Abstract
Drawing on theories of social norms, we study the relative influence of female and male students using a year-long, network-based field experiment of an anti-harassment intervention program in a high school. A randomly selected subset of highly connected students participated in the intervention. We test whether these highly connected females and males influenced other students equally when students and teachers considered the problem of “drama”—peer conflict and harassment—to be associated with girls more than with boys. Exposure to male, but not female, intervention students caused decreased perceptions of the acceptability of harassment and decreased participation in negative behavior. Status beliefs became activated through the intervention program: gender differences in influence stem from higher levels of respect afforded to highly connected males in the program. The results support an account of social influence as it occurs across time in conjunction with other group processes.

Keywords
social influence, gender, social networks, norms, adolescents, bullying

Patterns of peer harassment in middle and high schools have received enormous public and scholarly attention over the past several years (e.g., Bazelon 2013; Juvonen and Graham 2014). In contrast to accounts that focus on students’ character or inherent features of their developmental stage, we adopt a context-based explanation based on theories of collective norms and social influence in a network. This focus directs our attention to the social influence processes that generate aggressive behavior and harassment.

According to theories of collective norms, individuals’ behavior is driven by their perceptions of what behaviors are considered typical or desirable by their peer groups (e.g., Cialdini and Goldstein 2004; Hechter and Opp 2001). Perceptions of typical and desirable behaviors within a group are perceived collective norms. Individuals quickly and easily perceive cues regarding what others do (e.g., drinking on a college campus) and adjust their behavior in light of these cues (Prentice and Miller 1993). Behavioral patterns in groups result from a cycle of influence.
among group members that unfolds over time as group members interact with each other, observe the behavior of others, and adjust their behavior accordingly.

Adolescents conform to the behaviors that they perceive to be desirable and typical in order to facilitate friendships in the group (Dornbusch 1989). Students with friends who participate in bullying or aggressive behavior are more likely to participate themselves as they model their own behavior on that of their peers (Mouttapa et al. 2004). Girls and boys who participate in peer conflict and harassment at school are responding to cues about the social standards of their peers at school. A collective norms account is also consistent with findings that peer aggression is driven by struggles over status, where the majority of students participate in order to establish their place in the school hierarchy (Faris and Felmlee 2011). Students observe others participating in these status struggles and make inferences about the acceptability of those behaviors at the school. In sum, harassment is not driven primarily by individual characteristics but by the “rules of the game” that prevail in a given social environment.

Given the importance of the “rules of the game,” who influences these rules? That is, which students have the most influence over the perceived collective norms and thus behaviors in a school? We investigate this question in the case of “drama,” which students and staff defined as gossip, rumors, exclusion, mean looks in the hallway, and occasional physical confrontations. In line with wider stereotypes (e.g., Males and Chesney-Lind 2010), students and adults at the school associated drama with girls more than with boys. At the same time, previous research highlights the differential ability of males to exert influence in groups (see Carli 2001). We conducted a novel, year-long field experiment in which we randomize a subset of all highly connected male and female students in a high school network to lead an in-school intervention in which the students publicly indicate their opposition to peer harassment. We test whether male and female students affect their peers’ perceptions of school norms about harassment and their actual harassment behavior to the same extent.

The innovative experimental design moves beyond correlation-based insights into social influence and change in behavioral patterns in social networks (see also Aral and Walker 2011; Centola 2010; Valentе 2012). We extend current research on gendered influence to examine how gender affects social influence processes in a network in a natural setting. The study allows for a more expansive conceptualization of social influence that includes influence from multiple sources (e.g., many students whose behavior changes and who interact with each other) that extends across time (in this case, a school year) as relationships change. We provide an explanation for gendered influence—differential respect afforded to male and female students—as students develop opinions about whom they respect across time. Studies of single instances of influence in laboratory settings cannot account for the dynamics of social influence that unfold over time and as individuals interact with each other (Mason, Conrey, and Smith 2007).

Inferring Collective Norms and Desirable Behavior from the Cues of Social Referents

Using data from this project, we previously established that social referents—highly connected individuals whose behavior can serve as a reference point for group members to infer the norms of the group as a whole (Sherif and Sherif...
1964)—can exert particular influence on the perceptions and behavior of other group members. Specifically, using a randomly assigned subset of social referent students, we show that exposure to the behavior of highly connected individuals influences individuals’ perceptions of collective norms and their own behaviors (Paluck and Shepherd 2012). These social referents provide strong cues to typical and desirable behaviors within the overall group because of their high observability in the network and the awareness that other group members also pay attention to their behaviors (cf. Watts and Dodds 2007 on the role of influentials in diffusion). Do male or female social referents exert greater influence over the perceived norms and behavior of their peers over time, in particular when the behavior in question is associated with females?

**Gender and Social Influence**

A variety of theoretical perspectives suggest that we should expect females to exert more influence than males in the school examined here. First, most studies examining the relative influence of males and females over others find that men exert greater influence over others than do women, except in domains that are associated with women, where women exert more influence (Carli 2001). Because the behavior under study is typed as female behavior within the group, we would expect the behavior of female social referents to serve as a stronger cue to peers toward perceived norms of drama, compared to male social referents. According to this prediction, female social referents would be perceived as more informative on the norms for drama in the school than male social referents (e.g., Feldman-Summers et al. 1980; Gerrard, Breda, and Gibbons 1990).

The proportional gender composition of the group also affects the degree of influence of males and females. When the composition of the group is gender-balanced, women are more able to exert influence on others than when they are in the minority (Craig and Sherif 1986; Taps and Martin 1990). When girls are in the majority in a school, they become the reference group for attitudes and behaviors (Demanet et al. 2013). In the school under study, the majority of students were female: 57 percent of the 291 students. Females’ status as the majority may lead students to be more able and likely to call to mind female students when thinking of a “typical” student. Female students may therefore serve as the prototype—a representative exemplar—of students at the school (e.g., Mullen 1991). Students may use and remember the behavioral cues from prototypical students (girls) more than those from other students when forming representations of collective norms (Hogg and Reid 2006). Thus, both the gendered nature of drama and the majority status of females in the school network lead to the prediction that female social referents will exert more influence over the norm perceptions and behaviors of their peers than would male social referents.

By contrast, a number of theories might suggest that male social referents may be more influential than female social referents with respect to drama in the school. In some studies, men who are in the minority exert disproportionate influence on the group (Craig and Sherif 1986; Williams 1995). An extension of this finding to this case is that boys’ behavior may become more salient by virtue of their small numbers, in both absolute and relative terms. A related explanation is that the behavioral cues from male social referents may stand out more to their peers by virtue of being incongruent with expectations about who participates in typical peer harassment behaviors. Findings from cognitive
psychology establish that individuals pay more attention to unexpected or novel stimuli and under some conditions are more likely to remember them (Fiske and Taylor 1991; Griffiths, Johnson, and Mitchell 2011). Boys may be just as involved in harassment as girls, but because they are not cognitively associated with drama and harassment to the same extent as are girls, male social referents’ behavior or publicly expressed beliefs about harassment may draw greater attention.\(^1\)

Additionally, males are typically afforded more respect than are females (e.g., Ridgeway 2011). According to expectation states theory, when differential status becomes associated with social categories, as in the case of gender, beliefs about competence and authority become associated with the social categories themselves (Berger, Cohen, and Zeldich 1972). Gender-based status beliefs—widely shared beliefs about the relative competence, authority, and esteem of individuals in particular social categories—guide individuals to infer that males are more competent and worthy of respect than females in many domains, not only those associated with their gender (Ridgeway and Correll 2004). For example, individuals in groups more often listen to males when they speak and take their comments more seriously, which makes their points of view more influential in group decision making (Kollock, Blumstein, and Schwartz 1985; Ridgeway et al. 2009). This perspective suggests that even when a behavior is associated more with girls, male social referents would exert more influence over their peers because they are afforded greater respect.

The majority of the studies described previously involve deliberate persuasion attempts that occur once or twice in small groups or in dyadic interactions, often within a laboratory setting. What these studies cannot address is how gendered influence unfolds across a longer period of time, where individuals have access to a much broader scope of information about the topic of persuasion (e.g., peer harassment) and much more interaction with those doing the persuading (the social referent students). An experiment that occurs over months, instead of over hours or days, sheds light on how gender differences in social influence occur alongside developing status beliefs about gender.

### Hypotheses

Existing evidence provides a number of arguments regarding why highly connected students who were randomly selected to participate in the intervention program (hereafter called intervention referents) may exert differential influence over their peers, depending on their gender. One hypothesis is that female intervention referents will exert greater influence on their network neighbors’ normative perceptions than will male intervention referents, such that:

**Hypothesis 1:** The more exposure a student has to female intervention referents, the less they will perceive drama to be acceptable and desirable.

We also predict influence on the behavior of network neighbors:

**Hypothesis 2:** Greater exposure to female intervention referents will lead students to participate less often in

---

\(^1\)Girls committed 28 percent of peer harassment disciplinary infractions, but they were brought into the office for conflict resolution sessions far more frequently than were boys (school counselor, personal communication, June 2011). However, because we are interested in how students arrive at the perception that drama is a widespread and desirable behavioral strategy, whether girls in the school we study actually contributed to harassment at higher rates than boys is not our central concern.
The three reasons that we might observe these relationships (female students are perceived as more informative regarding drama, are seen as more representative of the school, and have the opportunity to provide proportionally more frequent cues as the majority in the school) may contribute differentially to these effects or may work together to produce these outcomes.

Alternatively, male intervention referents may exert greater influence than female referents on their network neighbors’ norm perceptions, such that:

**Hypothesis 3:** The more exposure a student has to male intervention referents, the less they will perceive drama to be acceptable and desirable.

Additionally, in terms of behavior, we would predict that:

**Hypothesis 4:** Greater exposure to male intervention referents will lead students to participate less often in harassment-related behavior and more often in prosocial behavior.

The three possible sources of these effects (male behavior being more salient because they are not the gender group most commonly associated with the behavior, because they are a numerical minority in the school, and because of the greater social status commonly afforded to males) may also contribute differentially to the effects or all work together.

**METHOD**

**Experimental Context**

We studied a small public high school (N = 291) that taught grades 10th to 12th; it drew students from urban and suburban areas of Connecticut. The school was a magnet school; many of these students moved to the school after poor experiences in another school either due to their own behavior or due to how they were treated by other students. At this school, both staff and students perceived that girls were the most frequent participants in drama. For example, in response to a question regarding how to explain the school to a new student, one student said,

Girls are always starting drama. . . . You fight with friends here [if they are fighting, you fight alongside them], back them up. Girls say what is on their minds. For guys, they are more laid back, but they have quick tempers and more instant confrontations. (Student KP, personal communication, May 2011)

Other research has identified this form of peer harassment, often called relational conflict, as occurring particularly among girls and physical aggression as occurring more frequently among boys (Archer and Coyne 2005; Bjorkqvist 1994). In response to observations regarding high levels of harassment among students and particularly among girls, the school commissioned the Anti-Defamation League (ADL) to run an intervention program called “Names Can Really Hurt Us” (NAMES). The NAMES intervention prepares a small group of selected students to present their experiences of and reasons to oppose harassment (verbal or physical abuse and social ostracism) in a school-wide assembly.

Prior to the intervention program, we analyzed the school’s social network using surveys at the beginning of the school year to identify a pool of social referent students. We then randomly assigned a subset of those social referents to participate in the program. Participation included leading a school-wide assembly in the fall and providing reminders about
the themes of the program through publicity campaigns during the remainder of the school year. We first describe the school-wide survey to illustrate how we identified the social referent students within the school’s social network while measuring students’ perceived norms, beliefs, and experiences of harassment prior to the intervention.

School-Wide Survey: Social Network and Norms Measurement

One week after school began in September, we administered a survey during a single class period to every student in the school (N = 260 due to absences on the survey day and in the days following; 57 percent female; 42 percent African American, 27 percent Latino, 22 percent white; 59 percent of students were new to the school that year; between 40 percent and 50 percent of students received free or reduced-price lunches). Parents signed a consent form for their child to participate, and students also provided informed consent. The survey consisted of four parts: demographic information, questions about relationships with other students (social network questions), personal beliefs about and experiences of harassment-related events at the school, and perceptions of collective social norms regarding harassment at the school. Based on previous work on harassment in schools and on initial qualitative work in which students referred to harassment as “drama,” we adopted the term starting drama to refer to these issues in all relevant survey questions. We defined drama in the survey as “talking behind the backs of other students or to their faces in a mean or rude way; spreading rumors by text, Facebook, MySpace posts, or IMs; giving other students mean or rude looks in the hall.”

Network questions. Six questions investigated students’ relationships with other students. Four asked about friendships at the school in behavioral terms (“With whom did you spend time in the last week?”; “With whom did you communicate online last week?”; “Who would defend you if you were having drama with other students?”; “Who would you talk to if something bad or upsetting happened to you?”). Two questions elicited nominations of high prestige peers (i.e., students “who you really respect” and “who you think are most popular”). We provided each student with a complete roster of students in the school, arranged by grade, sorted alphabetically by first name. Each student was assigned a number based on the order of the list, and students used the numbers to nominate an unrestricted number of students for each question.

We use the spending time together question, measured two more times throughout the year, to approximate the frequency with which a student is exposed to the behavior of their peers. In particular, we use this question alone to map the social interactions through which we trace the influence of the intervention referents. We use the number of respect nominations, also measured two more times throughout the year, as a measure of relative status of students in the school.

Identification of social referent students. We used social network questions from the first survey to construct the complete network of relationships among students at the school. We combined the results of the four friendship-related questions to form one network matrix. If two participants, p_i and p_j, shared a relationship based on any of the four friendship-related questions, the value of their intersecting cell, p_ij, was one. In a different network matrix, we followed the same procedure to combine the results of the two prestige questions. We combined questions in order to create robust
networks that captured several dimensions of relationships and because some questions elicited higher response rates than others.

Within the prestige matrix, we identified two types of students: students who were nominated as being high prestige by many other students (i.e., high indegree in the prestige matrix) and whose prestige matrix associates did not have many prestige connections to one another (i.e., low local clustering coefficient in the prestige matrix). This metric verifies that widely known students serve as social referents for a wide swath of students, rather than for a smaller, interconnected group of students; we refer to these students as “widely known.” Second, we identified students who received many friendship nominations (high indegree in the friendship matrix) and whose connections also shared friendship connections to each other (high local clustering coefficient). This method of designating students, referred to as “clique leaders,” indicates that they served as leaders of tightly interconnected groups in the school.2

Random assignment of social referents to intervention. The final pool contained 83 eligible social referent students: 42 widely known students and 41 clique leaders. We stratified the pool by gender and grade level and used a random number to select 30 students, 15 of each type, to participate in the intervention. Six of these students refused participation in the program, leaving 24 social referent students who participated in the intervention (16 girls and 8 boys; equal numbers of sophomores, juniors, and seniors; 4 students who identified as white, 11 who identified as black, 7 who identified as Latino, 1 student who identified as biracial, and 1 student who declined to report his race) and 53 control social referent students who did not participate in the intervention (31 girls and 22 boys; 19 sophomores, 11 juniors, and 23 seniors). In the following, when we refer to control social referents, we mean those social referent students who were not randomly selected for the program; by intervention referents, we mean social referent students who were randomly assigned to participate in the intervention from the pool of eligible social referents.

Collective norms. We used a series of five questions to measure perceptions of prescriptive norms regarding drama, specifically perceptions of student approval of drama, and of behavior that can deescalate drama. We averaged the values for these five questions to create one composite measure. Cronbach’s alpha for the five questions was .70. We assessed an overall prescriptive norm of harassment: “How many students at [school] believe it’s normal when students start drama or any other kind of conflict with other students?” For norms regarding behavior to deescalate conflict, we asked: “How many students at [school] believe it’s wrong, or would criticize you, if you tried to stop other students from starting drama?”; “. . . believe it’s wrong, or would criticize you, if you did not defend your friends when someone else was making drama for them?”; “. . . believe it is important to defend your friends when someone is making drama for them?”

2See Appendix A of the online supplemental materials for more details about the selection of intervention-eligible students.

3Interviews with students indicated that behaviors in which students defend their friends or step into the middle of conflict do not necessarily de-escalate conflict. Instead, teachers and students identified behaviors like stepping back and not grouping together to defend friends as behaviors that prevent future conflict.
Students responded to these questions using a pictogram with six options, each of which featured a collection of outlined figures. The proportion of shaded figures in each picture represented the percentage of students who believed or supported the statement (e.g., 1 = nobody, no shaded figures; 3 = about 50 percent, half shaded figures; 6 = almost everybody, all shaded figures). By using percentages, we capture students' perceptions of the sentiments of the collective as opposed to sentiments of each individual's idea of a prototypical group member.

A successful intervention would create the perception of less widespread approval for behaviors supporting drama and therefore smaller scores on the composite norms scale at later waves. Students would perceive behaviors that deescalate conflict such as stopping others from "starting drama," refusing to participate in the conflicts of friends, and ignoring instead of engaging rumors about oneself to be more desirable and normal. We consulted with school staff and students in order to capture school-specific issues and appropriate language.

**Second and Third Wave Surveys**

We administered the social network and norms survey described previously to the entire school two more times: one week after the intervention described in the following (N = 250, 57.2 percent female) and at the end of the school year (N = 220, 58.18 percent female). A total of 278 unique students completed at least one wave of the survey; 190 students completed all three waves.4

---

4Pretest levels of perceived norms do not predict missingness in Waves 2 and 3, and posttest levels of perceived norms did not predict missingness in the first wave. Given the lack of association between missingness and our key outcome variables, we rule out selectively missing data as a driver of our findings.

**Behavioral Outcome Measures**

**Behavior reported by teachers.** In a survey administered before the start of the NAMES program and at the end of the year, teachers and administrators used the school roster to nominate students "who defend other students when they are being harassed or picked on," which we consider a prosocial measure of attempting to support other students and stop the cycle of drama, and students who "cause the most trouble or contribute to a negative school environment." Students received a point in these categories for each nomination from a teacher or administrator. To the extent that teachers and administrators pay attention to the everyday behaviors of their students, these measures provide an overall assessment of student behavior that captures even subtle harassment behaviors.

**Disciplinary records.** We obtained the complete school records of all disciplinary events receiving administrative attention throughout the school year. We combined codes for "peer harassment" and "disruptive behavior" toward peers to create a variable for whether a student was disciplined for peer harassment at school.

To test whether the intervention had an effect on poor behavior more generally, we also combined all reported disciplinary events for each student into a count variable. As mirrors previous research, boys were significantly more likely to be cited for disciplinary infractions overall and for harassment-related events than were girls (e.g., McFadden and Marsh 1992). While formal disciplinary reports capture only a small percentage of harassment behavior, they provide one measure of how frequently students are participating in peer harassment broadly defined and poor behavior generally.

**Intervention**

The NAMES assembly program functions as a platform to broadcast certain
students’ experiences with and reactions to harassment to the student body and to facilitate public discussion about harassment among students. Intervention referents first participated in two training sessions to prepare for the assembly. A facilitator from the ADL led activities that prompted reflection on the nature and effects of harassment at their school. Teachers from the school, in consultation with ADL facilitators, selected five essays written by the intervention referents about their own experiences of harassment to be read by the student authors at the assembly. The essays were selected to represent the perspective of students who had been both targets and perpetrators. The other intervention referents wrote a skit illustrating common types of harassment at the school and ways to speak out against them.

On the day of the assembly in October, the intervention referents performed the skit about a rumor spreading about a girl being a “slut,” something that the intervention students identified as a common form of harassment at the school. In the concluding scene, another girl defends the girl who has been targeted. After the skit, the five intervention referents selected to read their essays, three girls and two boys, did so. One girl’s essay described switching elementary schools because a girl had mobilized her group of friends to continuously harass her, while another girl spoke of her own participation in making fun of other students. One of the boys talked about getting in a physical fight at school, which perpetuated a cycle of aggression. In between the intervention referents’ performances, an ADL representative spoke about the effects of harassment.

At the end of the assembly, there was an open microphone session in which any student could share their own experiences, and dozens of students did so. After the assembly, all students were divided into small groups to discuss the assembly, supervised by intervention referents and adults. Follow-up events during the school year reinforced the association between the intervention referents and anti-drama messages. Intervention referents read announcements about the consequences of harassment over the loudspeaker during morning announcements, designated a special “NAMES” table at lunch period two times during the year where they spoke with other students about ways to report harassment, and created a series of posters featuring photos of intervention referents and one of several anti-drama and pro-inclusion slogans such as “Whatever your story, I’ll listen” and “People who spread rumors are no friends of mine.”

Though all students in the school were exposed to the initial assembly program, we are interested in the causal effect of everyday exposure, represented by spend time network ties (ties measured by the question “with whom did you spend time in the last week” described previously), to intervention social referents on norm perceptions and behaviors.

**ANALYTIC APPROACH**

We test the effects of the randomly assigned female and male intervention referents on their peers in the school using linear fixed effects regressions. In order to measure the effect of the number of social network ties to female and male intervention referents on students’ perceptions of collective norms and their behavior, we use students’ reports of how many female or male intervention referents

---

5We observed no systematic differences between male and female intervention referents in the content of their contributions in the assembly, in the small group sessions, or in participation in follow-up intervention activities (self-reported means of 7.5 activities for boys and 6.06 activities for girls, \( p = .19 \)).
referents they spent time with in the last week. We account for their ties to control students. See Figure 1 for a representation of the design. For the purposes of the analyses, we consider all treatment social referents’ participation in the intervention to be equivalent.\(^6\)

In our regressions, we examine the effect of each student’s treatment dosage, measured in terms of the number of the student’s direct spend time ties to randomly assigned female and male intervention referents, controlling for the student’s total number of ties. We use students’ own nominations of the students they spent time with (their outdegree) because this measure best represents the relevance, both in terms of attention and in terms of interaction, of the intervention referents to other students. We identify the frequency of social interactions using spend time ties in order to causally relate the dosage of exposure to intervention referents, from zero to n number of ties, to perceptions of norms or anti-harassment behavior.\(^7\)

Many students in the sample were directly exposed to both intervention and control social referents because they spent time with both. Indeed, some control social referents themselves are directly exposed to intervention referents.\(^8\) Random assignment within a social network does not create isolated treatment and control groups; rather, random assignment creates different degrees of exposure to intervention referents depending on the structure of social interactions within the network; not all students have equal probability of exposure to intervention referents. A standard way to address this problem is to use fixed effects regression based on each individual’s total number of ties to intervention and control social referents; this controls for heterogeneous assignment probabilities induced by the fact that treatment was randomized over a network. Specifically, our fixed effects regression includes seven dummy variables that index whether an individual’s total number of ties to intervention and control referents.

---

\(^6\)Four girls and two boys declined the invitation to participate in the intervention. In order to avoid biased inference due to selection, the analyses consider ties to all intervention referents, including those who did not participate in the intervention program (intent-to-treat effects) (Gerber and Green 2012).

\(^7\)See Appendix B in the online supplemental materials for descriptive statistics on egocentric network nominations.

\(^8\)For information on the number of nominations of intervention and control referents, see Appendix C in the online supplemental materials.
control social referents at Wave 1 was zero through six (Equation 1).

\[ Y = \alpha + \beta_1 X_m + \beta_2 X_f + \beta_3 Z + \varepsilon, \]  

where \( \alpha \) is a fixed effects estimator of whether an individual has zero to six or more total network ties to intervention and control social referents; \( X_m \) are the number of spend time ties to male intervention referents; \( X_f \) are the number of spend time ties to female intervention referents; \( Z \) are the terms for covariates (respondent’s gender, grade point average, whether attended during previous year, total ties to all students, and Wave 1 measure of \( Y \) if applicable); \( \varepsilon \) is the error term.

The dummy variables of a fixed effects regression analysis account for each individual’s baseline levels of exposure to intervention and control students and thus their probability of being treated within the network since the probability of being treated is not equal for every individual. This approach is an appropriate estimation technique if there is an underlying linear relationship between exposure and outcomes (Angrist and Pischke 2008), and it reflects our hypothesis that greater exposure to intervention referents who model anti-drama behaviors should result in greater shifts toward anti-drama norms and behavior. The fixed effects model corrects for heterogeneity in the initial probability of exposure to intervention students based on network structure and thus is a conservative test of our effects. The results of the models without fixed effects are substantively identical to those reported here.

**RESULTS**

**Gendered-Perceived Collective Norms**

Both males and females viewed the norms at the school for the key behavior of “starting drama” as particularly strong among girls. At Wave 3, students were significantly more likely to indicate that it was normal for girls to start drama than to indicate it was normal for boys to start drama (\( M = 4.22 \) and 3.50, respectively; \( t = 5.39, p < .0001 \)). Males and females saw no difference in their judgment of how normal it was for boys to start drama. However, females had a tendency to report higher rates of girls starting drama than did males (\( p = .08 \)).

We first test the hypotheses that female and male intervention referents are differentially effective at communicating collective social norms regarding harassment for a gender-coded behavior.

---

9For a discussion and analysis of selection effects, see Appendix D and Table D1 in the online supplemental materials.
Lower scores on the collective norms measure indicate perceptions that other students in the school do not support harassment-related behaviors. In support of hypothesis 3, that males have a relatively stronger influence on norm perceptions, following the assembly intervention at Wave 2, students perceived significantly less support for drama when they had more first-degree ties to male intervention referents (see Table 1). We control for norm perceptions at the first wave, number of ties to all male students and number of ties to all female students, in addition to gender, whether the student was new to the school, and academic performance.

There is no significant effect of exposure to female intervention referents on perceived norms of drama, disconfirming hypothesis 1. A Wald test of the difference between the coefficients was statistically significant, $F(2, 157) = 3.00, p = .05$. Thus, male intervention referents are significantly better than their female referent counterparts at influencing their peers’ normative perceptions regarding drama. We do not observe a further change in the perceptions of these norms from Wave 2 to Wave 3 as a result of ties to either female or male intervention referents.

**Behavior.** In line with our results for perceived collective norms, we find that, in support of hypothesis 4, male intervention referents significantly shifted the harassment behaviors of those who spent time with them (see Table 1). Using data from year-end teacher nominations of students who create a negative environment, we find a significant influence of male intervention referents on the behavior of students who spent time with them, controlling for preintervention nominations. $t(282) = 2.71, p = .004$, and boys for creating a negative school environment, $t(282) = 4.91, p = .000$.

**Respect and Social Influence**

Even within gender groups, not all intervention referents’ behavioral signals were equal; students participated in follow-up activities throughout the year, 10Teachers differentially nominated students for particular behaviors based on gender. Teachers were more likely to nominate girls for defending other students, $t(282) = 4.91, p = .000$. 11For a discussion of the effects of the interaction between perceiver gender and intervention referent gender, see Appendix E in the online supplemental materials.
Table 1. Estimates of Harassment-Related Norms and Behaviors by Ties to Female and Male Intervention Students

<table>
<thead>
<tr>
<th>Predictor variables</th>
<th>Outcome variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Norm composite, T2</td>
</tr>
<tr>
<td>T1 value (norm or nomination)</td>
<td>.35***</td>
</tr>
<tr>
<td>Ties to female intervention students</td>
<td>.02</td>
</tr>
<tr>
<td>Ties to male intervention students</td>
<td>–.25*</td>
</tr>
<tr>
<td>Total number of ties to female students</td>
<td>.07</td>
</tr>
<tr>
<td>Total number of ties to male students</td>
<td>.12</td>
</tr>
<tr>
<td>Male</td>
<td>.05</td>
</tr>
<tr>
<td>Attended last year</td>
<td>.03</td>
</tr>
<tr>
<td>GPA</td>
<td>–.07</td>
</tr>
<tr>
<td>N</td>
<td>172</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.20</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>–39.56</td>
</tr>
</tbody>
</table>

Note: Estimates are standardized regression coefficients except where noted as odds ratios from logit models. Male is a dummy variable (0 = female, 1 = male), as is attended last year (0 = no, 1 = yes). GPA is a continuous variable of students' grade point average. The regression includes fixed effects (not shown). †p < .10. *p < .05. **p < .01. ***p < .001.
took on the cause of the program as their own, and participated in peer harassment themselves to varying degrees. Differences between intervention referents likely affect how influential they are. One of these differences, which may have consequences for the extent of the influence of intervention referents, is how widely respected they are at the school. “Respect” is a status concept distinct from being popular. According to interviews, respected students were social and “put themselves out there,” but they also stood up for other students when they were having trouble. To be respected, students had to be respectful themselves.

At the beginning of the year, students who were nominated by their peers as someone they respected were also the students who spent time with more students ($r = .54$) and who defended other students when they were being harassed or picked on ($r = .55$). To a lesser extent, those who were respected were also considered popular ($r = .34$). We test the reasoning that male intervention referents exert greater influence because they are more respected first by examining the importance of respect in social influence. We assess whether intervention referents who are more respected by their peers exert greater influence over the norm perceptions and behaviors of those with spend time ties to them. Second, we examine the extent to which gender and respect are correlated and specifically whether male intervention students receive more respect nominations from their peers than do female intervention students.

Influence of high and low respect intervention referents. At each wave of measurement, we consider intervention students to be “highly respected” when they are at or above the median number of respect nominations all intervention students received from their peers (three or more nominations at Wave 1; five or more at Wave 2; four or more at Wave 3). We repeated our aforementioned analyses, but rather than predicting perceived norms and behavior based on ties to social referents distinguished by gender, we distinguished between ties to referents with relatively high and relatively low numbers of respect nominations. See results in Table 2.

We find consistent support for the claim that having more ties to high respect intervention students leads to perceiving norms that do not support harassment at Wave 2 and to more prosocial behaviors: less participation in harassment, less participation in any events meriting discipline, more teacher nominations as someone who defends other students when they are getting harassed, and fewer teacher nominations as someone who contributes to a negative school environment. Wald tests indicated a significant difference between the coefficients for ties to high and low respect intervention students for Wave 2 norm perceptions, all disciplinary behavior events, and teacher nominations for creating a negative school environment. Having more ties to low respect intervention students did not change students’ norm perceptions or behaviors. These results demonstrate that not all intervention referents were equally influential: those who were considered respected by their peers exerted more influence on their network neighbors.

Respect nominations by gender. To what extent is the experimental effect of ties to male intervention referents a result of male intervention referents being more respected in the school than female intervention referents? Intervention females received more respect nominations than did intervention males at Wave 1, though the difference is not significant ($p = .11$);
<table>
<thead>
<tr>
<th>Predictor variables</th>
<th>T2 norm composite</th>
<th>T3 norm composite</th>
<th>Disciplinary action: peer harassment (odds ratio)</th>
<th>Disciplinary action: all reports</th>
<th>Teacher nomination: defend others</th>
<th>Teacher nomination: create negative environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 norm composite</td>
<td>.36***</td>
<td>.25**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1 teacher nomination</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ties to high respect intervention students</td>
<td>-.27**</td>
<td>-.20</td>
<td>-.21†</td>
<td>.16*</td>
<td>.18*</td>
<td>-.28**</td>
</tr>
<tr>
<td>Ties to low respect intervention students</td>
<td>.06</td>
<td>.04</td>
<td>-.02</td>
<td>1.23</td>
<td>.042</td>
<td>-.04</td>
</tr>
<tr>
<td>Total number of ties</td>
<td>.18</td>
<td>.11</td>
<td>.17</td>
<td>1.10</td>
<td>-.06</td>
<td>.19†</td>
</tr>
<tr>
<td>Male</td>
<td>.06</td>
<td>-.08</td>
<td>.15*</td>
<td>5.78*</td>
<td>-.11†</td>
<td>.23***</td>
</tr>
<tr>
<td>Attended last year</td>
<td>.04</td>
<td>.05</td>
<td>.00</td>
<td>1.01</td>
<td>-.01</td>
<td>-.00</td>
</tr>
<tr>
<td>GPA</td>
<td>-.08</td>
<td>.07</td>
<td>-.13</td>
<td>.32*</td>
<td>.14*</td>
<td>-.20**</td>
</tr>
<tr>
<td>N</td>
<td>172</td>
<td>162</td>
<td>227</td>
<td>227</td>
<td>227</td>
<td>227</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.20</td>
<td>.15</td>
<td>.11</td>
<td>.25</td>
<td>.34</td>
<td>.26</td>
</tr>
<tr>
<td>Log likelihood</td>
<td></td>
<td></td>
<td>-.39.05</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Estimates are standardized regression coefficients, except where noted as odds ratios from logit models. All other variables are the same as described in Table 1. †$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$. 

Table 2. Estimates of Norms and Harassment-Related Behaviors by Ties to High and Low Respect Intervention Students.
however, at Wave 2, intervention males received more respect nominations than intervention females \((p < .05)\). Again at Wave 3, male intervention students receive more respect nominations, though the difference is not significant \((p = .18)\). See Table 3 for means of respect nominations by gender and intervention treatment status.

We test for gender differences among intervention students in the likelihood of being nominated by peers as someone they respect at each survey wave. Results are provided in Table 4. Compared to the general school population, intervention students regardless of gender have more respect nominations at Wave 1. At Wave 2, we find an interaction effect whereby intervention males have many more respect nominations than their general school population counterparts, but intervention females have the same number of respect nominations as do their general school population counterparts (controlling for Wave 1 respect nominations). This pattern is replicated at Wave 3, though intervention females at Wave 3 have even fewer respect nominations than do their general school population counterparts.

Finally, to what extent are the differential respect nominations for male and female intervention referents due to the intervention itself, as opposed to a trend in the school unrelated to the intervention? We examine the number of respect nominations by gender, comparing those students selected for the intervention to control students who were in the eligible pool of social referents but who were not randomly selected to participate. These students provide a rigorous test of whether the males randomly selected to participate in the intervention had more respect nominations across the year as a result of participating in the program itself.

We find suggestive evidence that participation in the intervention itself positively affected the number of respect nominations of intervention males more so than of intervention girls (see Table 4). There are no differences in the number of respect nominations between intervention and control referents based on gender at Wave 1. There is a marginally significant effect such that the randomly selected intervention students received more respect nominations than did the control students, a chance imbalance.

At Wave 2, male intervention students received more respect nominations than did their control counterparts, while female intervention students received a similar number of nominations as their control counterparts, though the interaction effect is not statistically significant \((p = .11);\) means in Table 3). At Wave 3, male intervention students had more respect nominations than did their control counterparts, while female intervention students had fewer respect nominations than their control counterparts. Male intervention referents received a bump in respect nominations compared to their control counterparts at the two waves after the intervention, while female intervention referents received fewer respect nominations, particularly at Wave 3, as a result of their participation in the intervention.\(^\text{12}\)

A possible explanation for the relatively greater respect afforded to male intervention referents compared to female intervention referents after the intervention is that female intervention referents were more involved in harassment than intervention males, and as a result, they became less credible and less respected over the course of the school year. We find no evidence for this explanation: female intervention referents participate in significantly fewer

\(^{12}\text{For a discussion of the relationship between spending time and respect ties, see Appendix F in the online supplemental materials.}\)
Table 3. Means and Standard Deviations of Number of Respect and Spend Time Nominations by Intervention Group

<table>
<thead>
<tr>
<th></th>
<th>Wave 1 respect nominations</th>
<th>Wave 2 respect nominations</th>
<th>Wave 3 respect nominations</th>
<th>Wave 1 spend time nominations</th>
<th>Wave 2 spend time nominations</th>
<th>Wave 3 spend time nominations</th>
</tr>
</thead>
<tbody>
<tr>
<td>General girl</td>
<td>2.32 (1.81)</td>
<td>3.54 (2.43)</td>
<td>4.43 (3.07)</td>
<td>3.81 (2.71)</td>
<td>6.71 (4.21)</td>
<td>6.14 (4.12)</td>
</tr>
<tr>
<td>General boy</td>
<td>1.94 (1.49)</td>
<td>3.22 (2.66)</td>
<td>3.45 (2.14)</td>
<td>4.03 (3.07)</td>
<td>6.83 (4.18)</td>
<td>6.54 (4.30)</td>
</tr>
<tr>
<td>Control girl</td>
<td>2.72 (1.56)</td>
<td>4.55 (2.60)</td>
<td>5.77 (3.57)</td>
<td>4.42 (2.50)</td>
<td>8.71 (3.52)</td>
<td>8.10 (3.87)</td>
</tr>
<tr>
<td>Control boy</td>
<td>2.00 (1.11)</td>
<td>4.80 (3.86)</td>
<td>4.52 (2.79)</td>
<td>5.24 (2.26)</td>
<td>8.95 (4.51)</td>
<td>9.52 (5.53)</td>
</tr>
<tr>
<td>Intervention girl</td>
<td>3.50 (1.85)</td>
<td>4.80 (3.14)</td>
<td>4.05 (2.26)</td>
<td>5.25 (2.51)</td>
<td>8.75 (3.89)</td>
<td>6.15 (3.94)</td>
</tr>
<tr>
<td>Intervention boy</td>
<td>2.56 (1.81)</td>
<td>7.20 (3.52)</td>
<td>5.2 (4.73)</td>
<td>5.00 (2.49)</td>
<td>12.00 (5.06)</td>
<td>9.20 (4.92)</td>
</tr>
</tbody>
</table>

Note: Standard deviations are provided in parentheses.

Table 4. Regression Estimates of Number of Respect Nominations by Treatment Status and Gender

<table>
<thead>
<tr>
<th></th>
<th>Comparison: all students at school and intervention students</th>
<th>Comparison: control students and intervention students</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of respect nominations T1 T2 T3</td>
<td>Number of respect nominations T1 T2 T3</td>
</tr>
<tr>
<td>Number of T1 respect nominations</td>
<td>.39*** .27***</td>
<td>.21† .26*</td>
</tr>
<tr>
<td>Intervention students</td>
<td>.22** .04 −.11</td>
<td>.23† −.02 −.30*</td>
</tr>
<tr>
<td>Male</td>
<td>.11 −.01 −.12†</td>
<td>−.22† .07 −.11</td>
</tr>
<tr>
<td>Intervention students × male</td>
<td>−.06 .23** .17*</td>
<td>−.04 .25 .26†</td>
</tr>
<tr>
<td>N</td>
<td>243 227 228</td>
<td>80 76 79</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.06 .23 .09</td>
<td>.11 .11 .10</td>
</tr>
</tbody>
</table>

Note: Coefficients are standardized regression coefficients. Nonintervention eligible students are the default category for the comparison on the left side of the table. Control students are the default category for the comparison on the right side of the table. †p < .10. *p < .05. **p < .01. ***p < .001.
disciplinary events than control girls and intervention referent males, and teachers are marginally less likely to nominate them for creating a negative school environment compared to female control referents and intervention referent males. Intervention referent males are disciplined slightly less often than control referent males, though this difference is not significant \((p = .17)\), and there is no difference in how frequently the two groups are nominated as creating a negative school environment. Furthermore, there is no relationship between negative peer behaviors and respect ties, among either males or females in the school.

The evidence we consider here is consistent with a status beliefs explanation that boys’ participation in the intervention was interpreted more favorably than girls’ participation by other students, indicated by the respect “bump” that boys received following the intervention and by the decrease in respect nominations among intervention girls relative to control girls. This effect appears to be unrelated to changed behavior from intervention males compared to control males or from intervention females compared to control females.

**DISCUSSION**

Using a field experiment randomizing which students participated in an anti-harassment program, we find that male students who participated in the intervention program influenced their network neighbors’ harassment-related behaviors and perceptions of collective norms about drama, even when the behaviors were most commonly associated with girls. Among students who reported spending time with male intervention referents, the perceived acceptability of harassment and participation in harassment behaviors declined. Female students exerted influence over fewer harassment behaviors of their network neighbors compared to male students in the program. More respected intervention students influenced the norm perceptions and behaviors of their network neighbors while low respect intervention students did not. Males who participated in the intervention gained respect nominations more than females who participated in the program, leading to the conclusion that boys’ relative influence was at least in part a product of their increased respect after the intervention program.

Although we examine influence in a manner very different from studies of deliberate influence attempts in small group settings (Carli 2001), we find converging evidence of male influence. Our findings diverge from these largely lab-based studies of influence in groups in that we find evidence of male influence even in a female-typed domain such as drama and in a majority female setting. The divergent findings may be the result of differential instantiation of female-typed domains in earlier studies compared to this study. Previous studies examined topics of female “expertise” about which men and women may have relatively little regular interaction (e.g., birth control or women’s fear of crime). The nature of peer harassment is such that boys were necessarily involved in conflicts even when the main actors were girls (they would often “stir the pot,” as one informant told us). Thus, boys may be more relevant as possible influencers in the domain of peer harassment. Additionally, though we cannot rule out the possibility of unique results, the school’s characteristics provide a strict test of the influence of males: a female majority school where generally girls were vocal and seen as positive forces in the school by teachers and administrators, despite the prevalence of gender-coded drama. These features lead to an a priori prediction that girls would be
equally or more influential than boys. Our findings suggest that even when in a numerically and culturally dominant position as in this school, girls’ public participation did not improve their standing among other students as did boys’ public participation. Instead, the relative influence of male intervention students stems at least in part from generalized status beliefs about the greater status and importance of men and boys compared to women and girls that the intervention program made relevant, as predicted by expectation states theory.

Thus, a more important source of the difference between our results and those of previous research may be that studies over short time periods underestimate the importance of simultaneous processes and thus may underestimate male influence. Specifically, in this case, the changing nature of respect for male and female intervention students affected relative social influence. The pattern of respect nominations across the school year is consistent with an account that the intervention males were more influential because they were afforded respect following their program participation. General population boys were not more respected than general population girls; girls in general received more respect nominations across the year. Instead, boys—but not girls—who were randomly selected to become highly salient through their participation in the NAMES intervention became more respected in the school. Status beliefs about the relative competence of males in the school were not generalized to all males but were activated by the intervention, and this contributed to the greater influence of male intervention students over their peers’ behavior. The results provide a more nuanced account of how status beliefs become important in real-world settings over time.

This work draws on a network-based account of social influence and behavioral patterns: individuals’ perceptions of group norms develop from the behavioral cues they get from others in the group, which are structured by patterns of interaction within the group, both network-level characteristics and individual network position. These findings have implications for practitioners interested in changing the “culture” or “climate” of a group and suggest that targeting the public behavior of social referents can change their peers’ behavior and perceptions.

**ACKNOWLEDGMENTS**

The authors gratefully acknowledge the collaboration, generosity, and support of our high school research site, its staff and students, and in particular Kelly Bocuzzo and Alan Paluck, whose energy and dedication made this project possible. We received important feedback on this article from Deborah Carr, Paul DiMaggio, David Gibson, Amir Goldberg, Patrick Ishizuka, Jennifer Lee, Emily Marshall, and Sarah Thebaut. Izzy Gainsburg, David Mackenzie, and Ani Momjian provided superior research support.

**FUNDING**

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: A research grant from Princeton University to E. L. Paluck provided support for this research. Shepherd was funded by a National Science Foundation Graduate Research Fellowship during this project.

**REFERENCES**


Bazelon, Emily. 2013. *Sticks and Stones: Defeating the Culture of Bullying and*


BIOS

Hana Shepherd is an assistant professor of sociology at Rutgers University.

Her research explores social processes that create or impede change and shape patterns of inequality, particularly the role of social networks, social norms and cognition, and organizations. She studies peer conflict in middle and high schools, social networks and school characteristics, school implementation of new laws, and culture and cognition. She uses diverse methods including network analysis, lab and field-based experiments, interviews, and archival research.

Elizabeth Levy Paluck is associate professor in the Department of Psychology and in the Woodrow Wilson School of Public and International Affairs at Princeton University. Her research is concerned with the reduction of prejudice and conflict, including ethnic and political conflict, youth conflict in schools, and violence against women. She uses large-scale field experiments to test interventions that target individuals’ perceived norms and behavior about conflict and tolerance, including mass media and social network peer-to-peer interventions.