The Effectiveness of Policy Interventions for School Bullying: A Systematic Review

William Hall  University of North Carolina at Chapel Hill

ABSTRACT  Objective: Bullying threatens the mental and educational well-being of students. Although anti-bullying policies are prevalent, little is known about their effectiveness. This systematic review evaluates the methodological characteristics and summarizes substantive findings of studies examining the effectiveness of school bullying policies. Method: Searches of 11 bibliographic databases yielded 489 studies completed since January 1, 1995. Following duplicate removal and double-independent screening based on a priori inclusion criteria, 21 studies were included for review. Results: Substantially more educators perceive anti-bullying policies to be effective rather than ineffective. Whereas several studies show that the presence or quality of policies is associated with lower rates of bullying among students, other studies found no such associations between policy presence or quality and reductions in bullying. Consistent across studies, this review found that schools with anti-bullying policies that enumerated protections based on sexual orientation and gender identity were associated with better protection of lesbian, gay, bisexual, transgender, and queer (LGBTQ) students. Specifically, LGBTQ students in schools with such policies reported less harassment and more frequent and effective intervention by school personnel. Findings are mixed regarding the relationship between having an anti-bullying policy and educators’ responsiveness to general bullying. Conclusions: Anti-bullying policies might be effective at reducing bullying if their content is based on evidence and sound theory and if they are implemented with a high level of fidelity. More research is needed to improve on limitations among extant studies.

KEYWORDS: school, bullying, policy, law, effectiveness
doi: 10.1086/690565

Bullying in schools is a pervasive threat to the well-being and educational success of students. Bullying refers to unwanted aggressive behaviors enacted intentionally over time by an individual or group using some form of power to cause physical and/or psychological harm to another individual or group in a shared social context (Gladden, Vivolo-Kantor, Hamburger, & Lumpkin, 2014; Olweus, 2013). Bullying is also a widespread phenomenon. A meta-analysis of 82 studies conducted in 22 countries in North America, South America, Europe, South-
ern Africa, East Asia, and Australia and Oceania found that 53% of youth were involved in bullying as bullies, victims, or both bullies and victims (Cook, Williams, Guerra, & Kim, 2010).

Negative Outcomes Connected with Bullying

Involvement in bullying as perpetrators, victims, bully–victims, and bystanders has been linked with deleterious outcomes by both cross-sectional and longitudinal studies. Youths who are bullied can experience immediate negative effects that include physical injury, humiliation, sadness, rejection, and helplessness (Kaiser & Rasminsky, 2009). Over time, a number of mental and behavioral health problems can emerge, including low self-esteem, anxiety, depression, suicidal ideation and behavior, conduct problems, psychosomatic problems, psychotic symptoms, and physical illness (Arseneault, Bowes, & Shakoor, 2010; Dake, Price, & Telljohann, 2003; Gini & Pozzoli, 2009; Kim & Leventhal, 2008; Klomek, Sourander, & Gould, 2010; Reijntjes et al., 2011; Reijntjes, Kamphuis, Prinzie, & Telch, 2010; Ttofi, Farrington, Lösel, & Loeber, 2011a). In addition, students who have been bullied may not feel safe at school and may disengage from the school community due to fear and sadness, which may, in turn, contribute to higher rates of absenteeism and lower academic performance (Arseneault et al., 2006; Buhs & Ladd, 2001; Buhs, Ladd, & Herald, 2006; Glew, Fan, Katon, Rivara, & Kernic, 2005; Juvonen, Nishina, & Graham, 2000; Nakamoto & Schwartz, 2010).

Youths who bully also face psychosocial difficulties. These youths often grow up in harsh social environments with few resources (Hong & Espelage, 2012), and bullies often lack impulse control and empathy for others (O’Brien, Bradshaw, & Sawyer, 2009; van Noorden, Haselager, Cillessen, & Bukowski, 2015). Students who bully are more likely to skip school, perform poorly, and drop out (Jankauskiene, Kardelis, Sukys, & Kardeliene, 2008; Ma, Phelps, Lerner, & Lerner, 2009). Bullying perpetration also is associated with depressive symptoms, suicidal ideation and behavior, and violent and criminal behavior (e.g., assault, robbery, vandalism, carrying weapons, and rape; Dake et al., 2003; Kim & Leventhal, 2008; Klomek et al., 2010; Ttofi, Farrington, & Lösel, 2012; Ttofi, Farrington, Lösel, & Loeber, 2011b). Compared to nonperpetrators, students who bully have an increased risk of violent and criminal behaviors into adulthood. A meta-analysis of longitudinal studies found that school bullies were 2.5 times more likely to engage in criminal offending over an 11-year follow-up period (Ttofi et al., 2011b).

Other youths involved in bullying include bully–victims and bystanders. Bully–victims are students who have been bullied but also engage in bullying others. Bully–victims can experience a combination of internalizing and externalizing problems (Cook, Williams, Guerra, Kim, & Sadek, 2010). Student bystanders are present in up to 90% of bullying incidents (Atlas & Pepler, 1998; Craig & Pepler, 1995; Glew et al., 2005; Hawkins, Pepler, & Craig, 2001). Youths who witness bul-
lying often report emotional distress, including increased heart rate and higher levels of fear, sadness, and anger when recalling bullying incidents (Barhight, Hubbard, & Hyde, 2013; Janson & Hazler, 2004). Thus, across the literature, bullying is associated with problematic outcomes for perpetrators, victims, bully–victims, and bystanders alike.

Policy as an Intervention for Bullying
Perspectives vary on how to best address bullying in schools. Intervention strategies have included suspending and expelling bullies, training teachers on intervening, teaching empathy and respect to students through classroom lessons, maintaining constant adult supervision throughout school settings, collaborating with parents about student behavior, and enacting school-wide policies about bullying. In the United States, policies addressing bullying emerged in 1999 following the Columbine High School shootings. These policies have spread due to increased awareness and concern about student violence and school safety (Birkland & Lawrence, 2009).

A policy is a system of principles created by governing bodies or public officials to achieve specific outcomes by guiding action and decision making. Policy is an umbrella term that refers to various regulatory measures, including laws, statutes, policies, regulations, and rules. These terms vary based on the jurisdiction and legal authority of the individual or group who established the policy. In the United States, K–12 education policy, which includes school bullying policy, can be established at the federal, state, and local levels (Mead, 2009).

One advantage of policy interventions for bullying is that they can influence student, teacher, and administrator behavior as well as school organizational practices. For example, school bullying policies typically prohibit certain behaviors, such as threatening and harassing other students or retaliating against students who witness and then report bullying incidents. Policies may also require behaviors, such as requiring teachers to report bullying incidents to administrators and requiring administrators to investigate reports of bullying. Further, policies may promote certain behaviors by explicitly stating positive behavioral expectations for students or discourage behaviors by explicitly stating punishments associated with aggressive behaviors. At the school level, policies can guide organizational practices, such as establishing bullying incident reporting procedures and creating school-safety teams tasked with developing and executing school-safety plans. Thus, bullying policies can influence individual and organizational behaviors.

Another advantage of bullying policies is that they are upstream interventions that provide a foundation for downstream interventions. In other words, policies are systems-level interventions that typically require more targeted intervention programs, practices, and services at the organizational, group, and individual levels (McKinlay, 1998). For example, a bullying policy may be adopted within a state or district; the policy then applies to all schools within the state or district. This policy
may require training all school employees on bullying prevention strategies, integrating bullying awareness and education into classroom lessons and curricula, and providing counseling for students involved in bullying. Thus, policy lays the groundwork for an array of more specific and targeted interventions to be deployed in schools by outlining goals and directives in the policy document.

Policy design is important because the content influences a cascade of actions throughout school systems, which may result in positive or negative outcomes. For example, a bullying policy that requires schools to provide counseling services and positive behavioral reinforcement to students who perpetrate bullying is markedly different than a policy that requires schools to suspend or expel students who have carried out multiple acts of bullying. Research shows that overly harsh and punitive policies (e.g., “three strikes and you’re out” policies or “zero-tolerance” policies) are not effective at reducing aggression or improving school safety (American Psychological Association Zero Tolerance Task Force, 2008). Thus, bullying policies should be crafted and revised using evidence-based strategies.

Anti-bullying laws have been enacted in a number of countries, including Canada, the Philippines, the United Kingdom, and the United States. Although the United States does not have a federal law against school bullying currently, all states have enacted anti-bullying laws (U.S. Department of Health and Human Services, 2015). The content of these laws was reviewed in a U.S. Department of Education report, which shows some consistency but also variability in the inclusion of policy components (see Table 1; Stuart-Cassel, Bell, & Springer, 2011). These state laws apply to approximately 98,000 K–12 public schools and have a goal of protecting more than 50 million students from involvement in bullying (Snyder & Dillow, 2013; Stuart-Cassel et al., 2011).

Despite the widespread adoption and application of anti-bullying policies within the United States and in other countries, relatively few studies have examined the effectiveness of these interventions. Instead, research has focused on programmatic interventions (e.g., Cool Kids Program, FearNot!, Friendly Schools, KiVa, and Steps to Respect). Numerous systematic or meta-analytic reviews have been completed on the effectiveness of programmatic interventions for school bullying (e.g., Baldry & Farrington, 2007; Evans, Fraser, & Cotter, 2014; Ferguson, San Miguel, Kilburn, & Sanchez, 2007; Lee, Kim, & Kim, 2013; Merrell, Gueldner, Ross, & Isava, 2008; Ttofi & Farrington, 2011). However, a systematic review of the literature on the effectiveness of policy interventions for school bullying has not been completed.

Purpose of the Current Review
Given the proportion of students directly or indirectly involved in bullying, the array of educational and psychological problems associated with bullying, the extensive adoption of anti-bullying policies, and the absence of a review of the research on these policy interventions, the need for a systematic review on this topic is im-
The following questions drove this review: Are school policies effective in reducing or preventing bullying behavior among students? What is the state or quality of the research on school bullying policy effectiveness? What additional research is needed on school bullying policy effectiveness? Given these questions, the objectives of this review were threefold: to systematically identify, examine, and evaluate the methodological characteristics of studies investigating the effectiveness of school bullying policies; to summarize the substantive findings from these studies; and to provide recommendations for future research.

**Method**

In preparation of this review, the author adhered to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) criteria (Moher, Liberati, Tetzlaff, & Altman, 2009). Before undertaking the search for relevant studies, the author developed protocols for bibliographic database searches, study inclusion and exclusion criteria, and a data extraction tool. In addition, this review was registered with PROSPERO, an international database of systematic reviews regarding health and social well-being.

### Table 1

**Percentage of State Anti-Bullying Laws That Included Key Policy Components Identified by the U.S. Department of Education**

<table>
<thead>
<tr>
<th>Policy Component</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose of the policy</td>
<td>85</td>
</tr>
<tr>
<td>Applicability or scope of the policy</td>
<td>96</td>
</tr>
<tr>
<td>Prohibition of bullying behaviors</td>
<td>94</td>
</tr>
<tr>
<td>Enumeration of protected social classes or statuses</td>
<td>37</td>
</tr>
<tr>
<td>Requirement for districts to implement policies</td>
<td>98</td>
</tr>
<tr>
<td>Review of district policies by the state</td>
<td>43</td>
</tr>
<tr>
<td>Definition of bullying behaviors prohibited</td>
<td>63</td>
</tr>
<tr>
<td>Procedure for reporting bullying incidents</td>
<td>78</td>
</tr>
<tr>
<td>Procedure for investigating bullying incidents</td>
<td>67</td>
</tr>
<tr>
<td>Procedure for maintaining records of bullying incidents</td>
<td>39</td>
</tr>
<tr>
<td>Consequences for bullying perpetrators</td>
<td>91</td>
</tr>
<tr>
<td>Mental health services for victims and/or perpetrators</td>
<td>28</td>
</tr>
<tr>
<td>Communication of the policy to students, parents, and employees</td>
<td>91</td>
</tr>
<tr>
<td>Training for school personnel on bullying intervention and prevention</td>
<td>85</td>
</tr>
<tr>
<td>Data collection and monitoring bullying of incidents</td>
<td>39</td>
</tr>
<tr>
<td>Assurance of right to pursue legal remedies for victims</td>
<td>39</td>
</tr>
</tbody>
</table>

*Note.* The percentages are based on 46 state bullying laws passed between 1999 and 2011. Source: Stuart-Cassel, Bell, & Springer, 2011.
Search Procedure
A behavioral and social sciences librarian was consulted to assist with developing a search string and identifying relevant computerized bibliographic databases in which to search. The following search string was used to search all databases for studies published between January 1, 1995, and November 8, 2014: school AND bullying AND (law OR policy OR policies OR legislation OR statute) AND (effect OR effects OR effectiveness OR efficacy OR impact OR influence). The search of multiple databases increased the likelihood of identifying all possible studies falling within the scope of the review; thus, the author searched 11 databases, some of which included gray literature sources (e.g., conference papers, government reports, and unpublished papers). Searches were performed in the following databases via EBSCO using terms searched within the abstracts: CINAHL (Cumulative Index to Nursing and Allied Health Literature), Educational Full Text, ERIC (Education Research Information Center), PsycINFO, and Social Work Abstracts. The following databases were searched via ProQuest using terms searched within the titles, abstracts, and subject headings: ASSIA (Applied Social Sciences Index and Abstracts), Dissertations & Theses Full Text, and Social Services Abstracts. In addition, the Conference Proceedings Citations Index was searched using terms searched within titles, abstracts, and keywords. Finally, PubMed was searched using terms searched within titles and abstracts. These more formal bibliographic database searches were supplemented with internet searches using Google Scholar.

Inclusion Criteria
Studies were included in the review if they met the following criteria: (a) collected data and reported results on the effectiveness of policy interventions for bullying in school settings; (b) written in English; and (c) completed since January 1, 1995. Policy interventions for bullying were defined as statutes, policies, regulations, or rules established at the national, state, district, or school levels with the goal of reducing bullying in K–12 schools. Effectiveness referred to the extent to which a policy intervention prevented or reduced student bullying behavior. Given that school bullying policy is a nascent area of empirical inquiry with relatively few empirical investigations and evaluations, the author did not use stringent exclusion criteria in terms of study designs and methods. Only studies written in English were included due to the researchers’ language proficiency. Finally, the time period selected allowed for a comprehensive and contemporary review of the empirical literature completed in this area over the past 20 years.

Study Screening
After performing the bibliographic database searches, 481 results were imported into the RefWorks program to assist with organization and duplicate removal. Fol-
Following duplicate removal, 414 studies remained. An additional 8 studies were added from Google Scholar searches that were not present among the 414 studies. The author and a trained research assistant independently screened each of the 422 studies to determine eligibility. A checklist of the inclusion criteria was created prior to the search and was used for eligibility assessment. Studies had to meet all three inclusion criteria to be screened in. Most studies were included or excluded after reading the title and abstract; however, it was also necessary to examine the full source document of some studies to determine eligibility. To examine interrater agreement, the decisions of the two screeners were compared, and Cohen’s kappa was calculated with SPSS (Version 21), which showed excellent agreement: kappa = 0.97, p < .05 (Landis & Koch, 1977). There were only six disagreements between the screeners, which were resolved by the author examining the source documents. After screening, 401 studies were excluded because they did not meet all of the inclusion criteria. The most common reasons for exclusion included papers that were not empirical, lack of evaluation of effectiveness, lack of evaluation of policy, and studies that were not conducted in schools. After completing the search and screening processes, 21 studies were included for extraction and review (Figure 1).

Figure 1. Flow diagram depicting the identification, screening, and inclusion of studies.
Data Extraction
A data extraction sheet was developed to assist with identifying and collecting relevant information from the 21 included studies. Information extracted included the citation, purpose of the study, study design, sampling strategy and location, response rate, sample size and characteristics, measurement of relevant variables, analyses performed, and results and findings. The author extracted this information and a research assistant then compared the completed extraction sheets with the source documents to assess the accuracy of the extractions. There were only six points of disagreement between the extractor and checker, which they then resolved together by examining the source documents and extractions simultaneously.

Data Synthesis
Initial review of the included studies revealed that a quantitative synthesis, such as a meta-analysis, was not advisable due to the methodological heterogeneity of the studies and differences in approaches to evaluating policy effectiveness. Thus, a narrative thematic synthesis approach was used (Thomas, Harden, & Newman, 2012). The substantive findings on policy effectiveness were first categorized based on the outcome evaluated and then synthesized within each category.

Results
A total of 21 studies were included in this review: 9 peer-reviewed journal articles, 6 research reports that were not peer-reviewed, 5 doctoral dissertations, and 1 master’s thesis. A summary of the methodological characteristics of these studies is presented—including a synthesis of the substantive findings regarding the effectiveness of school bullying policies—in Table S1 (available online).

Methodological Characteristics of the Studies

**Designs.** Of the 21 studies, 12 (57%) used mixed methods, 8 (38%) used quantitative methods, and 1 (5%) used qualitative methods. All studies relied on cross-sectional designs. Most studies (65%) used convenience sampling, whereas the remaining studies used some form of probability sampling. More than half (57%) of studies used national samples, whereas 24% used samples from a single city or local region, 15% used statewide samples, and 5% used samples from areas in multiple countries. Over 80% of studies sampled participants in the United States, with other studies drawing participants from Europe, Australia, East Asia, and Southwest Asia. The most common recruitment sites were schools, followed by listservs, websites, community groups or organizations, professional associations, and personal contacts. Most studies reported participant response rates which varied from 21% to 98%, and the average response rate across studies was 57% ($SD = 29\%$). Eight studies did not report response rates.
**Samples.** Across studies, sample sizes varied from 6 to 8,584 participants. Only the qualitative study had fewer than 50 participants, and two studies had between 50 and 100 participants. Most studies had relatively large samples with more than 500 respondents. The most commonly used participants were students, followed by teachers. Other respondents included administrators, school psychologists, school counselors, education support professionals, and parents. About one third of studies included multiple participant groups (e.g., students and teachers). Most studies (62%) recruited participants from K–12 settings, whereas other studies recruited participants from a single school level: elementary, middle, or high school. Among adult participants, about 75% were female and 90% were White. These percentages are similar to those reported by the U.S. Department of Education, which show that 76% of teachers are female and 82% are White (Snyder & Dillow, 2013).

Samples of students were diverse in terms of race/ethnicity, with most studies consisting of about two-thirds White participants as well as Black, Hispanic/Latino/Latina, Asian, Native American Indian, Middle Eastern, and multiracial students. In addition, student samples were closer to having equal proportions of males and females. Five studies included student participants who were exclusively lesbian, gay, bisexual, transgender, or queer (LGBTQ), whereas 6 studies did not report information about student sexual orientation or gender identity. In addition, studies typically did not measure or report participant national origin, immigrant/citizenship status, religious identity, socioeconomic status, or ability/disability status. Finally, most students were high school students.

**Evaluation methods.** All studies relied on self-report data to evaluate school bullying policy effectiveness. However, studies varied based on the outcome used in their evaluations: Eight studies examined school members’ perceptions of policy effectiveness, 5 studies examined student bullying perpetration and/or victimization behaviors, 6 studies investigated anti-LGBTQ bullying victimization, and 2 studies considered educator intervention in bullying. The level of policies evaluated also varied: Eleven studies examined school-level policies, 3 studies examined district-level policies, 3 studies examined state laws, 3 studies examined both state laws and school-level policies, and one study examined a national policy.

Studies also varied in terms of the analytic approaches used to evaluate effectiveness: Nine studies used bivariate analyses, 8 studies used descriptive statistics of perceived effectiveness, 3 studies used multivariate analyses, and one study used both bivariate and multivariate analyses. Studies that used a bivariate analytic approach compared measures of teachers’ responsiveness to bullying or measures of student bullying between those in schools with and without anti-bullying policies or between schools with high- versus low-quality anti-bullying policies. In these studies, distinctions between high- and low-quality policies were made by the researchers in each study using content analyses of policy strategies that were theoretically and
empirically associated with effectiveness in the bullying literature (e.g., having a
definition of bullying, ensuring adult supervision of students, and outlining conse-
quences for bullies; Ordonez, 2006; Woods & Wolke, 2003). Policy content analysis
scores were then used to distinguish between high- and low-quality policies. De-
scriptive statistical analyses of effectiveness entailed participants responding to a
single self-report item about their perceptions of policy effectiveness (e.g., “How ef-
efective do you feel that your school’s anti-bullying policy is in reducing bullying?”),
with Likert-type response options related to agreement/disagreement or categorical
response options (e.g., yes or no). Multivariate analytic approaches primarily used
student bullying scores as the dependent variable and either a continuous anti-
bullying policy score or a dichotomous variable indicating whether or not the
school had an anti-bullying policy as the independent variable. Continuous school
bullying policy scores were based on either a set of items about the perceived pres-
ence of an anti-bullying policy (e.g., “I think my school clearly set forth anti-bullying
policies and rules”) or a content analysis of policy documents to identify the presence
of criteria or strategies associated with effectiveness (e.g., having a definition of bul-
lying, establishing procedures and consequences for bullies, having educational
events about the school’s bullying guidelines, ensuring adult supervision in school
areas prone to bullying, and formulating a school task group to coordinate anti-
bullying efforts).

The measures used to assess bullying among students varied; some studies used
established scales (e.g., Olweus Bullying Questionnaire), whereas other studies
used items developed by the researchers. The number of items used to measure
bullying varied from 3 to 23 (M = 18.2, SD = 6.1). Of the 11 studies that measured
bullying, the majority measured bullying victimization (n = 8). Only 2 studies
measured both bullying victimization and perpetration, and one study measured
just perpetration. In terms of the types of bullying measured, 5 studies measured
physical, verbal, social, electronic, and sexual bullying; 3 studies measured phys-
ical, verbal, and social bullying; one study measured physical, verbal, social, and
electronic bullying; one study measured physical, verbal, social, and property bul-
laying; and one study measured verbal bullying. In addition to student bullying, ed-
ucators’ responsiveness to bullying was another outcome variable that was used in
8 studies. Only one study used a scale to measure educator responsiveness, and the
remaining 7 studies used one to four items regarding educators responding to stu-
dent bullying.

Results on Policy Effectiveness
Given that the 21 studies differed on the outcomes used in their evaluations of
school bullying policy effectiveness, substantive results are presented by each out-
come category: school members’ perceptions of policy effectiveness, student bully-
ing perpetration and/or victimization, anti-LGBTQ bullying victimization, and educator intervention in bullying.

**Perceptions of policy effectiveness.** Eight studies reported results on participants’ perceptions of policy effectiveness. Results showed that 5% to 88% (M = 49.4%, SD = 33.4%) of educators perceived school bullying policies to be effective to some degree, 4% to 79% (M = 24.5%, SD = 23.6%) of educators perceived policies to be ineffective, and 16% to 70% (M = 51.3%, SD = 30.6%) of educators were uncertain about policy effectiveness (Barnes, 2010; Bradshaw, Waasdorp, O’Brennan, & Gulemetova, 2013; Hedwall, 2006; Isom, 2014; Sherer & Nickerson, 2010; Terry, 2010). Only one study measured students’ perceptions of policy effectiveness, and results showed that they perceived policies to be moderately effective (Ju, 2012). In addition, only one of the 21 studies collected multiple waves of data, although different sets of respondents were used at each of the two waves (Samara & Smith, 2008). In this study, researchers examined perceived effectiveness before and after the passage of an anti-bullying policy; however, there were no significant changes in perceived effectiveness.

**Student bullying perpetration and victimization.** Five studies reported findings on the influence of policy on general student bullying outcomes. Two of these 5 studies examined policy content in relation to effectiveness. One study found that students in schools with high-quality bullying policies reported lower rates of verbal and physical bullying victimization than students in schools with low-quality policies; however, no differences were found for social/relational or property bullying victimization (Ordonez, 2006). In this study, policy quality was evaluated based on the inclusion of the following elements: a definition of bullying; procedures and consequences for bullies; plans for disseminating the policy to students, school personnel, and parents; programs or practices that encourage acceptance of diversity, empathy for others, respect toward others, peer integration, and responsible use of power; supervision of students in school areas prone to bullying (e.g., playground, cafeteria, and hallways); and socio-emotional skills training for victims and bullies (Ordonez, 2006). Similarly, another study found lower rates of verbal, physical, and property bullying victimization among students in schools with high-quality bullying policies, yet higher rates of social/relational bullying perpetration (Woods & Wolke, 2003). In this study, policy quality was evaluated based on the inclusion of the following elements: a definition of bullying; recognition of negative outcomes associated with bullying; discussion of locations where bullying can occur; evaluation of the prevalence of bullying; involvement of stakeholders in policy development; supervision of students in school areas; formulation of a school task group to coordinate anti-bullying efforts; classroom rules about bullying; classroom sessions about bullying; discussion of bullying at PTA/PTO meetings; involvement of parents in bullying prevention efforts; and follow-up with victims and bullies after incidents (Woods & Wolke, 2003).
Other studies examined associations between policy presence and bullying outcomes. Three significant or marginally significant ($p \leq .095$) associations were found: the presence of an anti-bullying policy was inversely related to general bullying victimization, social/relational bullying perpetration, and verbal bullying perpetration (Farrington & Ttofi, 2009; Lee, 2007). Conversely, eight nonsignificant associations were found between school bullying policy presence and scores of general, physical, verbal, and social/relational bullying perpetration, as well as physical, verbal, and social/relational bullying victimization (Farrington & Ttofi, 2009; Khoury-Kassabri, 2011; Lee, 2007). In addition, having a bullying policy was not associated with increases in general bullying perpetration or victimization (Farrington & Ttofi, 2009).

**Anti-LGBTQ bullying.** Six studies with rather large samples of primarily LGBTQ students consistently found that compared to students in schools without an anti-bullying policy or with an anti-bullying policy that did not explicitly prohibit bullying based on sexual orientation and gender identity, students in schools with comprehensive anti-bullying policies that included protections based on sexual orientation and gender identity reported lower rates of anti-LGBTQ bullying, more school personnel frequently intervening when anti-LGBTQ comments were made in their presence, and more school personnel being effective in their anti-LGBTQ bullying responses (Kosciw & Diaz, 2006; Kosciw, Diaz, & Greytak, 2008; Kosciw, Greytak, Diaz, & Bartkiewicz, 2010; Kosciw, Greytak, Bartkiewicz, Boesen, & Palmer, 2012; Kosciw, Greytak, Palmer, & Boesen, 2014; Phoenix et al., 2006). These differences were consistent in analyses of both local anti-bullying policies and state anti-bullying laws.

**Educator intervention in bullying.** Educators play a key role in reducing bullying behavior among students. One study found that compared to those in schools without a bullying policy, educators in schools with bullying policies were more likely to enlist the help of parents and colleagues in responding to a bullying incident and were less likely to ignore bullying (Bauman, Rigby, & Hoppa, 2008). Conversely, a large, national study of educators found no relationship between having an anti-bullying policy and educators’ comfort intervening in both general and discriminatory bullying (O’Brennan, Waasdorp, & Bradshaw, 2014).

**Discussion**

The findings are discussed according to the research questions that drove the review.

**Are Policies Effective at Reducing Bullying?**

Educators were divided in their perceptions of the effectiveness of policies for school bullying; however, on average, about twice as many educators reported that policies were effective to some degree as those who reported that they were not effective. Nonetheless, descriptive summaries of perceptions of effectiveness are typically not viewed as compelling sources of evidence for the effectiveness of
an intervention (Petticrew & Roberts, 2003). However, educators are considered key informants who know what goes on in schools.

Two studies found lower rates of verbal and physical bullying in schools with high- rather than low-quality policies; however, in terms of social/relational bullying, one study found no difference, and another study found higher rates of social/relational bullying in schools with high-quality policies (Ordonez, 2006; Woods & Wolke, 2003). This tentative finding suggests that improving the quality of bullying policies may be effective for direct and overt forms of bullying (e.g., hitting and name-calling) but may not effect social/relational bullying. Across the two studies, elements of policy quality associated with decreases in verbal and physical bullying included a comprehensive definition of bullying; school and classroom rules and procedures about bullying; plans for communicating the policy within the school community; supervision of students across school areas; involvement of parents in anti-bullying efforts; involvement of multiple stakeholders in school-wide anti-bullying actions; and working with and educating students around social, emotional, and behavioral issues to prevent bullying. Extant policies may overemphasize traditional notions of what bullying is (i.e., physical and verbal harassment) and underemphasize or neglect to address more recent understandings of social/relational aggression as bullying. In addition, direct and overt forms of bullying may be more amenable to policy interventions because educators can directly observe these behaviors and then proceed with their response, whereas social/relational bullying often occurs away from the direct supervision of educators (Young, Nelson, Hottle, Warburton, & Young, 2013). Educators have reported difficulty in responding to bullying incidents that they did not witness (Mishna, Pepler, & Wiener, 2006). Similarly, although many educators are aware of cyberbullying, few take steps to address it and many are uncertain about how to confront cyberbullying, which often occurs outside of school (Cassidy, Brown, & Jackson, 2012; Stauffer, Heath, Coyne, & Ferrin, 2012; Vandebosch, Poels, & Deboutte, 2014). Nonetheless, educators can address cyberbullying occurring on or off school grounds if the aggression creates a hostile school environment and substantially disrupts a student’s learning environment (Stuart-Cassel et al., 2011).

Findings among the few studies that examined associations between policy presence and student bullying were mixed, although more nonsignificant than significant associations were found. At first glance, one may conclude from these findings that the presence of bullying policies does not influence bullying among students; however, the presence of a policy is necessary but is not sufficient to affect student behavior. Indeed, after a policy has been adopted, it must be put into practice. The mere adoption or presence of a policy does not mean that it will be immediately and consistently put into practice exactly as intended. The implementation of a policy is a complex, dynamic, and ongoing process involving a vast assortment of people,
resources, organizational structures, and actions. No study that examined the implementation of school bullying policies found that the policies were being implemented precisely as intended (Hall & Chapman, 2016a, 2016b; Hedwall, 2006; Holmgreen, 2014; Jordan, 2014; LaRocco, Nestler-Rusack, & Freiberg, 2007; MacLeod, 2007; Robbins, 2011; Schlenoff, 2014; Smith-Canty, 2010; Terry, 2010). Indeed, the extent of faithful implementation in these studies varied considerably by location and policy component. Therefore, fidelity of implementation (i.e., the extent that a policy is put into practice as intended based on the directives expressed in the policy document) may mediate the relationship between policy adoption or presence and the targeted policy outcome of student bullying. However, none of the studies reviewed measured policy implementation fidelity. Thus, one can conclude from this evidence that in some cases, policy presence was associated with decreases in bullying; in other cases, however, there were no such associations. Because data on implementation were not collected in any study, it is not known if the lack of significant associations was related to lack of faithful implementation of policies.

One area of consistent agreement in the findings relates to the benefits for LGBTQ students who are in schools with anti-bullying policies that explicitly provide protections based on sexual orientation and gender identity. These benefits included lower rates of victimization and higher rates of intervention by educators. Numerous studies have demonstrated that LGBTQ youths experience high rates of bullying victimization (Berlan, Corliss, Field, Goodman, & Austin, 2010; Espelage, Aragon, Birkett, & Koenig, 2008; Kosciw & Diaz, 2006; Kosciw et al., 2008; Kosciw et al., 2010; Kosciw et al., 2012; Kosciw et al., 2014; McGuire, Anderson, Toomey, & Russell, 2010; Varjas et al., 2008). However, only 20 states (40%) have enumerated protections based on sexual orientation and gender identity/expression in their anti-bullying laws (Human Rights Campaign, 2015). Given the evidence for the effectiveness of enumerated policies, all policies should prohibit harassment and bullying based on sexual orientation and gender identity.

Aside from the LGBTQ-focused studies, only two other studies examined educators’ responsiveness to bullying. Findings from these studies were somewhat contradictory, as one found a connection between having a bullying policy and responding to a bullying incident, whereas the other study found no relationship between having a policy and educators’ comfort in responding to bullying. However, the study that found no relationship included several other relevant independent variables (i.e., receiving training on how to implement the school’s bullying policy and having resources available in the school to help educators intervene), which were significantly associated with increased comfort in responding to bullying (O’Brennan et al., 2014). Thus, the relationship between the presence of a school bullying policy and educators’ responsiveness to bullying incidents may be mediated by training about putting the policy into practice and having resources available for intervention.
Finally, there was no evidence that one level of policy was more effective than another. Across the studies, school, district, and state policies all showed evidence for effectiveness as well as ineffectiveness. Policies do vary in terms of their weight in law. For example, a state statute has more legal force than an informal school policy established by a principal. Nonetheless, a school policy set by a principal is more proximal than a state policy, and therefore, the proximity may facilitate implementation of the policy at the school. Policy level may not be related to effectiveness. What likely matters more in terms of effectiveness are the strategies contained within a policy and the ways they are implemented.

What is the State of the Research on School Bullying Policy Effectiveness?
Systematic reviews summarize what is substantively known about a topic area and also provide a state of the research on a particular topic. Research to date on school bullying policy effectiveness has several strengths. In terms of designs, most studies have used a mixed-methods approach, which is advantageous because it capitalizes on the strengths of both quantitative and qualitative research and offsets weaknesses of using one or the other. Including quantitative methods allows for precise, numerical estimates related to distribution or the strength and direction of relationships, and including qualitative methods allows for rich, in-depth data related to context or complexity. Other strengths are related to sampling: More than one third of the studies used some form of probability sampling, over half of the studies used national samples, and many studies reported high response rates. These sampling strengths are beneficial in terms of generalizing findings. Also, almost all studies had sample sizes greater than 200, and two thirds of studies had large samples (i.e., approximately 500 to 8,500 participants). Larger samples can be more representative of a population and are beneficial in terms of statistical power. A final strength was that many studies collected data from multiple participant groups (e.g., teachers and students). Having multiple participant groups allows for a more comprehensive assessment and the triangulation of data sources, which can be used to compare and contrast findings and may help researchers corroborate findings.

On the other hand, several prominent methodological limitations were identified among the studies reviewed. First, the studies relied on evidence from cross-sectional surveys, which are vulnerable to selection bias and confounding. In addition, cross-sectional studies cannot examine a key criterion of causality: a temporal relationship wherein an anti-bullying policy was adopted and implemented, which then led to decreases in bullying over time. Second, most studies used convenience sampling. Although convenience sampling may be highly feasible and efficient, it can lead to the underrepresentation or overrepresentation of particular groups within a sample. Thus, convenience samples may not be representative of the populations of interest, which undermines the generalizations that can be made from
the findings. Third, most of the studies used descriptive statistics or bivariate analyses to evaluate the effectiveness of bullying policies. Such analyses can be oversimplified and leave out relevant explanatory or contextualizing variables. In addition, some of the studies that used bivariate analyses did not report the exact statistical test used (e.g., independent groups t-test and chi-square test) or effect sizes and instead focused on substantive findings. Although these reports seemed to be aimed at a more general, nonscholarly audience, the omission of this information can become problematic in understanding the methods used and drawing conclusions about the results. Fourth, many studies asked participants to report whether their school had an anti-bullying policy. This question might be problematic for student respondents because they might not know about the policies in their schools.

A final limitation involved the measurement of bullying. The main goal of policy interventions for bullying is to prevent and reduce bullying behavior among students. Thus, studies evaluating the effectiveness of these interventions should measure bullying among students as a primary outcome. Nonetheless, only half of the studies directly measured student bullying, and most of these studies did not measure both bullying perpetration and victimization. Policies are aimed at influencing multiple actors involved in the bullying dynamic, which includes bullies, targets, victims, bully–victims, bystanders, parents, and school personnel. Thus, studies that do not measure bullying perpetration and victimization among students are not assessing the two main targeted behavioral outcomes of anti-bullying policies. In addition, bullying behaviors can manifest in many forms, including physical bullying, verbal bullying, social/relational bullying, cyberbullying, property bullying, and sexual bullying (Hall, 2016). However, none of the studies in this review measured all of the dimensions of bullying.

What Future Research is Needed on School Bullying Policy Effectiveness?

Undoubtedly, research on the effectiveness of policy interventions for school bullying will continue to expand. In order to build upon and address gaps and limitations in the extant literature, six recommendations are presented for future research on school bullying policy effectiveness. These recommendations are based on the critical analysis of studies in this systematic review.

First, future studies should employ more rigorous designs to evaluate the effectiveness of policy interventions for bullying. The randomized controlled trial (RCT) is the “gold standard” approach for measuring the impact of an intervention; however, RCTs are often infeasible for evaluating public policy interventions due to the political and legal nature of policies, which are implemented across large organizational systems and typically with prescribed timelines (Oliver et al., 2010). Thus, researchers may need to rely on other rigorous and feasible designs for evaluating policy effectiveness: pretest/posttest cohort designs, pretest/posttest matched comparison group designs, and interrupted time series designs (Oliver et al., 2010;
Shadish, Cook, & Campbell, 2002). These study designs are superior to cross-sectional studies in determining the effectiveness of interventions (Coalition for Evidence-Based Policy, 2003; Petticrew & Roberts, 2003; Pilcher & Bedford, 2011).

Second, studies should collect data on outcomes and the implementation of policy components. None of the studies assessed implementation fidelity. When bullying policies do not successfully achieve targeted outcomes, we do not know whether those policies were implemented as intended and failed or whether lack of implementation fidelity is to blame. Implementation data, if collected, could be used to ensure that policies are being activated as intended with high levels of fidelity and reported along with outcome evaluation data in the study designs mentioned previously. These data also could be used to examine the predictive relationship between implementation fidelity and outcomes. Theory would suggest an inverse relationship where higher levels of implementation fidelity are associated with lower levels of bullying among students; however, this remains an untested hypothesis. Also, bullying policies are comprised of an array of directives to be put into action. Data on the fidelity of implementation of all components of an anti-bullying policy would allow researchers to examine the relative or combined impact of policy components on outcomes.

Third, analyzing policy content—versus only considering the presence of absence of a bullying policy—is needed for more nuanced understanding of which policies work, for whom, and why. A national review of state anti-bullying laws showed broad inclusion of some policy components (e.g., outlining the consequences for students who bully) and limited inclusion of other components (e.g., providing mental health services to perpetrators or victims of bullying; Stuart-Cassel et al., 2011). Theoretically and empirically based guidance about specific actions that can be prescribed in bullying policies is small but growing (Cornell & Limber, 2015; Nickerson, Cornell, Smith, & Furlong, 2013). Future research should analyze the relationships between policy content and bullying outcomes, which could help identify the most influential policy components. Examining only policy presence or absence is insufficient because a school district may indeed have an anti-bullying policy, but its content may not be evidence-based. Policies can also vary in the way they are written, as some policies are lengthy, vague, and contradictory, whereas other policies are clear, concise, and specific. This area of content could also be analyzed and may relate to educators’ comprehension of policies, which would influence implementation actions by educators, and subsequently, policy outcomes.

Fourth, future studies should use multivariate and multilevel analyses. The effectiveness of policy interventions for bullying are influenced by several variables, including policy content, fidelity of implementation, and school environmental factors. By using more complex statistical methods (e.g., regression modeling, structural equation modeling, propensity score matching, and hierarchical linear mod-
eling), researchers will be able to examine the influence of multiple variables, examine moderating and mediating relationships, control for extraneous variables, match intervention participants with control participants, and account for clustered data (e.g., students or teachers nested within schools). These statistical methods will be essential to execute the recommended study designs and analytic methods described previously. The use of these statistical methods will help ensure the integrity of future findings on policy effectiveness.

Fifth, studies should improve sampling practices. To attain more representative samples, researchers should partner with school districts, state departments of education, and departments of public instruction, and they should employ some form of probability sampling. Many of the studies in this review that used probability sampling involved data collection collaborations with state- and district-level educational agencies. Educational agencies have a vested interest in the implementation and success of bullying policies, especially those codified as law. In addition, future studies should sample from multiple respondent groups—such as administrators, teachers, school mental health professionals, and students—to gain a more comprehensive and multiperspective understanding of the implementation and effectiveness of school bullying policies. Researchers also should sample across the K–12 spectrum because state and district policy guidelines typically apply across these grade levels. Yet, there may be differences in policy effectiveness between elementary, middle, and high school. Certain policy strategies also may need to be tailored based on student developmental differences and differences in school structure across the K–12 system.

Finally, future studies should use scales to measure both bullying perpetration and victimization, and these measures should assess all of the dimensions of bullying: physical, verbal, social/relational, electronic, sexual, and property bullying. Researchers may find that policies are more effective at addressing certain types of bullying than others (e.g., direct vs. indirect bullying). Multifactor scales with a sufficient number of items are needed to measure the full range of bullying behaviors. The Centers for Disease Control and Prevention created a compendium of bullying measures that is available to the public (see Hamburger, Basile, & Vivolo, 2011). However, caution should be taken in selecting instruments because some measures have low internal consistency reliability values (i.e., $\alpha < .70$), low test-retest reliability coefficients (i.e., $r < .70$), no recall time frames, overly long and complex definitions of bullying, limited evidence of construct validity, limited evidence of criterion validity, and limited evidence regarding respondents’ understanding of the measure’s instructions and items (Hall, 2016). In addition, as opposed to questionnaires about bullying behaviors, peer and/or teacher nomination methods to identify students who are bullying victims or perpetrators may be more developmentally appropriate for elementary school-age children.
Strengths and Limitations of the Review
This review used a rigorous approach to identify relevant studies by searching 11 databases using an expert-informed search string. In addition, search records were independently screened by two screeners based on a priori inclusion criteria. Further, research reports and dissertations (forms of gray literature) were included to minimize publication bias. Nonetheless, unpublished research may be underrepresented in this review. Another limitation relates to the variability of studies: Studies varied in the respondents, sample locations, the types of policies examined, and the ways effectiveness was evaluated. This variability presented challenges for combining and comparing results. Another limitation of this review relates to the methodological limitations of some of the included studies. However, by presenting the methodological characteristics and substantive findings by study in Table S1 (available online), readers are able to assess the methodological rigor and trustworthiness of findings accordingly.

Conclusion
Bullying is a widespread problem in which about half of students are directly involved and up to 90% of students are indirectly involved (Atlas & Pepler, 1998; Cook, Williams, Guerra, & Kim, 2010; Craig & Pepler, 1995; Glew et al., 2005; Hawkins et al., 2001). Policy interventions are an approach to bullying that establishes legal mandates for schools, influences the behavior of students and school personnel, and guides the implementation of other targeted interventions within schools. Findings on the effectiveness of policy interventions for bullying are primarily mixed, and there are limitations in the evaluation methods used. Research on school bullying policy will undoubtedly continue to expand with the growing understanding of the need for evidence-based education policies and as bullying policies continue to be introduced and revised in schools across the globe. Future research must use more rigorous methods and designs and may indeed find that policy interventions play a key role as one of a constellation of intervention strategies for preventing and reducing school bullying.

Author Notes
William Hall, PhD, MSW, is a postdoctoral fellow at the University of North Carolina at Chapel Hill. Correspondence regarding this article should be directed to William James Hall, 325 Pittsboro Street, CB #3550, Chapel Hill, NC 27599-3550 or via e-mail to wjhall@email.unc.edu

Acknowledgments
I would like to thank Mimi Chapman, Natasha Bowen, Barbara Fedders, Mark Fraser, and Kathleen Rounds for their advice and feedback regarding this paper. I also thank Rachele McFarland for her research assistance. The author was supported by the National Research Ser-
vice Award Postdoctoral Traineeship from the National Institute of Mental Health, sponsored by Cecil G. Sheps Center for Health Services Research, University of North Carolina at Chapel Hill, and the Department of Psychiatry and Behavioral Sciences, Duke University School of Medicine (grant number T32 MH019117).

References

* Asterisks indicate studies that were included in the systematic review.


*Barnes, C. M. (2010). Bullying behavior: Perspectives on implementation of policy from building level administrators and school counselors (Unpublished doctoral dissertation). University of Memphis, Memphis, TN.


Smith-Canty, V. (2010). An analysis of the implementation of the South Carolina anti-bullying legislation in the middle schools involved in the Abbeville, South Carolina, school district lawsuit (Unpublished doctoral dissertation). South Carolina State University, Orangeburg, SC.


Manuscript submitted: March 30, 2016
Revision submitted: July 7, 2016
Accepted: July 27, 2016
Electronically published: January 26, 2017