OUTSIDERS’ PERSPECTIVES IN RESPONSE TO DIFFERENT TYPES OF TRAUMA

Male and Female Observers’ Evaluations of a Bullying Case as a Function of Degree of Harm, Type of Bullying, and Academic Level

RACHEL H. GENTRY and KERRI L. PICKEL
Department of Psychological Science, Ball State University, Muncie, Indiana, USA

Acting as disciplinary committee members, participants listened to a school bullying case that varied in terms of type (relational or verbal), degree of harm (low or high), and academic level of the victim and defendant (high school or university). Participants’ judgments (e.g., verdict, recommended sentence, seriousness, perceptions of both students) generally favored the victim when he experienced more rather than less harm, regardless of bullying type, and when the incident took place in a high school rather than a university. Additionally, women’s judgments supported the victim more than men’s. We propose that previous results suggesting that observers downgrade relational bullying occurred because no harm was specified. Moreover, we contend that observers relied on a “bullying schema” that includes the component that bullying occurs in primary and secondary schools, which led them to make less punitive judgments in the university case.

KEYWORDS aggression, bullying, peer victimization, school violence

Received 12 July 2013; revised 26 November 2013; accepted 26 November 2013.
Address correspondence to Rachel H. Gentry, Department of Psychological Science, Ball State University, Muncie, IN 47506. E-mail: rhgentry@bsu.edu
School bullying has recently received considerable attention both as the subject of media reports and as the focus of empirical investigations. Estimates of prevalence rates differ depending on the methodology and sample used, but data suggest that 24% to 45% of U.S. school children are bullied during the course of a year, and up to 20% are victimized several times per week (Swearer & Cary, 2003). Nansel et al. (2001) found that almost 30% of students in Grades 6 through 10 reported moderate or frequent involvement in bullying as victims, perpetrators, or both.

Whatever the exact rates, experts agree that bullying is a significant concern that is associated with increased problems for students who are involved in any way (Limber & Small, 2003; Swearer & Cary, 2003). In the United States and elsewhere, victims are more likely than other students to experience physical health issues, anxiety, depression, and suicidal ideation (Baldry, 2004; Bond, Carlin, Thomas, Rubin, & Patton, 2001; Gini & Pozzoli, 2009; Kaltiala-Heino, Rimpelae, Martunnen, Rimpelae, & Rantanen, 1999; Kim, Koh, & Leventhal, 2005; Klomek, Marrocco, Kleinman, Schonfeld, & Gould, 2007; Rigby, 2003). Moreover, their grades decline, and they might avoid going to school to escape victimization or drop out of school altogether (Cook, Williams, Guerra, Kim, & Sadek, 2010; Juvonen, Wang, & Espinoza, 2011; Rigby, 2003). Bullies are also at risk. Compared to other students, they more often suffer from psychological problems and substance abuse (Cook et al., 2010). Additionally, they are more likely to carry a weapon and to be convicted of a crime during adolescence or adulthood (Farrington & Ttofi, 2011; Nansel, Overpeck, Haynie, Ruan, & Scheidt, 2003; Olweus, 2011; Rigby, 2003). Several researchers have noted that school shootings covered in the media are frequently linked to bullying (Chapell et al., 2006; Hazler, Miller, Carney, & Green, 2001; Jordan & Austin, 2012).

Bullying is considered so serious that almost all U.S. states have enacted laws to address it, usually by requiring school districts to develop antibullying policies (American Educational Research Association, 2013; Brank, Hoetger, & Hazen, 2012; Limber & Small, 2003). Some statutes define bullying, whereas others leave it to individual districts to create a definition, but all either require or strongly recommend that schools establish procedures for reporting and documenting incidents of bullying, disciplining perpetrators, and implementing programs to prevent (or at least reduce) bullying in the first place. Arguably, the most critical element of any antibullying policy is the definition of the offense; inappropriate conceptualizations could allow genuine cases of bullying to go unreported, unpunished, and unchecked. The definitions currently used by legislators and school districts vary greatly, and unfortunately most conflict with those preferred by researchers (Brank et al., 2012).

Within the scholarly literature, bullying has been traditionally characterized as aggressive behavior that includes three key elements: It is intended
to cause harm, it arises within relationships in which there is an imbalance of power or physical strength between perpetrators and victims, and it occurs repeatedly over time (Brank et al., 2012; Limber & Small, 2003; Olweus, 1995). Researchers recognize multiple types of bullying (Bauman & Del Rio, 2006; Brank et al., 2012; Cook et al., 2010; Crick, Ostrov, & Werner, 2006; Jordan & Austin, 2012; Rigby, 2003; Yoon & Kerber, 2003). Physical bullying refers to any bodily assault on the victim, such as hitting, kicking, or pushing. Verbal bullying involves statements made directly to the victim by the perpetrator, for example, name calling, threats, abusive language, humiliation, and mockery. In contrast to these direct types, relational bullying is indirect, consisting of attempts to damage the victim’s relationships with others by manipulating others’ feelings or actions toward the victim. For instance, the perpetrator might spread rumors or gossip about the victim, deliberately ignore him or her, or enlist others to isolate the victim socially.

State legislators and school officials create antibullying laws and policies partly in response to public opinion, and judges might consider community attitudes when sentencing offenders, as Stalans (1993) noted. Once enacted, laws and policies require observers to recognize bullying and report it to authorities. Thus, it is important to learn more about how people evaluate episodes of aggression between students and how they want perpetrators treated within the system. How do they decide whether a specific incident constitutes bullying, how serious the transgression was, whether the perpetrator deserves punishment, and what the penalty (if any) should be? Although a number of studies have looked at the prevalence and consequences of bullying, relatively few have asked participants to evaluate cases of alleged bullying. In those that have, the participants were either school children, because the researchers were interested in the attitudes of individuals who are likely to experience or witness bullying directly (Maunder, Harrop, & Tattersall, 2010; Swearer & Cary, 2003) or teachers, preservice teachers, and school staff because these individuals must decide whether and how to intervene in cases of possible bullying that occur on school property (Bauman & Del Rio, 2006; Ellis & Shute, 2007; Hazler et al., 2001; Jacobsen & Bauman, 2007; Maunder et al., 2010; Nesdale & Pickering, 2006; Yoon, 2004). Previous research has not sampled adults over 18 years of age who are not employed at a school and are therefore not responsible for maintaining order in a school environment. Although these adults might sometimes observe school bullying directly, they would more importantly be in a position to influence antibullying efforts developed by legislators and administrators. Thus, one goal of this research was to examine evaluations made by participants who are not students or employees at primary or secondary schools. We asked them to listen to a disciplinary hearing involving a case of alleged bullying at a school and then make several judgments, including a verdict, a recommendation about punishment for the perpetrator, and a rating of the seriousness of the case.
In addition, we investigated certain contextual variables that might affect observers’ judgments. As mentioned previously, researchers have distinguished between different types of bullying. Although relational bullying is pervasive (Wang, Iannotti, & Nansel, 2009) and associated with long-lasting emotional distress and depression in victims (Bauman & Del Rio, 2006; Crick et al., 2006), studies have consistently demonstrated that teachers and school personnel tend to rate this type as less serious than the others. Additionally, participants report being less likely to intervene when they see incidents of relational bullying, they express less empathy for the victims, and they might not even define this type as “bullying” (Bauman & Del Rio, 2006; Jacobsen & Bauman, 2007; Maunder et al., 2010; Yoon & Kerber, 2003). However, evidence of observers’ lenient attitudes could be misleading because the vignettes participants read in past studies have not specified any harm that results from relational bullying, so the participants might have concluded that little or no damage was done. Moreover, as Bauman and Del Rio (2006) noted, it is comparatively more difficult to assess the consequences for victims of relational bullying. Whereas a broken limb is obviously a more severe injury than a scratch, and using an obscenity to address a classmate is worse than calling him or her a “jerk,” the harm resulting from relational bullying “is not directly observable; it must be inferred from the victim’s behaviors or the observer’s own experiences” (Bauman & Del Rio, 2006, p. 226). In addition, the victim could choose to hide his or her feelings, and the observer’s experiences might not have prepared him or her to understand the victim’s situation.

Therefore, in this experiment we independently manipulated the type of bullying and the resulting harm to the victim. We included two types of bullying (relational and verbal) but excluded the physical type for two reasons. First, the latter is less prevalent than the other types, especially among older students (Bauman & Del Rio, 2006; Chapell et al., 2006; Jordan & Austin, 2012; Wang et al., 2009) and, as explained later, we wanted the setting in one condition to be a university, where physical bullying might seem less plausible. Second, and more important, physical bullying that causes bodily injury would be prosecuted in court as assault rather than dealt with solely under a school’s antibullying policy (Brank et al., 2012), and we were interested in investigating bullying specifically rather than behaviors that also qualify as other kinds of offenses.

To vary the harm caused by the bullying, we created one condition with a relatively lower level and one with a higher level. Previous data suggest that observers will take harm into account when evaluating a case of alleged bullying. Hazler et al. (2001) discovered that school personnel were more likely to judge an incident as bullying and to rate its severity as greater as the apparent physical harm increased. In addition, Bauman and Del Rio (2006) reported that the vignette perceived as significantly less serious than the other five judged by the teachers in their study was the only one that
did not reveal any information about the harm done to the victim. These authors also found that experts on the subject of bullying endorsed amount of harm as one criterion that should be considered to judge the seriousness of an incident.

We therefore hypothesized that observers do not simply treat relational bullying leniently as a matter of course, but instead they consider the degree of harm. Specifically, we predicted that observers would make harsher judgments against the perpetrator and higher ratings of seriousness when more as opposed to less harm occurred, regardless of the type of bullying. In other words, we expected to obtain a main effect of harm but not of type.

Our third independent variable was the academic level of the students involved. In public forums, bullying is commonly discussed within the context of primary and secondary schools as opposed to college, and most existing research concentrates on students in high school and earlier grades (Chapell et al., 2004; Chapell et al., 2006; Olweus, 1995; Violence Prevention Works, n.d.). However, Chapell and colleagues observed that bullying is actually widespread among undergraduates, with 20% to 25% of these students reporting that they have been victimized by a classmate. In other words, although it is commonly believed that prevalence rates decrease with the age of those involved, bullying does not stop at the end of high school, but instead it “graduates to college” (Chapell et al., 2004, p. 59). How do observers react to cases of alleged bullying among undergraduates? We propose that they will be less likely to define an incident as bullying if it occurs among undergraduates rather than high school students. Our hypothesis is based on research demonstrating that people possess schemas for different crimes (Finkel & Groscup, 1997; Smith, 1991). These schemas might depart from the legal definitions. For example, Smith showed that a typical “kidnapping” schema includes the components “the victim is a child” and “there is a ransom demand,” even though neither detail is legally necessary for a defendant to be convicted of this charge. Moreover, mock jurors rely on their crime schemas to determine verdicts; Smith’s participants were more likely to find defendants guilty in scenarios that were consistent versus inconsistent with the schema for the crime in question. In effect, the task of rendering a verdict is a category membership decision. Jurors must decide whether the case presented to them is or is not an instance of the crime category (e.g., kidnapping). The more features the case shares with a juror’s crime schema, the higher the probability of a guilty verdict.

Translating Smith’s findings to our research, we predicted that an alleged case of bullying that occurs among undergraduates would be less likely to be defined as bullying than a case involving high school students because people typically think of bullying as a problem in primary and secondary schools rather than colleges and universities. In other words, incidents involving
undergraduates do not fit the “bullying schema” well. Consequently, we anticipated that, compared to observers given a high school case, those evaluating a case at the university level would treat the defendant more leniently and would rate the case as less serious.

In addition to the predictions already discussed, we expected female participants to make harsher judgments than males. This pattern would be consistent with data from two previous studies in which teachers read and evaluated vignettes (Ellis & Shute, 2007; Maunder et al., 2010). In both, women were more likely than men to define aggressive behaviors as bullying, to rate them as more serious, and to report that they would intervene in a mildly serious incident.

METHOD

Participants

The participants (N = 328) were undergraduate (n = 275) and graduate students (n = 25) from various departments at a medium-sized university in Indiana, as well as local community members (n = 28). The undergraduates received either course credit or extra credit for participating, and the graduate students and community members volunteered for the study. The participants ranged in age from 18 to 50 years (M = 20.70, SD = 4.82); 65% were female and 87% were White.

Materials and Procedure

The participants were asked to adopt the role of members of a disciplinary committee and to listen to a recording (approximate length 6 min) of a hearing involving a case of alleged bullying in a school environment. Undergraduate participants listened to the recording and filled out questionnaires in a testing room on campus in groups of up to 10 individuals. After completion of the session, the experimenter asked the students to help recruit additional participants by giving a stimulus packet to someone they knew who was not affiliated with the university. The packet contained instructions, a consent form, a questionnaire, and a website link to the disciplinary case audio recording. These community participants could then follow the instructions to read and sign the consent form, access the link, listen to the recording, complete the questionnaire, and mail the consent form and questionnaire to the researchers in a prepaid envelope. Graduate students were recruited by a research assistant who visited classes in 14 different departments within the university, explained the purpose of the study, and asked for volunteers. Interested students were given a packet like the ones prepared for community members.
The disciplinary case was specifically written to include the three critical elements of bullying that experts (e.g., Olweus, 1995) have identified (the behavior is intended to cause harm, it occurs repeatedly, and it involves an imbalance of power between the perpetrator and victim) and that appears in the legal definition we presented to participants (see later), although it also includes some mitigating factors. In the case, an 18-year-old male student accuses a male classmate of the same age of bullying him in the chemistry class they shared and elsewhere on school property over a period of several weeks. The defendant acknowledges his actions but minimizes them, saying they are being taken out of context, that he was just “playing around,” and that the victim never expressed being upset before bringing charges against him. Moreover, the victim admits behaving in ways that invited disapproval from his peers, such as asking so many questions during class one day that the instructor dismissed the students later than she had promised. The school’s Dean of Students presides over the hearing.

We manipulated three independent variables by altering some of the details within the case. First, we varied the academic level of the school the students attended. In different versions, the Dean of Students identifies the setting as either a high school or a university during her opening remarks. Second, we manipulated the type of bullying that occurred. In the verbal condition, the victim reports that the defendant often made hostile remarks to him, and as proof he offers a recording that he made with his voice recorder of a specific incident that occurred one day in chemistry class. The Dean of Students plays the recording as part of the disciplinary hearing so that the participants can listen to it for themselves. On it, the defendant can be heard insulting the victim, calling him “a complete idiot” and pointing out some of his failures in the course. In the relational condition, the defendant speaks to the rest of the class rather than to the victim, urging them to exclude and socially isolate him. Specifically, the defendant is heard on the recording telling the other students not to partner with the victim on lab days because “he’ll wreck your grade if you work with him” due to his incompetence.

The third independent variable was the degree of harm caused by the bullying. We recognize that this variable is continuous, and the amount of harm experienced by a victim is relative rather than absolute. For convenience, we refer to the two conditions as the “low” and “high” harm conditions. In the low harm condition, the victim testifies that the defendant’s actions made him feel depressed and miserable and that he dreaded going to school to the point that he skipped classes a couple of times. In the high harm condition, the victim describes these consequences but additionally testifies that he withdrew from the school because he was so upset. He emphasizes that he was reluctant to take this drastic step but decided he had no choice. He makes it clear that he is not currently attending any school.
The participants were randomly assigned to listen to one of the eight versions of the recording that were created by crossing the three independent variables. After listening, they were given a written definition of “bullying,” which was adapted from the statutory definition used in Indiana (Indiana Code 20-33-8-0.2), where the data were collected:

“Bullying” means overt, repeated acts or gestures by a student or group of students against another student with the intent to harass, ridicule, humiliate, intimidate, or harm the other student. “Acts or gestures” includes verbal or written communications, physical acts, or other behaviors that take place [on school or university property, at any school- or university-sponsored function, or (for the high school condition) on a school bus]. Bullying involves an imbalance of power or strength between the bully/bullies and the victim.

The participants were next asked to fill out a written questionnaire that included several items commonly used in juror decision-making research. The participants chose a verdict (either guilty or not guilty of bullying), estimated the probability of the defendant’s guilt, rated their confidence that the verdict they selected was appropriate from 0 (no confidence at all) to 10 (complete confidence), recommended a sentence for the defendant if the disciplinary committee were to find him guilty by choosing from a list of options that varied in severity (no punishment, formal meeting with an administrator, transfer to a different chemistry section to finish the class this semester, withdraw from the chemistry class and retake it next semester, suspension from all classes for the rest of the current semester, suspension from all classes for the rest of the current semester plus next semester, expulsion), and rated the seriousness of the incident from 0 (not at all) to 10 (extremely). The participants also made some judgments about the victim and the defendant. They were asked to report the level of responsibility that they would assign to the defendant and the victim for the incident, and they rated the amount of empathy they felt for each individual. All ratings were made on an 11-point scale.

Following these judgments, the participants responded to three manipulation checks asking them to remember what type of bullying occurred, the effects of the bullying on the victim, and what level of school the victim and defendant attended, according to the information presented in the case. All three were multiple-choice items.

The questionnaire concluded with demographic items asking participants to report their age, sex, race, and status as a student (i.e., undergraduate, graduate student, not in school). After finishing the questionnaire, they were thanked and debriefed.
RESULTS

Participants who failed one or more of the three manipulation checks were excluded from the analyses. A total of 21 participants, including 19 undergraduates (8 men, 11 women) and 2 community participants (1 man, 1 woman) were dropped. No graduate students failed to answer the checks correctly. After these 21 cases were removed, 307 participants remained in the sample for the analyses described next.

Preliminary Analyses

Preliminary analyses were conducted to determine whether any of the dependent measures varied as a function of participant group (undergraduates, graduate students, community members). We used a chi-square analysis to compare verdicts, a Kruskal–Wallis test for recommended sentences, and a simple analysis of variance to examine all other variables. We found no significant differences across groups for verdicts, $\chi^2(2, N = 307) = 2.00, p = .37, \text{Cramer’s } V = .08$; for sentences, $\chi^2(2, N = 307) = .58, p = .75$; or for any other variables ($Fs \leq 2$). Therefore, we combined the data from undergraduates, graduate students, and community participants for all further analyses. An additional preliminary analysis revealed that participant age was not significantly related to any measure ($ps \geq .09$). Except where noted, the examinations of the dependent variables discussed here used factorial analyses of variance, with academic level, type of bullying, amount of harm, and participants’ sex as factors.

Verdict, Confidence, and Probability of Guilt

A hierarchical log-linear analysis was conducted to test the effects of academic level, bullying type, degree of harm, and participants’ sex on verdict. Three main effects were identified (see Table 1). First, a main effect of academic level emerged, revealing that participants who learned that the events took place at a high school rather than a university were more likely to choose a guilty verdict, $\chi^2(1, N = 307) = 10.14, p = .001$. Second, participants were more likely to find the defendant guilty if they heard the victim testify that he experienced more rather than less harm, $\chi^2(1, N = 307) = 23.69, p < .001$. Finally, women were more likely than men to render a guilty verdict, $\chi^2(1, N = 307) = 3.73, p = .05$. No other effects were significant ($\chi^2$s $\leq 3$).

Three main effects emerged pertaining to participants’ confidence in their chosen verdict. Participants who heard the high school case reported more confidence than those who heard the university case, $F(1, 291) = 9.73, p = .002, \eta^2 = .03$. Additionally, participants in the high harm rather than low harm condition were more confident, $F(1, 291) = 31.34, p < .001, \eta^2 = .10,$
and women were more confident than men, $F(1, 291) = 4.72, p = .03, \eta^2 = .02$. No other significant effects were obtained ($F$s $\leq$ 2).

Regarding probability of guilt judgments, we discovered a main effect of academic level, with participants who listened to the high school case reporting higher probabilities than participants who listened to the university case, $F(1, 291) = 11.65, p = .001, \eta^2 = .04$. We also found that participants in the high harm condition assigned a greater probability of guilt than those in the low harm condition, $F(1, 291) = 113.57, p < .001, \eta^2 = .28$. Although there was a nonsignificant trend for women to give higher probability of guilt estimates than men, $F(1, 291) = 3.11, p = .08, \eta^2 = .01$, there were no other significant effects ($F$s $\leq$ 2).

### Sentence

Each participant was asked to recommend a sentence for the defendant in the event that the disciplinary committee finds him guilty. An ordinal logistic regression was conducted to test the effects of academic level, bullying type, harm, and participants’ sex on sentence choices, revealing two main effects.
First, participants in the high harm condition recommended harsher sentences than those in the low harm condition, $\chi^2(1, N = 307) = 14.88, p < .001$. Among participants who heard the victim describe experiencing less harm, the most common sentence was to require the defendant to transfer to another section of chemistry (recommended by 51%), followed by making him withdraw from the class and retake it next semester (24%), and having him attend a formal meeting with an administrator (18%). In contrast, the top choices of the participants in the high harm condition were transferring to another section (38%), retaking the class next semester (29%), and being suspended from school for the rest of the semester (18%).

The second main effect was that women’s sentences were more punitive than men’s, $\chi^2(1, N = 307) = 12.55, p < .001$. The most common recommendation from men was for the defendant to transfer to another section (preferred by 53%), followed by a meeting with an administrator (20%), and retaking the class next semester (17%). The women’s top choices were transferring to another section (40%), retaking the class (32%), and being suspended for the rest of the semester (14%).

A significant interaction between bullying type and sex also emerged, $\chi^2(1, N = 307) = 3.99, p = .05$. Follow-up Mann–Whitney $U$ tests demonstrated that, although women recommended harsher sentences than men for both types, this difference was significant in the relational condition, $U \geq 3.30, p < .001$, but only marginally significant in the verbal condition, $U = 1.88, p = .06$. To elaborate, the sentence most preferred by women in the relational condition was to require the defendant to withdraw from the class and retake it next semester, whereas the men’s modal response was to transfer him to another section. In the verbal condition, both women and men chose transferring to another section as their top sentence recommendation. Besides this interaction, there were no other significant effects ($\chi^2s \leq 2.5$).

**Seriousness**

As shown in Table 1, participants in the high school condition rated the incident as more serious than those in the university condition, $F(1, 291) = 12.22, p = .001, \eta^2 = .04$. Furthermore, higher ratings were reported by participants in the high harm rather than low harm condition, $F(1, 291) = 18.17, p < .001, \eta^2 = .06$. However, an interaction between academic level and harm was also significant, $F(1, 291) = 4.94, p = .03, \eta^2 = .02$. To clarify this interaction, we conducted simple effects analyses, which indicated that greater harm led to higher seriousness ratings when the incident took place at a high school (low harm: $M = 5.10, SD = 2.19$; high harm: $M = 6.73, SD = 1.71$), $F(1, 303) = 25.62, p < .001, \eta^2 = .08$, but harm had no significant effect in the university case (low harm: $M = 4.87, SD = 2.14$; high harm: $M = 5.41, SD = 1.97$), $F(1, 303) = 2.69, p > .05, \eta^2 = .01$. 
Besides these effects, we discovered that women assigned a higher level of seriousness than men, $F(1, 291) = 7.11, p = .008, \eta^2 = .02$. No other significant effects were found ($Fs \leq 2$).

**Judgments of Defendant and Victim**

Participants were asked to rate the defendant’s and victim’s responsibility for what happened. Concerning the responsibility of the defendant, there were main effects of both academic level and harm (see Table 2). Participants who heard the high school case assigned more responsibility to the defendant than those who heard the university case, $F(1, 291) = 10.89, p = .001, \eta^2 = .04$. In addition, compared to participants in the low harm condition, those in the high harm condition thought the defendant was more responsible, $F(1, 291) = 5.90, p = .02, \eta^2 = .02$. The interaction between academic level and harm approached significance, $F(1, 291) = 3.66, p = .06, \eta^2 = .01$; harm had a clear effect on participants’ judgments when the setting was a high school (low harm: $M = 7.73, SD = 2.40$; high harm: $M = 8.75, SD = \ldots$)

**TABLE 2** Ratings of Responsibility and Empathy for the Defendant and Victim as a Function of Academic Level, Harm, and Participants’ Sex

<table>
<thead>
<tr>
<th>Condition</th>
<th>Defendant responsibility</th>
<th>Victim responsibility</th>
<th>Defendant empathy</th>
<th>Victim empathy</th>
</tr>
</thead>
<tbody>
<tr>
<td>High school</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low harm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>7.59 (2.26)</td>
<td>4.31 (2.75)</td>
<td>3.44 (2.59)</td>
<td>4.84 (2.76)</td>
</tr>
<tr>
<td>Women</td>
<td>7.82 (2.52)</td>
<td>3.24 (2.28)</td>
<td>2.42 (1.62)</td>
<td>6.51 (2.67)</td>
</tr>
<tr>
<td>High harm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>9.04 (1.12)</td>
<td>3.17 (2.39)</td>
<td>2.79 (2.15)</td>
<td>6.21 (2.27)</td>
</tr>
<tr>
<td>Women</td>
<td>8.62 (1.68)</td>
<td>2.45 (2.16)</td>
<td>2.00 (1.76)</td>
<td>7.36 (1.80)</td>
</tr>
<tr>
<td>Total high school</td>
<td>8.25 (2.07)$_A$</td>
<td>3.17 (2.43)$_A$</td>
<td>2.54 (2.03)$_A$</td>
<td>6.43 (2.52)$_A$</td>
</tr>
<tr>
<td>University</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low harm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>7.34 (2.37)</td>
<td>3.66 (2.65)</td>
<td>2.21 (2.06)</td>
<td>5.38 (2.38)</td>
</tr>
<tr>
<td>Women</td>
<td>7.55 (1.99)</td>
<td>3.98 (2.10)</td>
<td>2.49 (1.85)</td>
<td>5.34 (2.23)</td>
</tr>
<tr>
<td>High harm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Men</td>
<td>7.35 (1.92)</td>
<td>3.43 (2.76)</td>
<td>3.61 (2.31)</td>
<td>5.30 (2.58)</td>
</tr>
<tr>
<td>Women</td>
<td>7.82 (1.57)</td>
<td>3.02 (1.85)</td>
<td>2.57 (2.08)</td>
<td>5.92 (2.31)</td>
</tr>
<tr>
<td>Total university</td>
<td>7.57 (1.92)$_b$</td>
<td>3.51 (2.26)$_A$</td>
<td>2.63 (2.07)$_A$</td>
<td>5.54 (2.34)$_b$</td>
</tr>
<tr>
<td>Total low harm</td>
<td>7.60 (2.26)$_A$</td>
<td>3.77 (2.42)$_A$</td>
<td>2.61 (2.03)$_A$</td>
<td>5.59 (2.56)$_A$</td>
</tr>
<tr>
<td>Total high harm</td>
<td>8.23 (1.70)$_b$</td>
<td>2.90 (2.21)$_b$</td>
<td>2.55 (2.07)$_b$</td>
<td>6.40 (2.31)$_b$</td>
</tr>
<tr>
<td>Total men</td>
<td>7.80 (2.10)$_1$</td>
<td>3.69 (2.65)$_1$</td>
<td>3.00 (2.34)$_1$</td>
<td>6.59 (2.53)$_1$</td>
</tr>
<tr>
<td>Total women</td>
<td>7.98 (1.98)$_1$</td>
<td>3.14 (2.16)$_1$</td>
<td>2.36 (1.84)$_2$</td>
<td>6.32 (2.37)$_2$</td>
</tr>
</tbody>
</table>

*Note.* Means are reported with standard deviations in parentheses. Ratings were made on an 11-point scale with higher numbers reflecting greater quantities. Values in the same column that do not share the same capital letter alphabetical subscript (total academic level rows), lowercase alphabetical subscript (total harm rows), or numerical subscript (total sex rows) differ significantly, $p < .05$. 

Downloaded by [Ball State University], [Keri L. Pickel] at 07:46:05 December 2014
1.54), $F(1, 303) = 10.62, p < .01, \eta^2 = .04$, but no significant effect when the setting was the university (low harm: $M = 7.47, SD = 2.13$; high harm: $M = 7.68, SD = 1.69$), $F(1, 303) = .40, p > .05, \eta^2 < .01$. No other significant effects emerged ($Fs \leq 2$).

Regarding responsibility ratings for the victim, participants in the high harm condition assigned less responsibility than those in the low harm condition, $F(1, 291) = 7.20, p = .008, \eta^2 = .02$. Also, the main effect of sex approached significance, $F(1, 291) = 3.08, p = .08, \eta^2 = .01$, such that women attributed somewhat less responsibility to the victim than men did. There were no other significant effects ($Fs \leq 3$).

The participants were asked to report the amount of empathy they felt for the defendant and for the victim. Regarding empathy for the defendant, a main effect of sex was found, with women reporting less empathy than men, $F(1, 291) = 8.62, p = .004, \eta^2 = .03$. The academic level by harm interaction was also significant, with participants who heard the high school case feeling more empathy when less harm occurred (low harm: $M = 2.84, SD = 2.12$; high harm: $M = 2.24, SD = 1.90$), whereas the pattern was reversed for participants who heard the university case (low harm: $M = 2.38, SD = 1.93$; high harm: $M = 2.89, SD = 2.19$), $F(1, 291) = 8.68, p = .003, \eta^2 = .03$. Simple effects analyses clarified that there was no significant effect of harm in the high school, $F(1, 303) = 3.48, p > .05, \eta^2 = .01$, or in the university condition, $F(1, 303) = 2.35, p > .05, \eta^2 = .01$. No other significant effects were obtained regarding empathy for the defendant ($Fs \leq 2$).

Concerning empathy for the victim, three main effects emerged. First, participants in the high school condition felt more empathy than those in the university condition, $F(1, 291) = 7.75, p = .006, \eta^2 = .03$. Second, participants in the high harm condition reported more empathy than those in the low harm condition, $F(1, 291) = 4.90, p = .03, \eta^2 = .02$. Additionally, compared to men, women rated their empathy as higher, $F(1, 291) = 8.56, p = .004, \eta^2 = .03$. However, we also found a significant interaction between bullying type and sex, $F(1, 291) = 5.92, p = .02, \eta^2 = .02$. Simple effects analyses showed that women were more empathetic than men toward the victim when the bullying was relational (women: $M = 6.56, SD = 2.29$; men: $M = 4.92, SD = 2.64$), $F(1, 303) = 16.12, p < .01, \eta^2 = .05$, but not when it was verbal (women: $M = 6.06, SD = 2.45$; men: $M = 5.84, SD = 2.36$), $F(1, 303) = .31, p > .05, \eta^2 < .01$. No other significant effects were identified ($Fs \leq 3$).

**DISCUSSION**

Although existing data have suggested that observers treat relational bullying with less concern than other types (Bauman & Del Rio, 2006; Jacobsen & Bauman, 2007; Maunder et al., 2010; Yoon & Kerber, 2003), we hypothesized...
Evaluations of a Bullying Case

that those findings emerged because previous researchers did not specify any particular harm that resulted from the bullying in the vignettes participants read, and the participants found it difficult to infer what injury the victims might have suffered. Therefore, we independently manipulated bullying type and harm. As predicted, we obtained no main effects of bullying type for any of our dependent measures, whereas more rather than less harm led to a higher proportion of guilty verdicts, greater confidence in the verdict, higher estimates of the probability of the defendant’s guilt, and higher ratings of seriousness. Participants in the high harm condition also felt more empathy for the victim, and they attributed more responsibility for the incident to the defendant and less to the victim. It is reassuring that, in accordance with experts’ recommendations (Bauman & Del Rio, 2006), observers seem to realize that harm is an important factor to consider when evaluating a bullying incident and that they might not see relational bullying as inherently less serious than the verbal type. In fact, if the harm experienced by the victim is made explicit to individuals who must review a case of alleged bullying (e.g., teachers, school administrators, jurors, disciplinary committee members), then their judgments might be fair and appropriate. Moreover, educating people about the damage caused by relational bullying might encourage them to intervene when they observe incidents of this type, to offer support to the victims, and to report such incidents to authorities.

Besides harm and bullying type, we also manipulated the academic level of the victim and perpetrator. Our participants generally made harsher judgments when the incident took place in a high school rather than a university. Specifically, participants who heard the high school case were more likely to find the defendant guilty, to rate their confidence in their verdict as higher, and to give higher estimates of the probability of guilt. They also reported feeling more empathy for the victim. In addition, we obtained one nonsignificant difference in the expected direction; participants in the high school condition rated the victim as less responsible for what happened. For two measures, academic level interacted with harm to qualify the significant main effect of harm. Specifically, greater harm led to higher ratings of seriousness and defendant responsibility only in the high school case. Collectively, these results suggest that participants drew important distinctions between incidents occurring in a high school and those occurring at a university.

Because both students were 18 years old in each of the two academic level conditions, these effects cannot be attributed to age. Instead, we propose that the results emerged because of observers’ conceptualizations of what “bullying” is and the circumstances under which it arises. Bullying is strongly associated with elementary, middle, and high school; most public discussions about bullying, proposed programs to combat it, and existing studies of its prevalence and effects focus on students in those grades (Chapell et al., 2004; Chapell et al., 2006; Olweus, 1995). Thus, just as jurors
hearing a kidnapping case might find a defendant not guilty if some of the elements of the case depart from their schema for that crime, regardless of the legal definition (Finkel & Groscup, 1997; Smith, 1991), our participants might have downgraded the university case because one component of their “bullying schema” is that it occurs among students in primary and secondary schools. In contrast, when the incident happened at a high school rather than a university, participants were more likely to decide that it matched the official definition of “bullying” and to return a guilty verdict. They also seemed more concerned about the high school victim, as reflected in their empathy ratings. Moreover, as shown by the academic level by harm interaction for seriousness ratings, the participants concluded that dropping out of school in response to the bullying, versus merely feeling upset, greatly elevated the gravity of the situation only if the victim attended high school. In other words, the participants apparently did not fully appreciate the impact of bullying on the university victim, who felt forced to leave the school.

The truth, however, is that bullying does occur among university students, including both the verbal and relational types (Chapell et al., 2004; Chapell et al., 2006), which means that researchers should work to instruct school administrators, faculty, students, parents, lawmakers, and the general public so they will recognize that school bullying happens at all academic levels and will be in a better position to help develop and support policies for preventing and responding to this problem. On a more positive note, after hearing the case we created, which was written to fit experts’ definition of bullying, most of our participants in both academic level conditions appropriately decided the defendant was guilty.

We found several differences in judgments as a function of the participants’ sex. Women were more likely than men to choose guilty verdicts, they were more confident in their verdicts, and they thought the incident was more serious. They also felt less empathy for the defendant. Three of the remaining dependent measures (probability of guilt, ratings of responsibility for the defendant and the victim) did not show significant main effects of sex, but across all three women favored the victim more than men did. These results are consistent with previous research (Ellis & Shute, 2007; Maunder et al., 2010), but few studies exist that have examined sex differences in adult observers. More research is needed to determine whether men and women understand and evaluate bullying differently, and if so, why. Although Maunder et al. (2010) suggested that empathy or other attributes linked to gender should be further investigated, no clear explanation has been offered for discrepancies between male and female observers’ judgments.

For two measures, we discovered sex differences that occurred only when the bullying was relational. Specifically, with relational but not verbal bullying, women recommended more punitive sentences than men and expressed greater empathy toward the victim. Although no interaction
between sex and bullying type emerged for the other dependent variables, it could be that women and men view the two types somewhat differently, perhaps in connection with the higher prevalence of relational bullying among female versus male students (Wang et al., 2009). Future research could investigate further.

Researchers should also explore whether altering the details within the stimulus case would produce results that diverge from ours. For example, observers’ gender stereotypes might lead them to weigh harm differently when the victim and the defendant are female rather than male or are an opposite-sex pair. Moreover, the participants’ own experiences pertaining to bullying or their exposure to antibullying campaigns in their schools or communities could have an effect on their responses. Researchers could also manipulate other variables such as the age of the parties involved, whether the victim or defendant has a disability, or the number of times the defendant bullied the victim.

The hypothesis that observers’ judgments are informed by their schema of what bullying is leads to some interesting predictions that could be tested in future studies. For instance, “typical” bullies are often physically bigger and older than their peers, whereas victims are younger, slight in stature, socially awkward, or members of a marginalized group (American Educational Research Association, 2013; Jordan & Austin, 2012). Observers relying on a schema that incorporates these components might be less likely to define an incident as bullying if the defendant is gay rather than heterosexual or if the victim is physically imposing rather than diminutive. Another feature of the bullying schema is that it occurs in primary and secondary schools. Observers might downgrade the seriousness of a case that occurs between adult employees at work instead of between children or teenagers in an academic setting.

This study represents a step toward understanding how adult observers evaluate cases of alleged bullying. Although a few previous studies have examined the perceptions of preservice teachers, teachers, and school staff, it is also important to learn about the views of adults who are not affiliated with school systems. These individuals indirectly influence statutes and school policies that define bullying and prescribe punishments by electing state legislators and school board members, who in turn will take their constituents’ opinions into account as they create regulations. Our findings indicate that adult observers recognize that degree of harm is one criterion that should be used when evaluating bullying cases. They do not, however, see the relational type as inherently less serious than the verbal type, assuming harm is made clear. Also, our data imply that the general public needs to be educated that bullying can occur at any academic level. Ideally, policies and laws should be explicitly written to protect students from all types of bullying in all learning environments.
ACKNOWLEDGMENTS

We thank Chris Altman, Heather Bauer, Carolyn Humbert, Tully Roll, Brad Root, and Jon Whitney for their assistance with this project.

REFERENCES


