Differential effects of the KiVa anti-bullying program on popular and unpopular bullies

Claire F. Garandeau [a,*], Ihno A. Lee [b], Christina Salmivalli [c,d]

[a] *University of Turku, Department of Psychology, Assistentinkatu 7, 20014 Turku, Finland*
[b] *Stanford University, Department of Psychology, Stanford University, 450 Serra Mall, Stanford, CA 94305, USA*
[c] *University of Turku, Department of Psychology, Assistentinkatu 7, 20014 Turku, Finland*
[d] *Center for Behavioral Research, Edith Cowan University, Perth, Western Australia, Australia*

ABSTRACT

This study utilized data from the evaluation of the Finnish KiVa program in testing the prediction that school bullies' high perceived popularity would impede the success of anti-bullying interventions. Multiple-group structural equation modeling (SEM) analyses were conducted on a subsample of 911 third-, fourth-, and fifth-graders identified as perpetrators of bullying. They belonged to 77 Finnish schools, including 39 schools implementing the KiVa program and 38 control schools. Data on peer-reported bullying and perceived popularity were collected before program implementation and one year later. Controlling for sex, age, and initial levels of bullying, KiVa participation resulted in lower rates of bullying (indicated by fewer peer nominations) after one year for bullies of low and medium popularity. However, there was no significant effect for those high in popularity, suggesting that popular bullies are less responsive to anti-bullying interventions than less popular bullies.

© 2013 Elsevier Inc. All rights reserved.

School bullying – the repetition of intentional, aggressive behaviors that involve an abuse of power – is a pervasive phenomenon in many countries (Currie et al., 2008) and has long-lasting damaging consequences on the mental and physical health of victims (e.g., Isaacs, Hodges, & Salmivalli, 2008; Rigby, 2003). However, attempts to reduce levels of bullying in schools have been met with mixed success. Meta-analyses on the effectiveness of anti-bullying programs have shown that approximately half of them have no or limited effect (Baldry & Farrington, 2007; Ferguson, San Miguel, Kilburg, & Sanchez, 2007; Merrell, Gueldner, Ross, & Isava, 2008). Recognizing that bullying can be highly rewarding, several theorists have proposed that bullies' high popularity in the peer group accounts for anti-bullying programs' mitigated success (see Volk, Camilleri, Dane, & Marini, 2012). To our knowledge however, this assumption has never been tested.

Bullying incidents generally take place in a group context where the perpetrators are socially powerful (e.g., Salmivalli, 2010). A large body of research shows that many aggressive children enjoy high status among their peers. Although young bullies tend to be low in social preference – an indicator of peers' personal liking usually associated with desirable traits such as kindness and cooperativeness – they often have high levels of perceived popularity (Caravita, Di Blasio, &

Salmivalli, 2009; Cillessen & Mayeux, 2004), which reflects social power, or visibility, and is linked to antisocial tendencies (e.g., Lease, Kennedy, & Axelrod, 2002).

Anti-bullying interventions should be less effective for popular bullies compared to their unpopular counterparts for three reasons: 1) popular bullying perpetrators reap valued social rewards for their behavior, 2) a position of power in the peer group facilitates bullying behaviors in various ways, 3) bystanders should be more motivated to join in the bullying and less motivated to intervene against it when it is initiated by popular classmates. This positive reinforcement may encourage bullies to pursue their actions.

The present study tests whether the effects of KiVa, a nationwide anti-bullying program in Finland, vary depending on the perceived popularity status of bullying students. Despite consistent findings of a positive association between bullying and perceived popularity, bullies do not form a uniformly popular group (Peeters, Cillessen, & Scholte, 2010; Rodkin, Farmer, Pearl, & Van Acker, 2000). We expected that KiVa-related decreases in bullying behaviors would be significantly smaller for popular bullies, compared to bullies lower in popularity.

Bullying brings coveted status rewards

Gaining social power among peers appears to be children's main motivation for bullying others (Olthof, Goossens, Vermande, Aleva, & van der Meulen, 2011). Interviews of early adolescents who had been excluded from school for causing bullying incidents reveal that initial

* Corresponding author at: Department of Psychology, University of Turku, Assistentinkatu 7, 20014 Turku, Finland. Tel.: +358 466247116; fax: +358 23335060.
E-mail addresses: garandeauclaire@hotmail.com (C.F. Garandeau), ihno@stanford.edu (I.A. Lee), tiina.salmivalli@utu.fi (C. Salmivalli).

bullying is a deliberate choice aimed at gaining recognition and respect, and subsequent bullying behaviors are then perpetuated to promote and maintain the reputation (Houghton, Nathan, & Taylor, 2012). Similarly, studies investigating social goals show that aggressive youth are more likely than their non-aggressive counterparts to endorse agentic goals, which aim towards power, mastery and status (Caravita & Cillessen, 2012; Ojanen, Gronroos, & Salmivalli, 2005; Sijtsema, Veenstra, Lindenberg, & Salmivalli, 2009). Young bullies are highly interested in being popular or socially dominant among peers, and bullying proves effective in reaching this goal: not only do cross-sectional studies document positive correlations between bullying and popularity among children and youth (Berger & Rodkin, 2012; Caravita et al., 2009; de Bruyn, Cillessen, & Wissink, 2010), but also longitudinal studies further demonstrate that aggression and bullying allow young students to increase their popularity over time (Cillessen & Borch, 2006; Cillessen & Mayeux, 2004).

As many popular bullies obtain highly coveted social rewards for their negative conduct, they should have little motivation to curb their behavior. Conversely, bullying students who are lower in perceived popularity should find it easier to alter their behavior as they have little to lose in terms of social standing. In addition, becoming less aggressive could allow them to improve their likeability, which is also an important goal for children (Veenstra, Lindenberg, Munniksma, & Dijkstra, 2010).

High popularity facilitates aggression

There is evidence of a bi-directional association between popularity and aggression: aggression positively predicts future popularity, but the achievement of high popularity status also promotes aggression (Cillessen & Mayeux, 2004). Popular bullies may feel increased pressure to maintain their rank and thus resort to coercive means. Being in a powerful position in the peer group may also lead to a sense of entitlement causing school bullies to abuse their power without any fear of negative consequences. Experimental studies with adults indicate that holding a position of power may significantly affect individuals' attitudes and behaviors in a way that facilitates aggressive conduct: it can reduce people's distress in response to others' suffering (Van Kleef et al., 2008) and their ability to take the perspective of others (Galinsky, Magee, Inesi, & Gruenfeld, 2006); it can also lead them to objectify others (Gruenfeld, Inesi, Magee, & Galinsky, 2008).

Furthermore, being socially powerful makes it easier to engage in relational forms of bullying, such as rumor spreading or social exclusion. Ethnographic studies of children and adolescent peer groups have shown that carrying out such forms of bullying requires a central and powerful position in that group (e.g., Merten, 1997). This was further demonstrated by research on social network centrality, another indicator of peer status. Faris and Felmlee (2011) found that most adolescents tend to become more aggressive as their position in the peer network becomes more central. Having numerous connections to multiple peer groups puts highly central individuals in an ideal position to exclude others as well as diffuse information and thus spread malicious rumors. As suggested by Garandeau and Cillessen (2006), indirect aggression also involves the skillful manipulation of the whole peer group and the capacity to exert influence on one's peers. High popularity indeed confers students the ability to influence peers' aggressive attitudes and conduct (Cohen & Prinstein, 2006). Therefore, anti-bullying interventions may be less likely to succeed among popular bullies, because these students are already in a position that fosters bullying.

Bullies' popularity and bystanders' behaviors

The perpetuation of bullying heavily depends on the behavior of bystanders in bullying situations: self-reported frequency of bullying was found to be lower in classrooms where children tend to defend the victim and avoid reinforcing the bully, as reflected by proportions of peer

nominations for these two behaviors (Salmivalli, Voeten, & Poskiparta, 2011). High classroom rates of bully reinforcement and low rates of victim defending also increase the likelihood that anxious and rejected children will become bullied (Kärnä et al., 2010). Both defending and reinforcing behaviors may be influenced by the popularity of the lead bully, primarily because of the contagious nature of popularity (Marks, Cillessen, & Crick, 2012). Studies show that being closely affiliated with highly popular peers is associated with high popularity for oneself, both concurrently (Dijkstra, Cillessen, Lindenberg, & Veenstra, 2010) and longitudinally (Marks et al., 2012). As popularity is highly valued by students, especially in early adolescence (LaFontana & Cillessen, 2010), one might expect most youth to attempt to befriend popular peers in order to "bask in reflected glory" (Cialdini & Richardson, 1980), or at least maintain these relationships in order to preserve their own status.

These preferences should have important implications for their behaviors in bullying situations. Defending a victimized peer is a risky behavior that implies a confrontation with a bully and sometimes his or her supporters (Pozzoli & Gini, 2010). Opposing a popular bully may be perceived by potential defenders as even riskier, as it is more likely to result in rejection from the peer group and a loss in status. A recent study suggests that defending, as indicated by proportions of peer nominations, is indeed less frequent in classrooms where bullies are more popular (Pöyhönen, Juvonen, Peets, & Salmivalli, 2013). On the other hand, openly supporting a popular bully may be particularly tempting as a way to enhance one's own popularity. Therefore, popular bullies may be more resistant to anti-bullying interventions, not only because their negative behaviors enable them to satisfy their need for dominance, but also because their peers are less likely to counteract the bullying.

The KiVa anti-bullying program

KiVa is a nationwide anti-bullying program in Finland. It was developed in 2006 at the request of the Finnish Ministry of Education, following the release of a World Health Organization (WHO) report revealing Finnish children's low liking of school. After an evaluation phase in 2007 and 2008 with a sample of 78 intervention and 78 control schools, the program was disseminated across the country in 2009. KiVa is based on the notion that the behavior of bystanders – reinforcing bullies, defending victimized peers, or remaining a passive observer – is essential in the continuation or cessation of bullying, and therefore intervention efforts should focus on all students, and not only the perpetrators and targets of bullying. The program includes universal actions directed at all students. The main components of these actions are lessons about the mechanisms of bullying, which consist of lectures as well as practical exercises and group discussions. A total of ten 90-minute lessons are delivered throughout the school year. Universal actions also include a computer game designed to match the themes raised during the lessons. The aim is to increase empathy for victims, raise awareness of the role played by the group in bullying incidents and of the various strategies that any child can adopt to thwart bullying. In addition to these components designed for every student, KiVa includes indicated actions directed at bullies and victims when a case of bullying comes to the attention of the school staff. These actions consist of a series of separate discussions with both bullies and victims aimed at stopping the bullying immediately.

The evaluation of the program as a whole has shown it to be effective at reducing self- and peer-reported victimization as well as self-reported bullying among fourth- to sixth-graders (Kärnä et al., 2011). During the nine months of implementation evaluated by randomized controlled trial, the implementation rate was high (see Haataja et al., 2013): the dose of tasks (i.e., the proportion of classroom activities delivered to the students), which included interactive exercises and coverage of the discussion topics described in the teacher manual, was almost 70%. This rate was equal or higher than 60% for 74% of the teachers in the sample. On average, teachers spent 79 min delivering the lessons and found the KiVa material easy to implement.

This study examines whether the effectiveness of the KiVa program at decreasing peer-reported bullying depends on bullies' peer status. We hypothesized that the effects of the program would be significantly smaller for bullies enjoying high levels of popularity, in comparison to bullies of average or low popularity.

Method

Sample

Data were collected as part of a large randomized controlled trial designed to evaluate the effectiveness of the KiVa program. Of the 78 schools originally taking part in this evaluation in 2007 and 2008, one school had to drop out before the first assessment due to facility-related issues. The data analyzed in this study were collected at two time points: before program implementation, at the end of one school year in May 2007 (T1) and 9 months after implementation, at the end of the subsequent school year in May 2008 (T2). At T1, participants were in grades 3, 4 and 5, with approximately 33% of the sample at each grade level. At T2, they were in grades 4, 5 and 6 (sixth grade is the last year of elementary school in the Finnish school system).

Between the two waves of data collection, 7 control schools (including 32 classrooms) dropped out of the study for they considered the data collection process to be highly time-consuming and unrewarding for them. In addition, 16 classrooms (from 8 control schools) and 15 classrooms (from 8 KiVa schools) dropped out between the two time points. The reason provided was that teachers could not find the time to administer the survey to their students. At T2, information was therefore available for only 712 of the 911 bullies identified at T1, as 148 came from schools or classrooms which had dropped out of the study and 51 had moved to another school or city.

The full sample consisted of 7975 third-, fourth-, and fifth-graders (50% male; 2% immigrants) ranging from 8 to 14 years of age ($M_{age} = 11.03$, $SD = 0.93$). They belonged to 401 classrooms in 77 schools (16% Swedish-speaking), including 39 intervention schools and 38 control schools. Schools were recruited across the country so as to be representative of the Finnish student population. The mean participation rate was 90.67%. From the initial sample of 7975 children, we selected 911 bullies who were defined as follows: scoring one standard deviation or higher on a peer-reported measure of bullying at T1 (see below). The variable had been z-standardized across the whole sample. Half of the 911 bullies were in intervention schools. Eighty-nine percent of them were boys and their age ranged from 9 to 14 years old ($M = 11.37$, $SD = 0.92$). Our analyses were conducted with this subsample of children.

Procedure

Students filled out internet-based questionnaires in the schools' computer labs during regular school hours. Anonymity was ensured by the use of individual passwords to log in to the surveys. They were supervised by the teachers, to whom detailed instructions about the data collection process had been provided. Prior to the survey, a definition of bullying from the Olweus' bully/victim questionnaire (Olweus, 1996) appeared on the students' computer screens and was read out loud to them: *Bullying occurs when students repeatedly perform any of the following behaviors directed towards another*: say "*mean and hurtful things*" or call him/her names, purposefully "*ignore or exclude him or her from their group of friends*", "*hit, kick, push, shove*", or "*tell lies or spread false rumors.*" The imbalance of power between bully and victim that characterizes bullying was also specified: *It is not bullying when two students of about equal strength or power argue or fight.* In addition, students were told that behaviors such as teasing a peer in a friendly way was not bullying. Participants were informed of the strict confidentiality of their answers.

Measures

Bullying. We used a peer-reported measure of bullying from the Participant Role Questionnaire (Salmivalli, Lagerspetz, Bjorkqvist, Osterman, & Kaukiainen, 1996), in which participants were asked to nominate classmates who fit the description provided. They could nominate an unlimited number of classmates for the three following items: a) starts bullying; b) makes the others join in the bullying; c) always finds new ways of harassing the victim. Participants could also indicate that no classmate exhibited the behavior. For each item, we computed proportion scores by dividing the number of received nominations by the number of participants. The three items formed an internally consistent scale ($\alpha = .91$ at T1, $\alpha = .93$ at T2). In the sample of 7975 children, bullying proportion scores ranged from 0 to 1 ($M = .07$, $SD = .12$) at T1 and from 0 to 1 ($M = .06$, $SD = .11$) at T2. The T1 bullying proportion score was used to select the subsample of bullies. The bullying proportion scores of the bullies selected for the analyses were .19 and higher. The mean of their bullying proportion scores was .36 ($SD = .15$) at T1 and .25 ($SD = .18$) at T2.

Although we are using a peer nomination procedure to assess a child's "level" of bullying, one should keep in mind that this method only reflects the proportion of classmates nominating this child as bullying others and is not a direct assessment of the behavior's frequency and/or intensity (see Olweus, 2013). Logically, high levels of bullying should translate into high proportions of peer nominations (and low levels of bullying into small proportions of peer nominations). However, two classmates can receive an equally high number of bullying nominations even though one may bully others more often and/or more severely than the other. Our results should be interpreted accordingly.

Perceived popularity. Each participant was presented with a roster of their classmates and requested to answer the question "Who are the most popular in your class?" by checking the names of the classmates who fit the description. The number of peer nominations was not restricted. Proportion scores for popularity were then computed by dividing the number of received nominations by the number of respondents. Only popularity at T1 was included in the analyses. In the full sample, popularity scores ranged from 0 to 1, with a mean of .16 ($SD = .19$) at T1. In the subsample of bullies, scores ranged from 0 to 1, with a mean of .23 ($SD = .23$). Popularity was positively correlated with bullying proportion scores at T1 in the full sample ($r = .16$, $p < .001$) as well as within the subsample of bullies ($r = .09$, $p = .008$).

The subsample of bullies was divided into 3 groups according to levels of perceived popularity.[1] The categorization was based on the z-standardized perceived popularity variable, which ranged from $-.85$ to 4.55 ($M = 0.34$, $SD = 1.22$); the standardization was done across the whole sample – not on the subsample of bullies – so as to reflect bullies' actual popularity in the whole peer group, and not simply in comparison to other bullies. Bullies in the *high popularity* group were those who scored 0.5 standard deviation or higher on the standardized variable. Their perceived popularity proportion scores ranged from .25 to 1 ($N = 335$). Bullies in the *medium popularity* group had a standardized perceived popularity score falling between -0.5 and 0.5 standard deviation, and their proportion scores ranged from .07 to .24 ($N = 290$). Bullies scoring lower than -0.5 on the standardized perceived popularity variable were classified into the *low popularity* group with proportion scores ranging from 0 to .06 ($N = 267$). The cut-off criterion of ± 0.5 standard deviation was chosen based on the distribution of perceived popularity scores, in order to obtain groups of relatively balanced sizes (30%, 34% and 36% of the distribution for the low-, medium-, and high-popularity groups, respectively).

---

[1] While categorizing a continuous variable has various drawbacks, this approach allows for direct comparison of KiVa effects between low, medium, and high popularity groups in a multiple-group framework. This categorization was necessary to facilitate testing of our hypotheses.

We conducted a one-way ANOVA to test whether mean bullying proportion scores at T1 differed across the three groups. The result was significant: $F(2, 889) = 6.79$, $p = .001$. Post-hoc comparisons using the Tukey test indicated that the mean bullying proportion score at T1 for the low-popularity bully group ($M = 0.33$, $SD = 0.12$) differed from the mean bullying proportion score of the medium-popularity group ($M = 0.36$, $SD = 0.15$, $p = .031$) and the high-popularity group ($M = 0.37$, $SD = 0.15$, $p = .001$); there was no significant difference in mean bullying proportion scores between the medium and the high-popularity groups ($p = .564$). To account for these differences, we controlled for T1 levels of bullying in our primary analysis.

Demographic variables. Age, sex, and intervention status were included as covariates of T1 bullying. Age was a continuous variable, ranging from 9 to 14 ($M = 11.37$, $SD = 0.92$) in the subsample of bullies. Sex was a dummy variable coded as boy $= 1$ and girl $= 0$. Intervention status was a binary predictor coded as KiVa $= 1$ and control $= 0$.

Results

Analytical strategy

We conducted multiple-group SEM analyses on the subsample of 911 bullies. A robust maximum-likelihood estimation method was used to account for the nested data structure (i.e., bullies nested in classrooms). Bullying at both time points was modeled as a latent factor, with correlated residuals estimated for corresponding indicators at T1 and T2. Prior to evaluating the effects of KiVa participation, we tested for measurement invariance across time (T1 and T2) and group (low, medium, high popularity) using well-known procedures and criteria (Brown, 2006; Meredith, 1993). A sequence of models was fit, from an unconstrained model (i.e., configural invariance model) to a more restricted model, imposing equality constraints on the corresponding factor loadings (i.e., weak invariance model). The resulting change in model fit was evaluated by comparing CFI and RMSEA indices (Cheung & Rensvold, 2002; Little, Preacher, Selig, & Card, 2007). In the primary analysis, bullying at T2 was predicted by KiVa intervention status, controlling for sex, age and bullying at T1 (see Fig. 1).

Some of the variables included in the analyses had missing data. The proportion of missing data was 11.7% for age and 2.3% for popularity. For
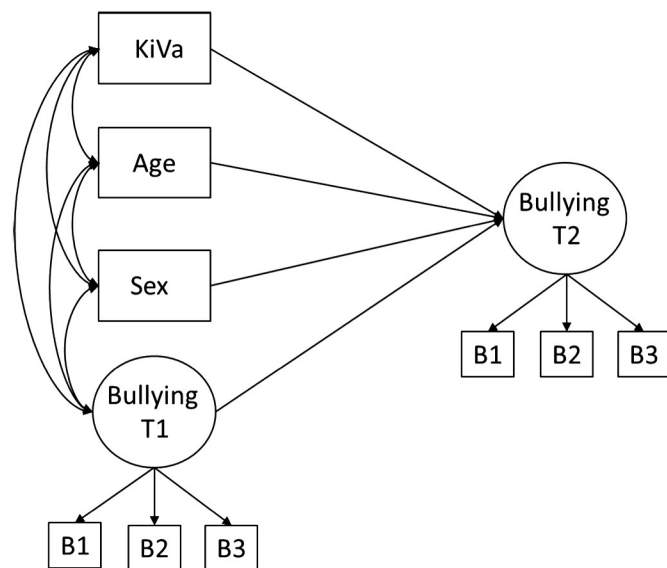

Fig. 1. Bullying at Time 2 (T2) as predicted by KiVa intervention status, age, sex and bullying at Time 1 (T1) within low, medium, and high popularity groups. Residual variances and correlated residuals were included in the model, but are not depicted for clarity.

bullying at T1, only the third item ("always finds new ways of harassing the victim") had some missing data (10.8%); these missing values are considered to be missing at random, as the missingness was due to a technical problem in the administration of the survey resulting in 36 classrooms (in the full sample, 33 classrooms in the subsample of bullies) not being presented with the item. For bullying at T2, the three items had 21.8% of missing data, which were primarily due to entire classrooms dropping out of the study. A *t*-test indicates that both T1 bullying and popularity scores do not differ among bullies missing at T2 versus bullies not missing at T2 (bullying: $t = .240$, $p = .810$; popularity: $t = .021$, $p = .983$). The data can also be considered missing at random.

Analyses were conducted using Mplus v.6.1 (Muthén & Muthén, 1998-2010), which utilizes full information maximum-likelihood procedures to account for missing data considered missing at random (MAR). These procedures use all available data in generating the final parameter estimates, which are not biased by the proportion of missing data in our study (Enders, 2010; Graham, Olchowski, & Gilreath, 2007).

Measurement model

We tested measurement invariance for the bullying factors across time (T1 and T2) and group (low, medium, high popularity). The initial, freely estimated model had excellent fit, $\chi^2(15) = 23.20$, RMSEA $= .043$ (95% CI $= .000$, $.075$), CFI $= .995$, TLI $= .985$, SRMR $= .022$. Following standard procedures to evaluate measurement invariance, we equated factor loadings across both time and group and found no significant changes in fit based on the RMSEA (i.e., the nested model RMSEA was included within the 90% RMSEA confidence interval of the constrained model; see Little et al., 2007) and CFI (i.e., changes in CFI were less than .01, see Cheung & Rensvold, 2002). This result shows that the bullying constructs, as measured by the three items, were qualitatively equivalent across school years and popularity groups (i.e., the same bullying constructs are being assessed in students of low, medium, and high popularity, and at T1 and T2).

The standardized factor loadings for the bullying measures ranged from .51 to .83 ($ps < .001$) at T1 and from .87 to .93 at T2 ($ps < .001$) across low-, medium- and high-popularity groups. Residual variances were generally small to moderate (from .15 to .47, with a maximum of .74); residual covariances between corresponding indicators were low (from .01 to .28). Overall, the measurement model estimates show that the bullying indicators are valid representations of the underlying latent constructs.

Structural model

The model had good fit, $\chi^2(77) = 180.04$, RMSEA $= .067$ (95% CI $= .054$, $.080$), CFI $= .948$, TLI $= .926$, SRMR $= .110$, based on existing guidelines (Hu & Bentler, 1999). Results are presented in Table 1. KiVa participation resulted in lower proportions of peer nominations for bullying behavior for bullies in the low-popularity group ($p = .035$) and in the medium-popularity group ($p < .001$). The effect size for the KiVa intervention effects was small to moderate, as reflected by the standardized estimates (see Cohen, 1988). However, KiVa participation did not significantly decrease the proportion of peer nominations for bullying behavior for bullies in the high-popularity group ($p = .740$). Comparing the KiVa intervention effect across popularity groups resulted in the following: the effect of KiVa participation on bullying was significantly stronger in the medium popularity group compared to the high popularity group, Satorra–Bentler[2] scaled $\chi^2(1) = 6.05$, $p = .014$; however, the strength of the KiVa effect in the low popularity group did

---

[2] Implementing a scaling correction for chi-square difference tests.

Table 1
Effects of KiVa intervention on bullying at T2, controlling for age, sex and bullying at T1 for low, medium and high popularity groups ($N = 911$).

	Unstandardized estimate	S.E.	Standardized estimate
Low popularity (N = 267)			
KiVa intervention	−0.182*	0.086	−0.146
Age	−0.068	0.076	−0.054
Sex (boy)	0.092	0.072	0.073
Bullying T1	0.708***	0.148	0.567
Medium popularity (N = 290)			
KiVa intervention	−0.330***	0.093	−0.231
Age	−0.013	0.100	−0.009
Sex (boy)	−0.094	0.142	−0.066
Bullying T1	0.989***	0.110	0.693
High popularity (N = 335)			
KiVa intervention	−0.030	0.089	−0.023
Age	−0.246**	0.096	−0.188
Sex (boy)	−0.012	0.091	−0.009
Bullying T1	0.876***	0.132	0.668

Note. Popularity data were missing for 19 of the 911 bullies selected for the analyses.
*$p < .05$. **$p < .01$. ***$p < .001$.

not differ from that in the medium popularity group, Satorra–Bentler scaled $\chi^2(1) = 1.59$, $p = .207$, or high popularity group, Satorra–Bentler scaled $\chi^2(1) = 2.06$, $p = .151$.

Across popularity groups, bullying scores at T1 strongly and positively predicted bullying scores at T2 ($ps < .001$); the higher the proportion of peer nominations received for bullying behavior before the start of the program, the higher the proportion of peer nominations received for bullying behavior after one year of program implementation. Age and sex were not significant predictors of bullying at T2 in the low-popularity group ($p = .372$ and $p = .201$, respectively), nor in the medium-popularity group ($p = .897$ and $p = .507$, respectively). In the high-popularity group, age had a significant effect on T2 bullying ($p = .010$); older children received lower bullying scores. There was no effect of sex ($p = .894$).

Discussion

In line with our hypothesis, results show that a nine-month exposure to the KiVa anti-bullying program led to a decrease in peer nominations for bullying behavior for bullies of medium or low popular status but not for highly popular bullies. This suggests that KiVa may be less effective at reducing bullying among perpetrators enjoying high social power in the peer group in comparison to their less popular counterparts. The finding that popular bullies may be the most resistant to anti-bullying intervention efforts is particularly worrisome as the KiVa program was developed with an eye towards bullies' powerful position in the peer group, by targeting bystanders and the support – either manifest or implicit – that they provide to their bullying peers during bullying incidents.

Nonetheless, our study indicates that targeting the peer group as a whole – as a growing number of programs now do – and no longer focusing solely on victims and perpetrators of bullying is a step in the right direction. Popularity is not an inherent characteristic of individuals but is determined by the peer group. It is strongly dependent on the social context; in particular, the association between popularity and aggression does vary as a function of group features (e.g., Garandeau, Ahn, & Rodkin, 2011). In order to be successful, anti-bullying programs will need to break the bullying–popularity cycle, and that involves altering the attitudes of all peers towards the bullies in their classrooms.

This study is the first, to our knowledge, to test whether the effectiveness of an anti-bullying program differs depending on bullies' level of popularity. However, there are several limitations that need to be considered when interpreting the findings. As mentioned earlier, the frequency and intensity of bullying behavior can only be assumed from the peer nomination procedure that we used but are not directly assessed by it. While we believe that it is reasonable to infer that a child displays a behavior less frequently when the proportion of classmates nominating him for this behavior decreases, we cannot be certain that it is always the case. Peer nominations also have the disadvantage of relying on what classmates can observe (see Olweus, 2013). Some bullies may resort to sophisticated, almost invisible, aggressive means, which may be difficult for most to detect. For instance, a decline in the peer-reported bullying score of a child may simply reflect the fact that the child switched from engaging in overt forms of bullying, such as beating someone up on the playground, to talking others into doing it. There is also a possibility that classmates keep nominating as a bully a peer who no longer engages in the behavior but retains this reputation in others' minds. Nevertheless, since such limitations likely apply to all bullying students and the objective of the study was to make comparisons across bully groups, the adverse effects of these limitations on the validity of our results should be relatively small.

The main question raised by the present finding is: how can we increase the effectiveness of anti-bullying interventions with high-status bullying perpetrators? Various approaches with different underlying ideologies can be considered. Consistent with the view that bullying is evolutionarily adaptive, one possible strategy is to offer dominance-aspiring youth opportunities to gain status through prosocial means (Ellis et al., 2012; Volk et al., 2012). A large body of research, drawing from adult and youth populations, suggests that enacting prosocial behaviors, such as giving or helping, allows individuals to enhance their status among their peers (e.g., Bereczkei, Birkas, & Kerekes, 2010; Hardy & Van Vugt, 2006; Olthof et al., 2011). Similarly, studies on resource–control theory have shown that children can achieve social dominance via prosocial strategies as much through coercive means (Hawley, 2003; Pellegrini & Bartini, 2001; Roseth, Pellegrini, Bohn, Van Ryzin, & Vance, 2007). The objective of such intervention would be to convince bullies that aggressive, bullying behaviors are not essential to maintain their social status. This approach acknowledges individuals' desire for status and aims at redirecting their status-enhancing strategies from violent abuse of power to less harmful and less risky ways to satisfy those needs.

Another possible approach is to make bullying less rewarding by challenging the established status hierarchy and striving to promote more equal relationships at the group level. Higher levels of status hierarchy in children and adolescents peer groups have been found to be associated with higher rates of bullying and victimization (Garandeau, Lee, & Salmivalli, 2013; Wolke, Woods, & Samara, 2009). Further studies suggest that young bullies are better accepted and enjoy higher status in more hierarchical classrooms (Ahn, Garandeau, & Rodkin, 2010; Garandeau et al., 2011; Schäfer, Korn, Brodbeck, Wolke, & Schulz, 2005). Bullying behaviors are favored by stable hierarchies where low-status victims have no possibility to escape. In addition to teaching children about the importance of treating everybody equally, it could therefore be advantageous to provide children with more opportunities to interact and form bonds with peers other than their classmates. An abundance of social resources (i.e., potential relationships) should make children less competitive and therefore less likely to bully (see Pellegrini, Roseth, Van Ryzin, & Solberg, 2011). Low-status children, who are the most vulnerable to victimization, may find it easier to escape the influence of a bully if they can create friendships in a new network. When peers have a choice of relationships, they should be less apt to accept the asymmetrical ties imposed by high-status bullies. In support of this strategy are studies showing that a) the absence of transition between primary and secondary (i.e., stability of peer groups) is associated with higher prevalence of bullying issues (Farmer, Hamm, Leung, Lambert, & Gravelle, 2011), b) rates of bullying may be higher in smaller classrooms (Garandeau et al., 2013) and c) changing school or classroom is, according to former victims, one of the most efficient ways to make bullying stop (Frisén, Hasselblad, & Holmqvist, 2012).

The present study points to the limited effectiveness of current anti-bullying interventions, which may fail to decrease bullying in highly

popular bullies. Going forward, we propose that anti-bullying strategies incorporate curricula that teach bullies socially acceptable ways to gain or maintain their status, or target the hierarchical setting of peer groups in order to make bullying less rewarding. Further empirical evidence is needed to determine which of these approaches would be most effective.

Another avenue for future intervention research is the examination of the effects of anti-bullying programs on the peer status of bullies. Children who keep engaging in bullying behaviors despite intervention efforts may nevertheless suffer a loss in popularity if the program successfully modified peers' assessments of aggressive classmates. In the long run, a status decline may encourage these bullies to cease their behaviors. Similarly, it is possible that, regardless of the results obtained in overall reductions in bullying, intervention programs help improve the status of victims. In turn, a gain in status would certainly alleviate the negative consequences of peer harassment for the targeted children. The evaluations of anti-bullying programs should go beyond the measurement of decreases in the undesired behavior, by examining for whom the intervention is most or least effective and by taking into account changes in the perceptions of perpetrators and targets of such behaviors.

References

Ahn, H. -J., Garandeau, C. F., & Rodkin, P. C. (2010). Effects of classroom embeddedness and density on the social status of aggressive and victimized children. *Journal of Early Adolescence, 30*, 76–101.

Baldry, A.C., & Farrington, D. P. (2007). Effectiveness of programs to prevent school bullying. *Victims and Offenders, 2*, 183–204.

Bereczkei, T., Birkas, B., & Kerekes, Z. (2010). Altruism towards strangers in need: Costly signaling in an industrial society. *Evolution and Human Behavior, 31*, 95–103.

Berger, C., & Rodkin, P. C. (2012). Group influences on individual aggression and prosociality: Early adolescents who change peer affiliations. *Social Development, 21*, 396–413.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York, NY: Guilford Press.

Caravita, S.C. S., & Cillessen, A. H. N. (2012). Agentic or communal? Developmental differences in the associations among interpersonal goals, popularity, and bullying. *Social Development, 21*, 376–391.

Caravita, S.C. S., Di Blasio, P., & Salmivalli, C. (2009). Unique and interactive effects of empathy and social status on involvement in bullying. *Social Development, 18*, 140–163.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255.

Cialdini, R. B., & Richardson, K. D. (1980). Two indirect tactics of impression management: Basking and blasting. *Journal of Personality and Social Psychology, 39*, 406–416.

Cillessen, A. H. N., & Borch, C. (2006). Developmental trajectories of adolescent popularity: A growth curve modeling analysis. *Journal of Adolescence, 29*, 935–959.

Cillessen, A. H. N., & Mayeux, L. (2004). From censure to reinforcement: Developmental changes in the association between aggression and social status. *Child Development, 75*, 147–163.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.)Hillsdale, NJ: Lawrence Earlbaum Associates.

Cohen, G. L., & Prinstein, M. J. (2006). Peer contagion of aggression and health-risk behavior among adolescent males: An experimental investigation of effects on public conduct and private attitudes. *Child Development, 77*, 967–983.

Currie, C., Gabhainn, S. N., Godeau, E., Roberts, C., Smith, R., Picket, W., et al. (2008). Inequalities in young people's health. Health behaviour in school-aged children. International report from the 2005/2006 survey. http://www.euro.who.int/data/assets/pdf_file/0005/53852/E91416.pdf (Available from URL:)

de Bruyn, E. H., Cillessen, A. H. N., & Wissink, I. B. (2010). Associations of peer acceptance and perceived popularity with bullying and victimization in early adolescence. *Journal of Early Adolescence, 30*, 543–566.

Dijkstra, J. K., Cillessen, A. H. N., Lindenberg, S., & Veenstra, R. (2010). Basking in reflected glory and its limits: Why adolescents hang out with popular peers. *Journal of Research in Adolescence, 20*, 942–958.

Ellis, B. J., Del Giudice, M., Dishion, T. J., Figueredo, A. J., Gray, P., Griskevicious, V., et al. (2012). The evolutionary basis of risky adolescent behavior: Implications for science, policy, and practice. *Developmental Psychology, 48*, 598–623.

Enders, C. K. (2010). *Applied missing data analysis.* New York, NY: Guilford Press.

Faris, R., & Felmlee, D. (2011). Status struggles: Network centrality and gender segregation in same- and cross-gender aggression. *American Sociological Review, 76*, 48–73.

Farmer, T. W., Hamm, J. V., Leung, M., Lambert, K., & Gravelle, M. (2011). Early adolescent peer ecologies in rural communities: Bullying in schools that do and do not have a transition during middle grades. *Journal of Youth and Adolescence, 40*, 1106–1117.

Ferguson, C. J., San Miguel, C., Kilburg, J. C., & Sanchez, J. (2007). The effectiveness of school-based anti-bullying programs: A meta-analytic review. *Criminal Justice Review, 32*, 401–414.

Frisén, A., Hasselblad, T., & Holmqvist, K. (2012). What actually makes bullying stop? Reports from former victims. *Journal of Adolescence, 35*, 981–990.

Galinsky, A.D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006). Power and perspectives not taken. *Psychological Science, 17*, 1068–1074.

Garandeau, C. F., Ahn, H. -J., & Rodkin, P. C. (2011). The social status of aggressive students across contexts: The role of classroom peer status hierarchy, academic achievement, and grade. *Developmental Psychology, 47*, 1699–1710.

Garandeau, C. F., & Cillessen, A. (2006). From indirect aggression to invisible aggression: A conceptual view on bullying and peer group manipulation. *Aggression and Violent Behavior, 11*, 612–625.

Garandeau, C. F., Lee, I. A., & Salmivalli, C. (2013). *Inequality matters : Classroom status hierarchy and adolescents' bullying.* : Journal of Youth and Adolescence Advance online publication.

Graham, J. W., Olchowski, A. E., & Gilreath, T. D. (2007). How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prevention Science, 8*, 206–213.

Gruenfeld, D. H., Inesi, M. E., Magee, J. C., & Galinsky, A.D. (2008). Power and the objectification of social targets. *Journal of Personality and Social Psychology, 95*, 11–127.

Haataja, A., Voeten, M., Boulton, A. J., Ahtola, A., Poskiparta, E., & Salmivalli, C (submitted for publication). *The KiVa antibullying curriculum and outcome: Does fidelity matter?* (submitted for publication).

Hardy, C. L., & Van Vugt, M. (2006). Nice guys finish first: The competitive altruism hypothesis. *Personality and Social Psychology Bulletin, 32*, 1402–1413.

Hawley, P. H. (2003). Prosocial and coercive configurations of resource control in early adolescence: A case for the well-adapted Machiavellian. *Merrill-Palmer Quarterly, 49*, 279–310.

Houghton, S. J., Nathan, E., & Taylor, M. (2012). To bully or not to bully, that is not the question: Western Australian early adolescents' in search of a reputation. *Journal of Adolescent Research, 27*, 498–522.

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Isaacs, J., Hodges, E. V. E., & Salmivalli, C. (2008). Long-term consequences of victimization by peers: A follow-up from adolescence to young adulthood. *European Journal of Developmental Science, 2*, 387–397.

Kärnä, A., Voeten, M., Little, T. D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011). A large-scale evaluation of the KiVa anti bullying program: Grades 4–6. *Child Development, 82*, 311–330.

Kärnä, A., Voeten, M., Poskiparta, E., & Salmivalli, C. (2010). Vulnerable children in varying classroom contexts: Bystanders' behaviors moderate the effects of risk factors on victimization. *Merrill-Palmer Quarterly, 56*, 261–282.

LaFontana, K. M., & Cillessen, A. H. N. (2010). Developmental changes in the priority of perceived status in childhood and adolescence. *Social Development, 19*, 130–147.

Lease, A.M., Kennedy, C. A., & Axelrod, J. L. (2002). Children's social constructions of popularity. *Social Development, 11*, 87–109.

Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development, 31*, 357–365.

Marks, P. E. L., Cillessen, A. H. N., & Crick, N. R. (2012). Popularity contagion among adolescents. *Social Development, 21*, 501–521.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58*, 525–543.

Merrell, K., Gueldner, B., Ross, S., & Isava, D. (2008). How effective are school bullying intervention programs? A meta-analysis of intervention research. *School Psychology Quarterly, 23*, 26–42.

Merten, D. E. (1997). The meaning of meanness: Popularity, competition and conflict among junior high school girls. *Sociology of Education, 70*, 175–191.

Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user's guide* (6th ed.). Los Angeles: CA: Muthén & Muthén.

Ojanen, T., Gronroos, M., & Salmivalli, C. (2005). An interpersonal circumplex model of children's social goals: Links with peer-reported behavior and sociometric status. *Developmental Psychology, 41*, 699–710.

Olthof, T., Goossens, F. A., Vermande, M. M., Aleva, E. A., & van der Meulen, M. (2011). Bullying as strategic behavior: Relations with desired and acquired dominance in the peer group. *Journal of School Psychology, 49*, 339–359.

Olweus, D. (1996). The Revised Olweus Bully/Victim Questionnaire. Mimeo. Bergen, Norway: Research Center for Health Promotion (HEMIL Center), University of Bergen.

Olweus, D. (2013). School bullying: Development and some important challenges. *Annual Review of Clinical Psychology, 9*, 751–780.

Peeters, M., Cillessen, A. H. N., & Scholte, R. H. J. (2010). Clueless or powerful? Identifying subtypes of bullies in adolescence. *Journal of Youth and Adolescence, 39*, 1041–1052.

Pellegrini, A.D., & Bartini, M. (2001). Dominance in early adolescent boys: Affiliative and aggressive dimensions and possible functions. *Merrill-Palmer Quarterly, 47*, 142–163.

Pellegrini, A.D., Roseth, C. J., Van Ryzin, M. V., & Solberg, D. (2011). The place of social dominance and aggression in children's and adolescents' popularity. In A. Cillessen, D. Schwartz, & L. Mayeux (Eds.), *Popularity in the peer system.* New York, NY: Guilford Press.

Pöyhönen, V., Juvonen, J., Peets, K., & Salmivalli, C. (submitted for publication). *Defending the victimized classmates: What explains variation between classrooms?* (submitted for publication).

Pozzoli, T., & Gini, G. (2010). Active defending and passive bystanding behavior in bullying: The role of personal characteristics and perceived peer pressure. *Journal of Abnormal Child Psychology, 38*, 815–827.

Rigby, K. (2003). Consequences of bullying in schools. *Canadian Journal of Psychiatry, 48*, 583–590.

Rodkin, P. C., Farmer, T. W., Pearl, R., & Van Acker, R. (2000). The heterogeneity of popularity in boys: Antisocial and prosocial configurations. *Developmental Psychology, 36*, 14–24.

Roseth, C. J., Pellegrini, A.D., Bohn, C. M., Van Ryzin, M., & Vance, N. (2007). Preschoolers' aggression, affiliation, and social dominance relationships: An observational, longitudinal study. *Journal of School Psychology, 45*, 479–497.

Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior, 15*, 112–120.

Salmivalli, C., Lagerspetz, K., Bjorkqvist, K., Osterman, K., & Kaukiainen, A. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior, 22*, 1–15.

Salmivalli, C., Voeten, M., & Poskiparta, E. (2011). Bystanders matter: Associations between reinforcing, defending, and the frequency of bullying behavior in classrooms. *Journal of Clinical Child & Adolescent Psychology, 40*, 668–676.

Schäfer, M., Korn, S., Brodbeck, F. C., Wolke, D., & Schulz, H. (2005). Bullying roles in changing contexts: The stability of victim and bully roles from primary to secondary school. *International Journal of Behavioral Development, 29*, 323–335.

Sijtsema, J. J., Veenstra, R., Lindenberg, S., & Salmivalli, C. (2009). Empirical test of bullies' status goals: Assessing direct goals, aggression, and prestige. *Aggressive Behavior, 35*, 57–67.

Van Kleef, G. A., Oveis, C., Van der Lowe, I., LuoKogan, A., Goetz, J., & Keltner, D. (2008). Power, distress, and compassion: Turning a blind eye to the suffering of others. *Psychological Science, 19*, 1315–1322.

Veenstra, R., Lindenberg, S., Munniksma, A., & Dijkstra, J. K. (2010). The complex relation between bullying, victimization, acceptance, and rejection: Giving special attention to status, affection, and sex differences. *Child Development, 81*, 480–486.

Volk, A., Camilleri, J., Dane, A., & Marini, Z. (2012). Is adolescent bullying an evolutionary adaptation? *Aggressive Behavior, 38*, 222–238.

Wolke, D., Woods, S., & Samara, M. (2009). Who escapes or remains a victim of bullying in primary school? *British Journal of Developmental Psychology, 27*, 835–851.