauzer, E.J. & Schmelkin, L.P. (1991). *Measurement, design, and analysis: An integrated approach.*

    Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Pellegrini, A. D., Bartini, B., & Brooks, F. (1999). School bullies, victims, and aggressive victims: Factors relating o group affiliation and victimization in early adolescence. *Journal of Educational Psychology, 91,* 216-224.

Pellegrini, A. D. & Bartini, M. (2000). An empirical comparison of methods of sampling aggression and victimization in school settings. *Journal of Educational Psychology, 92,* 360-366.

Pepler, D. J. & Craig, W. M. (1995). A peek behind the fence: Naturalistic observations of aggressive children with remote audiovisual recording. *Developmental Psychology, 31*, 548-553.

Perry, D. G., & Bussey, K. (1977). Self-reinforcement in high- and low- aggressive boys following an act of aggression. *Child Development, 48,* 653-658.

Perry, D. G., Kusel, S. J., & Perry, L. C. (1988). Victims of peer aggression. *Developmental Psychology, 24,* 807-814.

Perry, D. G., Perry, L. C., & Rasmussen, P. (1986). Cognitive social learning mediators of aggression. *Child Development, 57*, 700-711.

Reynolds, W. M. (2003). *Reynolds bully victimization scales for schools manual.* San Antonio, TX: The Psychological Corporation.

Rigby, K. & Slee, P. T. (1993). Dimensions of interpersonal relation among Australian children and implications for psychological well-being. *Journal of Social Psychology, 133,* 33-42.

Roderick, C., Pitchford, M., & Miller, A. (1997). Reducing aggressive playground behaviour by means of a school-wide 'raffle'. *Educational Psychology, 13,* 57-63.

Ross, D.M. (1996). *Childhood bullying and teasing: What school personnel, other professionals, and parents can do.* Alexandria, VA: American Counseling Association.

Scholte, R.H.J., Engels, R.C.M.E., Overbeek, G., de Kemp, R.A.T., Haselager, G.J.T. (2007). Stability in bullying and victimization and its association with social adjustment in childhood and adolescence. *Journal of Abnormal Psychology, 35,* 217-228.

Schwartz, D., Dodge, K. A., & Coie, J. D. (1993). The emergence of chronic peer victimization in boy's playgroups. *Child Development, 64*, 1755-1772.

Steiger, J.H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87,* 245-251.

Swearer, S. M., & Cary, P. T. (2003). Perceptions and attitudes toward bullying in middle school youth: A developmental examination across the bully/victim continuum. *Journal of Applied School Psychology, 19,* 63-79.

Wiggins, J. S., & Winder, C. L. (1961). The peer nomination inventory: An empirically derived sociometric measure of adjustment in preadolescent boys. *Psychological Reports, 9,* 643-677.

CLASSROOM BULLYING AND VICTIMIZATION CHECKLIST

Please read these instructions before completing this checklist:
On the attached sheet a paper is a list of all of the students in your classroom. For each student, please make two ratings. First, check the one box that best describes how often the student bullies other children in your opinion (please see examples below). Next, check the one box that best describes how often the student is a victim of bullying in your opinion (please see examples below). Please note that a child would be considered a perpetrator of a bullying incident if he/she displays one of these behaviors on his/her own or as part of a group of students. Similarly, a student would be considered a victim if he/she is the target of one of these behaviors as an individual or as part of a group.

Examples of Bullying and Victimization:
• Saying mean and hurtful things
• Making fun of other students
• Calling students mean and hurtful names
• Ignoring or excluding another student from a group
• Leaving another student out of things on purpose
• Hitting, kicking, pushing, shoving, or locking another student inside a room
• Telling lies or spreading rumors about another student
• Sending a mean or hurtful note to another student
• Trying to make other students dislike a student

Protecting Student's Confidentiality: After completing the ratings for each student in your classroom, please cut off the column that contains the students' names. The remainder of the sheet will include only the students' participant numbers and your ratings. Please return the remainder of the sheet to the OSU research team.

School: _____

Grade: _____

Name	Code	Bullying Incidents in the Past Month					Victimization Incidents in the Past Month				
		Never	Rarely	Some-times	Often	Very often	Never	Rarely	Some-times	Often	Very often

TABLES 1-3

Table 1. Bullying and Victimization Scores from BVS.

	3rd Graders			4th Graders			5th Graders			Total Sample			ANOVA		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*p*	η^2
Bullying	218	3.20	5.70	177	3.50	8.28	207	2.97	7.07	602	3.21	7.00	0.28	.760	0.00
Victimization	218	17.08	15.08	177	13.17	12.16	207	8.81	10.66	602	13.09	13.28	22.06	.000	0.07

Notes: Both subscales included 23 items rated from 0 to 3. Therefore, the possible range of scores was 0 to 69.

Table 2. Bullying and Victimization Scores from OBVQ.

	3rd Graders			4th Graders			5th Graders			Total Sample			ANOVA		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*p*	η^2
Bullying	197	1.37	3.26	166	1.32	2.51	211	1.47	3.41	574	1.39	3.12	0.13	.883	0.00
Victimization	197	6.97	7.12	166	5.87	5.85	212	4.61	5.86	575	5.78	6.38	7.19	.001	0.25

Notes: Both subscales included 9 out of 10 items rated from 0 to 4. Therefore, the possible range of scores was 0 to 36.

Table 3. Bullying and Victimization Scores from Teacher Ratings.

	3rd Graders			4th Graders			5th Graders			Total Sample			ANOVA		
	n	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*n*	*M*	*SD*	*F*	*p*	η^2
Bullying	178	0.84	1.06	151	0.86	1.05	216	0.72	1.06	545	0.80	1.06	0.94	.390	0.00
Victimization	177	0.66	0.80	151	0.58	0.80	215	0.73	0.88	543	0.66	0.83	1.41	.244	0.01

Notes: Both subscales included 1 item rated from 0 to 4.

TABLE 4

Table 4. Multi-Trait, Multi-Method Matrix, Total Sample

	OBVQ Bullying	BVS Bullying	OBVQ Victimization	BVS Victimization
OBVQ Bullying	(.846)			
BVS Bullying	Within-trait, cross-method .599** (N = 549)	(.932)		
OBVQ Victimization	Cross-trait, within-method .294** (N = 575)	Cross-trait, cross-method .228** (N = 550)	(.864)	
BVS Victimization	Cross-trait, cross-method .274** (N = 549)	Cross-trait, within-method .324** (N = 602)	Within-trait, cross-method .654** (N = 550)	(.941)

Notes: Cronbach's alpha values are shown in parentheses on the diagonal.

TABLE 5

Table 5. Multi-Trait, Multi-Method Matrix, Third Grade

	OBVQ Bullying	BVS Bullying	OBVQ Victimization	BVS Victimization
OBVQ Bullying	(.856)			
BVS Bullying	Within-trait, cross-method .522** (N = 193)	(.886)		
OBVQ Victimization	Cross-trait, within-method .450** (N = 197)	Cross-trait, cross-method .384** (N = 193)	(.883)	
BVSVictimization	Cross-trait, cross-method .388** (N = 193)	Cross-trait, within-method .530** (N = 218)	Within-trait, cross-method .637** (N = 193)	(.944)

Notes: Cronbach's alpha values are shown in parentheses on the diagonal.

TABLE 6

Table 6. Multi-Trait, Multi-Method Matrix, Fourth Grade

	OBVQ Bullying	BVS Bullying	OBVQ Victimization	BVS Victimization
OBVQ Bullying	.785			
BVS Bullying	Within-trait, cross-method .632** (N = 160)	.951		
OBVQ Victimization	Cross-trait, within-method .328** (N = 166)	Cross-trait, cross-method .267** (N = 160)	.831	
BVS Victimization	Cross-trait, cross-method .235** (N = 160)	Cross-trait, within-method .277** (N = 177)	Within-trait, cross-method .668** (N = 160)	.930

Notes: Cronbach's alpha values are shown in parentheses on the diagonal.

TABLE 7

Table 7. Multi-Trait, Multi-Method Matrix, Fifth Grade

	OBVQ Bullying	BVS Bullying	OBVQ Victimization	BVS Victimization
OBVQ Bullying	.866			
BVS Bullying	Within-trait, cross-method .666** (N = 196)	.944		
OBVQ Victimization	Cross-trait, within-method .129* (N = 211)	Cross-trait, cross-method .046 (N = 197)	.853	
BVS Victimization	Cross-trait, cross-method .200** (N = 196)	Cross-trait, within-method .200** (N = 207)	Within-trait, cross-method .614** (N = 197)	.935

Notes: Cronbach's alpha values are shown in parentheses on the diagonal.

TABLE 8

Table 8. Correlations with the teacher measure.

	Third Grade		Fourth Grade		Fifth Grade		Total Sample	
	Teacher Bullying	Teacher Victimization	Teacher Bullying	Teacher Victimization	Teacher Bullying	Teacher Victimization	Teacher Bullying	Teacher Victimization
OBVQ Bullying	.124 $p = .052$	.099 $p = .098$	.323 $p = 000$	.144 $p = .043$	.183 $p = .004$	.034 $p = .314$	.185 $p = .000$	.078 $p = .038$
OBVQ Victimization	.159 $p = .018$	.100 $p = .095$	.214 $p = .005$	.134 $p = .055$	.251 $p = .000$	.274 $p = .000$	.210 $p = .000$	.167 $p = .000$
BVS Bullying	.221 $p = .002$	.108 $p = 173$	.405 $p = .000$	.173 $p = .019$	.265 $p = .000$	.113 $p = .056$	.293 $p = .000$	.128 $p = .002$
BVS Victimization	.214 $p = .002$	.181 p = .009	.190 $p = .011$	.276 $p = .000$	.404 $p = .000$	.340 $p = .000$	.272 $p = .000$	.237 $p = .000$
Teacher Victimization	.376 $p = .000$	1.00	.299 $p = .000$	1.00	.503 $p = .000$	1.00	.404 $p = .000$	1.00
Teacher Bullying	1.00	.376 $p = .000$	1.00	.299 $p = .000$	1.00	.503 $p = .000$	1.00	.404 $p = .000$

TABLE 9

Table 9. Discriminant Validity Comparisons, Total Sample.

Correlation 1	Correlation 2	t-value	Significance
BVS Bullying and OBVQ Bullying	BVS Victimization and OBVQ Bullying	8.00	.0005
BVS Bullying and OBVQ Bullying	OBVQ Victimization and OBVQ Bullying	7.04	.0005
BVS Bullying and OBVQ Bullying	OBVQ Victimization and BVS Bullying	8.96	.0005
BVS Bullying and OBVQ Bullying	BVS Victimization and BVS Bullying	6.58	.0005
BVS Victimization and OBVQ Victimization	BVS Bullying and OBVQ Victimization	11.00	.0005
BVS Victimization and OBVQ Victimization	OBVQ Bullying and BVS Victimization	9.57	.0005
BVS Victimization and OBVQ Victimization	OBVQ Bullying and OBVQ Victimization	8.98	.0005
BVS Victimization and OBVQ Victimization	BVS Victimization and BVS Bullying	8.03	.0005

TABLE 10

Table 10. Discriminant Validity Comparisons, Third Grade.

Correlation 1	Correlation 2	t-value	Significance
BVS Bullying and OBVQ Bullying	BVS Victimization and OBVQ Bullying	2.24	.025
BVS Bullying and OBVQ Bullying	OBVQ Victimization and OBVQ Bullying	1.08	0.2
BVS Bullying and OBVQ Bullying	OBVQ Victimization and BVS Bullying	2.14	.025
BVS Bullying and OBVQ Bullying	BVS Victimization and BVS Bullying	-0.12	Non-Significant
BVS Victimization and OBVQ Victimization	BVS Bullying and OBVQ Victimization	4.61	.0005
BVS Victimization and OBVQ Victimization	OBVQ Bullying and BVS Victimization	4.20	.0005
BVS Victimization and OBVQ Victimization	OBVQ Bullying and OBVQ Victimization	3.44	.0005
BVS Victimization and OBVQ Victimization	BVS Victimization and BVS Bullying	1.80	.05

TABLE 11

Table 11. Discriminant Validity Comparisons, Fourth Grade.

Correlation 1	Correlation 2	t-value	Significance
BVS Bullying and OBVQ Bullying	BVS Victimization and OBVQ Bullying	5.19	.0005
BVS Bullying and OBVQ Bullying	OBVQ Victimization and OBVQ Bullying	3.99	.0005
BVS Bullying and OBVQ Bullying	OBVQ Victimization and BVS Bullying	4.97	.0005
BVS Bullying and OBVQ Bullying	BVS Victimization and BVS Bullying	4.53	.0005
BVS Victimization and OBVQ Victimization	BVS Bullying and OBVQ Victimization	5.44	.0005
BVS Victimization and OBVQ Victimization	OBVQ Bullying and BVS Victimization	6.10	.0005
BVS Victimization and OBVQ Victimization	OBVQ Bullying and OBVQ Victimization	4.52	.0005
BVS Victimization and OBVQ Victimization	BVS Victimization and BVS Bullying	5.27	.0005

TABLE 12

Table 12. Discriminant Validity Comparisons, Fifth Grade.

Correlation 1	Correlation 2	t-value	Significance
BVS Bullying and OBVQ Bullying	BVS Victimization and OBVQ Bullying	6.60	.0005
BVS Bullying and OBVQ Bullying	OBVQ Victimization and OBVQ Bullying	6.89	.0005
BVS Bullying and OBVQ Bullying	OBVQ Victimization and BVS Bullying	8.44	.0005
BVS Bullying and OBVQ Bullying	BVS Victimization and BVS Bullying	6.60	.0005
BVS Victimization and OBVQ Victimization	BVS Bullying and OBVQ Victimization	7.76	.0005
BVS Victimization and OBVQ Victimization	OBVQ Bullying and BVS Victimization	5.35	.0005
BVS Victimization and OBVQ Victimization	OBVQ Bullying and OBVQ Victimization	6.56	.0005
BVS Victimization and OBVQ Victimization	BVS Victimization and BVS Bullying	5.12	.0005

Oklahoma State University Institutional Review Board

Date: Thursday, March 08, 2007

IRB Application No   AS0714

Proposal Title: Bullying Measurement Study

Reviewed and   Expedited (Spec Pop)
Processed as:

Status Recommended by Reviewer(s): Approved   Protocol Expires: 3/7/2008

Principal
Investigator(s

Carrie Little   Cynthia Hartung
215 North Murray   215 N. Murray
Stillwater, OK 74078   Stillwater, OK 74078

The IRB application referenced above has been approved. It is the judgment of the reviewers that the rights and welfare of individuals who may be asked to participate in this study will be respected, and that the research will be conducted in a manner consistent with the IRB requirements as outlined in section 45 CFR 46.

[X] The final versions of any printed recruitment, consent and assent documents bearing the IRB approval stamp are attached to this letter. These are the versions that must be used during the study.

As Principal Investigator, it is your responsibility to do the following:

1. Conduct this study exactly as it has been approved. Any modifications to the research protocol must be submitted with the appropriate signatures for IRB approval.
2. Submit a request for continuation if the study extends beyond the approval period of one calendar year. This continuation must receive IRB review and approval before the research can continue.
3. Report any adverse events to the IRB Chair promptly. Adverse events are those which are unanticipated and impact the subjects during the course of this research; and
4. Notify the IRB office in writing when your research project is complete.

Please note that approved protocols are subject to monitoring by the IRB and that the IRB office has the authority to inspect research records associated with this protocol at any time. If you have questions about the IRB procedures or need any assistance from the Board, please contact Beth McTernan in 219 Cordell North (phone: 405-744-5700, beth.mcternan@okstate.edu).

Sincerely,

Sue C. Jacobs, Chair
Institutional Review Board

VITA

CAROLYN SIMS LITTLE

Candidate for the Degree of

Master of Science

Thesis: BULLYING MEASUREMENT STUDY

Major Field: Clinical Psychology

Biographical:

       Personal Data: Born in Birmingham, Alabama on November 1, 1982 to Chuck and
            Margie Little.

       Education:
       Completed the requirements for the Master of Science in Psychology at Oklahoma
       State University, Stillwater, Oklahoma July, 2007.

       Experience: Carolyn completed an honors thesis at Auburn University titled, *Alcohol
            and Choice Behavior Among College Students*. Since attending Oklahoma State
            University she has helped with research projects in the Disruptive Behavior
            Disorders laboratory.

       Professional Memberships:
            APA Division 53, Psychology Graduate Student Association

Name: Carolyn Little   Date of Degree: December, 2007

Institution: Oklahoma State University   Location: Stillwater, Oklahoma

Title of Study: BULLYING MEASUREMENT STUDY

Pages in Study: 71   Candidate for the Degree of Master of Science

Major Field: Clinical Psychology

Scope and Method of Study: Although some states have mandated the use of bullying prevention
        programs in schools, there is very little research available on the psychometric properties
        of the instruments that evaluate these programs. The current study evaluated the
        psychometric properties of two instruments in third through fifth grade students and their
        teachers. The psychometric properties were examined through internal consistency
        reliability, concurrent validity, discriminant, and convergent validity.

Findings and Conclusions: For the combined sample and each individual grade the psychometric
        properties of the BVS were stronger than the OBVQ. Children report higher levels of
        victimization than admitted bullying. When considering an instrument to measure a
        bullying intervention, the BVS will provide a more accurate portrayal of the bullying and
        victimization problems in 3rd graders. This will enable researchers or school personnel to
        gain accurate insight into the change in bullying and victimization behaviors over time.

ADVISER'S APPROVAL: Cynthia M. Hartung