Willingness to Intervene in Bullying Episodes Among Middle School Students: Individual and Peer-Group Influences

Dorothy Espelage¹, Harold Green², and Joshua Polanin³

Abstract
This study examined the associations among gender, empathy, attitudes toward bullying, willingness to intervene, and bullying within peer groups in a sample of sixth and seventh-grade students (N = 346; M Age = 12.22 years). Peer groups were identified via social network analysis using NEGOPY (Richards, 1995) and peer-group predictors were evaluated with multi-level modeling. Male peer-group willingness to intervene results indicated significant between-group variation (i.e., high ICC). Perspective-taking was associated with greater willingness to intervene within male peer groups after controlling for initial levels of willingness to intervene. Greater bullying perpetration within one’s peer group was highly predictive of less individual willingness to intervene. For females, willingness to intervene scores was not dependent on friendship group. This study suggests that bullying prevention
programs that encourage students to intervene on behalf of victims might be efficacious for male students with friends who bully others at low rates.

**Keywords**
bullying, friendship, peers, socialization

Interactions among friends during early adolescence become more frequent and their relationships are described as more intimate than the level of interactions among younger children (Furman & Buhrmester, 1992). Adolescent friendship groups heavily influence members’ attitudes and behaviors (Rubin, Bukowski, & Parker, 1998). Indeed, a plethora of developmental and sociological research has documented the strong influence of friends on one another (Alexander, Piazza, Mekos, & Valente, 2001; Cairns, Leung, Buchanan, & Cairns, 1995; Houtzager & Baerveldt, 1999; Kandel, 1978), and the difficulty that early adolescents have when they oppose their friends’ opinions, attitudes, and behaviors (Brown, 2004; Steinberg & Monahan, 2007).

As a result, members of adolescent friendship groups share remarkably similar attitudes and behaviors. This within-group similarity has been referred to as homophily (McPherson, Smith-Lovin, & Cook, 2001). Kandel (1978) argued that homophily involved two separate processes; selection and socialization. The selection process suggested that children who were similar to each other were more likely to form groups, and socialization referred to the tendency for adolescents to adopt attitudes and behaviors of other group members, an influence process (Kandel, 1978). The homophily hypothesis has been examined in relation to numerous characteristics including sex (Ibarra, 1992; Kandel, 1978), race (Mollica, Gray, & Trevino, 2003), attitudes, and behaviors such as achievement (Lin, 1999; Portes, 1998; Ryan, 2001; Wentzel & Caldwell, 1997), aggression (Espelage, Holt, & Henkel, 2003; Espelage, Green, & Wasserman, 2007), homophobia (Poteat, Espelage, & Green, 2007), and smoking and drug use (Alexander et al., 2001; Ennett & Bauman, 1994; Kandel, 1978). This study tested the homophily hypothesis in relation to willingness to intervene in bullying episodes among middle school students.

**School Bullying and Bystander Intervention**

Scholars suggest that bullying prevention programs that emphasize and encourage bystander intervention are likely to be effective in reducing bullying rates in schools (Newman, Horne, & Bartolomucci, 2000; Olweus, 1993; Rigby & Johnson, 2006; Salmivalli, Karna, & Poskipart, 2010). These prevention
researchers hypothesize that an increase in the amount of bystander intervening behavior results in a decrease in bullying. Thus, bystanders are encouraged to either report an incident of bullying or to confront students who are bullying other students. Data often suggest, however, that adolescents rarely intervene to assist victims (O’Connell, Pepler, & Craig, 1999). Further, self-declared bullies and bystanders sometimes report feeling sorry after bullying their peers, although they rarely intervene in bullying episodes (Borg, 1998). In the most comprehensive observational study of bullying, O’Connell et al. (1999) videotaped first through sixth graders (n = 120) during recess. The researchers found that 54% of peers spent their time reinforcing bullies by passively watching, 21% actively modeled bullies, and only 25% intervened. Older boys (Grades 4-6) were more likely to join actively with the bully than were younger boys (Grades 1-3) and older girls. Younger and older girls intervened on behalf of victims more often than older boys. Additional analyses of these videotapes showed that peers were present during 88% of bullying episodes, but only intervened 19% of the time. The researchers indicated that 57% of the interventions effectively stopped the bullying (Hawkins, Pepler, & Craig, 2001). This observational study clearly supports the current study’s investigation of how middle school students’ attitudes and behaviors are associated with their friend’s attitudes and behaviors.

Individual-Level Predictors of Willingness to Intervene

Given the significant cost and ethical issues with conducting observational studies, recent studies of bystander intervention have employed less intrusive assessments such as vignette methodology to identify correlates of bystander intervention. For example, participants are provided with scenarios of bullying and asked a series of questions related to the feelings of the characters, their own reactions to the scenarios, and how they might respond as a witness to the bullying. Results of these studies have provided further insight into the predictors of defender behavior (e.g., Rigby & Johnson, 2006).

Gender and age. Rigby and Johnson (2006) attempted to elucidate the characteristics of youth that are willing to intervene in bullying situations. Four hundred Australian primary and secondary students viewed a videotape of a bullying situation and were subsequently asked how they would respond if they witnessed the bullying. Primary students were more willing to intervene than secondary students, and females were more likely than males to intervene. Similarly, Gini and colleagues found that empathy was associated with defending a victim only for boys, not for girls (Gini et al., 2007), and
greater attitudes supportive of victims were found for children than preadolescents and girls. Taken together, we hypothesized that sixth graders would report higher levels of willingness to intervene than seventh graders and girls would report higher levels of willingness to intervene than boys.

**Attitudes, empathy, and bullying experiences.** Rigby and Johnson (2006) found that greater willingness to intervene was associated with having a positive attitude toward victims, having rarely or never bullied others, and having been victimized. Although decades of research point to the role of empathy in promoting prosocial behavior and inhibiting antisocial behavior, only recently have studies specifically extended empathy to willingness to intervene in bullying scenarios or defender behavior (Caravita, DiBlasio, & Salmivalli, 2009; Endresen & Olweus, 2001; Gini, Albiero, Benelli, & Al toe, 2007; Gini, Pozzoli, & Haiser, 2011; Gini, Pozzoli, Borghi, & Franzoni, 2008; Nickerson, Mele, & Princiotta, 2008; Pöyhönen, Juvonen, & Salmivalli, 2010; Pozzoli & Gini, 2010; Stavrinides, Georgiou, & Theofanous, 2010). Taken together, these studies find that among early adolescent samples, defending behavior is associated with greater empathy (Gini et al., 2007, 2008; Nickerson et al., 2008; Stavrinides et al., 2010) and bullies appear to be morally competent but lack in morally compassionate behavior in comparison to victims or defenders (Gini et al., 2011). Thus, we hypothesized that both empathy and perspective-taking, both components of empathy, would be related to increases in willingness to intervene in bullying situations.

**Peer-Level Predictors of Willingness to Intervene**

Consistent with Rigby and Johnson’s study, Pozzoli and Gini (2010) found that perceived positive peer pressure to defend a victim interacted with personal responsibility to predict defending. That is, students who held moderate or high levels of personal responsibility were more likely to defend a victim if they perceived their peers to hold a positive view toward defender behavior. Thus, in addition to individual attitudes predictive of willingness to intervene (attitude toward bullying, empathy), the goal of the current study was to systematically examine bullying within friendship networks to determine if membership in particular peer groups predicts willingness to intervene. Given the findings of extant literature, we hypothesized that greater bullying perpetrated by members of a group (directed either toward one another or to members outside of the group) would be associated with less willingness to intervene.
Current Study

The current study extended the scholarship on bystander intervention in several areas. First, this study is the first investigation of individual and peer-level influences in a longitudinal and multilevel design, which has been a limitation of many studies (Salmivalli, 2010). Second, the investigation analyzed male and female peer group samples separately to decompose the association of friendship groups by sex. Third, this study assesses the impact of both cognitive and emotional dimensions of empathy in relation to willingness to intervene. Recent studies of early adolescents suggest that cognitive and affective dimensions of empathy differentially predicted willingness to intervene in bullying situations (Gini et al., 2011; Stavrinides et al., 2010). More specifically, Gini and colleagues (2011) found that bullies had similar rates of morally competent (cognitive dimension) behavior as victims and defenders, but bullies had lower morally compassionate or caring (emotional dimension) attitudes than victims or defenders. Further, Pöyhönen and colleagues (2010) found that defending was associated with vicarious understanding (emotional dimension) of victims rather than a cognitive understanding of the victim.

Finally, no empirical studies have considered how willingness to intervene in bullying situations is potentially influenced by the amount of bullying among a student’s peers or within their friendship group (see review of peers and bullying research, Salmivalli, 2010). Some of the extant literature included the application of social network analysis to test similarity of friends on measures of bullying perpetration (Cairns, Cairns, Neckerman, Gest, & Gariépy, 1988; Mouttapa, Valente, Gallaher, Rohrbach, & Unger, 2004; Veenstra et al., 2007), but none have assessed how bullying perpetration among friends negatively influences adolescent’s willingness to intervene. It follows from classic behavioral theory that the behaviors of bullying perpetration and intervening to help victims/targets are not likely to co-occur in peer groups given that they are incompatible. From a social cognitive perspective (Bandura, 1989), if a peer group engages in high amounts of bullying, then this behavior is likely to be adopted by individuals within the group. Thus, this study employs network analytic techniques to identify friendship groups and employs multilevel modeling to examine the similarity in willingness to intervene among friends. Multilevel analyses are also used to test our hypotheses that being female and younger (sixth grader versus seventh grader), having greater proattitudes toward bullying and less empathy would be associated with less willingness to intervene. Finally, it was hypothesized that greater bullying perpetration by members of the individual’s peer group...
would predict less willingness to intervene over time after taking into account individual predictors.

Method

Participants

Data from participants were drawn from Waves 1 and 2 of a longitudinal investigation of bullying during early adolescence. In early spring 2002, parental permission forms were sent to all students \( n = 631 \) registered at a central Illinois middle school (Grades 6-8) and parents were asked to sign and return the consent form only if they did not want their child to participate in the study. Of the 631 enrolled students, 35 (5.5%) of the students did not complete Wave 1 surveys because their parents opted them out of participation, 19 (3%) students with disabilities did not complete the survey, and 12 (2%) either opted out of the survey themselves or randomly answered the items. Thus, a total of 565 students completed measures at Wave 1 for a total participation rate of 90%. Of these 565 students, 378 were sixth- and seventh-grade students and therefore were students in the school during Wave 2. In spring 2003, the sample consisted of 346 (92%) of these 378 students ranging in age from 11 to 14 years at Wave 1 \( M = 12.22 \text{ years}; \ SD = .76 \). Attrition from Wave 1 to Wave 2 among these students was attributed to 12 (3%) students leaving the district, parents of 6 (1.5%) of the students opted their child out of survey, 4 (1%) students were absent on the day of data collection, 10 (2.5%) did not want to complete the survey or randomly answered the survey items. Students completed surveys across two waves separated by 1 year. Participants who were not in Wave 2 did not differ from students in Waves 1 and 2 on the study’s measures \( ps > .05 \). Of the remaining 346 students, 51% were females \( n = 178 \) and 49% were males \( n = 168 \), with 51% sixth graders \( n = 178 \) and 49% seventh graders \( n = 168 \). Approximately 94.5% were White, 0.5% were African American, 0.5% were Asian, 0.7% were Hispanic, 0.4% were Native American, 2.3% were biracial, and 2.7% reported other racial backgrounds. Seven percent of the students at this school were on free or reduced lunch (retrieved from school report card data).

Procedure

Participants completed the study survey during a 45-minute free period. Surveys were administered to groups ranging in size from seven to twenty-five students. The facilitators informed the students that the survey would
ask them about aggression, their feelings, and their friends. Students consented by signing their name on the front colored coversheet. Names were collected to allow for matching students’ data across the longitudinal study. Names were converted to numbers within three hours of the data collection. An excel file was created with names and survey numbers that was stored on a secure university server that only the principal investigator had access to.

In each classroom, one of two trained examiners read each item and response option aloud while a second team member monitored students’ progress. Students were provided with a list of websites and hotlines in case they needed assistance for bullying situations and students were encouraged to talk to their parents, teachers, or counselors if they felt that they were not safe because of bullying.

**Measures**

The survey at both Wave 1 and Wave 2 consisted of six sections: (a) demographic questions (sex, race, grade), (b) friendship nominations, (c) empathy scales, (d) positive attitude toward bullying scale, (e) willingness to intervene in bullying situation scale, and (f) bullying perpetration scale.

**Friendship network data.** Based on previous studies of adolescent friendship networks (Ennett & Bauman, 1994, 1996), students were asked questions about their friends. Specifically, they were asked to list up to eight friends similar in age (but not siblings) with whom they hang out most often in their school. Students were allowed to nominate as few or as many students as they wished, up to eight names. These names were then converted to participant numbers to identify friendship networks and matched with survey data to provide network attributes.

**Empathy—Interpersonal Reactivity Index (IRI).** Given the multidimensional nature of empathy, the authors utilized several scales to assess components of empathy that have consistently emerged in the literature. Two scales from the IRI (Davis, 1983) were included to assess cognitive and emotional aspects of empathy. Empathy is conceptualized as a set of constructs all related to a general concern for others. The Perspective-taking (PT) scale consisted of seven items that assessed the “tendency to spontaneously adopt the psychological point of view of others” (Davis, 1983, p. 114). An example item is “I try to look at everybody’s side of a disagreement before I make a decision.” A seven-item Empathetic Concern (EC) scale assessed empathy and concern for others, and an example item is “When I see someone being taken advantage of, I feel kind of protective towards them.” Response options ranged from “Does not describe me well” through “Describes me very well.”
consistency coefficients have ranged from .71 through .77 for the scales and test-retest reliabilities of .62 through .71 have been reported (Davis, 1980). A Cronbach’s alpha coefficient of .67 was found in this study for the Perspective-taking scale and .79 for the Empathetic Concern scale at Wave 1.

Positive attitude toward bullying. A four-item University of Illinois Positive Attitudes toward Bullying Scale was utilized to assess this construct. This measure was developed from in-depth interviews with middle school students (Espelage & Asidao, 2001). The researchers asked students how much they agree or disagree with statements related to their attitude toward bullying (e.g., “A little teasing doesn’t hurt anyone.”). Response options were strongly disagree, disagree, agree, and strongly agree. Higher scores on this scale were interpreted as having a favorable or positive view of bullying. A Cronbach’s alpha coefficient of .81 was found for this study at Wave 1.

Self-reported bullying behavior. The nine-item University of Illinois Bully Scale (UIBS; Espelage & Holt, 2001) was used to assess bullying behavior that included teasing, social exclusion, name calling, and rumor spreading. This scale does not assess physical aggression. Researchers developed this scale based on interviews with middle school students, a review of the extant bullying measures literature, and extensive factor analytic investigations (Espelage, Bosworth, & Simon, 2000; Espelage et al., 2003). Students were asked to indicate how often in the past 30 days they had engaged in each behavior (e.g., “I teased other students.” and “I upset other students for the fun of it.”). Response options included never, 1 or 2 times, 3 or 4 times, 5 or 6 times, and 7 or more times. These response options assessed bullying persistence. Higher scores indicated more self-reported bullying behaviors. Espelage and Holt (2001) found a Cronbach’s alpha coefficient of .87 and the Bullying Scale was found to be moderately correlated ($r = .65$) with the Youth Self-Report Aggression Scale (Achenbach, 1991), suggesting convergent validity. Concurrent validity of this scale was established with significant correlations with peer nominations of bullying. This scale converged with peer nomination data (Espelage et al., 2003). This scale was not significantly correlated with the Illinois Victimization Scale ($r = .12$), and thus provided evidence of discriminant validity (Espelage et al., 2003). A Cronbach’s alpha coefficient of .88 was found for the current sample at Wave 1.

Willingness to intervene in bullying episodes. The authors created the University of Illinois Willingness to Intervene in Bullying Episodes specifically for this study. The 5-item scale was developed from a series of interviews and surveys of students in grades third through eighth. The researchers asked
students the extent that they agree with statements about intervening directly or indirectly when they encounter bullying (e.g., “If a kid is being teased, I will stick up for him/her.,” “I will tell an adult if a kid is being teased a lot.”). Response options were *strongly disagree, disagree, agree,* and *strongly agree.* Cronbach’s alpha coefficients were .75 for Waves 1 and 2.

**Peer-group bullying.** To estimate the level of peer-group bullying, we aggregated self-reported bullying scale scores of individuals of mutually exclusive peer groups, and divided by the number of individuals per peer group. This provided an average bullying measurement per peer group.

**Analyses**

Given the documented sex and age differences in empathy and attitude toward bullying in the extant literature (Espelage, Mebane, & Adams, 2004; Hoffman, 2000), the impact of sex and grade was examined for all study variables, including the understudied willingness to intervene variable using ANOVA and MANOVA absent any friendship data. Recognizing that peer groups may have an even greater impact on children’s attitudes and behaviors than sex and grade, the authors then utilized social network analyses to identify friendship networks and estimated peer-group effects using hierarchical linear modeling.

**Identifying peer groups.** Friendship nomination data were subjected to social network analysis using NEGO PY (Richards, 1995). A total of 2,527 pairwise friendship nominations were made across the 378 Wave 1 study participants (sixth and seventh graders). Of these nominations, only 11% (n = 296) included names of students not enrolled in this study. Ninety-four (n = 350) percent of the students identified at least one friend that was enrolled in the study. The number of friendship nominations ranged from zero to eight. Students on average nominated 5.90 friends (SD = 2.05). In addition, only 9% of the nominations (n = 218) were pairs of students not within the same grade. Eighth graders that were nominated as friends were not included in these analyses. There were no significant differences between sixth and seventh graders on the number of friends nominated (p > .05).

Using peer nomination data, analyses identified each participant’s position in their social network using NEGO PY (Richards, 1995). Eigenvector decomposition approaches via the NEGO PY program were used to identify groups of students who reported having more contact with each other than with students in other groups. Students are identified as members of groups, dyads, as liaisons, or isolates as a function of the pattern and strength of their
friendship nominations. For this investigation, we based group identification on reciprocated nominations and shared friendships with others (Richards, 1995). Based on previous investigations that have used SNA to identify peer networks (Ennett & Bauman, 1994, 1996), parameters were set using the following criteria to identify peer groups: (a) at least 50% of a student’s reciprocated friendships must be in the peer group, (b) a direct (reciprocated friendship) or indirect (common friendship) link exists from each member to every other member of the peer group, and (c) indirect links must not exceed three links. These parameters were used to confirm group membership after initial eigenvector decomposition. Students were categorized as group members, liaisons, isolated dyads, and isolates, though only group members were the focus of subsequent analyses.

Group members belong to groups that consisted of at least three individuals who have most of their interaction with members of the same group and at least two links with others. Dyads consisted of two individuals and were not included in this analysis. Students with links to multiple groups were assigned to the peer group with the greatest number of friendship links (assumed their primary peer group); however, if the links to the multiple groups were equivalent then the student was considered a liaison and was not included in subsequent analyses. These criteria provide a strict methodology to identify clearly defined and cohesive groups of individuals, strengthening the ability to compare differentiated peer groups and to model peer group effects.

Hierarchical linear modeling. The authors utilized hierarchical linear modeling (HLM) to evaluate the association between individual characteristics and peer-level bullying on willingness to intervene in bullying episodes. HLM provides the unique opportunity to include both individual (Level 1) and peer group (Level 2) characteristics to predict the willingness to intervene outcome. Level 1 variables included perspective-taking, empathetic concern, and attitude toward bullying. The Level 2 predictor for this analysis was the peer-group bullying scale aggregate score.

The authors conducted separate analyses for males and females to capture unique experiences across the male and female peer groups and to examine the Level 1 and Level 2 variables’ effects with greater clarity. Separate gender analyses followed the guidelines described by Bryk and Raudenbush (1992). The authors initially conducted mixed-gender models (both boys and girls in a single model); however, the girls-only results revealed that little variation occurred between groups proportional to within-groups. Further, because only two of the peer groups included both males and females, the analysis reflected the reality that sixth- and seventh-grade peer groups consisted of
the same sex. For these two groups, one consisted of a majority of girls and was considered a female peer group and one consisted of a majority of boys was considered a male peer group. Taken together, the authors believe that separate boys and girls analyses provided a clear representation of the sample and the phenomenon.

Three steps constitute the development of a two-level hierarchical linear model. First, a fully unconditional (null) model is estimated. The fully unconditional model is analogous to conducting a one-way random effects ANOVA model and involves no Level 1 or Level 2 predictors. Using this first model, the within- and between-group variance in the outcome variable is calculated and is used to calculate the intraclass correlation (ICC). The ICC represents the proportion of between-group (Level 2 variance) to within-group variation. General social science research requires an ICC amount of at least .10, or 10% (Lee, 2000). Although other authors have advocated for the use of multilevel modeling or alternate modeling procedures when ICC fails to exceed 10% (Bliese, 2009; Bliese & Ployhart, 2002), we believed that the small number of peer groups utilized in this study required a greater intraclass correlation to warrant multilevel modeling. Further, establishing a conservatively high ICC a priori guards against possible Type 2 errors, and reinforces our hypothesis that peer groups contribute to an individual’s willingness to intervene beyond an individual’s latent traits.

For this initial procedure, the null model equation was as follows:

\[ \text{WILL TO INTERVENE}_{ij} = \beta_{0j} + e_{ij} \]

(1)

\[ \beta_{0j} = \gamma_{00} + \mu_{0j} \]

where \( \beta_{0j} \) is the intercept, or average level of willingness to intervene for students in peer group \( j \), and \( e_{ij} \) is error for student \( i \) in peer group \( j \), \( \gamma_{00} \) is the grand-mean outcome of the peer group, and \( \mu_{0j} \) is peer group \( j \)’s random effect.

Assuming significant between-group variation exists, the second step was to create a Level 1 model. As delineated by Heck and Thomas (2000), the authors tested a Level 1 random-intercept model with fixed slope coefficients; followed by a slope heterogeneity evaluation. In the Level 1 or within-group model, all variables were group-mean centered and the intercept and slope coefficients were specified as random (Heck & Thomas). This model assessed the relations among the individual-level variables and the outcome variable of
willingness to intervene at Wave 2 and indicated between-group intercept and slope variation. The within-group model equation was follows:

\[ WTI_{2ij} = \beta_0j + \beta_{1j} (EC1_{ij}) + \beta_{2j} (PT1_{ij}) + \beta_{3j} (POSATT1_{ij}) + \beta_{4j} (WTI1_{ij}) + e_{ij} \]  

(2)

where \( \beta_0j \) is the intercept, or average level of WTI2 for students in peer group \( j \) adjusted by the other predictors in each peer group, \( \beta_{1j} \) is the slope, or the association between empathetic concern (EC1) and WTI2, \( \beta_{2j} \) is the slope between perspective-taking (PT1) and WTI2, \( \beta_{3j} \) is the slope between positive attitude toward bullying (POSATT1) and WTI2 in peer group \( j \), and \( \beta_{4j} \) is the slope between willingness to intervene at Wave 1 and 2.

Accepting the hypothesis that significant intercept or slope variation between groups remains, the authors continued to the third step. The third model-building step was to specify a two-level model where Level 2 variables (i.e., peer-group mean bullying score) were hypothesized to explain Level 1 outcome variation (i.e., individual willingness to intervene at Wave 2).

Finally, the authors tested a between-group model to address the peer-level bullying effects on male students’ WTI2. In this model, the intercept that resulted from the within-group equation served as the dependent variable modeled as a function of the peer group bullying level. We continued to adjust for the individual predictors of Wave 1 empathetic concern, perspective-taking, positive attitudes toward bullying, and willingness to intervene. The Level 2 equation was as follows:

\[ \text{Avg. } WTI_{2ij} (\beta_{0j}) = \gamma_{00} + \gamma_{01} (\text{PEER GROUP BULLYING1})_j + U_{0j} \]  

(3)

where \( \gamma_{00} \) is the grand mean for student WTI2 and \( \gamma_{01} \) is the effect of peer-group bullying level after considering other predictors. Equation 2 described above continues to reflect the Level 1 model.

To evaluate improvement of model fit between models, the authors compared both deviances and between group variance across models (Kreft & De Leeuw, 1998). The difference between deviances of two models should be twice as large as the difference in the number of estimated parameters between the two models. In addition, each model has a \( \chi^2 \) goodness of fit distribution with specific degrees of freedom.
Results

Sex and Grade Difference Analyses

The authors conducted a two-way MANOVA to evaluate the hypothesis that willingness to intervene at Waves 1 and 2 would differ by sex and grade. A significant overall MANOVA effect was found for sex ($\Lambda = .75; F = 15.11, p < .001; \eta^2 = .25$), grade ($\Lambda = .90; F = 4.97, p < .001; \eta^2 = .10$), and sex by grade interaction ($\Lambda = .91; F = 4.55, p < .001; \eta^2 = .10$) on this set of variables. Follow-up ANOVAs for sex indicated that males reported less willingness to intervene in bullying episodes at both Waves 1 and 2 ($\eta^2$s = .12, .16, $p$s < .001). With respect to grade, follow-up ANOVAs revealed that seventh graders reported less willingness to intervene at Wave 1 and 2, but with relatively low effect sizes ($p$s < .01; $\eta^2$s = .08, .07). Main effects for sex and grade were qualified by significant interactions of sex and grade for willingness to intervene at Wave 2 ($p < .01; \eta^2 = .05$). Seventh-grade males had the lowest levels of willingness to intervene in comparison to sixth-grade males and females in both grades. Analyses for all study variables can be found in Table 1.

Peer-Group Membership

For the 178 sixth graders, eigenvector decomposition via NEGOPY yielded 13 peer groups (excluding dyads; $n = 91, 55\%$) ranging in size from three members to 41 members ($M = 6.57; SD = 10.05$). Twenty-five students (13%) were nominated by other students but none of their nominations were reciprocated and 16 (9%) did not nominate anyone, but were nominated by other students. These two groups were categorized as isolates and were not included in subsequent analyses. Forty students (22%) were defined as liaisons because they had indirect links to several groups but appeared to have no primary peer-group affiliation. Six students (3%) were defined as dyads and excluded as well.

For the 168 seventh graders, eigenvector decomposition via NEGOPY yielded 15 peer groups (excluding dyads; $n = 94, 55\%$) ranging in size from three members to 20 members ($M = 6.40; SD = 5.24$). Eighteen students (11%) were nominated by other students but none of their nominations were reciprocated and 13 (7%) did not nominate anyone but were nominated by other students. These two groups were categorized as isolates and were not included in subsequent analyses. Twenty students in dyads (11%) were identified by this procedure and were eliminated from the analysis. Twenty-three
Table 1. Mean (Standard Deviations) and ANOVA Statistics for Sex and Grade Differences on Empathy, Attitudes, and Willingness to Intervene Scales

<table>
<thead>
<tr>
<th>Scale</th>
<th>Male students</th>
<th>Female students</th>
<th>ANOVA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sixth</td>
<td>Seventh</td>
<td>Sixth</td>
</tr>
<tr>
<td>Empathetic concern Wave 1</td>
<td>3.15 (.80)</td>
<td>2.77 (.99)</td>
<td>3.92 (.56)</td>
</tr>
<tr>
<td>Perspective-taking Wave 1</td>
<td>3.05 (.76)</td>
<td>2.69 (.77)</td>
<td>3.44 (.69)</td>
</tr>
<tr>
<td>+ Attitude t/d bullying Wave 1</td>
<td>2.15 (.60)</td>
<td>2.35 (.73)</td>
<td>1.82 (.48)</td>
</tr>
<tr>
<td>Willingness to intervene Wave 1</td>
<td>3.00 (.49)</td>
<td>2.59 (.57)</td>
<td>3.28 (.58)</td>
</tr>
<tr>
<td>Willingness to intervene Wave 2</td>
<td>2.83 (.62)</td>
<td>2.25 (.76)</td>
<td>3.09 (.53)</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.
students (15%) were defined as liaisons because they had indirect links to several groups but appeared to have no primary peer-group affiliation. These liaisons were not included in subsequent analyses.

Bullying Within the Peer-Group Context: Multilevel Analysis

Initial mixed-gender analysis results indicated limited between-group variation (i.e., small ICC), and female results indicated relatively little between-group variation. Therefore, multilevel analyses results are presented for males and females separately. In addition, we utilized general data screening techniques to confirm that the data met appropriate standards (Raudenbush & Bryk, 2002; Tabachnick & Fidell, 2007). We addressed multicollinearity via estimation of Level 1 correlations (Table 2). The correlation between empathy and positive attitude toward bullying approached significant magnitude ($r = -0.64$), but failed to meet suggested a priori removal levels (Tabachnick & Fidell, 2007). Centering each variable during the modeling stage also decreased multicollinearity concerns (Miles & Shelvon, 2001).

Predicting willingness to intervene for males. A total of 82 male participants in 11 peer groups had data on all predictors. The outcome variable was willingness to intervene at Wave 2. To determine whether male peer groups differed in average Wave 2 willingness to intervene (WTI2) levels, the authors estimated a fully unconditional model. The ICC indicated similarity in scores among peer-group members and suggested that male students affiliate with

| Table 2. Pearson Correlation Matrix: Boys Only and Girls Only Models |
|-------------------|---------------|---------------|---------------|---------------|
|                   | 1             | 2             | 3             | 4             |
| **Boys model**    |               |               |               |               |
| 1. Empathetic concern | —             |               |               |               |
| 2. Perspective-taking | .61***       | —             |               |               |
| 3. Positive attitude toward bullying | — .64***       | — .39         | —             |               |
| 4. Willingness to intervene, Time 1 | .59***       | .52***        | — .52***      | —             |
| 5. Willingness to intervene, Time 2 | .41***       | .52***        | — .37***      | .61***        |
| **Girls model**   |               |               |               |               |
| 1. Empathetic concern | —             |               |               |               |
| 2. Perspective-taking | .39***       | —             |               |               |
| 3. Positive attitude toward bullying | — .39***       | — .38***      | —             |               |
| 4. Willingness to intervene, Time 1 | .46***       | .45***        | — .43***      | —             |
| 5. Willingness to intervene, Time 2 | .40***       | .33***        | — .33***      | .63***        |

***$p < .001$. 

**Table 2. Pearson Correlation Matrix: Boys Only and Girls Only Models**
Table 3. Hierarchical Linear Models for Predicting Willingness to Intervene at Wave 2 among Males

<table>
<thead>
<tr>
<th>Variables</th>
<th>Unconditional</th>
<th>Random coefficient</th>
<th>Slopes-as-outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (U0)</td>
<td>2.52 (.16)***</td>
<td>2.53 (.17)***</td>
<td>2.60 (.06)***</td>
</tr>
<tr>
<td>Peer group bullying</td>
<td></td>
<td></td>
<td>−.71 (.12)***</td>
</tr>
<tr>
<td>Empathy</td>
<td>−.12 (.16)</td>
<td></td>
<td>−.08 (.13)</td>
</tr>
<tr>
<td>Perspective</td>
<td>.33 (.13)*</td>
<td>.25 (.11)*</td>
<td></td>
</tr>
<tr>
<td>Bullying Attitude</td>
<td>.03 (.17)</td>
<td>−.03 (.10)</td>
<td></td>
</tr>
<tr>
<td>Willingness Wave 1</td>
<td>.39 (.19)†</td>
<td>.52 (.14)***</td>
<td></td>
</tr>
<tr>
<td>Deviance</td>
<td>169.10</td>
<td>147.64</td>
<td>142.58</td>
</tr>
<tr>
<td>U0</td>
<td>.18</td>
<td>.26</td>
<td>.01</td>
</tr>
<tr>
<td>R</td>
<td>.40</td>
<td>.24</td>
<td>.30</td>
</tr>
</tbody>
</table>

Note: () indicates standard error; italics indicate Level 2 variable.
†p < .10, *p < .05, **p < .01, ***p < .001

peers who have comparable levels of willingness to intervene. Furthermore, the ICC revealed that Wave 2 willingness to intervene scores are dependent on peer-group membership; specifically, 31.6% of the variance was between peer groups. The χ² value associated with the Level 2 variance component was 35.86 with 10 degrees of freedom, leading to a rejection of the null hypothesis that mean willingness to intervene scores of male adolescents in all peer groups were equal. These results suggested that further multilevel modeling was appropriate.

As Table 3 indicates, the average mean willingness to intervene score at Wave 2 adjusted for individual predictors was 2.53. Even after considering individual predictors of Wave 1, willingness to intervene, empathetic concern, perspective-taking, and a positive attitude toward bullying, significant mean score variation of willingness to intervene at Wave 2 remained across peer groups (χ² = 27.36, p < .001). Thus, knowing individual predictor scores at Level 1 did not provide enough information to explain group differences of willingness to intervene at Wave 2. The ICC indicated that 57% of variance in willingness to intervene at Wave 2 was between group, an increase from the null model, but deviance decreased from 169.25 to 147.64. This suggested a better fit between models. The results also suggested that significant variation remained for the slope coefficients between willingness to intervene Wave 1 and Wave 2, and perspective-taking slope coefficients at Wave 1 and willingness to intervene Wave 2. The reliability estimate of these
coefficients; however, remained quite low (.23 and .26, respectively). As such, all subsequent slope coefficients models remained fixed.

Finally, this second model indicated that perspective-taking was the only variable significantly related to willingness to intervene at Wave 2 ($\beta = 0.33$, $t = 2.50$, $p < .05$). Willingness to intervene at Wave 1 showed a marginally significant association with willingness to intervene at Wave 2 as well ($\beta = 0.39$, $t = 2.02$, $p = .07$). Further models included these Level 1 variables.

Results of the between-groups model indicated that peer-group bullying influenced individual willingness to intervene at Wave 2 levels ($\gamma = -.71$, $t = -5.76$, $p < .001$) after controlling for individual predictors (see Table 3). As hypothesized, higher peer-group bullying scores were associated with less individual willingness to intervene in bullying episodes. Controlling for peer-group and individual effects, perspective-taking was significantly associated with an individual’s willingness to intervene at Wave 2 ($\beta = 0.25$, $t = 2.27$, $p < .05$). These results indicated that an individual’s willingness to intervene was related to both the peer and individual level factors.

This model represented a significant improvement over the within-group model; deviance decreased significantly to 142.58, and the addition of the peer-group bullying variable accounted for 94% of the between-group variance. We discontinued the model-building process because little between-group variation remained ($\chi^2 = 10.47$, $p > .05$).

Predicting willingness to intervene for females. One hundred and three female participants within 17 peer groups had data on all predictors. To evaluate whether female peer groups differed in average willingness to intervene levels at Wave 2, the authors estimated a fully unconditional model that was analogous to the male null model delineated above (see Equation 1). The ICC indicated peer group heterogeneity because less than 2% of the variance occurred between peer groups on willingness to intervene. Further, the null hypothesis that mean scores of willingness to intervene were identical across female peer groups was retained ($\chi^2 = 19.35$, $p > .05$). This indicated that mean willingness to intervene scores did not differ across female peer groups. As such, multilevel modeling was not appropriate, and the authors conducted a basic regression analysis that utilized Equation 2. The four predictors explained 41% of the willingness to intervene Wave 2 variance, and the only significant predictor of willingness to intervene at Wave 2 was willingness to intervene at Wave 1 ($\beta = .52$, $t = 5.67$, $p < .001$).

The authors also evaluated female peer group bullying effects to illustrate gender differences. Results (Table 4) indicated that peer group bullying had no effect on females’ willingness to intervene ($\beta = -.09$, $t = -.44$, $p > .05$).
Table 4. Hierarchical Linear Models for Predicting Willingness to Intervene at Wave 2 Among Males and Females

<table>
<thead>
<tr>
<th>Variables</th>
<th>Boys model</th>
<th>Girls model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.60 (.06)****</td>
<td>3.07 (.05)****</td>
</tr>
<tr>
<td>Peer group bullying</td>
<td>-.71 (.12)****</td>
<td>-.09 (.21)</td>
</tr>
<tr>
<td>U0</td>
<td>.009</td>
<td>.028</td>
</tr>
<tr>
<td>R</td>
<td>.304</td>
<td>.122</td>
</tr>
</tbody>
</table>

Note: () indicates standard error.

These findings suggested that gender differences indeed existed and corroborated the need to compute gender analyses separately.

Discussion

This study employed multilevel modeling to examine both individual-level and peer-group level predictors of willingness to intervene in bullying situations. It represented a significant improvement over the extant literature because very few studies have focused on identifying factors associated with student’s willingness to intervene in bullying situations in longitudinal designs and none have examined this construct within-peer or friendship groups. This lack of attention to peer group influences on bullying attitudes and behaviors is an unfortunate phenomenon because an individual student’s decision to defend a victim has been associated with peer-group norms (Gini et al., 2007; Pöyhönen et al., 2010), and prevention programs emphasize bystander intervention (Newman et al., 2000; Olweus, 1993; Rigby & Johnson, 2006; Salmivalli et al., 2010).

Several specific findings should be mentioned. For boys only, the results revealed that higher peer-group bullying levels were significantly and negatively related to an individual’s willingness to intervene in a bullying situation, controlling for all other predictors. This finding represented a significant improvement compared to other models. Further results revealed that an individual’s perspective-taking was significantly and positively related to willingness to intervene, but empathy and a positive attitude toward bullying were not. Although we expected that either or both of those predictors would be related to willingness to intervene over time, the correlation analyses revealed relatively moderate relations that should temper expectations. At the request of one reviewer, we also conducted the analysis removing each of the
predictors one at a time. This procedure failed to improve fit significantly and empathy and a positive attitude toward bullying were not significant predictors of willingness to intervene.

Many scholars posit that modifying attitudes supportive of violence and empathy training positively influence bullying prevention. Numerous character education, bullying curricula, anger management, and social problem-solving prevention/intervention programs include empathy training and promote prosocial, nonviolent attitudes (e.g., Goldstein, Glick, & Gibbs, 1998; Newman et al. 2000; Pecukonis, 1990). These programs are predicated on the assumption that understanding negative behavior toward others (i.e., empathy) and engaging in prosocial behavior will decrease an individual’s bullying behavior. Our findings indicate that bullying prevention programs that focus solely on empathy training or bystander intervention without attention to decreasing the amount of bullying perpetration within one’s primary peer group will not yield positive effects. Programs routinely encourage individual children to be an effective bystander or an ally from an individual perspective, but fail to have a conversation with children about how their intervening is viewed by their friends. For example, Rigby and Johnson (2006) attempted to increase bystander intervention through video modeling, showing elementary and middle school students’ videos of peer-intervening behavior during simulated bullying situations. In addition, Frey, Hirschstein, Edstrom, and Snell (2009) have attempted to bolster peer support with the Steps to Respect Program. Although both programs observed treatment effects, we posit that these effects would intensify if peer group behaviors were also considered. This is supported with Pozzoli and Gini’s (2010) finding that peer normative pressure to defend a victim interacted with personal responsibility to increase the likelihood of bystander intervention. It is important to recognize that early adolescence is a time in which the opinions and attitudes of friends play a pivotal role in the individual decisions that kids make about their own behavior.

Our finding that willingness to intervene was associated with peer-group membership only for boys is consistent with previous work. For example, Gini and colleagues (2007) found the association between empathy and defending a victim was associated for boys only. Similarly, Salmivalli and Voeten (2004) found that group norms across 48 school classes (Grades 4-6) around students’ behaviors in bullying situations, including defending the victim, had a greater effect on girls than boys. It is important to note that other studies have found peer-level effects for bullying and other forms of aggression perpetration for both males and females. For example, in a similar study of middle school students, multilevel modeling was appropriate (i.e.,
high ICCs) when the outcome was not willingness to intervene but bullying perpetration (Espelage et al., 2003). These findings suggest that it is important to explore predictors of attitudes and behaviors across multiple levels, including individual and peer groups.

Despite the contributions this study provides to the current literature on willingness to intervene in bullying situations among middle school students, it has several limitations. First, the data were collected from one middle school, and we were not able to examine the extent that Level 3 variables, school-level factors, predicted attitudes toward bullying, and willingness to intervene. Therefore, ecological validity is a concern, and future studies should include more schools and assess factors such as school climate, teacher’s attitudes toward bullying, bullying policies, and so on. Second, data included in this article were self-report; future studies should incorporate behavioral measures of willingness to intervene (or actual intervention) to evaluate whether attitudes to intervene are correlated with behaviors, and further explore possible bystander effects (Darley & Latané, 1968; Latané & Darley, 1970). Moreover, we should also caution readers not to assume causality. Although the results indicated strong associations between peer-group bullying perpetration levels and willingness to intervene, one cannot say (with certainty) that increased peer group perpetration causes decreased willingness to intervene. Similar to other observational studies, these analyses merely represented relationships.

Our findings suggested a possible differential response between girls and boys, but we only briefly analyzed these differences and gender differences should be considered in future studies of peer influence on bystander intervention. Third, the lack of homophily among female groups in their willingness to intervene should not be interpreted until future studies are conducted. Next, this study should be extended to examine more specifically the characteristics of those students who are willing to intervene despite a peer group norm that discourages it. For example, a recent study of 356 school classes (Grades 3-5; n = 7,481) found 72% of victims had one defender, that was usually of the same gender and liked by the victims (Sainio, Veenstra, Huitsing, & Salmivalli, 2010). Extending our understanding of willingness to intervene to isolates, dyads, and liaisons might also indicate how students who do not reside in peer groups contribute to a culture of bystander intervention.

Finally, this study examined peer factors on willingness to intervene solely based on peer-group membership, but did not examine specific characteristics of each group, such as density, embeddedness, or concentration. Thus, social network analytic strategies that enable modeling at a greater level of detail (such as $p^*$/ERGM models) need employment to represent more
adequately the impact of an individual’s position in friendship networks with respect to their willingness to intervene (Espelage et al., 2007). That is, rather than merely relying on network techniques that identify structural roles such as isolates, liaisons, and group members, subsequent analyses should focus on an individual’s specific pattern of relationships and attributes and how those variables may coevolve. This would also overcome the potential limitation of aggregating individual-level bullying for peer-level analyses as employed in this current study.

Declaration of Conflicting Interests
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

References


Espelage et al. 25


Developmental Psychology, 43, 1531-1543. doi:10.1037/0012-1649.43.6.1531
Pearson Education.
Veenstra, R., Lindenberg, S., Zijlstra, B. J. H., De Winter, A. F., Verhulst, F. C., & 
2007.01102.x
membership: Relations to academic achievement in middle school. Child Devel-
opment, 68, 1198-1209.

Bios

Dorothy Espelage is a professor in the Department of Educational Psychology at the 
University of Illinois, Urbana-Champaign. Her research programs include investiga-
tions of bullying, sexual harassment, and dating violence among adolescents for 
almost two decades. She is engaged in a large randomized clinical trial of a school-
based bullying prevention program.

Harold Green, Jr. is a behavioral scientist at the RAND Corporation. His work 
focuses on the social determinants of health using social network analysis theory and 
method. His studies typically focus on high-risk groups both nationally and 
internationally.

Joshua Polanin is a doctoral student in the School of Education at Loyola University 
Chicago. His research interests include bullying prevention programs, multilevel 
modeling techniques, and methods of meta-analysis.