

Systematic Review Protocol

Title:

Effectiveness of Programs to Prevent School Bullying

Reviewers:

David P. Farrington

Institute of Criminology, Cambridge University

Anna C. Baldry

*Department of Psychology, Second University of Naples
and Intervict Institute, Tilburg University*

Britta Kyvsgaard

Ministry of Justice, Copenhagen, Denmark

Maria M. Ttofi

Institute of Criminology, Cambridge University

Lead Reviewer:

David P. Farrington,

*Institute of Criminology, Sidgwick Avenue,
Cambridge CB3 9DT, UK.*

Tel. (44) 1223 872555; Fax (44) 1223 335356;

email: dpf1@cam.ac.uk

Sources of Support

Nordic Campbell Center

Swedish National Council on Crime Prevention

Revised February 2008

Contents

1. Background for the Review
2. Objectives of the Review
3. Methods
 - 3.1 Criteria for inclusion and exclusion of studies
 - 3.2 Search strategies
 - 3.3 Description of methods used in the primary research
 - 3.4 Criteria for determination of independent findings
 - 3.5 Details of study coding categories
 - 3.6 Statistical procedures and conventions
 - 3.7 Treatment of qualitative research
4. Time Frame
5. Plans for Updating the Review
6. Acknowledgements
7. Statement Concerning Conflict of Interest
8. References and some Key Studies
9. Table 1: Key Bullying Prevention Projects
10. Appendix: Coding Protocol

1. Background

The definition of school bullying includes several key elements: physical, verbal, or psychological attack or intimidation that is intended to cause fear, distress, or harm to the victim; an imbalance of power (psychological or physical), with more powerful child (or children) oppressing less powerful ones; and repeated incidents between the same children over a prolonged period (Farrington, 1993; Olweus, 1993; Roland, 1989). School bullying can occur in school or on the way to or from school. It is not bullying when two persons of the same strength (physical, psychological, or verbal) victimize each other.

Bullying is different from aggression or violence; not all aggression/violence involves bullying, and not all bullying involves aggression/violence. For example, bullying includes being called nasty names, being rejected, ostracized or excluded from activities, having rumors spread about you, having belongings taken away, teasing and threatening (Baldry & Farrington, 1999). Our aim is to review programs that are specifically intended to prevent or reduce school bullying, not programs that are intended to prevent or reduce school aggression or violence.

School bullying is perceived to be an important social problem in many different countries. The nature and extent of the problem, and research on it, in 21 different countries, is reviewed in Smith *et al.* (1999). Special methods are needed to study bullying in different countries because of problems of translating the term “bullying” into different languages. Smith *et al.* (2002) have reviewed the meaning of bullying in 14 different countries. In studying bullying, researchers often ask about specific components such as “hit him/her on the face” or “excluded him/her from games” rather than use the actual word “bullying” in their interviews and questionnaires (Pateraki & Houndoumadi, 2001). However, they still use the word “bullying” in their reports.

Many school-based intervention programs have been devised and implemented in an attempt to reduce school bullying. These have been targeted on bullies, victims, peers,

teachers, or on the school in general. Many programs seem to have been based on commonsense ideas about what might reduce bullying rather than on empirically-supported theories of why children bully, why children become victims, or why bullying events occur.

The first large-scale anti-bullying program was implemented in Norway in 1983. A more intensive version of the national program was evaluated in Bergen by Olweus (1991). This program aimed to increase awareness and knowledge of teachers, parents, and students about bullying and to dispel myths about it. Booklets were distributed to schools, folders were distributed to families, and videos about bullying were shown and discussed in schools. Also, schools received information about the nature and extent of bullying (based on questionnaires completed by students), teachers were encouraged to develop explicit rules about bullying, and monitoring and supervision of children in the playground was improved.

The evaluation by Olweus (1991) showed a dramatic decrease in victimization of about half after the program. Since then at least 15 other large-scale anti-bullying programs, some inspired by Olweus and some based on other principles, have been implemented and evaluated in at least 10 other countries. Baldry and Farrington (2007) reviewed 16 major evaluations in 11 different countries and concluded that 8 produced desirable results, 2 produced mixed results, 4 produced small or negligible effects, and 2 produced undesirable results. Most programs were quite complex, and the effectiveness of different components of programs was not clear. These 16 evaluations are listed after the References.

The best single source of reports of anti-bullying programs is the book edited by Smith, Pepler, and Rigby (2004), which contains descriptions of 13 programs in 11 different countries. The most relevant existing reviews are by Smith *et al.* (2004), who summarized effect sizes in 14 whole-school anti-bullying programs, and by Vreeman and

Carroll (2007), who reviewed 26 school-based programs. These two prior reviews are of high quality. However, neither carried out a full meta-analysis measuring weighted mean effect sizes and correlations between study features and effect sizes. The Smith *et al.* (2004) review covered only 14 evaluations up to 2002, 6 of which were uncontrolled. The Vreeman and Carroll (2007) review covered 26 evaluations up to 2004, and was restricted to studies published in the English language. We hope to go beyond these previous reviews by (a) searching for evaluations up to the end of 2007, (b) searching for international evaluations, and (c) carrying out more extensive meta-analyses, as specified in this protocol.

American research is generally targeted on school violence or peer victimization rather than bullying. There are a number of existing reviews of school violence programs and school-based interventions for aggressive behavior (e.g. Howard *et al.*, 1999; Mytton *et al.*, 2006; Wilson *et al.*, 2003; Wilson & Lipsey, 2007). We will consult these, but we must emphasize that our research aims to review anti-bullying programs specifically.

2. Objectives of the Review

The main objective is to assess the effectiveness of school-based anti-bullying programs in reducing school bullying. Our aim is to locate and summarize all the major evaluations of programs in developed countries. Bullying has been studied in Australia, Austria, Belgium, Canada, Cyprus, Denmark, England and Wales, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Luxembourg, Japan, Malta, New Zealand, Northern Ireland, Norway, Portugal, Scotland, Spain, Sweden, Switzerland, The Netherlands, and the United States (Smith *et al.*, 1999). We aim (potentially) to include research in all these countries. We aim to measure effect sizes in each evaluation and to investigate which features (e.g. of programs, students, and schools) are related to effect sizes. We hope to make recommendations about which components of programs are most effective in which circumstances, and hence about how future anti-bullying programs

might be improved. We also hope to make recommendations about how the design and analysis of evaluations of anti-bullying programs might be improved in future.

3. Methods

3.1 *Criteria for inclusion and exclusion of studies*

We propose the following criteria for inclusion of studies in our systematic review:

- (a) The study described an evaluation of a program designed specifically to reduce school (Kindergarten to High School) bullying. Studies of aggression or violence will be excluded.
- (b) Bullying was defined as including: physical, verbal, or psychological attack or intimidation that is intended to cause fear, distress, or harm to the victim; and an imbalance of power, with the more powerful child (or children) oppressing less powerful ones. Many definitions also require repeated incidents between the same children over a prolonged period, but we do not require that, because many studies of bullying do not specifically measure or report this element of the definition.
- (c) Bullying (specifically) was measured using self-report questionnaires, peer ratings, teacher ratings, observational data, or school records.
- (d) The effectiveness of the program was measured by comparing students who received it (the experimental condition) with students who did not receive it (the control condition). We require that there must have been some pretest control of extraneous variables in the evaluation (establishing the prior equivalence of conditions) by (i) randomization, or (ii) pretest measures of bullying, or (iii) matching on pretest bullying or risk factors/scores for bullying, or (iv) pretest measurement of risk factors or risk scores for bullying. Because of low internal validity, we will exclude uncontrolled studies that only had before and after measures of bullying in experimental schools or classes. However, we propose to include studies that controlled for age. For example, in the Olweus (1991)

evaluation, all students received the anti-bullying program, but Olweus compared students of age X after the program with different students of the same age X in the same schools before the program. We will include this kind of evaluation.

- (e) Published and unpublished reports of research conducted in developed countries between 1983 and the present will be included. We believe that there was no worthwhile evaluation research on antibullying programs conducted before the pioneering research of Olweus carried out in 1983.

There is an additional criterion for inclusion in our meta-analysis (but not in our systematic review):

- (f) It was possible to measure the effect size. The main measures of effect size are the odds ratio, based on numbers of bullies/nonbullies (or victims/nonvictims), and the standardized mean difference, based on mean scores on bullying and victimization. Where the required information is not presented in reports, we will seek to obtain it by contacting the authors directly.
- (g) We have agreed not to set a minimum initial sample size for including studies. Initially, we wished to set a minimum initial sample size of 200, for the following reasons: First, larger studies are usually better-funded and of higher methodological quality (e.g. Jolliffe & Farrington, 2007). Second, we are very concerned about the frequently-found negative correlations between sample size and effect size (e.g. Farrington & Welsh, 2003). We think that this correlation probably reflects publication bias. Smaller studies that yield statistically significant results are published, whereas those that do not are left in the file drawer. In contrast, larger studies (often funded by some official agency) tend to be published irrespective of their results. Excluding smaller studies minimizes problems of publication bias and therefore yields a more accurate estimate of the true effect size. Third, we think that larger studies are likely to have higher external validity or

generalizability. Fourth, attrition (e.g. between pretest and posttest) is less problematic in larger studies. A study with 100 children that suffers 30% attrition will end up with only 35 boys and 35 girls: these are very small samples (with associated large confidence intervals) for estimating the prevalence of bullying and victimization. In contrast, a study with 300 children that suffers 30% attrition will end up with 105 boys and 105 girls: these are much more adequate samples. However, the reviewers of our protocol argued that useful information was contained in smaller studies, that the minimum sample size of 200 was arbitrary, that methodological quality and publication bias should be addressed in other ways, that attrition is equally important in smaller and larger studies, and that larger studies do not necessarily have higher external validity. Based on Vreeman and Carroll (2007), we estimate that about half of all evaluations have an initial sample size of 200 or more, and conversely half have an initial sample size less than this. Therefore, not setting a minimum sample size doubles the number of studies included in our systematic review. We plan to analyze larger and smaller studies separately to investigate whether the results obtained with the two types of studies are concordant or discordant.

3.2 Search strategies

- (a) We propose to search as many as possible of the following electronic databases using the following keywords:

Bully*/Bullies/Anti-Bullying

AND

School;

AND

Intervention*/Program*/Outcome*/Evaluation*/Effect*

We will not include Violence or Aggression as key words along with Bully/Bullies/Anti-Bullying because we know that this will identify many studies that are not

relevant to the present review. We have considered whether to experiment with key words such as Intimidation/Harrassment/Teasing/Victim* but, as mentioned above, our aim is to review studies of interventions designed specifically to reduce school bullying.

(b) *List of Databases*

Australian Criminology Database (CINCH)

Cochrane Controlled Trials Register

C2-SPECTR

Criminal Justice Abstracts

Database of Abstracts of Reviews of Effectiveness (DARE)

Dissertation Abstracts

Educational Resources Information Clearinghouse (ERIC)

EMBASE

Google Scholar

MEDLINE

National Criminal Justice Reference Service (NCJRS)

PsychInfo/Psychlit

Sociological Abstracts

Social Sciences Citation Index (SSCI)

Our experience is that electronic searches are very time-consuming and identify few previously unknown studies.

We propose to handsearch the following key journals since 1983:

(c) *Aggression and Violent Behavior*

Aggressive Behavior

Archives of Pediatrics and Adolescent Medicine

British Journal of Educational psychology

Child Development

Criminal Justice and Behavior

Developmental Psychology

Educational Psychology

International Journal on Violence and Schools

Journal of Adolescence

Journal of Educational Psychology

Journal of Interpersonal Violence

Journal of School Health

Journal of School Psychology

Journal of School Violence

Journal of Emotional Abuse

School Psychology International

School Psychology Review

Victims and Offenders

Violence and Victims

- (d) We propose to seek information from key researchers on bullying and from international colleagues in the Campbell Collaboration, including Catherine Blaya, Vincente Garrido, Peter van der Laan, Friedrich Lösel, Dan Olweus, Debra Pepler, Ken Rigby and Peter Smith. Our experience is that this method identifies the most relevant studies. Where international colleagues identify a report in a language that we cannot understand, we will ask them to provide us with a brief translation of key features that are needed for our coding schedule. We believe that, with the cooperation of colleagues in the Campbell Collaboration, we can include research in many different developed countries.
- (e) We will search reference lists of review articles and primary studies.

3.3 Description of methods used in the primary research

Very few studies randomly assign students to experimental or control groups. A few studies randomly assign schools or school classes (e.g. Cross *et al.*, 2004). Some studies have before and after measures of bullying in experimental and control schools or school classes (e.g. Alsaker & Valkenover, 2001). Some studies compare students of a particular age before the intervention with students of the same age in the same schools after the intervention (e.g. Olweus, 2005). Many studies have before and after measures of bullying in experimental schools but no control schools or school classes (e.g. Pitts & Smith, 1995). These latter studies will be excluded.

3.4 Criteria for determination of independent findings

Most studies measure bullying and victimization using student self-report questionnaires. Information is usually collected separately about whether students bully and whether students are bullied. Some studies also use teacher ratings, peer ratings, or (more rarely) playground observations or school records. We propose to calculate effect size measures of bullying and victimization, based on self-report questionnaires, in as many studies as possible. Our main meta-analysis will be based on this measure. Where self-report questionnaires are not available, we will calculate effect size measures based on other sources. Where short-term and long-term follow-up measures are available, we will calculate effect sizes separately for the shortest and longest follow-up periods.

3.5 Details of study coding categories

We propose to cover the following topics at least:

Author(s)

Dates of research

Date of publication(s)

Country and place of research

Age of students (Experimental, Control)

Gender composition (E, C)

Ethnic composition (E, C)

Initial sample size (E, C)

Final sample size (E, C)

Number of schools (E, C)

Number of classes (E, C)

Research design and methodological quality (taking account of randomization, generalizability and attrition)

Type and components of intervention (e.g. “whole-school”, curriculum; work with bullies, victims, peers, teachers, improving playground supervision, etc.)

Target of intervention (e.g. students, teachers, parents)

Duration of intervention

Follow-up time period

Outcome measures (self-report questionnaires, teacher ratings, peer ratings, systematic observation, school records)

The Appendix contains a draft coding schedule. A random sample of 20 studies will be coded by two persons in order to measure reliability.

3.6 *Statistical procedures and conventions*

The main measure of effect size will be the odds ratio (OR), calculated by comparing experimental and control conditions on bullies/nonbullies (and victims/nonvictims). Where scores are reported, the standardized mean difference d will be calculated and will be converted into LOR using the equation:

$d = 0.5513 * LOR$ (see Lipsey & Wilson, 2001, p. 202).

Where before and after information is provided, the OR and d will be adjusted for

the before data.

Effect size measures will be calculated for bullying and victimization separately. Weighted mean effect size measures will be calculated using the procedures described in Lipsey and Wilson (2001). Where appropriate, fixed effects or random effects models will be used (depending on the heterogeneity Q). Correlations between features of studies and effect sizes will be investigated, in an attempt to assess which components of programs might be the most effective in which circumstances. Meta-analytic regressions will also be carried out to investigate independent influences of program components, methodological quality, features of participants, and design features. As mentioned, we hope to make recommendations about how to make anti-bullying programs more effective and about how to improve the design and evaluation of such programs, including recommendations about analytic techniques that the researchers should use. Ideally, we could produce a “CONSORT” statement for reports of anti-bullying programs that specifies what information should always be reported.

Statistical problems arise when the unit of allocation (e.g. schools or school classes) is different from the unit of analysis (e.g. students). This problem has rarely been addressed in the bullying literature. We will address it using formulae to correct test statistics for clustering developed by Cathleen McHugh based on Hedges (2007). We will carry out a sensitivity analysis assuming intraclass correlations between .10 and .20, as used by the What Works Clearinghouse of the U.S. Department of Education (Lipsey, 2008). We will also address the possibility of publication bias by using techniques described by Rothstein, Sutton, and Borenstein (2005), such as the funnel plot.

3.7 *Treatment of qualitative research*

The review will be focussed on quantitative studies but will include qualitative

information (e.g. on the degree of implementation of interventions) where this is helpful in discussing explanations for findings or conflicting results.

4. Time Frame

Our tentative time schedule is as follows. We have already completed most of the searches.

February, 2008: Develop/test coding scheme

March-April: Code included studies, analyze results

May-June 2008: Complete first draft of review and submit it to Campbell Collaboration

Later in 2008: Receive feedback from Campbell Collaboration. Submit final draft of review for electronic publication on the Campbell website.

5. Plans for Updating the Review

The review will be updated every 3 years. The lead reviewer will take the lead in arranging this.

6. Acknowledgements

We are very grateful to David Wilson, Jeff Valentine and two anonymous reviewers for helpful comments.

7. Statement Concerning Conflict of Interest

The only possible (minor) conflict of interest is that one of the evaluations of anti-bullying programs was conducted by Baldry and Farrington (2004). Otherwise, none of us has any interest in the conclusions of the review and none of us stands to benefit in any way from any source that has any interest in the conclusions of the review. We are not aware of any personal, political, academic, or financial factors that might bias our judgment.

8. References and Some Key Studies

- Alsaker, F. D. (2004). Bernese program against victimization in kindergarten and elementary school. In P. K. Smith, D. Pepler, & K. Rigby (Eds.). *Bullying in schools: How successful can interventions be?* (pp. 289-306). Cambridge: Cambridge University Press.
- Alsaker, F. D. & Valkanover, S. (2001). Early diagnosis and prevention of victimization in kindergarten. In J. Juvonen & S. Graham (Eds.), *Peer harassment in school* (pp. 175-195). New York: Guilford.
- Baldry, A. C. & Farrington, D. P. (1999). Types of bullying among Italian school children. *Journal of Adolescence*, 22, 423-426.
- Baldry, A. C. & Farrington, D. P. (2004). Evaluation of an intervention program for the reduction of bullying and victimization in schools. *Aggressive Behavior*, 30, 1-15.
- Baldry, A. C. & Farrington, D. P. (2007). Effectiveness of programs to prevent school bullying. *Victims and Offenders*, 2, 183-204.
- Cowie, H, Smith, P. K, Boulton, M., & Laver, R. (1994). *Cooperation in the multi-ethnic classroom*. London: David Fulton.
- Cross, D., Hall, M., Hamilton, G., Pintabona Y., & Erceg, E. (2004). *Australia: The Friendly Schools project*. In P. K. Smith, D. Pepler, & K. Rigby (Eds.). *Bullying in schools: How successful can interventions be?* (pp. 187-210). Cambridge: Cambridge University Press.
- Eslea, M. & Smith, P. K. (1998). The long-term effectiveness of anti-bullying work in primary schools. *Educational Research* , 40, 203–218.
- Farrington D. P. (1993). Understanding and preventing bullying. In M. Tonry (Ed.), *Crime and Justice*, vol. 17 (pp. 381-458). Chicago: University of Chicago Press.
- Farrington, D. P. & Welsh, B. C. (2003). Family-based prevention of offending: A meta-

- analysis. *Australian and New Zealand Journal of Criminology*, 36, 127-151.
- Frey, K., Hirschstein, M. K., Snell, J. L., van Schoiack Edstrom, L., MacKenzie, E. P., & Broderick, C. J. (2005). Reducing playground bullying and supporting beliefs: An experimental trial of the Steps to Respect program. *Developmental Psychology*, 41, 479-491.
- Hanewinkel, R. (2004). Prevention of bullying in German schools: An evaluation of an anti-bullying approach. In P. K. Smith, D. Pepler, & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 81-97). Cambridge: Cambridge University Press.
- Hedges, L. V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*.
- Howard, K. A., Flora, J., & Griffin, M. (1999). Violence-prevention programs in schools: State of the science and implications for future research. *Applied and Preventive Psychology*, 8, 197-215.
- Jolliffe, D. & Farrington, D. P. (2007). *A rapid evidence assessment of the impact of mentoring on reoffending*. London: Home Office Online Report 11/07.
(homeoffice.gov.uk/rds/pdfs07/rdsolr1107.pdf)
- Limber, S. P., Nation, M., Tracy, A. J., Melton, G. B., & Flerx, V. (2004). In P. K. Smith, D. Pepler, & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 55-79). Cambridge: Cambridge University Press.
- Lipsey, M. W. (2008). Personal communication, January 29.
- Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Mytton, J., DiGiuseppi, C., Gough, D., Taylor, R., & Logan, S. (2006). School-based secondary prevention programs for preventing violence. *Cochrane Database of Systematic Reviews, Issue 3*. Art No. CD 004606.

- Olweus, D. (1991). Bully/victim problems among school children: Basic facts and effects of a school-based intervention program. In D. J. Pepler & K. H. Rubin (Eds.), *The development and treatment of childhood aggression* (pp. 411-448). Hillsdale, NJ: Erlbaum.
- Olweus, D. (1993). *Bullying at school: What we know and what we can do..* Oxford, England: Blackwell.
- Olweus, D. (2005). A useful evaluation design, and effects of the Olweus bullying prevention program. *Psychology, Crime and Law*, 11, 389-402.
- Olweus, D. & Alsaker, F. D. (1991). Assessing change in a cohort-longitudinal study with hierarchical data. In D. Magnusson, L. R. Bergman, G. Rudinger, & B. Torestad (Eds.), *Problems and methods in longitudinal research* (pp. 107-132). Cambridge: Cambridge University Press.
- O'Moore, A. M. & Minton S. J. (2004). Ireland: The Donegal primary school anti-bullying project. In P. K. Smith, D. Pepler, & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 275-288). Cambridge: Cambridge University Press.
- Ortega, R., Del Rey, R., & Mora-Merchan, J. A. (2004). SAVE Model: An antibullying intervention in Spain. In P. K. Smith, D. Pepler, & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?*(pp. 167-186) Cambridge: Cambridge University Press.
- Ortega, R. & Lera, M. J. (2000). The Seville anti-bullying in school project. *Aggressive Behavior*, 26, 113–123.
- Pateraki, L. & Houndoumadi, A. (2001). Bullying among primary school children in Athens, Greece. *Educational Psychology*, 21, 167-175.
- Pepler, D. J., Craig, W. M., O'Connell, P., Atlas, R., & Charach, A. (2004). Making a difference in bullying: Evaluation of a systemic school-based program in Canada.

- In P. K. Smith, D. Pepler, & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 125-140). Cambridge: Cambridge University Press.
- Pepler, D. J., Craig, W. M., Ziegler, S., & Charach, A. (1994). An evaluation of an antibullying intervention in Toronto schools. *Canadian Journal of Community Mental Health*, 13, 95–110.
- Peterson, L. & Rigby, K. (1999). Countering bullying at an Australian secondary school. *Journal of Adolescence*, 22, 481–492.
- Pitts, J. & Smith, P. (1995). *Preventing school bullying*. London: Home Office (Crime Detection and Prevention Series Paper 63).
- Rigby, K. (2002). *A meta-evaluation of methods and approaches to reducing bullying in preschools and early primary school in Australia*. Canberra: Attorney General's Department, Crime Prevention Branch.
- Rock, E. A., Hammond, M., & Rasmussen, S. (2004). School-wide bullying prevention program for elementary students. *Journal of Emotional Abuse*, 4, 225-239.
- Roland, E. (1989). Bullying: The Scandinavian tradition. In D. P. Tattum & D. A. Lane (Eds.), *Bullying in schools* (pp. 21–32). Stoke-on-Trent, England: Trentham Books.
- Roland, E. (1993). Bullying: A developing tradition of research and management. In D. P. Tattum (Ed.), *Understanding and managing bullying* (pp.15–30). Oxford, England: Heinemann Educational.
- Rothstein, H., Sutton, A. J., & Borenstein, M. (Eds.) (2005). *Publication bias in meta-analysis: Prevention, assessment, and adjustments*. Chichester, England: Wiley.
- Ruiz, O. (2005). Prevention programs for pupils. In A. Serrano (Ed.), *School Violence*. Valencia, Spain: Queen Sofia Center for the Study of Violence.
- Salmivalli, C. (2001). Combating school bullying: A theoretical framework and preliminary results of an intervention study. Paper presented at International Conference on Violence in Schools and Public Policies, Paris (April).

- Salmivalli, C., Lagerspetz, K., Bjorkqvist, K., Österman, 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.
- Smith, J. D., Schneider, B., Smith, P. K., & Ananiadou, K. (2004). The effectiveness of whole-school anti-bullying programs: A synthesis of evaluation research. *School Psychology Review*, 33, 548-561.
- Smith, P. K., Ananiadou, K., & Cowie, H. (2003). Interventions to reduce school bullying. *Canadian Journal of Psychology*, 48, 591-599.
- Smith, P. K., Cowie, H., Olafsson, R. F. & Liefoghe, A. P. D. (2002). Definitions of bullying: A comparison of terms used, and age and gender differences, in a 14-country international comparison. *Child Development*, 73, 1119-1133.
- Smith, P. K., Morita, J., Junger-Tas, D., Olweus, D., Catalano, R., & Slee, P. T. (Eds.) (1999). *The nature of school bullying: A cross-national perspective*. London:Routledge.
- Smith, P. K., Pepler., D., & Rigby, K. (Eds.) (2004). *Bullying in schools: How successful can interventions be?* Cambridge: Cambridge University Press.
- Smith, P. K. & Sharp, S. (Eds.). (1994). *School bullying: Insights and perspectives*. London: Routledge.
- Stevens, V., de Bourdeaudhuij, I., & van Oost, P. (2000). Bullying in Flemish schools: An evaluation of anti-bullying intervention in primary and secondary schools. *British Journal of Educational Psychology*, 70, 195–210.
- Vreeman, R.C. & Carroll, A. E. (2007). A systematic review of school-based interventions to prevent bullying. *Archives of Pediatrics and Adolescent Medicine*, 161, 78-88.
- Whitaker, D. J. Rosenbluth, B., Valle, L. A., & Sanchez, E. (2004). Expect Respect: A school-based intervention to promote awareness and effective responses to bullying and sexual harassment. In D. L. Espelage & S. M. Swearer (Eds.),

Bullying in American schools: A social-ecological perspective on prevention and intervention (pp. 327-350). Mahwah, NJ: Lawrence Erlbaum.

Whitney I. & Smith, P. K. (1993). A survey of the nature and extent of bullying in junior/middle and secondary schools. *Educational Research, 35*, 3-25.

Wilson, S. J., Lipsey, M. W., & Derzon, J. H. (2003). The effects of school-based intervention programs on aggressive behavior: A meta-analysis. *Journal of Consulting and Clinical Psychology, 71*, 136-149.

Wilson, S. J. & Lipsey, M. W. (2007). School-based interventions for aggressive and disruptive behavior: Update of a meta-analysis. *American Journal of Preventive Medicine, 33*, S130-S143.

9. Table 1

Key Bullying Prevention Projects (from Baldry & Farrington, 2007)

Project	Components of the program	Participants	Assessment methods	Research Design
Bergen, Norway (Olweus, 1991, 1993)	Individual, classroom and school components. Each school had to set up anti-bullying rules	2,500 students aged 10-15 in 42 primary and secondary schools.	Student self-report questionnaires, teacher ratings	Age cohort design with two post-tests after 8 and 20 months
Rogaland, Norway (Roland, 1989, 1993)	Similar to the Bergen evaluation but less detailed	7,000 students aged 8-16 in 37 primary and secondary schools	Student self-report questionnaires, teacher ratings	Pre-test, post-test design. Follow-up after 3 years. No control schools
Sheffield, England (Smith & Sharp, 1994; Eslea & Smith, 1998)	Whole school approach with several components.	6,500 students aged 8-16 from 16 primary and 7 secondary schools (intervention). 4 control schools	Student self-report questionnaires, interviews with teachers	Pre-test and post -test 18 months later. 3 year follow-up in 4 intervention schools
Liverpool and London, England (Pitts & Smith, 1995)	Anti-bullying policy, focusing on peer support and training in assertiveness skills	1,284 students aged 8-16 in 4 primary and secondary schools.	Student self-report questionnaires	Pre-test and post-test (follow-up after 2 years). No control schools.
Toronto, Canada (Pepler <i>et al.</i> , 1994)	Olweus components based on the individual, the classroom, parents, and the school	First study: 898 students aged 8-14 from 4 elementary schools. Second study: 922 students from 3 schools.	Student self-report questionnaires, playground observations	Pre-test, post test (after 18 and 30 months), with no control schools.
New South Wales, Australia (Peterson & Rigby, 1999)	Peer support and method of shared concern by staff.	758 students aged 12-17 in 1995, and 657 in 1997	Student self-report questionnaires	Pre-test, post-test design with no control group
Donegal, Ireland (O'Moore & Minton, 2004)	Training teachers, Information pack, work with students	527 students from 42 schools. For only 22 it was possible to match pre and post data	Student self-report questionnaires	Pre-test, post-test design with no control group
Turku and Helsinki, Finland (Salmivalli, 2001)	Individual, classroom, school and teacher interventions	1,220 students aged 9-12 in 16 schools	Student self-report questionnaires, peer nominations	Baseline, low, high implementation, age cohort design; follow-up after 6 months.

Flanders, Belgium (Stevens <i>et al.</i> , 2000)	Individual, classroom and school interventions	1,104 students aged 10-16 from 18 schools.	Student self-report questionnaires	Full, part implementation or control. Pre-test, 6 months later, and one year later.
Expect Respect, Texas, US (Whitaker <i>et al.</i> , 2004)	Classroom education, staff training, policy/procedure development, parent education, support services	929 experimental and 834 control students aged 11-12 in 12 schools	Student self-report questionnaires	Pre-test, post-test, intervention and control groups.
Berne, Switzerland (Alsaker & Valkanover, 2001)	Whole school approach to bullying focussed on teachers	152 experimental children aged 5-7 in 8 classes. 167 control children in 8 classes.	Teacher ratings, peer nominations	Pre-test, post-test, control group design
"Bulli & Pupe", Rome, Italy (Baldry & Farrington, 2004)	Work with the class, with booklets and videos on bullying and family violence	239 students aged 10-16 in 13 schools	Student self-report questionnaires	Intervention and control groups, random assignment, pre-test and post-test measures
SAVE Seville, Spain (Ortega & Lera, 2000)	Community approach focussing on democratic values, cooperative group work, empathy.	910 students aged 8-18 in primary and secondary schools.	Student self-report questionnaires	5 intervention schools had pre-test and post-test measures, compared to 3 control schools with only post-test measures. Follow-up after 4 years
Friendly Schools, Perth (Cross <i>et al.</i> , 2004)	Friendly schools approach addressing individuals, classes and schools, and teacher training.	2,068 students aged 9-10 from 29 schools	Student self-report questionnaires	Pre-test and post-test data from intervention and control schools, randomly assigned
Steps to Respect, Pacific Northwest, US (Frey <i>et al.</i> , 2005)	Staff training, promotion of prosocial beliefs	1,126 students aged 8-12 in 6 schools	Teacher and student ratings, playground observation	Pre-test, post-test, experimental and control groups, randomly assigned.
Norway (Olweus, 2005)	As for the Bergen evaluation	20,709 students aged 9-12	Student self-report questionnaires	Pre-test and post-test measures

Appendix
Coding Protocol

Use one coding sheet for each distinct research project that is included in the systematic review.

Throughout, zero = not known, 1 = Yes, 2 = No

Name of coder Coder _____

Date of coding Date _____

Study

Study no. identifier Studyid _____

Author(s) name Author _____

Author(s) affiliation Affil _____

Date(s) research was conducted Dateres _____

Date of primary publication Datepub _____

Place of research Place _____

Country of research Country _____

Source of funding

Details of primary report

Publication type:

- 1 book
- 2 book chapter
- 3 Journal article
- 4 technical report
- 5 unpublished conference paper
- 6 dissertation

Pub type _____

Details of other reports

2. Eligibility criteria

Are the following inclusion criteria present? (if not clear, attempt to find out from the author)

The report deals with bullying defined as physical, verbal, or psychological attack or intimidation that is intended to cause fear, distress, or harm to the victim; an imbalance of power, with the more powerful child (or children) oppressing less powerful ones.

Defn _____

Types of bullying measured _____

The report analyzes repeated incidents between the same children over a prolonged period.

Repeat _____

The report describes an evaluation of a program designed to reduce school bullying.

Bully _____

The report indicates which methods were adopted to gather information on bullying (self-report questionnaires, peer ratings, teacher ratings, observational data, or school records).

Meths _____

Before data on bullying and victimization as defined above (specifically) are reported.

Before _____

After data on bullying and victimization as defined above (specifically) are reported

After _____

The effectiveness of the program was measured by comparing students who received it (the experimental group) with students who did not receive it (the control group).

Control _____

The Study had some control of extraneous variables (establishing the prior equivalence of groups) by (i) randomization, or (ii) pretest measures of bullying, or (iii) matching, or (iv) pretest measurement of risk factors or risk scores for bullying.

Extran _____

Numbers of bullies/nonbullies and victims/nonvictims are reported.

Numbers_____

Scores on bullying and victimization are reported.

Scores_____

The research was conducted in one of the following countries since 1983:

Australia	01	Austria	02	Belgium	03
Canada	04	Cyprus	05	Denmark	06
England and Wales	07	Finland	08	France	09
Germany	10	Greece	11	Iceland	12
Ireland	13	Israel	14	Italy	15
Japan	16	Luxembourg	17	Malta	18
New Zealand	19	Northern Ireland	20	Norway	21
Portugal	22	Scotland	23	Spain	24
Sweden	25	Switzerland	26	The Netherlands	27
United States	28	Other	29		

Count1_____

3. Sample

Mean age of students:

Experimental

Eage_____

Control

Cage_____

Gender composition:

Experimental

% female _____

Epcf_____

Control

% female _____

Cpcf _____

Ethnic composition:

Experimental

Control

Initial sample size:

Experimental

En1_____

Control

Cn1_____

Sample size in short follow-up:

Experimental

En2_____

Control

Cn2_____

Sample size in long follow-up:

Experimental

En3_____

Control

Cn3_____

Number of schools:

Experimental

Eschools _____

Control

Cschools_____

Scales:

a) What scale was used to measure bullying? Bulscale_____

b) What scale was used to measure victimization? Vicscale_____

What reference period was used for bullying? (in days) Bulref_____

What reference period was used for victimization? (in days) Vicref_____

Was there any information about reliability or validity of measures? Rel_____

Val_____

6. Intervention (experimental condition)

Type and components of intervention: (Yes = 1, No = 2)

Whole-school policy and procedures	Intwhole	
Classroom rules	Intrules	
School conferences/providing information about bullying	Intconf	
Curriculum materials	Intcurr	
Classroom management	Intclass	
Cooperative group work	Intcoop	
Work with bullies (e.g. empathy, skills training)	Intbull	
Work with victims (e.g. assertiveness training)	Intvic	
Work with peers (e.g. peer support, mentoring)	Intpeers	
Work with teachers	Intteach	
Work with parents	Intparent	
Improving playground supervision	Intsup	
Disciplinary methods	Intdisc	
Non-punitive methods (e.g. Pikas)	Intnonpun	
Restorative justice approaches	Intrj	
Bully courts, school tribunals	Intcourt	
Other	Int other	

Was the intervention highly structured, that is did it follow a protocol or manual?
 (Yes = 1, No = 2) Struct_____

Mode of program delivery: (Yes = 1, No = 2)

Video _____
 Role play _____
 Booklets _____
 Other _____

Duration of intervention (No. of days): Durint _____

What time of the year (month) did the program start? Start _____

What time of the year (month) did the program end? End _____

Who delivered the intervention? (Yes = 1, No = 2)

external researchers Delres _____

teachers Delteach _____

psychologists or professionals within the school system Delpsych _____

other (specify _____) Deloth _____

Were there any problems of implementation?

(specify) _____

Was there a measure of treatment integrity? Integ _____

What happened to the control group?

1. No intervention
 2. Wait-list control
 3. Placebo control
 4. Given some information on bullying
 5. Management as usual
 6. Other (_____)
- Contgp _____

Were any materials delivered to control participants (video, booklets)?

(Yes = 1, No = 2) Contdel _____

7. Post-test measures

Follow-up time period (No. of days): Short FU _____

Long FU _____

What measures were used? (Yes = 1, No = 2)

	Short FU	Long FU
Self-report questionnaires	SSRQ	LSRQ
Teacher ratings	STR	LTR
Peer ratings	SPR	LPR
Systematic observation	SSO	LSO
School records	SSR	LSR
Other	SOTH	LOTH

Scales:

- a) What scale was used to measure bullying in the short follow-up? Ssbul_____
- b) What scale was used to measure bullying in the long follow-up? Lsbul_____
- c) What scale was used to measure victimization in the short follow-up? Ssvic_____
- d) What scale was used to measure victimization in the long follow-up? Lsvic_____

What reference period was used for bullying? Short FU: Srbul_____

Long FU: Lrbul_____

What reference period was used for victimization? Short FU: Srvic_____

Long FU: Lrvic_____

8. Effect Size Measures

Prevalence of Bullying

	Experimental	Control
No. of bullies before	EBBef	CBBef
No. of nonbullies before	ENBBef	CNBBef
Short: No. of bullies after	EBS	CBS
Short: No. of nonbullies after	ENBS	CNBS
Long: No. of bullies after	EBL	CBL
Long: No. of nonbullies after	ENBL	CNBL

Short: Odds Ratio ORBS_____ Confidence Interval CIBS_____

Long: Odds Ratio ORBL_____ Confidence Interval CIBL_____

Prevalence of Victimization

	Experimental	Control
No. of victims before	EVbef	CVbef
No. of nonvictims before	ENVbef	CNVbef
Short: No. of victims after	EVS	CVS
Short: No. of nonvictims after	ENVS	CNVS
Long: No. of victims after	EVL	CVS
Long: No. of nonvictims after	ENVL	CNVL

Short: Odds Ratio ORVS _____ Confidence Interval CIVS _____

Long: Odds Ratio ORVL _____ Confidence Interval CIVL _____

Bullying scores

	Experimental	Control
Before		
Mean	EMBBef	CMBBef
SD	ESDBBef	CSDBBef
N	ENOBef	CNOBef
After		
Short: Mean	EMBS	CMBS
SD	ESDBS	CSDBS
n	ENOBS	CNOBS
Long: Mean	EMBL	CMBL
SD	ESDBL	CSDBL
n	ENDOBL	CNOBL

Short: dBS _____ SEBS _____

Long: dBL _____ SEBL _____

Victimization Scores

	Experimental	Control
Before		
Mean	EMVBef	CMVBef
SD	ESDVBef	CSDVBef
N	ENOVbef	CNOVBef
After		
Short: Mean	EMVS	CMVS
SD	ESDVS	CSDVS
n	ENOVs	CNOVS
Long: Mean	EMVL	CMVL
SD	ESDVL	CSDVL
n	ENOVL	CNOVL

Short: dVS _____ SEVS _____

Long: dVL _____ SEVL _____