EFFECTIVENESS AND ACCEPTABILITY OF A BULLY PREVENTION PROGRAM WITH AND WITHOUT VIDEO SUPPLEMENTS

by

Benjamin Bruce Belnap

A thesis submitted to the faculty of The University of Utah in partial fulfillment of the requirements for the degree of

Master of Science

Department of Educational Psychology

The University of Utah

May 2009
SUPERVISORY COMMITTEE APPROVAL

of a thesis submitted by

Benjamin Bruce Belnap

This thesis has been read by each member of the following supervisory committee and by majority vote has been found to be satisfactory.

Robert E. O'Neill
FINAL READING APPROVAL

To the Graduate Council of the University of Utah:

I have read the thesis of Benjamin Bruce Belnap in its final form and have found that (1) its format, citations, and bibliographic style are consistent and acceptable; (2) its illustrative materials including figures, tables, and charts are in place; and (3) the final manuscript is satisfactory to the supervisory committee and is ready for submission to The Graduate School.

Date

Daniel E. Olympia
Chair: Supervisory Committee

Approved for the Major Department

Elaine Clark
Chair/Dean

Approved for the Graduate Council

David S. Chapman
Dean of The Graduate School
ABSTRACT

With a growing number of children affected by bullying and an increased public awareness of the issue in recent years, bullying prevention programs are becoming more widespread in school settings. Student involvement in multimedia and technology is also increasingly widespread. Therefore, the question is raised as to whether students would respond better to a bully prevention program with an audiovisual component than to a program with lessons taught solely from a manual presented by the classroom teacher.

This study compared the effectiveness of presenting lessons from the Tough Kid Bully Blockers curriculum when using only the manual compared with when using the manual and an audiovisual component. A pretest-posttest design involving two groups was used to assess content knowledge acquisition by third-grade students. While assessing the effectiveness of the program, the study also addressed whether there were differences between the manual-only and manual + video groups in student and teacher acceptability ratings of the program. Acceptability was assessed by checklist using the Children’s Intervention Rating Profile for the students and the Behavior Intervention Rating System for the teachers. A final element of the study focused on whether there were differences in treatment integrity between the manual-only and manual + video groups using checklists and observations. No differences were found between the manual-only and
manual+video groups, although both groups showed significant gains from pretest to posttest in content knowledge acquisition. In addition, no differences were found between groups in overall acceptability ratings. Implications for future research and practice are included.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Bullying Prevention and Intervention Programs</td>
<td>2</td>
</tr>
<tr>
<td>Use of Video Instruction</td>
<td>7</td>
</tr>
<tr>
<td>Acceptability of Treatment</td>
<td>10</td>
</tr>
<tr>
<td>Operationalizing Treatment Acceptability</td>
<td>12</td>
</tr>
<tr>
<td>Research Questions</td>
<td>15</td>
</tr>
<tr>
<td>METHODS</td>
<td>16</td>
</tr>
<tr>
<td>Development of the Study</td>
<td>16</td>
</tr>
<tr>
<td>Video Supplement for the Tough Kid Bully Blockers Program</td>
<td>17</td>
</tr>
<tr>
<td>Setting</td>
<td>18</td>
</tr>
<tr>
<td>Participants</td>
<td>18</td>
</tr>
<tr>
<td>Dependent Measures</td>
<td>19</td>
</tr>
<tr>
<td>Design and Data Analysis</td>
<td>20</td>
</tr>
<tr>
<td>Procedures</td>
<td>21</td>
</tr>
<tr>
<td>RESULTS</td>
<td>23</td>
</tr>
<tr>
<td>Research Question 1</td>
<td>23</td>
</tr>
<tr>
<td>Research Question 2</td>
<td>26</td>
</tr>
<tr>
<td>Research Question 3</td>
<td>32</td>
</tr>
<tr>
<td>Research Question 4</td>
<td>34</td>
</tr>
<tr>
<td>DISCUSSION</td>
<td>38</td>
</tr>
<tr>
<td>Limitations</td>
<td>44</td>
</tr>
<tr>
<td>Implications for Practice and Future Research</td>
<td>46</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>Page</td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>A.  UNIT AND LESSON TITLES WITH LESSON OBJECTIVES</td>
<td>49</td>
</tr>
<tr>
<td>B.  BULLYING KNOWLEDGE QUIZ</td>
<td>53</td>
</tr>
<tr>
<td>C.  TREATMENT INTEGRITY CHECKLIST</td>
<td>56</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>58</td>
</tr>
</tbody>
</table>
LIST OF TABLES

Table                                Page

1. Means and Standard Deviations of Knowledge Acquisition        23
   Scores From Pretest to Posttest

2. Item Ratings for Both Conditions                                  27


4. Mean Reliability Coefficients for Raters                        35

5. Mean Number of Checklist Items Completed With and Without       36
   an Observer Present
INTRODUCTION

On April 20, 1999, 2 high school students entered Columbine High School in suburban Denver, Colorado, carrying firearms and homemade bombs. In the course of approximately 1 hour, these 2 students killed 11 fellow students and 1 teacher while injuring 24 other students before committing suicide themselves. To more easily victimize the athletes in the school, the Columbine shooters commanded all “jocks” to stand up when they entered each classroom. After finding home-video and journal evidence from the shooters after the massacre, the public has widely accepted that many athletes at Columbine were targets as a result of previous bullying behavior toward the shooters (Fitzgerald, 2007).

Consequently, the Columbine massacre provoked much interest in the topic of bullying, which continues to be a topic of great interest among researchers and the general public alike (American Society of Educators, 2003; Baseline Research, LLC, 2000; Brunner & Lewis, 2007).

Early and current research on bullying has been effective in determining negative short-term and long-term effects for both bullies (Olweus, 1993; Ridgway, 1973) and their victims (Hugh-Jones & Smith, 1999; Olweus, 1978, 1994; Rivers, 2001; Wessler & DeAndrade, 2006). Anderson et al. (2001) found that perpetrators of homicide were more than twice as likely than victims of homicide to have been bullied by peers. Almost 30% of school-age youth report some
involvement in moderate or frequent bullying as aggressors, victims, or both (Nansel et al., 2001). Being involved in bullying is of great concern, as research shows that perpetrators and victims alike appear to be at greatest risk of loneliness, difficulty making friends, lack of success in school, and involvement in problem behaviors such as smoking and drinking alcohol (Ericson, 2001). Because of its long-term effects, bullying continues to be an issue of great concern (Elinoff, Chafouleas, & Sassu, 2004; Frey, Hirschstein, & Snell, 2005).

**Bullying Prevention and Intervention Programs**

With increased attention on this issue, many bully prevention programs have been developed for both educators and students in the school system (Hoover & Oliver, 1996; Olweus, 1991; Olweus, Limber, & Mihalic, 1999; Whitted & Dupper, 2005). Although many programs are available, only two meta-analyses indicated limited effects from bully prevention programs (Merrell, Gueldner, Ross, & Isava, 2008; Schneider, Smith, & Smith, 2004). Merrell et al. evaluated 16 studies on antibullying programs. They found some evidence supporting the notion that many school bullying interventions tend to enhance self-esteem, social competence, and peer acceptance of students but do not generalize to actual responses in bullying situations. Overall, Merrell et al. found limited positive effects of bully prevention programs in decreasing self-reported bullying behavior. Coinciding with the findings of Merrell et al., Schneider et al. concluded from their meta-analysis that the majority of programs available yielded nonsignificant effects according to self-report data. Schneider et al. indicated, however, that
programs with systematic monitoring at school or home yielded larger effect sizes than those with limited or no monitoring. A brief review is offered of several such programs.

Olweus et al. (1999) developed the most widely known bully prevention program called the Olweus Bullying Prevention Program, a universal intervention for the prevention and reduction of bullying for both victims and bullies that can be used at the schoolwide, classroom, or individual levels. By focusing primarily on bullying identification, prevention and early detection of bullying behaviors are the focal points. Because early identification is a key element of this intervention, schools can attempt to stop the problem before it becomes serious and unmanageable. This program also encourages regular meetings or assemblies to remind each student of the program and the individual bully prevention goals they have set. The Olweus Bullying Prevention Program has demonstrated significant reductions in reported bullying and general antisocial behavior and has resulted in improvements in the overall social climate of the classroom (Olweus et al.).

While the Olweus Bullying Prevention Program is used only in the school setting (Olweus et al., 1999), another program, the Bully-Proofing Series (Garrity et al., 1994), can be used schoolwide, in individual classrooms, or even in the home by parents. Focusing primarily on the bystander rather than on the victim or bully, the Bully-Proofing Series provides bystanders with feelings of empowerment so they can assist victims in seeking help. Included with the program is a set of five books designed to prevent bullying by targeting successive segments in the
time line of a child’s life, including early childhood, elementary school, middle school, and high school. The books, which are very readable to a wide audience, provide simple information to parents, children, and educators alike. Parents are able to participate and in the final books are taught how to teach bully prevention to their children. Allowing for parent involvement makes the Bully-Proofing Series unique because it can be implemented outside the school.

In contrast to this primary focus on bystanders, Slaby’s (1995) Aggressors, Victims, and Bystanders is a program developed to target victims and bullies as well. This program addresses the thought processes involved in problem solving or “the think-first model.” The program is composed of four steps that are learned and practiced in the classroom to translate effectively into real-world situations. These lessons include the following: (a) Step One—Keep Cool, (b) Step Two—Size Up the Situation, (c) Step Three—Think It Through, and (d) Step Four—Do the Right Thing. Based on principles of cognitive psychology, this program removes the element of impulsive action in bullying and replaces it with concrete thought processes. Field tested with nearly 700 students in urban, suburban, and small-city school districts, the program was found to be effective in decreasing bullying behavior (Slaby).

Similar to Slaby’s (1995) program in its focus on thought processes, Greenberg, Kusché, and Mihalic (1998) developed PATHS: Promoting Alternative Thinking Strategies. This program has proven to be an effective bully prevention program that discourages impulsive action by focusing on self-control, regulating
emotions, and improving thinking and planning skills. Using the techniques learned in this program can lead a victim to methodically come to his or her own conclusions on how to act when faced with a conflict; this allows children to gain feelings of empowerment while keeping calm and resolving conflicts peacefully.

One program, which is technology based, that differs from other bully prevention programs is the Get Real About Violence program (Meyer, Roberto, Boster, & Roberto, 2004). This program is a mixed media prevention program that uses computer training and Internet resources from kindergarten through 12th grade. This program targets a wide range of bullying behavior such as spreading rumors, teasing, and using physical aggression. The goal is to provide students with the skills necessary to keep themselves safe by learning conflict resolution without violence. The Get Real About Violence program is designed for use at the schoolwide level but can also be used with individuals. Matched treatment and control groups of approximately 250 junior high school students in a major metropolitan area completed pretest and posttest questionnaires to determine the effectiveness of the program. The survey and observation data demonstrated positive effects, as students who experienced the Get Real About Violence program were less likely to be verbally aggressive, less likely to passively watch a fight, and less likely to spread rumors.

The Quit It! program, which focuses strictly on younger grades (kindergarten through third grade; Froschl, Sprung, & Mullin-Rindler, 1998), targets the behaviors of teasing and name-calling. The Quit It! program is designed
for teachers to develop safe classrooms where students feel welcome and comfortable. The program is broken down into three components. First, rules are created for the classroom. Second, both the students and teacher discuss teasing and bullying by including opportunities for student discussion about where and why they feel unsafe. Opportunities to practice safe methods in response to bullying are also included. The final step in the Quit It! program explores courage in which students define courage as it relates to teasing and bullying.

Hoover and Oliver (1996) developed a bully prevention program directed at administrators and teachers. The *Bullying Prevention Handbook: A Guide for Principals, Teachers, and Counselors* teaches educators how to recognize bullying behavior in its early phase so as to prevent it. This program, which addresses how an administrator can mentor teachers and coaches, optimizes the educator’s role in a schoolwide bully prevention program. When implemented effectively, this program can lead to increased awareness schoolwide because of teacher and administrator participation. This program relies on an effective reinforcement system for students who actively participate in the early detection of bullying, consequently increasing school unity.

Another popular bully prevention program is called Bully Busters (Horne, Bartolomucci, & Newman-Carlson, 2000). This program consists of a set of two books designed for (a) kindergarten through fifth grade and (b) grades six, seven, and eight. The program is research based and contains materials that emphasize self-control and bully prevention by helping teachers increase their own awareness,
knowledge base, and intervention skills. The Bully Busters program includes 36 activities designed for grades kindergarten through fifth grade and 39 activities for grades six, seven, and eight, helping students attain bully prevention skills.

The Tough Kid Bully Blockers program (Bowen, Rhode, Jenson, & Ashcraft, 2008) is divided into six units, with several lessons for each unit intended for weekly implementation either in the classroom or at the schoolwide level. The units include the definition of a bully, how to report bullying, friendship skills, problem solving, respecting differences, and building self-confidence. The program is unique in its sensitivity to teachers’ busy schedules, using little teaching time. Supplemental activities are included for each lesson, providing students with hands-on practice in “blocking” bullies.

Use of Video Instruction

In reviewing previous research on the use of video as a means of instruction, McCoy and Hermansen (2007) indicated that video technology has been used successfully for 3 decades to teach a variety of social, academic, and functional skills. Video technology has grown in popularity because of its lack of expense and ease of standardization with clinicians and educators (Webster-Stratton, Kolpacoff, & Hollinsworth, 1988). Early studies on the use of video instruction reported it to be more effective than written materials, lectures, rehearsal of information (O’Dell et al., 1982), parent training, or behavior modeling (Webster-Stratton et al.). Video instruction, which is effective in helping individuals visualize difficult-to-understand concepts that arise in clinical and
school settings (Bransford, Brown, & Cocking, 2003), has been shown to have both positive short- and long-term effects of up to 1 year for a range of social skills (Webster-Stratton, Hollinsworth, & Kolpacoff, 1989).

One of the most common uses for video technology is the instruction of individuals with autism. By using adults, peers, or self in video modeling of salient social cues, specific communicative behavior, and sequences for task completion (Quill, 2000), individuals with autism were able to learn and execute these skills at a higher and more accurate rate than with live instruction (McCoy & Hermansen, 2007). These individuals generally responded well to video instruction because of the added stimuli (Charlop-Christy, Le, & Freeman, 2000).

Another area in which video instruction is effective is self-monitoring. Video modeling is a behavioral technique that uses video rather than live scenarios or role plays to demonstrate appropriate social behaviors. This method of self-modeling involves the video recording of target behaviors of the individual in order to expand the individual's capability to memorize, imitate, and even generalize those behaviors (Hitchcock, Dowrick, & Prater, 2003). Video modeling has resulted in faster rates of acquisition and increased generalization of a skill when compared with live modeling (Charlop-Christy et al., 2000; Dowrick, 1991).

Video technology is also used as a method of instruction in parent training procedures. Webster-Stratton et al. (1989) found that the use of videos in parent training is efficacious and even enhanced when coupled with live instruction. The benefits of using video instruction in parent training are similar to the benefits of
self-modeling; that is, video instruction is time efficient, cost effective, and more engaging than a live instructor alone. The positive effects of video instruction also make training staff less complicated and standardization procedures easier to follow with the uniformity of video instruction (Webster-Stratton, 1992).

In addition, video instruction has demonstrated positive effects in teaching social skills to individuals with and without disabilities. The relative lack of expensive instruction and uniformity of instruction makes video instruction a plausible method to enhance social skills curricula. A common critique of social skills programs is that learned skills often do not generalize beyond the group instruction setting (Kavale & Mostert, 2004). However, Hosford and Mills (1983) indicated that the use of video-modeling techniques coupled with live role plays can increase the potential for generalization of the skills. In fact, the use of video instruction can also enhance generalization to the home environment (Lasater & Brady, 1995).

Video instruction can be a beneficial alternative or supplement to live instruction because it is easier to train instructors, is more cost effective, is often more engaging and stimulating, and provides uniformity of instruction for standardization of treatment. In addition, video instruction can enhance acceptability by instructors. By appealing to instructors who find live instruction and modeling to be too time consuming (Charlop-Christy et al., 2000), the videos can be used repeatedly, requiring less time for lesson preparation. In addition, positively affecting acceptability is the fact that teachers can assign the instruction
to paraprofessionals with confidence that the program will be delivered with uniformity and integrity (Ayres & Langone, 2005).

Acceptability of Treatment

Difficulties often arise with the implementation of classroom and schoolwide behavior programs. A program can be empirically supported and proven effective, but if not implemented, students will not experience the positive effects. A program is effective only to the degree to which it is implemented. Kazdin (1977) emphasized this concept by saying: “It is not enough for behavioral procedures to be effective. They must also be accepted by the individuals with whom they are being implemented” (p. 428). Three years later, Kazdin (1980) defined treatment acceptability as “judgments of treatment procedures by the consumers of those treatments such as nonprofessionals, laypersons, clients, and others” (p. 259). Wolf (1978) also wrote: “If the participants don’t like the treatment, then they may avoid it, or run away, or complain loudly. And thus, society will be less likely to use our technology, no matter how potentially effective and efficient it might be” (p. 206). Thus began a focus of finding exactly what a teacher considers “acceptable.”

Teacher acceptability is a vital element of any program (Ghaith & Yaghi, 1997). The amount of time required, expense, inconvenience, lack of effectiveness, complexity, or low student interest levels (Bowen, Jenson, & Clark, 2004; Elliott, 1988; Kazdin, 1977) are the primary determining factors as to whether teachers will accept and implement a program. Researchers have found that several factors
contribute to treatment acceptability as perceived by teachers, with overall simplicity being one of them. If an intervention is easy to implement with little planning and preparation involved, it is rated as more acceptable by teachers (Elliott, Witt, Galvin, & Peterson, 1984). In addition, some teachers have identified materials needed, the amount of teacher skill, and the effects of the intervention on other children to be important considerations (Martens, Witt, Elliott, & Darveaux, 1985; Witt, Martens, & Elliott, 1984). Dake, Price, Telljohan, and Funk (2003) found that even teachers’ lack of understanding of the consequences of the target behavior can be a factor in whether a teacher will implement a program. If a teacher is unsure what will come about as a result of the behavior of the student, he or she may not implement the intervention.

While teachers deem these characteristics of an intervention to be important, none has been more significant in acceptability than time-and-effort commitments (Kallestad & Olweus, 2003). Time constraints were considered most critical in nearly every article on the subject of treatment acceptability for teachers (Alderman & Gimpel, 1996; Alderman & Nix, 1997; Elliott, 1988; Elliott et al., 1984; Johnson & Pugach, 1990; Martens et al., 1985; Miltenberger, 1990; Reimers, Wacker, & Koeppel, 1987; Witt, 1986; Witt, Elliott, & Martens, 1984; Witt & Martens, 1983). In fact, in a study focusing on the selection and use of interventions in schools (Nicholas, 1998), of 194 teachers surveyed from four U.S. urban areas, all indicated time commitment to be a major factor in selecting an intervention, while approximately 28% indicated time commitment to be the most
important factor in selecting an intervention. In the same study, approximately 26% of school psychologists surveyed from the same four urban areas determined time commitment to be the most important factor in selecting an intervention. It appears that time involved in a behavior program is significant for teachers because of the many obligations they must fulfill. Teachers are required to prioritize their time in order to cover the academic material that must be taught throughout the school year. Teachers are also required to keep their students on task so as to minimize behavior problems in the classroom while meeting individual student needs in large classes. Their list of obligations is lengthy. With all of these requirements, a schoolwide or classroom-level behavior program may seem difficult and low in priority. Clearly, time constraints are a major factor in acceptability of any school- or classroom-based intervention.

Operationalizing Treatment Acceptability

The most common method for assessing acceptability of behavior interventions is that of surveys completed by teachers, parents, and participants (Finn & Sladeczek, 2001). Many surveys are commonly used to accomplish this task. The Behavior Intervention Rating Scale, developed by Elliott and Von Brock Treuting (1991), is one of these surveys. The Behavior Intervention Rating Scale is a 24-question survey designed for teachers that is rated on a 6-point Likert-type scale ranging from 1 (strongly disagree) to 6 (strongly agree). The Behavior Intervention Rating Scale has proven to be particularly effective because of its ability to measure pretreatment conditions, teacher acceptability, and perceived
effectiveness. To determine the reliability of each scale within the Behavior Intervention Rating Scale, a coefficient alpha was used, yielding an alpha of .97. The three factors of the Behavior Intervention Rating Scale (Acceptability, Effectiveness, and Time Efficiency) yielded alphas of .97, .92, and .87, respectively.

Another frequently used survey for acceptability assessment is the Abbreviated Acceptability Rating Profile (Tarnowski & Simonian, 1992). The Abbreviated Acceptability Rating Profile is comprised of eight questions rated on a 6-point Likert-type scale. This survey is particularly useful when utilized with diverse populations of limited educational background; that is, one half of both validation samples were comprised of persons from diverse socioeconomic and cultural backgrounds. The Abbreviated Acceptability Rating Profile was also rated as having a higher readability level than other abbreviated scales used for the assessment of acceptability. In the psychometric assessment of the Abbreviated Acceptability Rating Profile, a significant positive correlation was found between scale items on the Abbreviated Acceptability Rating Profile and total scale items on the Abbreviated Acceptability Rating Profile scores. Results also yielded support for the reliability of the instrument with split-half and Cronbach-alpha coefficients of .95 and .97, respectively.

The Treatment Evaluation Inventory (Kazdin, French, & Sherick, 1981), which yields a high Cronbach alpha coefficient, is a 19-item questionnaire designed to assess perceived acceptability of interventions for behavior problems in children.
Respondents indicate their agreement or disagreement on a 5-point Likert-type scale. The items are coded such that higher scores indicate greater levels of acceptability. Cronbach's alpha for the Treatment Evaluation Inventory was calculated at .96.

In contrast to other surveys in which teachers answer the questions, the Children's Intervention Rating Profile is a survey that is completed by children involved in an intervention so as to assess the children's acceptability of the intervention (Witt & Elliott, 1983, 1985). The Children's Intervention Rating Profile is a seven-question, 6-point Likert-type rating scale, with one-factor items ranging from 1 (I agree) to 6 (I do not agree). A Cronbach's alpha of .75 was obtained for the Children's Intervention Rating Profile.

Finally, the Intervention Rating Profile-15 (Martens et al., 1985) includes items rated on a 6-point Likert-type scale in which higher scores are translated as greater levels of acceptability. Total scores on the Intervention Rating Profile-15 range from 15 (very low acceptability) to 90 (very high acceptability). The Intervention Rating Profile-15 assesses teachers' acceptability of individual treatments, focusing primarily on the perceived effectiveness of the intervention.

All of these surveys are similar to one another; that is, those who developed them share the goal of assessing teacher and student acceptability and effectiveness of behavioral interventions. Each survey targets characteristics of interventions that are vital in assessing how participants feel about them or the degree of participant acceptability.
While the prevalence of bully prevention programs has greatly increased, there is still a significant lack of supporting data with regard to teacher acceptability of these programs from educators, school psychologists, and parents. Bullying behavior continues to be a serious issue in schools. Solutions may lie in a closer look at the treatment acceptability of bullying programs and the difficulties that schools and teachers have in implementing them due to time constraints in the classroom.

Research Questions

1. Will students participating in the manual + video condition show greater differences in their scores than students in the manual-only condition in the pretest-posttest knowledge quiz?

2. Will teachers in the manual + video condition rate the program as more acceptable than teachers in the manual-only condition?

3. Will students in the manual + video condition rate the program as more acceptable than students in the manual-only condition?

4. Will differences be seen in treatment integrity between the manual + video and manual-only conditions?
METHODS

Development of the Study

During the 2005-2006 school year, graduate students in the school psychology program at the University of Utah began implementation of a program based on the Tough Kid Bully Blockers program (Bowen et al., 2008) at a local elementary school. The program consisted of a number of activities, including classroom presentations, teacher and staff in-service programs, and follow-up activities to encourage student participation.

Initially, the graduate students' role at the elementary school was that of trainers for teachers to assist them with implementation of the Tough Kid Bully Blockers program. Each week the graduate students spent time in the classrooms addressing concerns about implementing the Tough Kid Bully Blockers program and providing teachers with support. The goal was to gradually phase out their support, eventually allowing the teachers to implement the program independently. However, as the school year progressed, the graduate students found that when their support level was high, the program was implemented consistently and there was an increased awareness of bullying. Conversely, with few exceptions, when less support was provided, program implementation diminished. It appeared that teachers were much less likely to implement the program without additional support.
Video Supplement for the Tough Kid Bully Blockers Program

The next year new graduate students at the same elementary school developed six short video-based lessons to serve as a review of critical program elements from the Tough Kid Bully Blockers curriculum. Teaching assistants were trained to present each video to the students. Graduate students also developed specific activities that were linked to each video, actively engaging the students in reviewing the content after viewing each video. The goal of the video development was to facilitate implementation for teachers, maintain fidelity with the core concepts of the program, and, consequently, increase teacher acceptability of the program. Initial informal feedback was positive, suggesting that video development was an effective tool for transmission of program content.

Each video was planned to supplement the Tough Kid Bully Blockers (Bowen et al., 2008) curriculum and to focus on an array of topics in bully prevention, including defining bullying behavior, reporting bullying without tattling, friendshipping skills, solving problems peacefully, respecting differences, and building self-confidence (see Appendix A). Each lesson sequentially progresses and builds on the previous lesson. For example, after learning how to recognize a bully, students learn how to report bullying. Once students feel confident in their abilities to report bullying, they learn how to include and support victims of bullying. Moving through the sequence of bully prevention, students are more prepared to solve problems peacefully and to use the friendshipping skills learned in the previous video. Finally, the videos focus on teaching each student self-
confidence and the skills necessary to increase protective factors of victims.

Setting

The study was conducted in an elementary school located in a suburban school district in a large city in the western United States. This school serves students from kindergarten through sixth grade. The school is modernized, offering a combination of community resources (i.e., gymnasium, community classrooms, and fitness center), enclosed classrooms, and the latest technology (i.e., data projectors and wireless Internet access in all classrooms). Operating on a modified year-round schedule, the school serves approximately 1,100 students who are divided into four attendance tracks. Typically, each track is in school for 45 days followed by 15 days of vacation.

In addition, this school is the location of a unique partnership with a local state university. The university’s purpose is to promote a culture of inquiry and collaboration. Because this elementary school had been a site for implementation of the Tough Kid Bully Blockers program for the 2 years preceding this study (2005-2006 and 2006-2007), there was strong administrative support and familiarity with the program and graduate students from the local state university.

Participants

Teacher and student participants were recruited from four participating third-grade classrooms, with an average class size of approximately 26 students per classroom; this generated an initial sample size of 103 students (59 females, 44
males) and 4 teachers (all female), with a mean of 8.25 years of teaching experience. However, due to student absences throughout the intervention, the student sample size varied from 22 to 26 students in each classroom.

Two classrooms were randomly assigned to the condition, receiving the Tough Kid Bully Blockers lessons using a manual and supplemental videos. The remaining two classrooms were randomly assigned to the condition, receiving the lessons using only the Tough Kid Bully Blockers manual.

**Dependent Measures**

Content knowledge acquisition by students was assessed using a brief quiz developed for the current study, taking material directly from the Tough Kid Bully Blockers manual. The quiz contained 13 items (4 true/false, 9 multiple choice) worth 1 point each, providing a possible score of 13 (see Appendix B). Before receiving any instruction on the bullying program, the students completed a quiz based on the material covered in the manual and videos. The lessons were delivered by manual and video to one group and by manual only to the other group. Directly after implementation of the intervention, the students took the identical quiz administered during baseline.

To meet the conditions requested by the school district involved in the study, some constraints were imposed in the data collection. The district required assurance that no identifying information would be used. Therefore, participants were given tests marked “A” for the manual+video condition and “B” for the manual-only condition and asked not to include their names on the tests.
Design and Data Analysis

Respective gains in knowledge between groups from pretest to posttest were analyzed using an independent sample $t$ test. In keeping with the district requirements, a paired sample design was not utilized, thus limiting the statistical design. Instead, group means on the pretest and posttest measures were compared between the two conditions, and differences between group means were analyzed using independent sample $t$ tests. It was hypothesized that the classrooms in the manual+video condition would show greater difference in their pretest-posttest scores than the classrooms in the manual-only condition. Effect sizes of the program were also calculated, and a within-group design was used to analyze knowledge acquisition gains from pretest to posttest within each group.

Student acceptability of the intervention was assessed using students' total scores on the Children’s Intervention Rating Profile (Witt & Elliott, 1983), which was administered in the last week of the intervention after the last lesson was taught. The survey used for assessing teacher acceptability was the Behavior Intervention Rating Scale (Elliott et al., 1991). The teachers were also given the opportunity to informally report issues they had encountered, including student response to the lessons and whether they would be willing to implement a similarly formatted program in their classrooms.

Some of the wording on the Behavior Intervention Rating Scale and Children’s Intervention Rating Profile was modified slightly to change generic wording (e.g., “the intervention”) to wording more specific to the study (e.g.,
For the purpose of the study, that item was changed to: “I would suggest the use of Bully Blockers to other teachers.” The content on the two surveys was not modified.

To assess whether teachers would implement the program consistently and with integrity, a checklist was developed (see Appendix C). The checklist, containing nine items, outlined the necessary steps to be taken in teaching the lessons. As teachers completed each of the specified components of the lesson, they were asked to sign off on that step. To ensure that teacher ratings were accurate, I observed 20 (10 per group) out of the 92 (46 per group) lessons taught using the same checklist as the teachers. Ten observations of each group were conducted and 3 observations were completed by two observers. Agreements between the rater(s) and the teacher were compared. Reliability coefficients were calculated to determine whether ratings teachers gave themselves were accurate representations of their actual fidelity in treatment.

**Procedures**

All of the classrooms received lessons from the first three units of the Tough Kid Bully Blockers curriculum (Bowen et al., 2008). Six lessons were in Unit One, 7 lessons were in Unit Two, and 10 lessons were in Unit Three. The lessons were taught from the manual two or three times per week over a 9-week period. Along with the lessons from the manual, the classrooms in the
manual + video condition also viewed videos specifically developed for each unit of the Tough Kid Bully Blockers curriculum. One video for each unit was shown to the students twice: (a) once at the beginning of the unit and (b) once at the end of the unit. The first viewing was an introduction to each new unit, and the second viewing was a review of the unit after it had been taught. Appendix A contains lesson titles and objectives.

Each teacher participating in the study received a brief 45-minute training on the use of the materials and the importance of addressing bullying in the classroom. Training also included a question-answer session wherein concerns were addressed. The teachers were also given my contact information, and they were encouraged to contact me with any questions or concerns.

Each teacher received a hard copy of the lesson outlines as well as the content-appropriate activities to follow the lessons, enabling him or her to lead discussion at each lesson’s conclusion. After the intervention’s termination, the teachers completed the Behavior Intervention Rating Scale and the students completed the Children’s Intervention Rating Profile to assess their respective levels of acceptability of the program.
RESULTS

Research Question 1

Will students participating in the manual+video condition show greater differences in their scores than students in the manual-only condition in the pretest-posttest knowledge quiz? It was hypothesized that students in the manual+video condition would show greater gains in knowledge acquisition than students in the manual-only condition. A pretest-posttest design was used to determine knowledge acquisition differences between groups. The pretest and posttest contained the same questions, and both groups took identical tests.

To determine a relationship or difference between group means, the pretest-posttest data were analyzed using an independent sample t test. Table 1 represents mean scores of the manual+video and manual-only groups on both pretest and

<table>
<thead>
<tr>
<th>Groups</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual+video</td>
<td>8.09</td>
<td>10.10</td>
</tr>
<tr>
<td>Manual-only</td>
<td>8.88</td>
<td>10.73</td>
</tr>
</tbody>
</table>

Note. The reported scores represent correct answers out of the 13 items on the test.
posttest measures alongside their respective standard deviations.

The pretest mean scores for both groups were analyzed using an independent sample $t$ test. The $t$ test identified no statistically significant differences between groups on the pretest measure, $t(95) = -1.43, p = .155$. Group differences on the posttest measure were then compared using an independent sample $t$ test wherein no statistically significant differences were identified, $t(95) = -1.362, p = .176$. Since there were no significant differences between groups at baseline (pretest) or on the posttest measure, it can be concluded that the null hypothesis cannot be rejected.

While data do not show significant differences between groups, both groups showed significant increases in their scores within groups from pretest to posttest (see the Figure), indicating that both methods of lesson delivery were effective. The manual-only group showed significant gains from pretest ($n = 51, M = 8.88, SD = 2.99$) to posttest ($n = 45, M = 10.73, SD = 1.97$), $t(94) = -3.53, p = .001$. The manual+video group also showed significant gains from pretest ($n = 46, M = 8.09, SD = 2.40$) to posttest ($n = 52, M = 10.10, SD = 2.55$), $t(96) = -4.00, p = .000$.

Effect sizes were calculated for each condition by subtracting the pretest mean from the posttest mean and dividing by the standard deviation of the two conditions. This calculation resulted in a Cohen's $d$ of .75 for the manual-only group and a Cohen's $d$ of .81 for the manual+video group. Standardized mean effect sizes such as Cohen's $d$ can be interpreted similarly to $z$ scores. These effect
Gain scores on bullying quiz for manual+video group and manual-only group from pretest to posttest.
sizes indicate the mean difference between two scores expressed in standard deviation units. A score of zero represents no change, and effect sizes can be negative or positive. Effects sizes for classrooms in both conditions of the study can be interpreted as moderate to large (Cohen, 1988).

**Research Question 2**

Will teachers in the manual + video condition rate the program as more acceptable than the teachers in the manual-only condition? Data were collected using the Behavior Intervention Rating Scale (Elliott & Von Brock Treuting, 1991). The survey has 24 items rated on a 6-point Likert-type scale ranging from 1 (strongly disagree) to 6 (strongly agree). Mean ratings and standard deviations for both groups are reported in Table 2.

An item of interest on the Behavior Intervention Rating Scale is item 20 (The child’s behavior will remain at an improved level even after Bully Blockers is discontinued.). One of the teachers in the manual + video condition gave this item a score of 6 while the other scored it a 1. Teachers in the manual-only condition gave this item a score of 3 and 4, respectively. In relation to the high overall ratings on the other items, item 20 was rated lower by 3 out of the 4 teachers. These ratings indicate that there may be some reservations in this small sample of teachers with regard to whether the intervention will have an enduring effect.

Although some specific items received lower acceptability scores, overall, both pairs reported high levels of acceptability. On the 6-point Likert-type scale in which a rating of 1 indicates a very low level of acceptability and 6 indicates a
Table 2

*Item Ratings for Both Conditions*

<table>
<thead>
<tr>
<th>Behavior Intervention Rating Scale questions</th>
<th>Manual+video group</th>
<th>Manual-only group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Teacher A</td>
<td>Teacher B</td>
</tr>
<tr>
<td>1. Bully Blockers would be an acceptable intervention to combat bullying behaviors.</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>2. Most teachers would find Bully Blockers appropriate for bullying intervention and/or prevention.</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>3. Bully Blockers should prove effective in targeting bullying.</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>4. I would suggest the use of Bully Blockers to other teachers.</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>5. Bullying in my classroom is severe enough to warrant use of Bully Blockers.</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>6. Most teachers would find Bully Blockers suitable in targeting bullying.</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>7. I would be willing to use Bully Blockers in my classroom.</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>8. Bully Blockers would not result in negative side-effects for the child.</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>9. Bully Blockers would be an appropriate intervention for a variety of children.</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>10. Bully Blockers is consistent with other bullying programs I have used in classroom settings.</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>11. Bully Blockers is a fair way to handle bullying.</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>12. Bully Blockers is reasonable for difficulties that arise from bullying.</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>
Table 2 (continued)

<table>
<thead>
<tr>
<th>Behavior Intervention Rating Scale questions</th>
<th>Manual+video group</th>
<th>Manual-only group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Teacher A</td>
<td>Teacher B</td>
</tr>
<tr>
<td>13. I like the procedures used in Bully Blockers.</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>14. Bully Blockers is a good way to handle bullying in this classroom.</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>15. Overall, Bully Blockers would be beneficial for the children in my classroom.</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>16. Bully Blockers would quickly improve a bully’s behavior.</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>17. Bully Blockers would produce a lasting improvement in a bully’s behavior.</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>18. Bully Blockers would improve a bully’s behavior to the point that it would not noticeably deviate from other classmates’ behavior.</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>19. Soon after using Bully Blockers, the teacher would notice a positive change in the bullying behavior.</td>
<td>6</td>
<td>4</td>
</tr>
<tr>
<td>20. The child’s behavior will remain at an improved level even after Bully Blockers is discontinued.</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>21. Using Bully Blockers should not only improve the bully’s behavior in the classroom, but also in other settings (e.g., other classrooms, home, playground).</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>22. When comparing a bully with a nonbully peer before and after use of Bully Blockers, the bully’s and the peer’s behavior would be more alike after using Bully Blockers.</td>
<td>6</td>
<td>5</td>
</tr>
</tbody>
</table>
Table 2 (continued)

<table>
<thead>
<tr>
<th>Behavior Intervention Rating Scale questions</th>
<th>Manual + video group</th>
<th>Manual-only group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Teacher A</td>
<td>Teacher B</td>
</tr>
<tr>
<td>23. Bully Blockers should produce enough improvement in the bullying behavior so the behavior no longer is a problem in the classroom.</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>24. Other behaviors related to bullying also are likely to be improved by Bully Blockers.</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

*Note.* Ratings range on a Likert-type scale from 1 (*very low*) to 6 (*very high*).  

very high level of acceptability in the program, the manual + video group reported a mean of 5.17 ($SD = .565$) across all items. The manual-only group also reported high levels of acceptability with a mean rating of 5.21 ($SD = .896$).

After completing the Behavior Intervention Rating Scale, the teachers were asked to respond to two open-ended questions concerning the intervention. The first open-ended question asked the following: “What was the most useful part of the lessons?”

Teachers from the manual + video group generally enjoyed the activities and videos. Comments included the following:

I think the most useful part of the lessons was that we were able to do hands-on activities that kept the students’ interest level up the whole time. Then when they watched the videos, everything they had learned from the activities was reinforced.

The kids really liked the videos. It was a good opportunity for them to learn it in the lessons, practice the skills in role play, then see two guys that they thought were funny draw on principles from the lessons. I also found the lessons to be pretty time efficient, which was nice.

In the manual-only group, a teacher indicated that she enjoyed the program’s use of role play:

When the students had the opportunity to practice what they had just learned from the lesson, it really made the message hit home. I think that the examples given in the book for bullying situations were very appropriate to what our kids face at this school.

The second teacher in the manual-only group found high activity levels helpful:

The lessons were good for getting the students up and moving around. When third-graders have to sit for long periods of time, they get antsy. These lessons were active enough to where the kids never really got too bored.
Judging from the qualitative teacher responses to this question, teacher acceptability of the program was high regardless of using videos. Common to all responses within the manual + video group was that the videos enhanced the children’s learning atmosphere. Each teacher surveyed also mentioned the activities, role plays, or both associated with the lessons to be fun and beneficial for the students.

The second open-ended question asked the following: “Do you have any suggestions to improve the lessons or program?” A teacher from the manual + video group expressed the same concern raised by the meta-analysis of Merrell et al. (2008); that is, the knowledge might not generalize to actual responses. She wrote:

The lessons were very good for identifying problem behavior in regard to bullying, but it seemed like there wasn’t very much about what exactly to do when the student gets bullied. Maybe some [lessons] that are easier to grasp for younger children would be useful, something to walk them through each step one at a time.

The second teacher from the manual + video group showed continued enthusiasm for the program. She wrote: “I can’t think of anything offhand that you could do to make it better. I think you're doing great!”

From the manual-only group, a teacher suggested updating the literature list:

Your program uses a lot of children’s books to convey messages, which I like. But there have been many books that have come out in the past couple of years that would be very good to add to the list. It’s a great program though. Thank you!
The second teacher in the manual-only group wrote:

This was a great program, but I think that it would be tough to make it applicable to children older than about third grade. I don’t know what your target audience is, but it felt like the material would be too young for older elementary students.

Teachers suggested that animation specific to particularly higher grades would be more acceptable for elementary students.

Research Question 3

Will students in the manual + video condition rate the program as more acceptable than the students in the manual-only condition? This research question was addressed by the Children’s Intervention Rating Profile (Witt & Elliott, 1985). The Children’s Intervention Rating Profile, developed specifically for children, has seven items using a 6-point Likert-type scale ranging from 1 (strongly agree) to 6 (strongly disagree). Some reverse-scored items are on the scale to ensure that the children read the questions carefully, although these items could cause some confusion for children completing the survey. Mean ratings and standard deviations for classrooms in both conditions are found in Table 3.

Independent sample $t$ tests were conducted on each item to test for significant differences in student ratings of the intervention between the manual + video and manual-only groups. Of the seven questions, a statistically significant difference between difference ratings of item 6 ("I liked Bully Blockers.") was obtained. The manual + video group rated the item significantly higher than the manual-only group, $t(92) = -2.49, p = .015$. This difference could
Table 3

**Mean Ratings for Manual-Only and Manual + Video Groups**

<table>
<thead>
<tr>
<th>Children's Intervention Rating Profile questions</th>
<th>Manual + video group (n = 50)</th>
<th>Manual + only group (n = 44)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Bully Blockers is a fair intervention.</td>
<td>1.94 1.74</td>
<td>1.55 1.11</td>
</tr>
<tr>
<td>2. Bully Blockers is too harsh on bullies.</td>
<td>4.88 1.60</td>
<td>4.43 1.90</td>
</tr>
<tr>
<td>3. Bully Blockers may cause problems with the bully's friends.</td>
<td>2.90 2.03</td>
<td>3.39 1.98</td>
</tr>
<tr>
<td>4. There are better ways to handle a bully's problem than Bully Blockers.</td>
<td>3.30 2.23</td>
<td>3.45 2.04</td>
</tr>
<tr>
<td>5. Bully Blockers would be good to use with other children.</td>
<td>1.94 1.82</td>
<td>2.18 1.80</td>
</tr>
<tr>
<td>6. I liked Bully Blockers.</td>
<td>1.46* 1.15</td>
<td>2.18* 1.65</td>
</tr>
<tr>
<td>7. I think that Bully Blockers would help a bully do better in school.</td>
<td>1.94 1.50</td>
<td>2.48 1.82</td>
</tr>
</tbody>
</table>

*There was a significant difference between group means for this item.

**Note.** Items on the Children's Intervention Rating Profile (Witt & Elliott, 1985) are rated on a 6-point Likert-type scale ranging from 1 (*strongly agree*) to 6 (*strongly disagree*).
be attributed to a “likability factor” associated with the video content that enhances the learning content and environment for a student. Nevertheless, it should be noted that both groups rated the program with rather high scores of acceptability.

Research Question 4

Will differences be seen in treatment integrity between the manual + video and manual-only conditions? This research question addressed whether teachers would implement the intervention with integrity and whether there would be a difference in levels of treatment integrity between groups. To promote and assess fidelity in treatment for both groups, teachers completed a checklist taken from the Tough Kid Bully Blockers manual after completing each lesson (Bowen et al., 2008). This checklist highlights nine key tasks that the teachers needed to complete for each lesson (see Appendix C). The checklist, for example, asks whether the teacher reviewed everything taught in the lesson at the conclusion of the lesson. If the teacher completed that step, she would check that particular item on the checklist.

To determine whether the teacher ratings on the checklist were accurate, a researcher attended classrooms during the intervention periods to observe the lesson delivery during 20 out of the 92 sessions. The observer used the same checklist as used by the teacher to check off the systematic teaching steps as the teacher completed them. In addition, a second observer accompanied the first observer to lessons at random six times (three times for each group). Using a second observer made it possible to determine a reliability coefficient between
observers as well as another rater to check the teacher's reliability. Computing the number of agreements between the observer(s) and teacher and then dividing that number by the total number of items (agreements plus disagreements) on the checklist (nine) yielded a reliability coefficient of the percentage of steps on which the raters agreed. As suggested by Anastasi (1976), .80 was set as the target reliability coefficient for the raters. Since there are only nine items on the checklist, a level of .80 would allow for only one disagreement between raters, making .80 a difficult but attainable goal. Table 4 demonstrates the reliability coefficients for the observers and teachers in the study across observations.

Table 4

*Mean Reliability Coefficients for Raters*

<table>
<thead>
<tr>
<th>Raters</th>
<th>Paired observations</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$M$</td>
<td>$SD$</td>
</tr>
<tr>
<td>Observer 1 versus</td>
<td></td>
<td>.890</td>
<td>.070</td>
</tr>
<tr>
<td>Observer 2</td>
<td></td>
<td>.817</td>
<td>.090</td>
</tr>
<tr>
<td>Observer 2 versus</td>
<td></td>
<td>.798</td>
<td>.108</td>
</tr>
<tr>
<td>teacher</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unaccompanied observations (one observer)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observer 1 versus</td>
<td></td>
<td>.866</td>
<td>.107</td>
</tr>
<tr>
<td>teacher</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* These percentages were calculated by taking the number of agreements between raters (Observer 1 = ratings given by Observer 1, Observer 2 = ratings given by Observer 2, and teacher = ratings the teacher being observed gave himself or herself) and dividing them by the total number of items on the checklist (nine).
The reliability coefficient for the two observers was high (.89), as was that of the teachers and Observer 1 when unaccompanied by a second observer (.866). Reliability coefficients in all cases were above the target of .80, with the exception of Observer 2 compared with the teacher (.798).

After calculating reliability coefficients, it was necessary to determine whether having an observer present had a reactive effect on the teachers. Table 5 reports the mean number of items checked on the 9-item checklist across two conditions: (a) observed ($N = 10$) and (b) not observed ($N = 36$).

Teachers in both conditions completed more steps on average when an observer was present. Teachers in the manual + video condition completed .66 more steps on the 9-item checklist the 10 times they were observed, and the teachers in the manual-only condition completed .26 more steps when observed. With a sample size of only 2 teachers in each condition and only 10 observations for each group, it is not possible to determine a significant observer effect. It is

<table>
<thead>
<tr>
<th>Condition</th>
<th>Manual + video</th>
<th>Manual-only</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed ($N = 10$)</td>
<td>7.29</td>
<td>8.23</td>
</tr>
<tr>
<td>Not observed ($N = 36$)</td>
<td>6.63</td>
<td>7.97</td>
</tr>
</tbody>
</table>

*Note.* There were 10 observed sessions and 36 unobserved sessions for each group. Of the 10 observed sessions for each group, 3 were conducted with 2 observers.
important, however, to recognize that teachers in both conditions completed more steps on the checklist with an observer present. Data regarding the fidelity of treatment in this study should be approached with caution.
DISCUSSION

The primary goal of this study was to determine whether students who receive a bullying intervention by way of a manual plus an audiovisual component would learn and retain more information than students who receive the lesson only by manual. Treatment fidelity was also assessed to determine how closely teachers followed prescribed activities. Secondary analyses within the study were conducted to assess student and teacher acceptability.

Findings from the study demonstrated that students made gains in knowledge acquisition. Students in both groups made significant gains from pretest to posttest on the knowledge acquisition quiz. Results suggest that the Tough Kid Bully Blockers program (Bowen et al., 2008) is an effective means of teaching children regardless of whether there is an audiovisual component. Although large effect sizes and significant gains in test scores were observed from pretest to posttest within groups, there were no significant differences in pretest-posttest gain scores between groups. Knowledge acquisition was neither enhanced nor hindered by adding an audiovisual component to the bullying program. However, a finding on the Children’s Intervention Rating Profile (Witt & Elliott, 1983) indicated that children who received the intervention with an added audiovisual component liked the intervention significantly better than those who did not have the audiovisual element added to the intervention. It is possible that this “likability factor” could
have enhanced learning in the manual + video group.

Another indication that the intervention was successful was the acceptability reported by teachers and students. Results from the limited feedback obtained from teachers indicated that the Tough Kid Bully Blockers program requires little time or expense from teachers. Furthermore, the data from this study indicated that programs incorporating similar content could also be effective in improving student knowledge, another factor determining acceptability. For these reasons, the high acceptability ratings given by teachers on the Behavior Intervention Rating Scale are understandable (Elliott & Von Brock Treuting, 1991).

The high acceptability scores given by students on the Children’s Intervention Rating Profile (Witt & Elliott, 1983) from both groups, coupled with qualitative responses elicited from teachers in open-ended questions, suggest that the students in the study enjoyed the program. All 4 teachers indicated that the students enjoyed the role plays and activities associated with the lessons. The mean scores on the Children’s Intervention Rating Profile suggested that the students shared their teachers’ opinions and enthusiasm for the program.

Observers also provided anecdotal evidence that students enjoyed the program. The observers noted that when they entered the classroom, many students would pump their fists emphatically, saying, “Yeah! Bully Blockers time!” or they would celebrate less disruptively with peers by giving “high 5s” and gesturing excitedly towards one another. It was evident through anecdotal, qualitative, and even quantitative data that children evidenced high levels of acceptability of the
While acceptability ratings from teachers and students in both groups were high, there was no significant difference in acceptability ratings between groups, with the exception of one item from the Children’s Intervention Rating Profile. As noted earlier, students in the manual+video group rated item 6 of the Children’s Intervention Rating Profile (“I liked Tough Kid Bully Blockers.”) significantly higher ($M = 1.46$, $SD = 1.15$) than the manual-only group ($M = 2.18$, $SD = 1.65$). Otherwise, there were no significant acceptability rating differences between groups for teachers or students.

One possible explanation as to why teachers expressed no preferences for having an audiovisual component to the interventions they implement could be because the time, expense, and complexity of the intervention were not changed. The teachers in the manual+video condition were asked to administer all of the lesson content from the manual regardless of whether there was a corresponding video with the lesson. Perhaps adding an audiovisual element to the intervention may have required more teacher time, making the intervention more complex. Less time was required of the teachers in the manual-only condition than of the teachers in the manual+video condition.

The lack of differences between student ratings on the Children’s Intervention Rating Profile, with the exception of item 6, may also be explained by my observation that third-grade children, by nature, are enthusiastic. They often enjoy extracurricular activities and are eager to participate in new activities. They
were consistently excited when “Bully Blockers time” came and were uninhibited in showing their enthusiasm for classroom activities. That the children were so enthusiastic about the program regardless of their access to videos suggests that the program can be successful in either form. For children who are still enthusiastic in the classroom, getting instruction from a teacher and participating in activities and role plays may be enough to keep interest levels high. Perhaps significant between-group differences in the acceptability ratings would be found in older students who have lost a sense of unabashed enthusiasm in the classroom and may need age-appropriate adaptations of materials, role plays, and group activities to keep their interest levels high.

An unanticipated finding from this study concerns the disparity between groups regarding fidelity of treatment. Due to the small sample size of teachers ($N = 4$), direct conclusions about treatment integrity cannot be drawn. In this study, though, the teachers in the manual-only condition completed more steps on the teacher checklist, which was contrary to what was hypothesized. The hypothesis regarding treatment integrity was that the teachers in the manual+video condition would implement the program with higher levels of integrity than the teachers in the manual-only condition. Of the 9-item checklist that was provided as a guide to enhance fidelity and uniformity of lessons, teachers in the manual+video condition completed only 7.29 out of 9 steps when observed and 6.63 out of 9 steps when not observed. Teachers in the manual-only condition completed 8.23 out of 9 steps when observed and 7.97 out of 9 steps when not
observed. Thus, teachers in the manual-only condition completed more steps on the checklist than teachers in the manual+video condition. Even though there were high percentages of agreement between observers and teachers, suggesting that the scores reported were good indications of what was implemented, differences between groups should be interpreted with caution due to the small sample size and evidence of a potential reactive effect of observation.

Although results cannot be generalized because of these limitations, it is still interesting to look at the differences in levels of treatment integrity between the two conditions. Although students in the manual+video group were not taught as many steps on average as students in the manual-only group, they still made the same gains in their content knowledge as the manual-only group.

It is not clear why no differences were found between group knowledge acquisition scores with discrepancies in levels of treatment integrity, but there are several possible reasons. First, the manual+video group received more instruction. The videos reviewed most of the principles taught in each of the lessons. Since the videos always contain the same information, the children in the manual+video group received consistent and uniform instruction that reviewed key elements from the lessons. It is possible that the additional instruction compensated for the lower rates of reported treatment fidelity from teachers in the manual+video group.

The videos may have also provided an added element for the children that captured their attention and, thus, enhanced their learning environment. As the Children’s Intervention Rating Profile scores suggested, students in the
manual + video group reported liking the program more than the manual-only group, perhaps leading to a better grasp of lesson concepts.

Another explanation could be that the teachers in the manual + video group overlooked steps that are not critical to student knowledge acquisition. A consistently overlooked step on the checklist was the following: “Introduce new vocabulary or concept and discuss definition.” Teachers in the manual + video condition overlooked this step 21 out of 46 times, indicating that approximately 44% of the time teachers in the manual + video group did not complete this step. Perhaps introducing new vocabulary or concepts is not essential to student learning in this program. This step was the only step on the 9-item checklist that was consistently overlooked by teachers in the manual + video group.

The fact that the manual + video group made similar gains from pretest to posttest as the students in the manual-only group when manual + video teachers completed fewer steps on average could be a positive reflection on the bullying intervention. Somehow the manual + video group learned the material and made similar significant gains as the manual-only group from pretest to posttest. These gains could be a result of the added video element or teacher reinforcement of lesson principles outside of instruction, or students in the manual + video condition were able to learn enough from the lessons—even if their teachers completed fewer checklist items on average than teachers in the manual-only condition.
Limitations

This study contained several limitations. Students were nested in classrooms, and there were differences between groups in the number of checklist items completed, making uniformity of instruction an area of concern. The students did not show differences between groups in knowledge acquisition scores from pretest to posttest, but had there been more uniformity in the instruction, the outcome of the study might have been different.

Second, in order to meet the participating school district's request for confidentiality, the pretest and posttest knowledge acquisition quizzes were not paired, resulting in a weaker statistical design and making it impossible to compare difference scores between the two groups. Instead, independent sample t tests were used to compare pretest average scores between groups, finding no significant differences between the two groups. Independent sample t tests were used to compare posttest average scores between groups. Since there were no significant baseline differences between groups, a significant difference in average posttest scores between groups would have concluded that there was a treatment effect. However, no significant differences were found in posttest mean scores between groups; thus, it was concluded that there was no significant treatment effect resulting in retention of the null hypothesis. Had the pretest (baseline) mean scores between groups been significantly different from one another, this type of statistical design would not have been possible. Finding no significant differences in pretest scores between groups made it possible to use this statistical design that, though
more limiting, is still a valid method of analyzing the data.

A third limitation from this study concerns the relatively high pretest scores from both groups. The manual-only group reported an average score of 8.88 out of 13 (68%), and the manual+video group reported an average score of 8.09 out of 13 (62%). It is preferable for average scores on pretest measures to be 50% or less. Collecting respective pretest scores of between 62% and 68% created a low ceiling for this measure, making it difficult to discriminate the effect of the treatment. With the lower ceiling created by high pretest scores, it is difficult to determine that one form of treatment is better than the other.

In addition, this study included a small number of teachers reporting acceptability data. Using four classrooms provided an acceptable sample size for students but limited the sample size for teachers. Because of this limitation, I was constrained to collect qualitative data on acceptability and treatment integrity measures rather than the more valid and useful quantitative data that a larger sample size would have made possible. However, teacher acceptability rating was a secondary goal of this study and did not influence the study's primary goal.

Furthermore, this study concerns the nature of self-report data from children. The acceptability of the materials and lessons was assessed primarily by survey, assuming that children participating in the study would be able to read and understand the questions as well as the Likert-type rating scale. Teachers provided assistance to students when they completed the surveys, but some students may have been unclear about the questions or Likert-type scale.
Finally, the teachers in the study were monitored closely. The frequent monitoring and checking in with teachers may have created more incentive to implement the intervention. Without such pressure, the teachers may have implemented the intervention with less integrity. Frequent monitoring of the teachers is quite different from the naturalistic school setting in which such monitoring may not exist. Adding a video component to the program, however, creates standardization. With this standardization, there is less dependence on teachers to cover essential material from the lessons.

**Implications for Practice and Future Research**

The topic of bullying remains an important one in education research, and there is a need for effective programs in schools. Results from this study, although limited by a small sample size, demonstrate that material from the Tough Kid Bully Blockers program (Bowen et al., 2008) was generally well received and could be effective in instructing students in the elementary school setting. Kazdin (1977) said: "It is not enough for behavioral procedures to be effective. They must also be accepted by the individuals with whom they are being implemented" (p. 428). Thus, the higher level of acceptability for this program by teachers in this study is a good starting point to determine the Tough Kid Bully Blockers program as a practical option for elementary school teachers.

This study also demonstrates a need to further investigate possible benefits of audiovisual components such as whether instructional videos could be effective with other populations. Third-grade children are enthusiastic and eager to learn;
hence, adding a video element for this population may be unnecessary. However, as children get older and more conscious of their peers, their overt enthusiasm for school bully prevention activities may wane. Adding an audiovisual component to a bullying program may prove to be more critical with older children who are heavily influenced by peer behavior. Future studies could be conducted on the effectiveness of a video-supplemented bully intervention with an older population of students.

Future studies in this area may also address the effectiveness and acceptability of an all-video-formatted bullying intervention. In this study, the teachers in the manual + video group had more responsibilities than the teachers in the manual-only group. The teachers in the manual+video group taught all the lessons from the manual that the teachers in the manual-only group taught, with the additional responsibility of showing each video twice within the same time period. Therefore, the expectation that teachers in the manual + video group would rate the program as more acceptable than the teachers in the manual-only group was not entirely realistic. Teachers implementing an all-video program may rate that program as more acceptable than teachers teaching from a manual; that is, an all-video program would require less time of the teacher and would prove to be less complex—key elements of teacher acceptability (Bowen et al., 2004).

A study with a larger teacher sample size would be beneficial. A larger teacher sample size would allow the researcher(s) to use statistics in determining treatment acceptability and fidelity differences between groups. In this study, it
would be important to pair samples on the pretest-posttest design to offer more statistic sensitivity.

Finally, future research in this area could address the long-term impact of bully prevention education by using a more longitudinal study of the Tough Kid Bully Blockers program. The goal of the program is to generalize knowledge gained to behavior change. A longitudinal study determining both long-term retention of material from the lessons and behavior tracking would be useful in the area of bully intervention research. Office referrals, attitude surveys, and playground observation data could be used to assess generalization of the skills learned in the program to behavior.
APPENDIX A

UNIT AND LESSON TITLES WITH LESSON OBJECTIVES
### Unit One: Learning About Bullying

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Title</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Definition of Bullying</td>
<td>Discuss terms in definition Post definition in classroom</td>
</tr>
<tr>
<td>2</td>
<td>Using Literature To Learn About Bullying</td>
<td>Read and discuss associated story Discuss application to real-life situations</td>
</tr>
<tr>
<td>3</td>
<td>Bully Blockers Pledge</td>
<td>Discuss definition Sign pledge and post pledge in classroom</td>
</tr>
<tr>
<td>4</td>
<td>Writing About Bullying</td>
<td>Participate in writing or drawing activity relating bullying to a personal experience</td>
</tr>
<tr>
<td>5</td>
<td>Recognizing Types of Bullying</td>
<td>Play class game in which students decide whether a situation is bullying Learn to recognize the different types of bullying</td>
</tr>
<tr>
<td>6</td>
<td>Bully Poster Contest</td>
<td>Each student designs and creates a poster with a “no bullying” theme to display in the classroom</td>
</tr>
</tbody>
</table>
### Unit Two: Bully Blocker Strategies

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Title</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Reporting Bullying</td>
<td>Learn to recognize when to report bullying</td>
</tr>
<tr>
<td>2</td>
<td>Using Literature To Learn</td>
<td>Listen to a selected story involving bullying</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learn strategies to use when being bullied</td>
</tr>
<tr>
<td>3</td>
<td>How To Look/Act When Sticking Up To a Bully</td>
<td>Learn how to stay calm and be assertive when bullied</td>
</tr>
<tr>
<td>4</td>
<td>Teasing Blocks</td>
<td>Identify verbal teasing and name-calling as bullying behavior</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Learn two or more strategies to block name-calling</td>
</tr>
<tr>
<td>5</td>
<td>Rumor Blocks</td>
<td>Learn strategies to handle rumors</td>
</tr>
<tr>
<td>6</td>
<td>What To Do When Someone Else Is Bullied</td>
<td>Learn what bystanders can do to block bullying</td>
</tr>
<tr>
<td>7</td>
<td>Writing About “What I Can Do If I Am Bullied”</td>
<td>Write or draw what to do when bullied</td>
</tr>
</tbody>
</table>
## Unit Three: Friendship Builders

<table>
<thead>
<tr>
<th>Lesson</th>
<th>Title</th>
<th>Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>What Makes a Good Friend?</td>
<td>Discuss characteristics of a good friend&lt;br&gt;Identify examples of friendly and positive qualities</td>
</tr>
<tr>
<td>2</td>
<td>Give Put-Ups, Not Put-Downs</td>
<td>Learn to give put-ups&lt;br&gt;Create a positive classroom environment</td>
</tr>
<tr>
<td>3</td>
<td>Using Literature to Teach About Friendship</td>
<td>Learn and discuss qualities of a friend</td>
</tr>
<tr>
<td>4</td>
<td>Asking to Join</td>
<td>Learn friendship skills for asking to join in a group</td>
</tr>
<tr>
<td>5</td>
<td>Inviting Someone to Join</td>
<td>Learn friendship skills for inviting someone else to join in</td>
</tr>
<tr>
<td>6</td>
<td>Cooperation</td>
<td>Participate in an activity that encourages cooperation skills and teamwork</td>
</tr>
<tr>
<td>7</td>
<td>Giving Compliments</td>
<td>Learn how to give and accept compliments</td>
</tr>
<tr>
<td>8</td>
<td>Writing About Friendship</td>
<td>Write or draw personal experiences with friendship</td>
</tr>
<tr>
<td>9</td>
<td>Class Friendship Book</td>
<td>Highlight positive statements about each student in class</td>
</tr>
<tr>
<td>10</td>
<td>Name Game</td>
<td>Play a name game to encourage friendly interaction and to help students learn the names of all students in the class</td>
</tr>
</tbody>
</table>

APPENDIX B

BULLYING KNOWLEDGE QUIZ
(For collection of pretest and posttest data)

1. Bullying is
   ◦ not that big of a deal.
   ◦ repeated, intentional, and one sided.
   ◦ fighting, teasing back and forth, and playing together.
   ◦ something that happens only on the playground.

2. How can bullying happen?
   ◦ With words
   ◦ Physically
   ◦ From a group
   ◦ By leaving someone out
   ◦ All of these

3. What is “direct” bullying?
   ◦ Bullying that happens all the time
   ◦ Something that happens only once
   ◦ Teasing back and forth
   ◦ Face-to-face, usually aggressive

4. What is an example of “indirect” bullying?
   ◦ Spreading rumors
   ◦ Pushing someone
   ◦ Calling someone a bad name
   ◦ Throwing a snowball at someone

5. People who just watch someone get bullied are called
   ◦ bystanders.
   ◦ bullies.
   ◦ victims.
   ◦ tattle-tales.

6. Spreading rumors can be bullying.
   ◦ True
   ◦ False

7. It is NOT important to decide if the bullying is harmful or repeated before telling an adult.
   ◦ True
   ◦ False
8. What does it mean to be assertive?
   ◦ Turn around, run away, tell a teacher
   ◦ Hit the bully, call them names back, tell a teacher that they hurt you
   ◦ Look at the bully, stand tall, use a firm voice
   ◦ Get your big friend to scare the bully

9. How can you be a “superfriend?”
   ◦ Hit and push
   ◦ Tease others and play a lot
   ◦ Let others join in and give put-ups instead of put-downs
   ◦ Always play only what you want to play at recess

10. A good friend always treats you with respect.
    ◦ True
    ◦ False

11. It is important to make eye contact when you give someone a compliment.
    ◦ True
    ◦ False

12. Fill in the missing word: Just like physical bullying can hurt on the outside, ______ hurt on the inside.
    ◦ bullies
    ◦ put-ups
    ◦ bystanders
    ◦ put-downs

13. Before problem-solving with a person, it is important to
    ◦ remember how mad you are at them.
    ◦ take deep breaths and relax.
    ◦ tell them that you want it your way only.
    ◦ bring all of your friends so you can scare the other person.
APPENDIX C

TREATMENT INTEGRITY CHECKLIST
Please check the line when you complete the specified step:

- Explain what will be taught in the lesson.
- Review what has already been taught.
- Discuss how the skill discussed will help students.
- Introduce new vocabulary or concept and discuss definition.
- Ask students to discuss previous experiences they have had that relate to this concept or skill.
- Provide examples and elicit examples of the concept from students.
- Model or demonstrate skill for students.
- Ask students to demonstrate the skill through role play and verbal or written responses.
- Summarize concept and ask students to summarize what they have learned.

What was the most useful part of the lessons?*

__________________________
__________________________
__________________________

Do you have any suggestions to improve the lessons or program?*

__________________________
__________________________
__________________________

*Only on the checklist for the final lesson.


