KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland

Christina Salmivalli, Department of Psychology, University of Turku, Finland
Elisa Poskiparta, Centre for Learning Research, University of Turku, Finland

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland

Christina Salmivalli, Department of Psychology, University of Turku, Finland
Elisa Poskiparta, Centre for Learning Research, University of Turku, Finland

Vol. 6 (2) 2012
KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland

Christina Salmivalli, Department of Psychology, University of Turku, Finland
Elisa Poskiparta, Centre for Learning Research, University of Turku, Finland

The effects of a Finnish national school-based antibullying program (KiVa) were evaluated in a randomized controlled trial (2007–2009) and during nationwide implementation (since 2009). The KiVa program is found to reduce bullying and victimization and increase empathy towards victimized peers and self-efficacy to support and defend them. KiVa increases school liking and motivation and contributes to significant reductions in anxiety, depression, and negative peer perceptions. Somewhat larger reductions in bullying and victimization were found in the randomized controlled trial than in the broad rollout, and the largest effects were obtained in primary school (grades 1–6). The uptake of the KiVa program is remarkable, with 90 percent of Finnish comprehensive schools currently registered as program users.

Bullying, defined as repeated aggressive behavior against a victim who cannot readily defend him- or herself (Olweus 1991) is recognized as a recurrent and serious problem among school-age children across the world (Craig and Harel 2004). The targets of such treatment suffer psychosocial problems such as depression, anxiety, and peer rejection (Card and Hodges 2008; Hawker and Boulton 2003). For a number of victimized students, these experiences continue to affect their lives later on in adulthood (Isaacs, Hodges, and Salmivalli 2008). Besides the targets, bullying constitutes a risk for the healthy development of the perpetrators (Sourander et al. 2007) as well as for bystanders merely witnessing victimization (Rivers et al. 2009). The need for evidence-based interventions against bullying is therefore indisputable and urgent.

Numerous initiatives to prevent and tackle bullying have emerged during past decades, many of these taking the form of school-based programs. Although bullying may take place both in and out of school, school is a context where bullied students cannot escape their tormentors (school attendance being compulsory) while bullies are often socially rewarded by peers who join their behavior or just reinforce it by verbal or nonverbal signals (Salmivalli 2010). As bullying is typically related to peer group dynamics, such as norms emerging in classrooms, targeting any individual child’s behavior might not be enough: group problems need group solutions. The fact that large numbers of children and adolescents (i.e., whole cohorts) can be easily reached in the school context makes school-based programs potentially very cost-effective as well. On the other hand, schools often struggle with multiple expectations combined with limited resources; their readiness to implement prevention programs cannot be taken for granted. Program developers should provide them programs that are not only shown to be effective but that are also system-ready, possible to implement with fidelity in a school system.

The evidence on the effectiveness of school-based antibullying programs is overall much thinner than we would like it to be. The effects show great variability across programs and across studies, typically being small to modest in size (Ferguson et al. 2007; Merrell et al. 2008; Smith et al. 2004; Ttofi, Farrington, and Baldry 2008; Vreeman and Carroll 2007). Some of the antibullying programs even seem to bring about undesirable outcomes, i.e. increases in bullying problems. In the most recent meta-analysis on the
topic, however, Ttofi and Farrington (2011) conclude that school-based anti-bullying programs are effective. The programs selected for their analysis showed average reductions of 17–23 percent for bullying others and 17–20 percent for being bullied. Ttofi and Farrington point out that even relatively small effect sizes correspond to a substantial amount of bullying and victimization prevented and thus to huge amounts of suffering avoided.

Despite the conclusion of Ttofi and Farrington that school-based interventions work, it should be kept in mind that there are numerous programs “in the market” that 1) have not been evaluated at all; 2) have not been evaluated with methodological rigor; 3) have been evaluated but shown only minimal positive effects, no effects at all, or negative effects; or 4) have been shown to work in a small-scale trial but not during broad rollout, i.e., perhaps are not system-ready. Hopefully, with growing awareness of the importance of evidence among stakeholders and practitioners, the demand for rigorous evaluations of antibullying programs will increase – to the benefit of both research and good practice.

In Finland, the Ministry of Education and Culture has funded the development, evaluation, and national implementation of the KiVa antibullying program. Our exceptionally large data sets will enable us to continue studies on the effects of the program and its mechanisms for years to come. In the present article we summarize the studies evaluating KiVa, based on data sets from a large-scale randomized controlled trial, and from a nonrandomized trial during the first year of broad implementation.

1. KiVa Antibullying Program

In 2006, the Finnish Ministry of Education and Culture commissioned the University of Turku to develop and evaluate an anti-bullying program for Finnish comprehensive schools (the basic nine-year education from grades 1 to 9). From the very beginning, the vision was to develop a program that would be suitable for nationwide implementation.

The resulting KiVa antibullying program is predicated on the idea that how peer bystanders react when witnessing bullying is crucial for putting an end to (or sustaining) it (e.g., Salmivalli et al. 1996; Salmivalli 2010). Influencing the peer context is therefore an essential part of the KiVa program. The program is designed to produce its effects, first of all, by encouraging students to support victimized peers instead of providing social rewards to the bullies. In addition, adults (teachers as well as parents) are provided with information about bullying and efficacy to intervene and prevent it.

KiVa includes both universal and indicated actions. The core of the universal actions consists of theme lessons (primary school) and theme days (secondary school). The topics cover a variety of issues related to group interaction and group pressure, the mechanisms and consequences of bullying, and most importantly, what students can do together to counter bullying and support their victimized peers. Virtual learning environments (such as anti-bullying computer games) are an integral part of the universal actions. Their content is closely connected to the topics of the student lessons and themes, enhancing the learning process and motivating students to apply the learnt skills in everyday interactions with peers (Poskiparta et al. 2012). In addition, the universal actions include a parents’ guide, web resources for teachers, and materials reminding both students and school personnel of KiVa (posters, highly visible, bright-colored vests with the KiVa logo for teachers supervising recess time).

The indicated actions involve discussions with victims and bullies, as well with selected prosocial classmates, who are challenged to support the victimized classmate. The discussions with the bullies and victims are conducted by KiVa teams within the schools, while the classroom teacher organizes separate meetings with potential supporters of the victim (for a more detailed description of program content, see Salmivalli, Kärnä, and Poskiparta 2010a, 2010b).

2. Evaluation Strategy and Methods Used

Besides testing the effectiveness of the KiVa antibullying program, the evaluation studies were designed to answer the most pressing questions in the area of bullying interventions: when (under which circumstances) the program works, for whom does it work, and how does it produce its
effects. At the same time, the studies were designed to meet the highest methodological standards and to include assessments that have been neglected in much of the previous research (e.g., implementation fidelity). Many of these studies are still ongoing. In the present article we provide an overview of the findings concerning the main effects of KiVa after one school year (nine months of implementation) under two different circumstances: a randomized controlled trial and large-scale implementation. In addition, we briefly summarize the first results concerning implementation fidelity.

We put the effectiveness of the KiVa program to a stringent test in a randomized controlled trial across all grade levels in comprehensive education (grades 1 to 9). Students in these grade levels are aged seven to fifteen years. All schools in mainland Finland were informed about the evaluation study and invited to participate either as intervention or control schools. From the volunteering schools, 234 were chosen and assigned to the two conditions. The control schools were told that they could start implementing the KiVa program after serving as controls for one year – which most of them did.

Most data were collected by means of online questionnaires where students logged in using the same personal user ID at each assessment point – thus, individual students were followed over time. Bullying and victimization were assessed by self-reports, peer-reports, and dyadic data (e.g., “who has been bullying you?”) enabling us to link bullied children to their tormentors and to identify bully-victim dyads. In addition, data were collected on numerous variables related to bullying (attitudes, bystander behaviors, perceptions of teachers’ efficacy to tackle bullying) and defending of victimized peers (empathy towards victims, self-efficacy, outcome expectations regarding defending). Students’ sociometric status and psychosocial (e.g., self-esteem, peer perceptions, depression, social anxiety, loneliness) and academic (e.g., school liking, academic motivation, academic performance) adjustment were also assessed. Finally, data were collected from teachers who delivered the program lessons as well as from school staff who were members of KiVa teams tackling the cases of bullying that came to attention. The choice of variables was guided by our preliminary hypotheses on the mediating mechanisms of program effects, possible moderators of the effects, as well as constructs that might be positively influenced by reductions in bullying or victimization.

After the randomized controlled trial, KiVa became available for all Finnish schools. There were several reasons to continue collecting data in schools that adopted the program. First, since our control schools became intervention schools after the first year, continuing the randomized controlled trial was impossible. We wanted to know, however, what happened to the effects of the KiVa program in the long run. We also wanted to find out what happens to implementation fidelity when the program is implemented on a large scale. Moreover, ongoing data collection enabled us to provide annual feedback to the schools implementing the program: The annual online survey served as a tool to monitor their success, as well as the quality of implementation.

The first large pre-test survey for schools starting the implementation in fall 2009 took place in May 2009, followed by an annual assessment every May. Already after the first year we were able to estimate program effects utilizing a cohort-longitudinal design (Olweus and Alsaker 1991). In this design, post-test data from students in each age cohort were compared with baseline data from same-aged students from the same schools (i.e., in the previous cohort), who had not yet been exposed to the intervention. For instance, grade two students who had been targeted by KiVa for one year (post-test data collected in May 2010) were compared with grade two students from the same schools who had not yet been involved (pre-test data from May 2009).

3. Results of Evaluation Studies

3.1. Randomized Controlled Trial

The findings from the randomized controlled trial were promising: the first phase of evaluation involving grades 4–6 (Kärnä et al. 2011b) showed that KiVa significantly reduced both victimization and bullying. After the first school year (i.e., nine months of implementation of KiVa) the odds of being bullied systematically (at least two to three times a week) were about 1.5 times higher for a con-
control school student than for a student in an intervention school, whereas the odds of bullying others systematically were 1.3 times higher for a control school student than for an intervention school student. These effect sizes correspond to approximately 30 percent and 17 percent reductions in victimization and bullying. Importantly, peer-reported victimization and bullying were significantly reduced as well (effect sizes equal or larger than in the case of self-reporting). The program also resulted in reductions in negative bystander behaviors (reinforcing the bully), as well as increases in empathy towards victimized peers and self-efficacy to support and defend them.

In a further study focusing on different forms of victimization in grades 4–6 (Salmivalli, Kärnä, and Poskiparta 2011) KiVa was shown to reduce each of the examined nine forms (physical, verbal, social exclusion, social manipulation, threatening, racist, material, and sexual victimization, and cybervictimization), the reductions varying between -20 percent (threatening) and -63 percent (material victimization, such as taking or breaking the target peer’s belongings).

The second phase of the evaluation involving also younger (grades 1–3) and older (grades 7–9) students indicated that the effectiveness of the program varied considerably across grade levels (Kärnä et al. forthcoming). Overall, the effects were larger in primary than secondary grade levels. The average weighted odds ratios across all grade levels 1–9 were 1.28 for victimization and 1.30 for bullying, corresponding to reductions of about 20 percent in each case.

During the randomized controlled trial, we also tested two different approaches to dealing with children who had been involved in bullying others, which we refer to as confronting and nonconfronting approaches (Garandeau, Poskiparta, and Salmivalli 2012). In the former approach, the bully was firmly told that his/her behavior was not tolerated and had to cease immediately whereas in the latter, the adults did not blame the bully but rather shared their concern about the victim with him or her. In half of the schools involved in the randomized controlled trial as intervention schools, KiVa team members were instructed to use the confronting approach, whereas the other half used the nonconfronting approach. The outcome variable was the victim’s report (in the follow-up discussion with KiVa team members) on whether the bullying had stopped, decreased, increased, or remained unchanged. In 79 percent of all cases, victims reported that the bullying had stopped completely, and overall the two methods were equally effective at making the bullying stop. There were some moderators of their effectiveness, however: The nonconfronting approach was more successful than the confronting approach in cases of long-term victimization and in primary school. The confronting approach, on the other hand, was more effective in cases involving more than one bully.

Few anti-bullying program evaluations have investigated potential positive “side-effects” of the programs, in other words, beneficial outcomes that go beyond the initially intended reductions in bullying and victimization. In a study based on the randomized controlled trial data from grades 4 to 6 (Williford et al. 2012), the KiVa program was found to be effective in reducing students’ internalizing symptoms (anxiety and depression) and improving their peer-group perceptions. Furthermore, Salmivalli, Garandeau, and Veenstra (2012) found that KiVa increased school liking, academic motivation, and even academic performance among students in KiVa schools (compared to control schools).

3.2. Going to Scale: National Rollout

The national launch of the KiVa antibullying program started in the fall of 2009, when 1,450 schools started to implement it. In 2010 and in 2011 new schools joined, and at present there are about 2,500 schools implementing the program. They represent 90 percent of all schools providing comprehensive education in Finland. We continue collecting data in these schools, both from students and staff. Whereas most evidence concerning the effects of antibullying programs comes from short-term, small-scale efficacy studies, our aim is to study both implementation and effectiveness during national rollout over several years.

The evaluation of program effects during the national rollout is based on students’ responses to web-based surveys that are completed every spring in schools implementing the program. The effects were generally weaker during the
broad rollout than in the randomized controlled trial (Kärnä et al. 2011a). They were statistically significant at the primary school level (Grades 1–6) with respect to bullying others, as well as being bullied. At the secondary, or junior high school level (Grades 7–9), the effects for bullying others were in the right direction but not significant, and the effects for being bullied were just at the border of being significant (except in grade eight, where the reduction in victimization was significant). On average, the prevalence of children bullying others, as well as of those being bullied systematically (at least two to three times a week) were both reduced by about 15 percent during the first year of national rollout, the odds ratios being 1.22 for victimization and 1.18 for bullying. Effects of this magnitude mean that had all schools in Finland been implementing KiVa, the reductions would amount to about 7,500 bullies and 12,500 victims during the first one-year period. This demonstrates how even rather small effect sizes can make a huge difference in the lives of numerous children and adolescents.

### 3.3. Implementation Fidelity

Implementation fidelity, referring to the extent to which an intervention program is delivered as planned (Dusenbury et al. 2003), is a critical precondition for success of any prevention/intervention program. However, in studies evaluating antibullying programs, implementation has often been assessed at a very general level – if at all. In the evaluation of KiVa we placed concerted emphasis on the assessment of various aspects of implementation (such as preparation, dose, coverage, and student responsiveness) of the different program components.

In the randomized controlled trial we attempted to capture, besides the level of implementation, also the implementation process by gathering monthly data from teachers delivering the student lessons involved in the program. Furthermore, KiVa teams tackling acute cases of bullying reported back to us each step taken when handling a case of bullying. During the broad rollout of KiVa (since fall 2009) we continued gathering data on implementation via annual online surveys for the school personnel. There is already evidence of a positive association between the level of implementation and reduction in victimization, both from the randomized controlled trial (Haataja, Voeten, and Salmivalli 2011) and broad rollout (Kärnä et al. 2011a).

Although the level of implementing KiVa program was overall high (e.g., most teachers delivered the majority of program lessons), it tended to decrease already during the first academic year (from fall to spring; Haataja, Voeten, and Salmivalli 2011) and even more so in subsequent years. Overall, implementation fidelity was somewhat lower during the broad rollout than during the randomized controlled trial (Salmivalli, Haataja, and Poskiparta 2011). Whereas primary school teachers delivered on average 8.7 out of 10 lessons during the randomized controlled trial, the corresponding number was 7.8 lessons during the first and 7.2 lessons during the second year of broad implementation. An important future task is to identify individual and school-level factors enhancing the likelihood of high-quality implementation. For instance, support from the school principal for antibullying work seems to be a crucial precondition for the successful delivery of program lessons (Ahtola et al. 2012).

### 4. Discussion

Studies of our randomized controlled trial and evaluation of large-scale implementation of the KiVa antibullying program show that the program is effective in reducing bullying and victimization (Kärnä et al. 2011a, 2011b, forthcoming). During the randomized controlled trial, the average reductions in bullying and victimization amounted to -20 percent whereas the corresponding reduction during broad rollout of KiVa was -15 percent. The former number corresponds to the average effects found in the meta-analysis by Farrington and Ttofi (2009). It should be noted, however, that no previous study has included such a wide age range as ours. Even though the effects of KiVa were small (even minimal) at secondary school level (grades 7–9) they were clearly above average in primary school grades. Moreover, the effects of KiVa (even when calculated across all grade levels) were stronger than the effects obtained in previous randomized controlled trials, and they were confirmed by multiple informants (children themselves as well as their peers). It should also be remembered that all effects reported in the present paper concerned changes after only nine months of program implementation.
Besides reducing bullying and victimization, KiVa was shown to reduce internalizing problems (Williford et al. 2012). It also led to increases in well-being and school liking among a much wider group of children than just previous victims and bullies (Salmivalli et al. 2012). The indicated actions (i.e. discussions between schools’ KiVa team members and children involved in bullying) were highly effective (Garandeau et al. 2012).

We are not aware of any country where a school-based anti-bullying program has spread at the pace KiVa did in Finland. How can the willingness of Finnish schools to adopt the program be explained? One explanation is the support provided by the government, which enabled the schools that introduced the program during 2009–2011 get all the materials as well as the two-day pre-implementation training free of charge. Second, we believe that the KiVa program was regarded as feasible by school staff and its reputation as an effective and user-friendly program spread fast – partly because of the relatively broad media attention devoted to the program. Third, two school massacres took place in Finland in 2007 and 2008, just before the broad rollout of KiVa started. They were widely linked to bullying problems and thus generated a lot of discussion about bullying and the need for effective prevention.

4.1. Strengths and Limitations
Throughout the KiVa evaluation project, we made concerted efforts to meet the highest methodological standards by including an appropriate control condition, random assignment, multilevel modeling of hierarchical data, multitmethod and multi-informant outcome measures, systematic monitoring of implementation, adequate sample size, and accurate handling of missing data. We are not aware of any other evaluation of a bullying intervention with similar methodological rigor, an equally large and representative sample, and such numerous outcome variables (both self- and peer-reported).

Several limitations of our evaluation should be pointed out, however. It was unfortunate that our randomized controlled trial only lasted for one school year: after that, it was not possible to assess the effects of the KiVa program by comparing changes occurring in KiVa schools with those taking place in control schools. We can only study the long-term effects of KiVa utilizing the cohort-longitudinal design. However, as this design is potentially sensitive to history effects it is not considered as strong as the RCT design.

All our data were collected by (mostly online) questionnaires. It would have been desirable to collect additional observational data or perhaps interview students in a randomly selected subsample. Whereas observational data is free from response bias, interviews might have provided insight into how the students perceived the program lessons and what they felt made them think and behave differently. Assessing implementation by observation, in addition to teacher and student reports, might have provided a good addition to questionnaire data.

4.2. Future Directions
Many planned studies, for example concerning the mediators and moderators of the effects of the KiVa anti-bullying program, are in progress. Besides the already existing data sets, it will be exciting to follow the implementation and effectiveness of the KiVa program during years to come. Going to scale is not unproblematic: one of the challenges is the quality of implementation in a situation where schools get limited external guidance. As much as program developers would like school staff to put effort into delivering the program as recommended, reality often restrains their possibilities for doing so. Multiple projects running alongside the usual curriculum, lack of resources, organizational changes, and work stress are among the factors that might decrease school staff’s willingness to implement a program, even where they believe it has positive outcomes. The greater the standardization of a program, however, the easier it is for local agents to grasp the model and to identify deviations from it (Bradach 2004). As KiVa includes concrete guidelines and schedules for the different measures, it is relatively easy to implement with fidelity even with minimal external support. The motivation to implement the program may diminish over time, however, and support for schools is essential. It remains to be seen whether the effects obtained so far will be sustained and – hopefully – grow further.
References


Salmivalli, Christina, Anne Haataja, and Elisa Poskiparta. 2011. Implementation Fidelity of the KiVa Antibullying Program During Randomized Controlled Trial and Broad Dissemination. Presentation in symposium on large-scale implementation of school-based anti-bullying programs: strategies and some challenges, chaired by D. Olweus at the 15th European Conference of Developmental Psychology, Bergen, Norway.


Christina Salmivalli  
eijasal@utu.fi

Elisa Poskiparta  
elisa.poskiparta@utu.fi