Children’s Physiological and Emotional Reactions to Witnessing Bullying Predict Bystander Intervention

Lydia R. Barhight and Julie A. Hubbard
University of Delaware

Christopher T. Hyde
BioAssessments

Study goals were to explore whether children clustered into groups based on reactions to witnessing bullying and to examine whether these reactions predicted bullying intervention. Seventy-nine children (M = 10.80 years) watched bullying videos in the laboratory while their heart rate (HR) was measured, and they self-reported on negative emotion after each video. Bullying intervention was assessed by school peers. Two groups emerged based on reactions to the bullying videos: The Emotional group (43% of children) displayed HR acceleration and reported high negative emotion, whereas the Unemotional group (57% of children) showed HR deceleration and reported low negative emotion. Group membership predicted bullying intervention, with peers reporting that Emotional children were more likely to stop a bully than Unemotional children.

Bullying is defined as intentional actions repeated over time that harm, intimidate, or humiliate another person (Olweus, 1993). An aggressive act is only defined as bullying when the perpetrator is more powerful than the victim, either physically or socially (Baldry & Farrington, 2007; Olweus, 1993). Bullying is an alarmingly common problem in schools. Although prevalence estimates vary, in the most recent large-scale study in the United States, 13% of students reported being the victim of physical bullying and 37% of students reported being the victim of verbal bullying (Wang, Iannotti, & Nansel, 2009).

Children who are bullied are at increased risk for a host of negative outcomes. They are more likely than their nonbullied peers to be depressed (Hanish & Guerra, 2002) and to engage in self-harm (Barker, Arseneault, Brendgen, Fontaine, & Maughan, 2008). Bullied children are also more likely to be anxious and to experience psychosomatic complaints, such as headaches and dizziness (Natvig, Albrektzen, & Qvarnstrom, 2001). In addition, bullied children suffer academically, with poorer grades and higher rates of absenteeism than their nonbullied peers (Irving & Parker-Jenkins, 1995; Srabstein & Piazza, 2008). Of most concern, bullying can lead to tragic consequences; bullied children are at increased risk for suicidal behavior (Klomek, Marocco, Kleinman, Schonfeld, & Gould, 2008), and 71% of children who perpetrated school shootings were chronically bullied (Vossekuil, Fein, Reddy, Borum, & Modzeleski, 2002). For all these reasons, bullying has been identified as a public health risk, one that must be addressed primarily through intervention in schools.

The Importance of Bystanders in Bullying Episodes

A common myth about bullying is that it occurs covertly. In fact, Pepler and Craig (1995) found that bystanders are present in at least 85% of bullying episodes. These bystanders are usually peers, since most bullying occurs during unsupervised periods and in unsupervised areas at school (Atlas & Pepler, 1998; Olweus, 1993; Rigby, 2008). Children who witness bullying have great potential power. The reactions and behaviors of these bystander children may impact the outcome of the bullying event. Cheering the bully on, joining in the bullying, or even just ignoring the incident could encourage victimization, whereas telling an adult...
or defending the victim could end the incident and discourage the bully in the future. The Olweus Bullying Prevention Program (e.g., Olweus et al., 2007) describes how every bystander to a bullying episode is involved in some way. For example, the “follower” joins in, the “defender” actively supports the victim, and the “disengaged onlooker” is not affected by seeing the bullying.

Estimates vary concerning the percentage of bystander children who engage in each of these behaviors. Using self-report, Whitney and Smith (1993) found that about half of 8- through 11-year-old students reported that they would try to help a victim of bullying. In contrast, O’Connell, Pepler, and Craig (1999) observed actual bystander behavior during bullying episodes on the playground among first- through sixth-grade children. They found that, compared to the self-report estimates mentioned earlier, fewer children intervened (only 25%); most children ignored the incident (roughly 50%), or even joined in (25%). This discrepancy between children’s self-reports and their actual behaviors is consistent with the findings of Salmivalli, Lagerspetz, Björkqvist, Österman, and Kaukiainen (1996), which suggest that children tend to overestimate the likelihood that they would intervene. The authors compared peer- and self-report estimates of different “participant roles” in bullying episodes in a sample of sixth graders. Many children self-reported that they would intervene when they saw bullying happen; however, peer report indicated that a significantly lower number of children (only 17%) actually intervened.

Increasing bystander intervention is an important goal, because it has been shown to be effective. Pepler and Craig (1995) found that when a bystander actively expressed disapproval during a bullying incident, bullies stopped aggressing approximately 50% of the time. Together, this information highlights the need for bullying interventions that target the important role that bystander children can play in decreasing school bullying.

**Interventions: Inclusion of Components Aimed at Bystanders**

In fact, many school-based bullying interventions do include a component aimed at bystanders. This component usually involves parents and teachers emphasizing to children the necessity of intervening or seeking adult help if they see a peer being bullied. The positive outcomes of such interventions for bullied children are stressed, and role-plays are often included for children to model and practice the skills needed to intervene or tell an adult. Parents and teachers are encouraged to praise and reward children who engage in these behaviors (e.g., Olweus, 1993; Pepler, Craig, O’Connell, Atlas, & Charach, 2004; Rigby, 2008; Stevens, Van Oost, & De Bourdeaudhuij, 2004). The KiVa program, an antibullying curriculum recently evaluated in Finland, is notable for its more comprehensive focus on bystanders. This program actively works to enhance children’s empathy for victims, increase their self-efficacy about intervening in bullying episodes, and provide them with concrete support when they do try to help bullied peers (Karna et al., 2011).

Few evaluations of bullying interventions report specifically on change in bystander behavior. However, in several studies that have, the change was not dramatic. In one study, only 34% of children who indicated at preintervention that they would ignore a bullying episode reported at postintervention that they would now intervene (Evers, Prochaska, Van Marter, Johnson, & Prochaska, 2007). In another study, Pepler et al. (2004) reported no significant increases in peer report of fellow students’ bystander intervention. Even the Finnish KiVa program, which demonstrated strong effects for decreases in self- and peer-reported victimization, did not produce similar sustained effects for bystander intervention. Although children in schools receiving the KiVa intervention reported that their peers defended victims of bullying more than children in control schools at the mid-point of the intervention year, these effects were not maintained by the end of the intervention (Karna et al., 2011). Thus, despite efforts to encourage bystander children to intervene, evaluations of bullying interventions suggest that many children resist doing so.

**First Goal: Identification of Groups Differing in Reactions to Witnessing Bullying**

These findings suggest that basic research on children’s reactions to bullying is needed, as connections between these reactions and bystander intervention are not well understood. The degree to which children are upset by bullying may be a critical factor in predicting active bystander intervention. It is likely that children react in different ways when they are bystanders to bullying incidents. Some children may be strongly affected, demonstrating emotional and physiological arousal corresponding with empathy and concern for the victim. However, witnessing bullying may not particularly bother other children. These children may not have
an emotional or physiological reaction and instead may respond with indifference.

The first goal of the current project was to examine the prediction that children would cluster into identifiable “Emotional” and “Unemotional” groups based on their physiological and emotional reactions to bullying, as depicted in videos shown in a laboratory procedure. We hypothesized that an “Unemotional” group of children would not demonstrate significant physiological or emotional arousal, corresponding to the indifferent reaction described earlier. The second hypothesized group was an Emotional group, whom we predicted would be particularly affected by the bullying videos. We expected that these children would show an increase in heart rate (HR) and report feeling high levels of negative emotions, corresponding to the concerned and upset reaction described earlier.

Second Goal: Prediction of Bystander Intervention

Our second goal was to investigate the hypothesis that children’s physiological and emotional reaction to the bullying videos would predict their likelihood of intervening in bullying episodes at school. We expected that children in the Emotional group would be rated by peers in the classroom as more likely to “stop a bully” than children in the Unemotional group. Thus, our study was not designed to test whether children’s in-the-moment reactions to witnessing bullying would correspond to intervention in that same bullying episode. However, we wanted to explore associations between children’s reactions to witnessing bullying in a laboratory procedure and their tendency to intervene when bullying happens at school.

Little is understood about how children’s reactions to bullying relate to bystander behavior. Some previous work has suggested that individuals who are emotionally aroused may be less likely to engage empathically or prosocially, if their main goal is to reduce personal distress (Batson, Fultz, & Schoenrade, 1994). However, the theory of emotion utilization (Izard, Stark, Trentacosta, & Schultz, 2008) suggests that both positive and negative emotions can lead to adaptive action. For example, fear or worry may cause someone to take steps to ensure safety, or anger may give someone the strength to face a difficult confrontation. In the context of our study, the physiological and emotional arousal that emotional children feel when witnessing bullying may be associated with intervening to try to stop the bully. Arousal may serve as the catalyst that children need to have the strength and determination to confront a bullying peer.

Our goal was not only to explore whether children’s physiological and emotional reactions to bullying videos predicted the likelihood that they would intervene in bullying episodes at school, but to determine whether this prediction held over and above other factors hypothesized to influence the likelihood of intervention. Moreover, through this approach, we aimed to identify additional factors that may help determine whether children attempt to stop bullying peers.

The first additional factor that we included was the extent to which children are victimized by peers. Previous studies suggest that children who have a history of being bullied are less likely to intervene when they witness bullying than other children (e.g., Rigby, 2008; Rigby & Johnson, 2006). In this situation, these children may be more focused on their own safety than on that of their peers. Thus, we hypothesized that peer victimization would negatively predict children’s bystander intervention in bullying.

The second factor that we investigated was children’s sense of efficacy in confronting peers. In previous work, Gini, Albiero, Benelli, and Altoè (2008) demonstrated that social self-efficacy was an important predictor of bystander intervention in bullying. However, the broader construct of social self-efficacy incorporates the ability to perform a variety of social behaviors, such as making friends. In the current study, we focused specifically on the narrower construct of efficacy about confronting peers. To attempt to stop a bully, children likely need to believe that they will be successful in handling that particular and difficult social situation. This specific element of efficacy is quite distinct from a general sense of oneself as a socially skilled person. For this reason, we hypothesized that children’s perceived self-efficacy about confronting peers would be a strong positive predictor of bystander intervention in bullying incidents.

The third factor that we included was the extent to which children tend to be emotionally expressive across situations. Although some children are generally open and express emotions freely, other children are temperamentally cooler and more detached. Following from the theory of emotion utilization described earlier, we reasoned that general levels of emotional expressiveness might influence children’s likelihood of intervening to help bullied peers, with more emotionally expressive children being more inclined to get involved and attempt to help. However, we were interested in exploring whether children’s specific emotional reactions to the bullying videos versus their more
trait-like tendency to be emotionally expressive across situations would more strongly predict their intervention in bullying episodes. We hypothesized that both factors would be important but that emotional responses to the bullying videos in particular would emerge as a significant predictor over and above children's temperamental emotional expressiveness.

The fourth factor that we examined was affective empathy. Separate from the construct of emotional expressiveness, empathy refers to the extent to which children understand or experience others' emotions. Theorists have distinguished two forms of empathy, cognitive empathy and affective empathy (Jolliffe & Farrington, 2006). While cognitive empathy refers to the ability to understand how others are feeling, affective empathy represents the tendency to actually feel the same emotions as others. Previous work by Gini, Albiero, Benelli, and Altoè (2007) found that affective empathy in particular positively predicted helping a victimized classmate. The emotional pull of affective empathy may more strongly motivate children to intervene when others are being bullied, compared to the detached reasoning of cognitive empathy, and so we focused on this form of empathy in the current study. When children keenly experience the same fear or humiliation that a victim is feeling, these emotions may compel them to act to try to help the distressed peer. By including affective empathy in our model, we were able to explore whether children's specific emotional reactions to the bullying videos predicted the likelihood that they would intervene in bullying episodes beyond their more general tendency to be affectively empathic. We hypothesized that both factors would predict intervention, with emotional reactions to the bullying videos in particular making a significant contribution over and above the general tendency to be affectively empathic.

The final factors that we considered were the demographic variables of age and gender. Previous research suggests that younger children and girls are more likely to intervene when they see bullying happen than older children and boys (Rigby, 2008; Rigby & Johnson, 2006). We were interested to see whether these demographic effects were replicated in the current study.

In summary, when predicting children's likelihood of intervening in bullying at school, we included the following predictors: age, gender, peer victimization, efficacy in confronting peers, emotional expressiveness, affective empathy, and membership in the Emotional or Unemotional group (based on reactions to witnessing bullying in videos in the laboratory). Drawing from previous research and our theorizing, we hypothesized that: (a) age and peer victimization would negatively predict intervention, and (b) female gender, efficacy in confronting peers, emotional expressiveness, affective empathy, and membership in the Emotional group would positively predict intervention. Finally, we expected that membership in the Emotional group would predict bystander intervention in bullying episodes over and above all other factors.

The Importance of Groups in the Current Study

Notably, we expected that the combination of physiological and emotional reactions to the bullying videos would be important in predicting children's bystander behavior. This combination may more comprehensively and accurately capture children's reactions than any single physiological or emotional variable alone. Moreover, we hypothesized that children would cluster into distinct Emotional and Unemotional groups based on this combination of reactions to witnessing bullying, rather than demonstrating a more continuous range of responses. Thus, we predicted that membership in the Emotional or Unemotional group, based on a combination of variables assessing physiological and emotional reactions to the bullying videos, would predict children's tendency to intervene in bullying episodes at school. Other studies have used a similar approach, forming latent groups based on children's reactions to emotionally evocative situations and demonstrating that group membership predicted regulatory and behavioral constructs, even when individual reaction variables often did not (e.g., Wilson, Lengua, Tininenko, Taylor, & Trancik, 2009).

Method

Overview

Data collection took place in two phases, a classroom phase and a laboratory phase. Classroom data collection was conducted in 43 fourth- and fifth-grade classrooms. A fourth- and fifth-grade sample was chosen because bullying becomes increasingly common by this age (e.g., Olweus, 1993). During classroom data collection, we collected peer-report data on bystander behavior to bullying and self-report data on peer victimization, efficacy in confronting peers, emotional expressiveness, and affective empathy.
A subsample of 79 children subsequently participated in laboratory data collection. At this time, we assessed children’s physiological and self-reported negative emotional reactions as they watched videos depicting bullying episodes.

Participants

Classroom sample. Children were recruited for classroom data collection through parental permission letters sent home from school. To ensure a strong participation rate, three rounds of letters were distributed to children who had not returned a previous letter. Two incentives encouraged children to return their letters. First, each child who returned a letter (regardless of whether the parent granted consent) received a small prize (a colorful pencil). Second, children were told that if 80% of the letters were returned in a class (regardless of whether parents granted consent), children would receive a party with refreshments provided by our laboratory.

Eight hundred and forty-five children (78%) were given permission to participate. Classroom permission rates ranged from 50% to 100% (M = 79.2). At the beginning of classroom data collection, the experimenter began by reading an assent form aloud and giving children an opportunity to grant or decline assent. Of the children with parental permission, 2.5% (21 children) declined assent, and an additional 6.3% (53 children) were absent on the day of their classroom data collection.

Thus, our final classroom sample included 771 children (415 girls). Participants’ age ranged from 9.25 to 12.59 years (M = 10.58). Four hundred and sixty-four (60.2%) children were identified by their parent or guardian as European American, 140 (18.2%) as Latino American, 89 (11.5%) as African American, 28 (3.6%) as Asian American, 28 (3.6%) as being of mixed race/ethnicity, and 10 (1.3%) as Native American. Twelve parents or guardians (1.6%) declined to respond to the question about race/ethnicity on the parental permission form.

The 43 classrooms were located in nine elementary schools in a mid-Atlantic state. Four of the schools were parochial (N = 195, 25% of classroom sample), and five were public (N = 576, 75% of sample). Three schools (one parochial, two public) were in suburban areas (N = 323, 42% of sample), and the remainder were in urban areas (N = 448; 58% of sample).

Data on the peer victimization experienced by the classroom sample were collected using the self-report measure by Neary and Joseph (1994) described in the following sections. The mean for the classroom sample was 2.11 (SD = .76) on a 1–4 rating scale, with higher scores indicating greater peer victimization. Peer victimization did not vary by public versus parochial schools, F(1,769) = 1.02, ns. However, children in urban schools (M = 2.17, SD = .76) reported more peer victimization than children in suburban schools (M = 2.03, SD = .75), F(1,769) = 5.86, p < .05, Cohen’s d = .19.

Laboratory sample. Participants in the laboratory sample included a random subset of 80 children from the classroom sample, stratified by gender. One child’s physiological data were lost due to equipment malfunction. We dropped this participant, because assessment of physiological arousal was central to the first goal of the study (identification of groups differing in reactions to bullying videos). Thus, the final laboratory sample included 79 participants (40 girls).

Families were recruited through telephone contacts in which the purpose and procedures of the study were explained in full detail. At the beginning of the laboratory visit, the parent completed a permission form, which included demographic information such as age, race/ethnicity, and income. The child completed an assent form.

The laboratory sample did not differ significantly from the classroom sample on any demographic or assessed variable. Age ranged from 9.88 to 12.19 years (M = 10.80). Parents or guardians identified 54 children (67.5%) as European American, 11 (13.75%) as African American, 8 (10%) as being of mixed race/ethnicity, 5 (6.25%) as Latino American, and 1 (1.25%) as Asian American. One parent or guardian (1.25%) declined to report the child’s race/ethnicity. Income ranged widely, from $0 (unemployed) to $400,000 (Mdn = $70,000).

Classroom Data Collection Procedures, Questionnaires, and Data Reduction

An experimenter and approximately four undergraduate assistants conducted 1-hr visits to each of the 43 classrooms to collect self- and peer-report data. The experimenter group-administered paper-and-pencil measures to participating children. Children received a manila folder to stand upright on their desk as a “privacy shield.” Undergraduate assistants circulated throughout the room to ensure that children stayed on track, answer children’s questions, and maintain privacy. In addition, other assistants worked individually and privately with any children who required reading assistance to complete the measures validly, as determined.
beforehand through consultation with the teacher or as needed. Classrooms were compensated with $100 to be used for classroom supplies.

**Bystander intervention.** Bystander intervention in bullying was assessed through peer report during classroom data collection. Children were first provided with the following definition of bullying: “Bullying is when a kid does something on purpose to hurt or threaten another kid. It could involve hurting the kid’s body, like hitting or kicking. It also could involve hurting the kid’s feelings, like mean teasing or name calling.” Children then completed the peer nomination “When other kids are being bullied, who tries to stop the bullying?” Children nominated an unlimited number of classmates who fit this description by circling their names on a roster. The variable Stop the Bully was computed by standardizing the number of nominations received within classroom.

**Peer victimization.** Peer victimization was assessed through self report during classroom data collection. Children completed the six-item Peer Victimization Scale by Neary and Joseph (1994). For each item, participants first selected which of two statements (one indicative of peer victimization and the other not) was more like them. Next, they indicated if that statement was “really like me” or “sort of like me” (1 = really like me for phrase describing lack of victimization to 4 = really like me for phrase describing victimization).

To determine whether these data collected from this classroom sample formed a unidimensional scale, we conducted a principal components analysis (PCA) with a varimax rotation. Items were removed from the scale if the Kaiser–Meyer–Olkin Measure of Sampling Adequacy was less than .50 for that item or if the communality value for that item was less than .50. Iterations of the PCA procedure were repeated until all remaining items met these two criteria. On the final iteration, we checked to ensure that the overall Kaiser–Meyer–Olkin Measure of Sampling was at least .50, that the $p$ value for the Bartlett’s Test of Sphericity was significant, that one and only one component with an eigenvalue greater than 1.00 emerged, and that the proportion of variance accounted for by that component was at least 60%.

This PCA procedure resulted in three items being retained: “Some kids are often teased by other kids; other kids are not teased by other children,” “Some kids are often bullied by other children; other kids are not bullied by other children,” and “Some kids are often picked on by other children; other kids are not picked on by other children.” Internal consistency for this three-item scale for this sample was $\alpha = .77$. We averaged the three items to compute the variable Peer Victimization.

**Efficacy in confronting peers.** Efficacy in confronting peers was assessed through self report during classroom data collection. Children completed three items from the Social Self Efficacy Scale by Pastorelli, Caprara, and Bandura (1998; 1 = not at all to 5 = very) that best described the ability to confront peers in difficult situations. The scale was published by Pastorelli et al. (1998) in Italian; for the present study, the English translation published in Pastorelli et al. (2001) was used. After performing the PCA procedure described earlier, two items were retained: “Stand up for yourself when you feel you are being treated unfairly” and “Deal with situations where others are annoying you or hurting your feelings.” The correlation between these two items for this sample was .62, $p < .0001$. We averaged the two items to create the variable Efficacy in Confronting Peers.

**Emotional expressiveness.** Emotional expressiveness was assessed through self report during classroom data collection. Children completed the five-item Unemotional subscale from the Inventory of Callous-Unemotional Traits Youth Version (Kimonis et al., 2008; 1 = not at all true to 4 = completely true). After performing the PCA procedure described earlier, two items were retained: “It is easy for others to tell how I am feeling” and “I am emotional and show my feelings.” The correlation between these two items for this sample was .53, $p < .0001$. We averaged the two items to create the variable Emotional Expressiveness.

**Affective empathy.** Affective empathy was assessed through self report during classroom data collection. Children completed the 11-item affective empathy subscale of Jolliffe and Farrington’s (2006) Basic Empathy Scale (1 = strongly disagree to 5 = strongly agree). As the measure was developed in England, we made a small number of minor changes in item wording to reflect vocabulary that would be more familiar to our American sample. After performing the PCA procedure described earlier, five items were retained: “I get frightened when I watch characters in a good scary movie,” “Other people’s feelings do not bother me at all” (reverse scored), “I often become sad when watching sad things on TV or in movies,” “I tend to feel scared when I am with friends who are afraid,” and “My friend’s unhappiness does not make me feel anything” (reverse scored). Internal consistency for this five-item scale for this sample was $\alpha = .57$. 

We reverse scored items as necessary and averaged the five items to compute the variable Affective Empathy.

**Laboratory Data Collection Procedures and Measures**

An experimenter and an undergraduate assistant conducted a 1-hr laboratory visit for each parent and child. Parents and children were each compensated with $20. After the parent consent form and child assent form were completed, the experimenter placed the physiological equipment on the participant, with the parent observing. Next, the experimenter and child played games or completed an arts-and-crafts project, while the child habituated to wearing the equipment.

The experimenter then asked the participant to remain seated, still, quiet, and with his or her attention on a computer monitor while he or she watched a series of eight focal videos. The first and last videos depicted neutral interactions between children, whereas the other six videos depicted episodes of bullying. Videos were quite brief, lasting between 10 and 21 s. These videos were excerpts from an educational film and a commercial film on bullying designed for this age group (Adelson, Braunstein, Jaffe, & McLoughlin, 2005; Faull, 2007).

Bullying videos were selected to represent different types of bullying, with three videos depicting physical bullying and the other three showing verbal bullying. In each video, the bully(ies) and victim were of the same gender; three videos were of boys and the other three were of girls. Videos also depicted bullies and victims of different races or ethnicities. Examples of bullying videos include one in which two boys push a third boy in a locker room and another in which one girl verbally insults another girl in the cafeteria while her friends watch and laugh.

After watching each of the eight focal videos and answering questions (described in the following sections), the child saw a 20-s nature video depicting a flowing creek before the next focal video was presented. This nature video has been empirically demonstrated to be relaxing and soothing (Ulrich et al., 1991). Its purpose was to allow the child's physiological arousal to return to baseline before the introduction of the next focal video.

After giving instructions, the experimenter left the room while the child watched the videos and answered the questions. The experimenter and parent monitored the child for signs of distress via video camera from an adjacent room; none of the participants became overtly distressed. Children were informed during the child assent process that they were free to terminate participation at any time; however, no child did so. Thus, all children completed the entire session.

**Self report of emotion following videos.** After each focal video, the child was asked to rate on a 5-point scale how much he or she felt each of three negative emotions (sad, scared, and mad). Questions appeared on the computer monitor, and the child responded using a keypad. This approach yielded three variables indexing negative emotional reactions to the bullying videos (Sad Bullying, Scared Bullying, and Mad Bullying). These variables resulted from averaging children's self reports across the six bullying videos. Internal consistency was .90 for Sad Bullying, .84 for Scared Bullying, and .90 for Mad Bullying. Similar scores were calculated for children's self reports of emotion following the two neutral videos (Sad Neutral, Scared Neutral, and Mad Neutral).

**Physiological reactions to videos.** Physiological responding was measured through HR or electrocardiogram (ECG). The ECG was recorded from three Ag–AgCl disposable electrodes, with two active electrodes on the ribs and one reference electrode on the collarbone. Wires from these electrodes were connected to a computer in the adjacent room through a small opening in the wall, and this computer was synchronized with the computer on which the participant viewed videos. LabVIEW software (National Instruments, Austin, TX) collected, synchronized, and time-stamped physiological signals with an average accuracy of 5 ms or less.

Software designed by BioAssessments was used to clean and process HR data. Processing began with the identification of artifacts. IBI (interbeat interval, a measure of HR describing the time in milliseconds between consecutive R-waves in the ECG) artifacts were defined as excessively long or short intervals relative to a moving 30-s average. Short IBIs were combined and long IBIs were segmented as appropriate.

If more than 10% of HR data for any of the 15 videos (6 bullying videos, 2 neutral videos, 7 nature videos) for any child required more than minimal editing, we made the conservative decision to exclude data from the entire video. Using this rule, we excluded HR data for 2.7% of videos (32 of 1,185 videos). For the remaining videos, approximately 70% required no IBI editing, and the other 30% required three or fewer edits.

We generated average IBI scores for each child for each 2-s interval of each bullying video. Next, we calculated the slope of each child’s IBIs over each bullying video. Then, we averaged across the
six bullying videos to compute a final slope score for each child. Averaging across videos resulted in a more robust index of participants’ physiological reaction to the videos, because HR is sensitive not only to the bullying content of the videos, but also to factors such as movement, noise, and lighting that differed across videos. Internal consistency across the six videos was .50. Although this estimate of internal consistency is somewhat low, we considered it acceptable, given the sensitivity of HR to extraneous qualities of the videos.

In this initial format, positive slopes indicated that IBIs increased in length over the duration of the video, or that HR decreased. In contrast, negative slopes indicated that IBIs decreased, or that HR increased. For ease of interpretation, we calculated the inverse of IBI slope by multiplying the values by –1. This yielded a final HR Slope Bullying score for each child, in which positive values indicate increasing HR over the videos. Similar scores were calculated for HR slope during the two videos depicting neutral and during the seven nature videos (HR Slope Neutral, HR Slope Nature).

Results

Preliminary Analyses

The steps described earlier yielded the final 14 variables listed in Table 1. This table includes descriptive statistics (mean, standard deviation, range, and skewness) for each variable. For all variables, higher scores represent increased levels of the construct of interest.

We identified skewed variables using a cutoff of ±0.5 (Glass & Hopkins, 1996). We corrected them by performing log, square root, and inverse transformations; we reflected negatively skewed variables prior to transformation. For each construct, the transformed variable that reduced skewness the most was retained for all subsequent analyses.

Then, we examined bivariate correlations among nine of the final variables (see columns 2–10 of Table 2). Variables representing children’s reactions to the neutral or nature videos were not included (HR Slope Neutral, HR Slope Nature, Mad Neutral, Sad Neutral, and Scared Neutral). These variables were not of primary interest but were used simply as a baseline or manipulation check. Within variables representing children’s reactions to the bullying videos: (a) the correlations between Mad Bullying, Sad Bullying, and Scared Bullying were all positive and significant, and (b) HR Slope Bullying and Sad Bullying were positively correlated. Between the reaction variables and child characteristics, Scared Bullying was negatively correlated with Efficacy in Confronting Peers and positively correlated with Emotional Expressiveness. No significant correlations emerged among the four child characteristics or between the child characteristics and Stop the Bully.

Table 2 also includes correlations between the demographic variables of age and income and each of these final nine variables (columns 10–11). The only significant correlation was a positive relation between age and Stop the Bully.

Next, we examined gender differences and race/ethnicity differences in these final nine variables. In terms of gender: (a) girls (M = 2.94, SD = 1.07) reported higher levels of Sad Bullying than boys (M = 2.40, SD = 0.89), F(1, 77) = 6.08, p < .05, Cohen’s d = .55; (b) girls (M = 0.28, SD = 0.18) reported higher levels of Scared Bullying than boys (M = 0.19, SD = 0.16), F(1, 77) = 4.57, p < .05, Cohen’s d = .47; and (c) girls (M = 1.62, SD = .82) reported higher levels of Emotional Expressiveness than boys (M = 1.06, SD = .74), F(1, 77) = 9.88, p < .01, Cohen’s d = .72. For race/ethnicity, there was a significant effect for Peer Victimization, F(3, 73) = 4.31, p < .01. Post hoc Bonferroni comparisons indicated that African American chil-
dren \((M = 2.70, SD = .98)\) self-reported more Peer Victimization than Latino American children \((M = 1.53, SD = .38)\), \(p < .05\), Cohen’s \(d = 1.57\), or than children of mixed race/ethnicity \((M = 1.58, SD = .43)\), \(p < .05\), Cohen’s \(d = 1.48\).

In addition, we examined children’s overall reactivity to the bullying videos. For physiological reactivity, we compared HR slopes during the bullying videos to HR slopes during the neutral or nature videos. HR Slope Bullying differed significantly from HR Slope Neutral, \(F(1, 77) = 11.94, p < .001\), Cohen’s \(d = .45\), and HR Slope Nature, \(F(1, 77) = 15.41, p < .0001\), Cohen’s \(d = .61\). As can be seen in Table 1, during the bullying videos, the average HR slope across all children was positive, suggesting physiological reactivity. However, during the neutral and nature videos, the average HR slope was negative. Basic psychophysiological research has repeatedly shown that individuals exhibit decreasing HR when orienting to a novel stimulus (e.g., Öhman, Hamm, & Hugdahl, 2000). Thus, these data suggest that children were attending to the neutral and nature videos, but did not exhibit physiological signs of emotional arousal.

For emotional reactivity, we compared children’s self report of emotion following the bullying videos (Mad Bullying, Sad Bullying, and Scared Bullying) with their self report of emotion following the neutral videos (Mad Neutral, Sad Neutral, and Scared Neutral). Children reported feeling angrier, \(\chi^2(1) = 70.05, p < .0001\); sadder, \(\chi^2(1) = 55.35, p < .0001\); and more scared, \(\chi^2(1) = 56.53, p < .0001\), following the bullying videos than the neutral videos. The nonparametric test Friedman’s two-way analysis of variance (ANOVA) by ranks was used for these analyses, because the variables were skewed and we could not assume normality or equality of variance.

Finally, within the larger classroom sample, we explored the percentage of children who were consistently nominated by their peers as someone who tries to stop bullying. We found that many children were nominated by at least a few peers, a pattern that is common when using peer nominations for positive behaviors. However, a much smaller percentage of children were nominated by a majority of their peers. In fact, only 20% of children were nominated by more than 50% of their peers as someone who intervenes in bullying episodes.

First Goal: Categorizing Children Based on Reactions to Witnessing Bullying

We used latent profile analysis (LPA) to examine whether children in our sample clustered into groups based upon their physiological and emotional reactions to the bullying videos. Analyses were conducted using Mplus 5.1 (Muthén & Muthén, 1998–2007). LPA is a person-centered approach that groups individuals into latent categories, or classes, on the basis of their scores on two or more observed continuous variables (McCutcheon, 1987; Walrath et al., 2004). In our LPA, we included observed continuous variables to index physiological reactions (HR Slope Bullying) and emotional reactions (Sad Bullying, Scared Bullying, and Mad Bullying) to the bullying videos.

In LPA, the optimal number of groups is determined through fit statistics and tests of significance, whereas this determination is more subjective in traditional cluster analysis (McLachlan & Peel, 2000; Nylund, Asparouhov, & Muthén, 2007;
Vermunt & Magidson, 2002). We used a combination of fit statistics to determine the number of groups that best fit the data. The Bayesian information criterion (BIC; Kass & Wasserman, 1995) was used to estimate model fit; lower numbers represent better fitting models. The Vuong–Lo–Mendell–Rubin likelihood ratio test (VLMR) and the adjusted Lo–Mendell–Rubin likelihood ratio test (adjusted LMR) were used to compare models; for these tests, significant p values suggest that the estimated model fits the data better than a model with one fewer group (e.g., Lo, Mendell, & Rubin, 2001; McLachlan & Peel, 2000; Vuong, 1989). Finally, the normalized entropy criterion (NEC), proposed by Celeux and Soromenho (1996), was used to indicate how well the model classified individuals into groups. NEC values range from 0 to 1, with values closer to 1 suggesting better classification of individuals to groups and less overlap between groups.

The data were fit to models with four indicators (HR Slope Bullying, Mad Bullying, Sad Bullying, and Scared Bullying) and increasing numbers of groups. An a priori decision was made to continue testing models with additional groups until none of the fit statistics suggested better fit of the model in question (N groups) compared to a model with N-1 groups. In addition, we decided in advance that should there be discrepancies among the fit statistics, we would choose the model supported by the largest number of fit statistics, while also considering parsimony. Using these rules, models with one to five groups were fit to the data.

The two-group model emerged as the one that fit best, with three of the four fit statistics supporting this model (see Table 3). The VLMR and adjusted LMR were only significant for the two-group model. In addition, entropy was higher for the two-group model than for any other model. Although the remaining fit statistic (BIC) supported the three-group model, the BIC value decreased negligibly by 1.23 from the two-group to the three-group model, suggesting that the three-group model was only very slightly better than the two-group model. A final consideration in choosing the two-group model over the three-group model was parsimony; models with fewer groups are more parsimonious.

Within Mplus, the two latent Groups were compared on the four indicators (see Table 4). Significant differences between groups emerged for all four indicators. Group 1 (57% of children) displayed a decreasing HR and reported low levels of all three negative emotions in response to the bullying videos. In contrast, Group 2 (43% of children) displayed an increasing HR and reported significantly higher levels of anger, fear, and sadness. Based on these data, we labeled Group 1 the Unemotional group and Group 2 the Emotional group.

For the Emotional group, each of the four indicators significantly predicted group membership (HR Slope Bullying: estimate and SE = 2.62, p < .01; Mad Bullying: estimate and SE = 22.78, p < .0001; Sad Bullying: estimate and SE = 30.94, p < .0001; Scared Bullying: estimate and SE = 20.78, p < .0001). For the Unemotional group, three of the four indicators predicted group membership (Mad Bullying: estimate and SE = 15.48, p < .0001; Sad Bullying: estimate and SE = 21.58, p < .0001; Scared Bullying: estimate and SE = 6.51, p < .0001).

The Mplus output includes scores for the conditional probability that each child is a member of each group. Children were assigned to the group for which they had the highest conditional probability. These scores were quite high, with an average highest conditional probability score of .98. Fifty-nine children had a highest conditional probability score of 1.00, 14 children had a score of .95–.99, 2 children had a score of .90–.94, 2 children had a score of .80–.89, 1 child had a score of .70–.79, and 1 child had a score below .70.

Analyses were run to examine whether the Emotional and Unemotional groups differed in terms of
gender, race/ethnicity, age, or income. The groups differed marginally by gender, \( \chi^2(1) = 3.67, p < .06 \), but not by race/ethnicity, \( \chi^2(3) = .95, ns \); age, \( F(1, 77) = 1.43, ns \); or income, \( F(1, 77) = .00, ns \). The Emotional group included more girls (62%) than boys (38%), whereas the Unemotional group included more boys (60%) than girls (40%).

We also examined whether the Emotional and Unemotional groups differed on bystander intervention or the four child characteristics. Two significant effects emerged. The Emotional group (\( M = 2.38, SD = .93 \)) self-reported more Peer Victimization than the Unemotional group (\( M = 1.99, SD = .75 \)), \( F(1, 77) = 4.24, p < .05 \), Cohen’s \( d = .46 \). In addition, and not surprisingly, the Emotional group (\( M = 1.55, SD = .70 \)) self-reported more Emotional Expressiveness than the Unemotional group (\( M = 1.18, SD = .88 \)), \( F(1, 77) = 3.93, p < .05 \), Cohen’s \( d = .47 \). A slight marginal trend also emerged for bystander intervention in bullying episodes. Peers reported that the Emotional group (\( M = .09, SD = .91 \)) was more likely to Stop the Bully than the Unemotional group (\( M = .23, SD = .98 \)), \( F(1, 77) = 2.21, p = .14 \), Cohen’s \( d = .34 \).

Since one of the fit statistics (BIC) supported a three-group model, we did explore the three groups that emerged from this LPA analysis. They included a group with decreasing HR and low reported negative emotion, a group with increasing HR and high reported negative emotion, and a group with even more quickly increasing HR and even more reported negative emotion. Thus, the three-group model was essentially a more detailed version of the more strongly supported two-group model discussed earlier.

Second Goal: Predicting Bystander Intervention in Bullying Episodes

We conducted a linear regression to predict children’s bystander intervention in bullying episodes at school. Stop the Bully served as the dependent variable. Predictors included: (a) the demographic variables of gender and age; (b) the child characteristics of Peer Victimization, Efficacy in Confronting Peers, Emotional Expressiveness, and Affective Empathy; and (c) group membership (Emotional or Unemotional). In combination, these variables significantly predicted bystander intervention in bullying, \( F(7, 69) = 3.46, p < .01 \), \( R^2 = .26 \). Age was a significant predictor, with older children being more likely to intervene than younger children. Peer Victimization and Emotional Expressiveness were marginal predictors, with children who reported higher levels of Peer Victimization and lower levels of Emotional Expressiveness being less likely to try to stop bullying. Most importantly, group membership significantly predicted bystander intervention, after taking other predictors into account, with children being more likely to intervene if they were in the Emotional group than the Unemotional group (see Table 5).

### Discussion

The goal of the current study was to investigate children’s physiological and emotional reactions to bullying videos and to relate these reactions to children’s tendency to intervene in bullying episodes at school. Participants watched bullying videos in the laboratory while we assessed their HR, and they self-reported levels of negative emotions after each video. To determine if children clustered into groups based upon these reactions, four variables (the physiological variable of HR slope and three self-reported emotion variables for anger, sadness, and fear) were examined using LPA. Two groups of children emerged based upon their reactions to the bullying videos, an Emotional group and an Unemotional group. The two groups can be summarized as follows: The Emotional group (43% of children) displayed HR acceleration, whereas the Unemotional group (57% of children) showed HR deceleration. The Emotional group also reported higher levels of fear, sadness, and anger to the bullying videos than the Unemotional group.

Membership in the Emotional or Unemotional group predicted peer ratings of children’s likelihood of intervening in bullying episodes at school. This prediction held even when other demographic factors (gender, age) and child characteristics (peer

<table>
<thead>
<tr>
<th>Variable</th>
<th>B</th>
<th>SE B</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>-.08</td>
<td>.22</td>
<td>-.04</td>
</tr>
<tr>
<td>Age</td>
<td>.53</td>
<td>.14</td>
<td>.41</td>
</tr>
<tr>
<td>Peer Victimization</td>
<td>-.21</td>
<td>.12</td>
<td>-.18</td>
</tr>
<tr>
<td>Efficacy in Confronting Peers</td>
<td>.11</td>
<td>.10</td>
<td>.12</td>
</tr>
<tr>
<td>Emotional Expressiveness</td>
<td>.25</td>
<td>.13</td>
<td>.21</td>
</tr>
<tr>
<td>Affective Empathy</td>
<td>-.09</td>
<td>.15</td>
<td>-.07</td>
</tr>
<tr>
<td>Group Membership</td>
<td>.43</td>
<td>.22</td>
<td>.22</td>
</tr>
</tbody>
</table>

Note. \( R^2 = .26 \).

† \( p < .10 \), * \( p < .05 \), *** \( p < .0001 \).
victimization, efficacy in confronting peers, emotional expressiveness, and affective empathy) were considered. It should be remembered that physiological and emotional reactions to bullying and the tendency to intervene in bullying episodes were assessed in the different contexts of laboratory and school. Thus, we cannot conclude from this study that children's physiological or emotional reaction in the moment of a bullying episode would translate to action in that same episode. At the same time, it is important to note that children who were more physiologically and emotionally reactive to bullying in the laboratory were more likely to intervene in bullying at school, even though the laboratory procedure only involved watching short videos of unfamiliar children being bullied. However, a notable limitation of this finding is that while group membership emerged as a significant predictor of bystander intervention in the context of the larger regression model, the effect was only a slight marginal trend when group membership alone was used to predict intervention behavior in an ANOVA analysis.

These findings suggest that children in the Emotional group were “upset” by seeing bullying. As described in the Introduction, previous work has suggested that individuals who are emotionally aroused may be less likely to respond prosocially, if their main goal is to reduce personal distress (Batson et al., 1994). However, in our study, emotional upset in response to the bullying videos was associated with intervening when bullying happens at school. These findings lend support to the theory of emotion utilization (e.g., Izard et al., 2008), which suggests that both positive and negative emotions can lead to adaptive action. Arousal may be a critical factor in predicting which children will have the motivation and determination to confront a bullying peer.

Importantly, though, more children in our sample clustered into the Unemotional group than the Emotional group (although both groups were sizable at 57% and 43% of children, respectively). The size of the Unemotional group suggests that lack of a physiological or emotional response to the bullying videos was common and not indicative of pathological levels of callousness or insensitivity. Rather, several alternative interpretations are possible. First, Unemotional children exhibited HR deceleration. Basic psychophysiological research suggests that HR decreases when individuals orient to a new stimulus (e.g., Öhman et al., 2000). Thus, children in the Unemotional group may have simply been attending to the bullying videos, but not particularly physiologically or emotionally aroused. Second, the videos may not have been sufficiently evocative for many children. More children may become physiologically or emotionally aroused when witnessing real-life bullying episodes involving familiar peers. Third, these children may witness bullying events so frequently that they have become desensitized to them or have found it adaptive to keep their emotional response to a minimum.

These possible explanations notwithstanding, the children who did not react physiologically and emotionally to the bullying videos were in fact rated by their peers as less likely to intervene when a classmate is bullied than those children who reacted more strongly. Membership in the Emotional and Unemotional groups, however, predicted only a portion of the variance in bullying intervention behavior, even in combination with other demographic factors and child characteristics. Additional factors that may be important to examine in the future include fear of the bully, dislike or blame of the victim, and diffusion of responsibility.

Despite these limitations, the current study explored the association between children’s reactions to bullying and bystander intervention in a methodologically sophisticated study. When children are simply asked how they feel about bullying and whether they would intervene, they tend to provide socially desirable answers. In contrast, in the current study, children’s reactions to bullying were assessed in the laboratory through both physiological and self-reported emotional measures, and their tendency to intervene in bullying episodes in the classroom were indexed through peer report. This approach gave us a more nuanced and robust understanding of the association between these constructs.

Our findings suggest that it was important to combine both physiological and emotional variables in the assessment of children’s reactions to bullying. Furthermore, our results indicate that children’s reactions were better represented by a latent variable that grouped children into Emotional and Unemotional groups than by a continuous range of responses. Specifically, although the group membership predicted bystander intervention in the context of the regression, bivariate correlations between the four individual reaction variables used to form the latent groups and bystander intervention were nonsignificant. This finding, combined with the clarity and consistency of the results of the LPA, suggests that the combination of children’s physiological and emotions to the bullying videos divided them fairly cleanly into two groups.
If replicated in future studies, these findings may help guide interventionists as they work to improve the bystander component of bullying intervention programs. Our results suggest that one target of intervention should be bystanders’ reactions to bullying. A substantial group of children in the sample were not particularly aroused by witnessing bullying, and this same group was considered by peers to be less likely to intervene in bullying episodes. Thus, intervention programs may need to increase children’s awareness of the seriousness of bullying and help children to take the perspective of and recognize the suffering of victims. If children become more engaged, concerned, and upset by bullying episodes, they may be more likely to intervene.

Beyond group membership, two other predictors of bystander intervention trended in the hypothesized direction (peer victimization, emotional expressiveness). However, both effects were only marginal, and they only emerged in the context of the regression model and not in bivariate correlations; thus, they should not be overinterpreted. Furthermore, findings for gender, efficacy in confronting peers, and affective empathy as predictors of bystander intervention failed to emerge as hypothesized, either within the regression model or at the level of bivariate correlations or ANOVA analyses. With a sample size of only 79, low power may explain these null findings. In addition, discrepancies from previous studies in which significant effects for these relations emerged may be explained by differences the age and country of origin of the sample studied, the source and measure of bystander behavior used, or the way in which the child characteristic was assessed. For example, in a previous investigation in which peer victimization negatively predicted intervention in bullying (Rigby & Johnson, 2006), bystander behavior was assessed through self report rather than the peer report used in the current study. Another example concerns the two studies by Gini et al. (2007, 2008), in which affective empathy and social self-efficacy predicted bystander intervention; these studies used somewhat older Italian samples, a different peer-report measure of bystander intervention, and broader self-report measures of social self-efficacy or affective empathy. In any case, more work is clearly needed to gain a fuller understanding of the demographic factors or child characteristics that drive children’s behavior when they are bystanders to bullying episodes. This work is an important goal for future research; such knowledge will be critical as we work to improve the bystander component of bullying prevention programs.

Finally, one demographic variable emerged as a strong predictor of bystander intervention in bullying episodes. However, in contrast to previous work, age was a strong positive predictor of bystander intervention, both at the level of bivariate correlation and in the context of the regression model. This finding runs counter to a previous study suggesting that younger children are more likely to intervene (Rigby & Johnson, 2006). Importantly, in the current study, we assessed bystander intervention through peer report, whereas the previous study used self report. Perhaps younger children are more likely to report that they would try to stop a bully, but older children are actually more likely to do so, at least in the eyes of their peers.

An additional issue addressed in the current project concerns the percentage of children who tend to intervene in bullying episodes. As discussed earlier, although about half of students report that they would try to help a victim of bullying (Whitney & Smith, 1993), observational (O’Connell et al., 1999) and peer-report (Salmivalli et al., 1996) data suggest that much smaller percentages of children actually do so (25% and 17%, respectively). The classroom-based data set (N = 771) used here provides another source of peer-report data on the percentage of children who try to stop bullies. These data suggest that only 20% of children were nominated by at least half of their classmates as intervening in bullying episodes, a percentage that is in line with the previous peer-report and observational data. Given the tendency for most children in a classroom to receive at least a few peer nominations for positive behaviors, the use of this conservative criteria (nomination by at least half of peers) seems appropriate as a means of identifying those children who consistently work to help others who are the victims of bullying.

Limitations and Future Directions

The current study has a number of limitations and suggests several directions for future research. First, in future studies, researchers should use multiple measures of physiological arousal to more completely assess children’s reactions to bullying. Cardiac response can be difficult to interpret, as it is affected by both the sympathetic and parasympathetic branches of the autonomic nervous system. Pairing electrodermal measurement with cardiac measurement often helps in this interpretation, since skin conductance is accepted as a purely sympathetic autonomic response (Dawson, Schell, & Filion, 2000).
Second, observational coding of children’s facial expressions during the bullying videos would have allowed for more comprehensive assessment of participants’ emotional response. Recent articles (e.g., Cole, Martin, & Dennis, 2004) have called for multimethod assessment of emotional arousal (e.g., physiological, self-report, and observational) as these components do not typically correlate highly. Adding observation to future studies will allow for more thorough examination of children’s emotional reactions to bullying.

Third, future researchers may wish to provide children with a more expansive definition of bullying when they complete the peer nominations of bystander intervention. The definition used in the current study was limited to physical and verbal bullying and did not specifically mention relational bullying. More children may have been nominated as intervening in bullying episodes if relational bullying were explicitly included in the definition.

Fourth, it will be important for future researchers to explore connections between the constructs that were the focus of the current study (reactions to bullying, bystander intervention behavior) and children’s own bullying behavior and victimization experiences. Such work will greatly enhance our understanding of the complex interplay of these constructs and may prove crucial to the improvement of bullying prevention programs.

Fifth, in future studies, researchers should explore associations between reactions to bullying and bystander intervention in different developmental periods. The findings of the current study cannot be generalized beyond the fourth- and fifth-grade age group. Changes across development in children’s perspective-taking abilities, social skills, and the nature of bullying may make our findings particularly age specific.

Finally, we would call for greater collaboration among developmental and social psychologists in the investigation of bystander behavior in children. The adult social psychology literature has a rich tradition of studying bystander intervention (e.g., Latané & Darley, 1970; Latané & Nida, 1981; van den Bos, Muller, & van Bussel, 2009). Unfortunately, there are few data on children’s bystander behavior, even beyond the specific context of bullying. Basic research in this area will increase our understanding of the various cognitive, emotional, and physiological factors that contribute to children’s bystander behavior. With greater understanding of why children do or do not intervene in bullying situations, we will be able to do more to prevent the serious problem that bullying currently represents in our schools.

References


Barhight, Hubbard, and Hyde