Report prepared for Brå by Maria M. Ttofi, David P. Farrington and Anna C. Baldry

Effectiveness of Programmes to Reduce School Bullying

Effectiveness of Programmes to Reduce School Bullying

A Systematic Review

Brå – a centre of knowledge on crime and measures to combat crime

The Swedish National Council for Crime Prevention (Brottsförebyggande rådet – Brå) works to reduce crime and improve levels of safety in society by producing data and disseminating knowledge on crime and crime prevention work and the justice system's responses to crime.

Contents

Foreword

Bullying has probably been a problem among children throughout the history of mankind. The past few decades have witnessed a steady increase in the number of programmes to combat bullying that have been developed and tested in school settings. But how well do they work? What does the research tell us?

There are never sufficient resources to conduct rigorous scientific evaluations of all the crime prevention measures employed in individual countries like Sweden. Nor has a high quality evaluation been conducted in Sweden of any program to combat bullying. For these reasons, the Swedish National Council for Crime Prevention (Brå) has commissioned renowned researchers to carry out an international review of the research published in this field.

This report presents a systematic review, including a statistical meta-analysis, of the effects of anti-bullying programmes on bullies and their victims. The work has been carried out by PhD Candidate Maria M. Ttofi and Professor David P. Farrington of Cambridge University (United Kingdom) and Associate Professor Anna C. Baldry of the Second University of Naples (Italy), who have also written the report. The study follows the rigorous methodological requirements of a systematic review. The analysis combines the results from a number of evaluations that are considered to satisfy a list of empirical criteria for measuring effects as reliably as possible. The meta-analysis then uses the results from these previous evaluations to calculate and produce an overview of the effects that anti-bullying programmes do and do not produce. Thus the objective is to systematically evaluate the results from a number of studies in order to produce a more reliable picture of the opportunities and limitations associated with anti-bullying programmes in relation to crime prevention efforts.

The systematic review, and the statistical meta-analysis, in this case build upon a large number of scientific evaluations from different part of the world, producing highly relevant findings on the effects of these programmes. Even though important questions remain unanswered, the study provides the most accessible and far-reaching overview to date of anti-bullying programmes and their effects on bullies and victims.

Stockholm, October 2008

Jan Andersson
Director-General

Executive Summary

This report presents a systematic review and meta-analysis of the effectiveness of anti-bullying programmes. Systematic reviews use rigorous methods for locating, appraising and synthesizing evidence from prior evaluation studies in order to minimize bias in drawing conclusions. They have explicit objectives, explicit criteria for including or excluding studies, extensive searches for eligible evaluation studies from all over the world, careful extraction and coding of key features of studies, and include a detailed report of the methods and conclusions of the revIew. Meta-analyses summarize effect sizes of interventions and investigate factors that correlate with effect size.

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 a more powerful child (or children) oppressing less powerful ones; and repeated incidents between the same children over a prolonged period. 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.

Studies were included in this review if they evaluated the effects of an anti-bullying programme by comparing an experimental group who received the intervention with a control group who did not. Four types of research design were included: a) randomized experiments, b) experimental-control comparisons with before and after measures of bullying, c) other experimental-control comparisons and d) age-cohort designs, where students of age X after the intervention were compared with students of the same age X in the same school before the intervention. Also, studies were included if bullying was measured using a self-report questionnaire and if the initial sample size (the total number of children in experimental and control conditions) was at least 200.

Extensive searches were carried out to find reports on anti-bullying programmes in schools. A total of 593 reports were found, but only 59 of these (describing evaluations of 30 different programmes) were eligible for inclusion in our review. The number of reports on anti-bullying programmes and on the necessity of tackling bullying increased considerably over time.

Our meta-analysis showed that, overall, school-based anti-bullying programmes are effective in reducing bullying and victimization. The results indicated that bullying and victimization were reduced by about 17–23% in experimental schools compared with control schools.

The most important programme elements that were associated with a decrease in bullying were parent training, improved playground supervision, disciplinary methods, school conferences, information for parents, classroom rules, classroom management, and videos. In addition, the

total number of elements, and the duration and intensity of the programme for children and teachers, were significantly associated with a decrease in bullying. Also, programmes inspired by the work of Dan Olweus worked best. Regarding the design features, the programmes worked better with older children, in smaller-scale studies, in Norway specifically, and in Europe more generally. Older programmes, and those in which the outcome measure of bullying was two times per month or more, also yielded better results.

The most important programme elements that were associated with a decrease in victimization (i.e. being bullied) were videos, disciplinary methods, work with peers, parent training, cooperative group work and playground supervision. In addition, the duration of the programme for children and teachers, and the intensity of the programme for teachers, were significantly associated with a decrease in victimization. Regarding the design features, the programmes worked better with older children, in Norway specifically and in Europe more generally, and they were less effective in the USA. Older programmes, those in which the outcome measure was two times per month or more, and those with other experimental-control and age-cohort designs, also yielded better results.

The main policy implication of our review is that new anti-bullying programmes should be designed and tested based on our results. These could be grounded in the successful Olweus programme but should be modified in light of the key programme elements that we have found to be most effective.

In conclusion, results obtained so far in evaluations of anti-bullying programmes are encouraging. The time is ripe to mount a new long-term research strategy on the effectiveness of these programmes, based on our findings.

1. Introduction

1.1 Impetus for the Systematic Review

Given the serious short-term and long-term effects of bullying on children's physical and mental health (Ttofi & Farrington, 2008) it is understandable why school bullying has increasingly become a topic of both public concern and research efforts. Research on school bullying has expanded worldwide (Smith, Morita, Junger-Tas, Olweus, Catalano & Slee, 1999), with a variety of intervention programmes being implemented (Smith, Pepler, & Rigby, 2004a), and with some countries legally requiring schools to have an anti-bullying policy (Ananiadou & Smith, 2002). The cost of victimization in schools is considerable (Hawker & Boulton, 2000) and intervention strategies aiming at tackling school bullying and promoting safer school communities can be seen as a moral imperative (Smith, Ananiadou, & Cowie, 2003).

Despite the marked increase in anti-bullying research, there is still much that needs to be learned about how to design and implement effective intervention programmes, especially taking into account the varying results of intervention research across studies in different countries (Pepler, Smith, & Rigby, 2004; Smith & Ananiadou, 2003). In what ways, and why, is one anti-bullying programme more effective than another? What intervention elements can predict the success of a programme in reducing school bullying? What intervention elements, possibly against common sense and stereotypic expectations, may be harmful and should be avoided by policy makers and practitioners? These questions have inspired our research.

A systematic review aims to comprehensively locate and synthesize research that bears on a particular question, using organized, transparent, and replicable procedures at each step in the process (Littell, Concoran, & Pillai, 2008). It includes explicit criteria for inclusion or exclusion of studies in a highly structured way that aims to minimize bias in the conclusions. It allows 'decisions to be made on a transparent and potentially defendable basis, as it draws on all relevant scientifically sound research, rather than on single studies' (Petticrew & Roberts, 2005, p. 11). Reviews are essential tools for health care workers, researchers, consumers and policy makers who want to keep up with the evidence that is accumulating in their fields. Systematic reviews allow for a more objective appraisal of the evidence than traditional narrative reviews and may thus contribute to resolve uncertainty when original research, reviews, and editorials disagree (Egger, Smith, & O'Rourke, 2001, p. 23). Given their great promise to inform policy and practice, the marked increase in systematic reviews in both health and social sciences should come as no surprise. Our systematic review follows 25 years of intervention research (from 1983 to 2008) and is based on extensive literature searches. Our meta-analytic approach offers a quanti-

tative summary of effect sizes of anti-bullying programmes and standardizes the evaluation results across studies with the aim of making solid inferences about what works in preventing bullying, for whom and under what circumstances.

1.2 Definition of Bullying

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 a 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 a type of aggressive behaviour (Andershed, Kerr, & Stattin, 2001; Cowie, 2000; Leary, Kowalski, Smith, & Philips, 2003; Roland & Idsoe, 2001; Salmivalli & Nieminen, 2002). However, it should not be equated with aggression or violence; not all aggression or violence involves bullying, and not all bullying involves aggression or violence. For example, bullying includes being called nasty names, being rejected, ostracized or excluded from activities, having rumours spread about you, having belongings taken away, teasing and threatening (Baldry & Farrington, 1999). Our aim is to review programmes that are specifically intended to prevent or reduce school bullying, not programmes 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, has been reviewed by Smith and his colleagues (1999). Special methods are needed to study bullying in different countries because of the problem of capturing the term "bullying" in different languages. Smith, Cowie, Olafsson and Liefooghe (2002) have reviewed the meaning of bullying in 14 different countries in an attempt to examine how the use of global terms (such as 'bullying') can affect the prevalence of admitting bullying. Smith and his colleagues (2002, p. 1121) also give a nice example of how even similar terms within the same language (e.g. bullying, teasing, harassment, abuse) have different connotations and contexts and may be understood differently by persons answering questionnaires. An alternative to using global terms such as bullying in surveys is to ask for information about particular acts, such as "hit him/her on the face" or "excluded him/her from games" (Smith et al., 2002, p. 1131), and this is what researchers often do (Kalliotis, 2000, p. 49; Pateraki & Houndoumadi, 2001, p. 174).

1.3 Background

Many school-based intervention programmes 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 programmes 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 programme was implemented nationally in Norway in 1983. A more intensive version of the national programme was evaluated in Bergen by Olweus (1991). The evaluation by Olweus (1991) showed a dramatic decrease in victimization (being bullied) of about half after the programme. Since then at least 15 other large-scale anti-bullying programmes, 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 sixteen major evaluations in eleven different countries and concluded that eight of them produced desirable results, two produced mixed results, four produced small or negligible effects, and two produced undesirable results. Most programmes were rather complex, and the effectiveness of the different components of programmes was not clear.

American research is generally targeted on school violence or peer victimization rather than bullying. There are a number of existing reviews of school violence programmes and school-based interventions for aggressive behaviour (e.g. Howard, Flora, & Griffin, 1999; Mytton, DiGuiseppi, Gough, Taylor, & Logan, 2006; Wilson, Lipsey & Derzon, 2003; Wilson & Lipsey, 2007). We have consulted these, but we must emphasize that our research aims to review programmes that are explicitly designed to reduce bullying and that explicitly measure bullying.

The most informative single source of reports of anti-bullying programmes is the book edited by P.K. Smith and his colleagues (2004a), which contains descriptions of 13 programmes implemented in 11 different countries. There are also some reviews containing summaries of major anti-bullying programs (e.g. Rigby, 2002; Ruiz, 2005; Smith, Ananiadou, & Cowie, 2003). The most relevant existing reviews are by J. Smith, Schneider, Smith and Ananiadou (2004), who summarized effect sizes in 14 whole-school anti-bullying programmes, and by Vreeman and Carroll (2007), who reviewed 26 school-based programmes. 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.

J.D. Smith et al. (2004) reviewed 14 evaluations up to 2002, 6 of which were uncontrolled. Vreeman and Carroll (2007) reviewed 26 evaluations up to 2004, restricted to studies published in the English language. The latest meta-analytic review was published by Ferguson,

San Miguel, Kilburn and Sanchez (2007). However, this included searches in one database only (p. 406), for articles published between the years 1995 and 2006 (p. 406). It included outcome variables that measured 'some element of bullying behaviour or aggression toward peers, including direct aggressive behaviour toward children in a school setting' (p. 407).

In the present report, we go way beyond these previous reviews by
- doing much more extensive searches for evaluations such as hand-searching all volumes of 35 journals from 1983 up to the end of April 2008.
- searching for international evaluations in 18 electronic databases and in languages other than English.
- carrying out much more extensive meta-analyses (including correlating effect sizes with study features and research design).
- focusing only on programmes that are specifically designed to reduce bullying and not aggressive behaviour (i.e. the outcome variables specifically measure bullying).

1.4 Objectives of the Review

Our main objective is to assess the effectiveness of school-based anti-bullying programmes in reducing school bullying. Our aim is to locate and summarize all the major evaluations of programmes in developed countries. Bullying has been studied in (at least) 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 evaluations in all these countries. We aim to measure effect sizes in each evaluation and to investigate which features (e.g. of programmes and students) are related to effect sizes. We hope to make recommendations about which components of programmes are most effective in which circumstances, and hence about how future anti-bullying programmes might be improved. We also aim to describe the major programmes in detail. We also hope to make recommendations about how the design and analysis of evaluations of anti-bullying programmes might be improved in future. However, we are of course limited by the information that is available in published reports.

2. Methods

2.1 Measuring the Effects of a Programme

How can the effects of an anti-bullying programme on bullying and victimization (being bullied) be established? The highest quality studies are those that maximize statistical conclusion validity, internal validity, construct validity, external validity, and descriptive validity (Farrington, 2003).

Statistical conclusion validity is concerned with the effect size (and its associated confidence interval) measuring the effect of the intervention on bullying. Internal validity is concerned with whether it really was the intervention that had an effect on bullying. Construct validity refers to whether the intervention really was an anti-bullying programme and whether the outcome really was a measure of bullying. External validity refers to the generalizability of the results, and can be best established in a systematic review. Descriptive validity refers to the adequacy of the presentation of key features of the evaluation in a research report.

Internal validity is the most important. The main threats to internal validity are well known (Campbell & Stanley, 1966; Cook & Campbell, 1979; Shadish, Cook & Campbell, 2002):

1. *Selection*: The effect reflects pre-existing differences between experimental and control conditions.
2. *Aging/Maturation*: The effect reflects a continuation of pre-existing trends, e.g. in normal human development.
3. *History*: The effect is caused by some event occurring during the same time period as the intervention.
4. *Testing*: The pretest measurement causes a change in the posttest measure.
5. *Instrumentation*: The effect is caused by a change in the method of measuring the outcome.
6. *Regression to the Mean*: Where an intervention is implemented on units with unusually high scores (e.g. classes with high bullying rates), natural fluctuation will cause a decrease in these scores on the posttest which may be mistakenly interpreted as an effect of the intervention.
7. *Differential Attrition*: The effect is caused by differential loss of children from experimental compared to control conditions.
8. *Causal Order*: It is unclear whether the intervention preceded the outcome.

In addition, there may be interactive effects of threats. For example, a selection-aging effect may occur if the experimental and control conditions have different pre-existing trends that continue, or a selection-histo-

ry effect may occur if the experimental and control conditions experience different historical events (e.g. where they are located in different settings). Also it is important to eliminate the problem of seasonal variations in bullying by measuring it at the same time of the year before and after an intervention.

In maximizing internal validity, it is essential to compare the intervention condition with some kind of control condition, in order to estimate what would have happened in the absence of the intervention. If children are merely measured before and after receiving the intervention, it is impossible to disentangle the impact of the programme from aging, history, testing, regression and attrition effects. In particular, bullying decreases steadily with age from 7 to 15 (Olweus, 1991). Therefore, if experimental children are tested before and one year after the intervention, their bullying will have decreased because of aging effects alone.

According to Cook and Campbell (1979), the minimum design that is interpretable requires experimental and control conditions. The best way of eliminating selection, aging, history, testing and regression effects is to randomly assign children to experimental and control conditions. Providing that a sufficiently large number of children are randomly assigned (e.g. at least 100), those in the experimental condition will be similar to those in the control condition (before the intervention) on all measured and unmeasured variables that might influence bullying.

In research on anti-bullying programmes, schools or school classes, not children, are randomly assigned to receive the programme. No experimental study of bullying has yet randomly assigned enough classes (e.g. at least 100) to achieve the benefits of randomization in eliminating threats to internal validity. Therefore, it is not clear that randomized experiments on anti-bullying programmes are methodologically superior to quasi-experimental evaluations with before and after measures of bullying in experimental and control conditions. It is clear that these two designs are potentially the best methodologically. The main threat to internal validity in them is differential attrition from experimental and control conditions. In addition, if the experimental classes are worse than the control classes to start with, regression to the mean could be a problem.

Comparisons of experimental and control classes with no prior measures of bullying are clearly inferior to comparisons with prior measures. Where there are no prior measures of bullying, it is important to include some pretest measures that might establish the comparability of experimental and control children. Otherwise, this design is vulnerable to selection and regression effects in particular.

The age-cohort design, in which children of a certain age X in year 1 before the intervention are compared with (different) children of the same age X in the same school in year 2, was pioneered by Olweus (1991). It largely eliminates problems of selection, aging, regression and differential attrition, but it is vulnerable to history and testing effects.

13

Overall, the experimental-control comparisons and age-cohort designs might be regarded by some researchers as methodologically inferior to the randomized experiments and before-after, experimental control comparisons, but all designs have advantages and problems. These are the best four designs that have been used to evaluate the effects of anti-bullying programmes, and we will give credence to all of them in providing useful information about the effectiveness of anti-bullying programmes.

2.2 Criteria for Inclusion or Exclusion of Studies

In line with our coding book, we use the following criteria for inclusion of studies in our systematic review:

(a) The study described an evaluation of a programme designed specifically to reduce school (kindergarten to high school) bullying. Studies of aggression or violence are excluded. For example, the study by Woods, Coyle, Hoglund and Leadbeater (2007) was excluded because the researchers did not specify that they were studying bullying specifically. Some other reports were also excluded from the present review because their focus was the impact of a specific anti-bullying programme on some other outcome measures such as educational attainment (e.g. Fonagy, Twemlow, Vernberg, Sacco, & Little, 2005), knowledge about and attitudes towards bullying (e.g. Meraviglia, Becker, Rosenbluth, Sanchez, & Robertson, 2003) or children's safety awareness with regard to different types of potentially unsafe situations, including being bullied (e.g. Warden, Moran, Gillies, Mayes, & Macleod, 1997).

(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. We set this restriction for the current review so that all included evaluations are comparable. A meta-analysis requires comparable effect size data. Most evaluations use self-report questionnaires but some employ other measures such as peer ratings, teacher ratings, observational data or school records.

(d) The effectiveness of the programme 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 control of extraneous variables in the evaluation (establishing the equivalence of conditions) by (i) randomization, or (ii) pretest measures of bullying, or (iii) choosing some kind of comparable control condition. Because of low internal validity, we exclude uncontrolled studies that only had before and after measures of bullying in experimental schools or classes. However, we include studies that controlled for age. For example, in the Olweus (1991) evaluation, all students received the anti-bullying programme, but Olweus compared students of age X after the programme with different students of the same age X in the same schools before the programme. We include this kind of age-cohort design because arguably the experimental and control students are comparable. We will compare results obtained in the four types of included research designs, namely randomized experiments, before-after experimental-control comparisons, other experimental-control comparisons, and age-cohort designs.

(e) Published and unpublished reports of research conducted in developed countries between 1983 and the present are included. We believe that there was no worthwhile evaluation research on anti-bullying programmes conducted before the pioneering research of Olweus, which was carried out in 1983.

(f) It was possible to measure the effect size. The main measures of effect size are the odds ratio, based on numbers of bullies/non-bullies (or victims/non-victims), and the standardized mean difference, based on mean scores on bullying and victimization (being bullied). These measures are mathematically related (see later). Where the required information is not presented in reports, we have tried to obtain it by contacting the authors directly. Some studies that included a randomized or non-randomized experimental design (e.g. Heydenberk, Heydenberk, & Tzenova, 2006; Twemlow, Fonagy, Sacco, 2005; Wiefferink, Hoekstra, Beek, & Van Dorst, 2006) were not included in this report because they did not provide enough data to allow us to calculate effect size. Some other controlled studies are included (e.g. Salmivalli, Karna, & Poskiparta, 2007)[1] even though their final results have not yet been published. In this case, we use the available evaluation data with the caveat that the final evaluation results are liable to change.

(g) The minimum initial sample size (total in experimental and control conditions) was 200. We set this minimum for the following reasons: First, larger studies are usually better-funded and of higher methodological quality. Second, we are very concerned about the frequently-found negative correlations between sample size and effect size (e.g. Farrington & Welsh, 2003; Jolliffe & Farrington, 2007). We think that these correla-

---

[1] Personal communication with Dr Salmivalli via e-mail (June 18, 2008).

tions might reflect publication bias. Smaller studies that yield statistically significant results may be published, whereas those that do not may be left in the file drawer. In contrast, larger studies (often funded by some official agency) are likely to be published irrespective of their results. Excluding smaller studies reduces 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.

2.3 Searching Strategies

(a) We started by searching for the names of established researchers in the area of bullying prevention (e.g. Australia, K. Rigby; England, P.K. Smith; Spain, R. Ortega; Norway, D. Olweus). This searching strategy was used in different databases in order to initially obtain as many evaluations of known research programmes in different journals as possible.

Table 1. List of Databases Searched.

- Australian Criminology Database (CINCH)
- Australian Education Index
- British Education Index
- 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
- Index to Theses Database
- MEDLINE
- National Criminal Justice Reference Service (NCJRS)
- PsychInfo/Psychlit
- Sociological Abstracts
- Social Sciences Citation Index (SSCI)
- Web of Knowledge

(b) We then searched by using several keywords in different databases. In total, we carried out the same searching strategies in 18 electronic

databases (Table 1). In all databases, the same key words were used with different combinations. More specifically:
Bully/Bullies/Anti-Bullying/Bully-Victims

And: School

And: Intervention/Program/Outcome/Evaluation/Effect/Prevention/ Tackling

Table 2. List of journals searched from 1983 until end of April 2008.

- Archives of Pediatrics and Adolescent Medicine, 1983 [vol. 137; 1] until 2008 [vol. 162; 3]
- Aggression and Violent Behavior, 1996 [vol. 1] until 2008 [vol. 13]
- Aggressive Behavior, 1983 [vol.9; 1] until 2008 [vol. 34; 2]
- Australian Journal of Education, 2000 [vol. 44] until 2007 [vol. 51]
- Australian Journal of Educational and Developmental Psychology, 2001 [vol. 1] until 2007 [vol. 7]
- British Journal of Educational Psychology, 1983 [vol. 53] until 2008 [vol. 78; 1]
- Canadian Journal of School Psychology, 1985 [vol. 9] and the following volumes: 12 [1 +2]; 13 [1 +2]; 14 [2]; 15 [1]; 16 [1 +2]; 17 [1 +2]; 18 [1 +2]; 19 [1 +2]; 20 [1 +2]; 21 [1 +2]; 22 [1 +2]
- Child Development, 1983 [vol. 34; 1] until 2008 [vol. 79; 2]
- Criminal Justice and Behavior, 1983 [vol. 10; 1] until 2008 [vol. 35; 4]
- Crisis-The journal of Crisis Intervention and Suicide Prevention, 2001 [vol. 22] until 2007 [vol. 28; 4]
- Developmental Psychology, 1983 [vol. 19; 1] until 2008 [vol. 44; 2]
- Educational Psychology, 1983 [vol. 3; 1] until 2008 [vol. 28; 1+2]
- Educational Psychology in Practice, 1985 [vol. 1] until 2008 [vol. 24; 1]
- Educational Psychology Review, 1989 [vol. 1] until 2008 [vol. 20; 1]
- Educational Research, 1983 [vol. 25] until 2008 [vol. 50; 1]
- International Journal on Violence and Schools, January 2006 until December 2007 [vol. 4]
- Intervention in School and Clinic, 1999 [vol. 35] until 2008 [vol. 43; 4]
- Journal of Educational Psychology, 1983 [75; 1] until 2008 [vol. 100; 1]
- Journal of Emotional Abuse, 2004 [vol. 4; 1] until 2007 [vol. 7; 2]
- Journal of Experimental Criminology, 2005 [vol. 1] until 2008 [vol. 4]
- Journal of Interpersonal Violence, 1986 [vol. 1] until 2008 [vol. 23; 1]
- Journal of School Health, 2005 [75; 1] until 2008 [vol. 78; 4]
- Journal of School Violence, 2004 [vol. 3] until 2007 [vol. 6; 3]
- Journal of Youth and Adolescence, 1983 [vol. 12; 1] until 2008 [vol. 37; 6]
- Justice Quarterly, 1986 [vol. 1] until 2008 [vol. 25; 1]
- Pastoral Care in Education, 1983 [vol. 1] until 2007 [vol. 25]
- Psychology, Crime and Law, 1994 [vol. 1] until 2008 [vol. 14]
- Psychology in the Schools, 1983 [vol. 20] until 2008 [vol. 45; 4]
- Scandinavian Journal of Psychology, 1997 [vol. 38] until 2008 [vol. 49; 2]
- School Psychology International, 1983 [vol. 4] until 2008 [vol. 29; 1]
- School Psychology Review, 1983 [vol. 12; 1] until 2008 [vol. 37; 1]
- Studies in Educational Evaluation, 1983 [vol. 9] until 2008 [vol. 34; 1]
- Swiss Journal of Psychology, 1999 [vol. 58; 1] until 2007 [vol. 66; 4]
- Victims and Offenders, 2006 [vol. 1] until 2008 [vol. 3]
- Violence and Victims, 1997 [vol. 3 + 4]; 1998 [vol. 3] and 2002 [vol. 17] until 2008 [vol. 23]

We did not include 'violence' or 'aggression' as key words along with Bully/Bullies/Anti-Bullying/Bully-Victims because we knew that this would identify many studies that were not relevant to the present review, which focuses specifically on studies designed to reduce school bullying.

(c) Table 2 gives a list of the journals that we have hand-searched, either online or in print, since 1983. In total, 35 journals have been searched. For some journals, a hard copy was not available. In this case, we tried to obtain an online version of the journal. For some journals, an online version was available for a year later than 1983 and, if so, this is indicated in the table.

(d) We sought information from key researchers on bullying and from international colleagues in the Campbell Collaboration. In March, we had a meeting with key educational users of the information in Copenhagen, organized by the Nordic Campbell Centre. Where we identified a report in a language other than English (e.g. Martin, Martinez & Tirado, 2005; Sprober, Schlottke & Hautzinger, 2006), we asked colleagues to provide us with a brief translation of key features that were needed for our coding schedule. We believe that, with the cooperation of colleagues in the Campbell Collaboration, we are able potentially to include research in many different developed countries.

(e) A stipulation was made that the title or abstract of each paper would have to include one of the essential key words that were searched. However, some book chapters, mainly from edited books on bullying prevention, were included even though their titles and/or abstracts (if provided) did not include any of our key words.

3. Results of Searches

3.1 Studies Found

A total number of 593 reports that were concerned with interventions to prevent school bullying, as indicated in either the title or the abstract, are included in our systematic review. All studies were categorized based on a relevance scale that we constructed (Table 3). Figure 1 shows the percentage of studies within each category. The vast majority of reports (40.8%) were somewhat relevant, making general suggestions about reducing bullying or, more rarely, reviewing anti-bullying programmes. With regard to the reports that we were not able to obtain, most of them were Masters or PhD theses (11 theses, 3 unpublished manuscripts and 2 conference papers). The cost of ordering these theses through the service of interlibrary loans was time-consuming and expensive enough to make this option unavailable to us. It is possible that some of them would be eligible for inclusion in our meta-analysis. For instance, we understand, based on the review by Vreeman and Carroll (2007), that the thesis of Kaiser-Urley (2003) might be eligible for inclusion in category 5. Moving on to the obtained reports, only 12.8% of them were eligible to be included in the present report or in our forthcoming review for the Campbell Collaboration (categories 5 and 6).

Table 3. Categorization of Reports based on their Relevance to the Present Review.

1. Minor relevance; recommendations for integration of survey results into anti-bullying policies; and/or talk generally about the necessity for bullying interventions.
2. Weak relevance; talking more specifically about anti-bullying programmes [description of more than one anti-bullying programme]; and/or reviews of anti-bullying programmes; and/or placing emphasis on suggestions/recommendations for reducing bullying.
3. Medium relevance; description of a specific anti-bullying programme.
4. Strong relevance; evaluation of an anti-bullying programme, but not included because it has no experimental versus control comparison, or no outcome data on bullying.
5. Included in the Campbell review; evaluation of an anti-bullying programme that has an experimental and control condition [N may be < 200; teacher and peer nominations may also be included as outcome measures].
6. Included in the present review; evaluation of an anti-bullying programme that has an experimental and control condition [N > 200, self-reported bullying as outcome measure].

Category	Count [N]	Percentage
Not Obtained	16	2.7%
Category 1	87	14.7%
Category 2	242	40.8%
Category 3	94	15.9%
Category 4	78	13.2 %
Category 5	17	2.9%
Category 6	59	9.9%

Figure 1. Percentage of studies within each category.

The number of reports concerned with anti-bullying programmes has increased markedly over time, as indicated in figure 2. The total time period was divided into 5-year chunks as follows: 1983–1987, 1988–1992, 1993–1997, 1998–2002 and 2003–2008.

The most obvious increase of interest in implementing and evaluating bullying prevention programmes occurred in the latest period. In the last five years or so (up to April 2008), the number of studies in each category has doubled since the previous 5-year period. It is rather encouraging that studies with a large sample size and including an experimental versus control condition are most prevalent in the last time period.
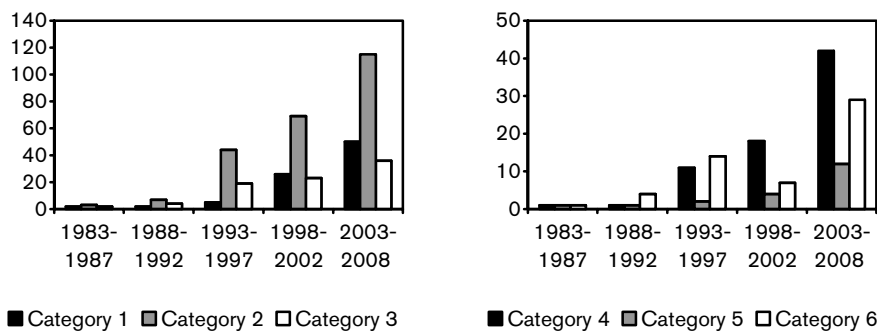
Figure 2. Number of reports in each category within year period.

A total number of 76 reports are eligible for inclusion in our Campbell review. However, 17 of these were excluded from the present review for the following reasons[2]:

- 5 measured bullying/victimization using peer nominations.
- 3 measured bullying/victimization using teacher nominations.
- 10 measured bullying/victimization using self-reports but had a sample size less than 200.
- 2 had an initial sample size of more than 200, but did not use self-reports as outcome measures of bullying.

---

[2] More than one of the following reasons could apply to some of these studies.

20

Consequently, 59 reports were included in the present review, some of which describe the same programme (see tables 4 and 5). These were evaluations of 30 different programmes.

3.2 Included Evaluations

The 59 reports were divided into four categories of research design: randomized experiments, before and after quasi-experimental designs, other quasi-experimental designs, and age-cohort designs. Table 4 lists the 59 reports included in the present systematic review. Within each of the four categories of research design, reports were grouped based on the programme they represent. It was quite possible for different reports from a particular project to be placed in different categories, depending on the content of the report. For example, the report on the Sheffield programme by Whitney, Rivers, Smith and Sharp (1994) was placed in category 6, because information was provided about bullying before and after in experimental and control conditions (schools). However, a later report on the same project by Eslea and Smith (1998) was placed in category 4, because it only presented before and after information about bullying in four experimental schools. As another example, whereas the report by Stevens, Van Oost and De Bourdeaudhuij (2001) was placed in category 6 because it contained outcome data on a specific project (the Flemish programme), the report by Stevens, De Bourdeaudhuij and Van Oost (2001) was placed in category 2 because it reviewed several anti-bullying programmes and did not present outcome data on one specific programme.

Table 5 summarizes key features of the 30 evaluations that are included in this report.

Table 4. 30 Included Evaluations of Anti-Bullying.

Programmes Randomized Experiments.
(1) Bulli & Pupe [Baldry, 2001; Baldry & Farrington, 2004]
(2) Friendly Schools [Cross et al., 2004; Pintabona, 2006]
(3) S.S.GRIN [De Rosier, 2004; De Rosier & Marcus, 2005]
(4) Dutch Anti-bullying Programme [Fekkes et al., 2006]
(5) Steps to Respect [Frey, Edstrom & Hirschstein, 2005; Frey et al., 2005; Hirschstein et al., 2007]
(6) Anti-bullying Intervention in Australian Secondary Schools [Hunt, 2007]
(7) Youth Matters [Jenson & Dieterich, 2007; Jenson et al., 2005a; 2005b; 2006a; 2006b]
(8) Expect Respect [Rosenbluth et al., 2004; Whitaker et al., 2004]
(9) Kiva [Salmivalli et al., 2007]

Before-After, Experimental-Control Comparisons
(1) Greek Anti-bullying programme [Andreou et al., 2007]
(2) Seattle Trial of the Olweus Programme [Bauer et al., 2007]
(3) Progetto Pontassieve [Ciucci & Smorti, 1998]
(4) South Carolina Programme; implementation of OBPP [Melton et al., 1998]
(5) 'Bullyproofing your School' programme [Menard et al., 2008]
(6) New Bergen Project against Bullying; 'Bergen 2' [1997-1998]
(7) Toronto Anti-bullying programme [Pepler et al., 2004]
(8) Ecological Anti-bullying programme [Rahey & Craig, 2002]
(9) Short Intensive Intervention in Czechoslovakia (Rican et al., 1996]
(10) Flemish Anti-bullying programme [Stevens, De Bourdeaudhuij & Van Oost, 2000; Stevens, Van Oost & De Bourdeaudhuij, 2000; Stevens et al, 2001; Stevens et al., 2004]
(11) Sheffield Anti-bullying programme [Whitney et al., 1994; Smith, P.K., 1997; Smith et al., 2004b]

Other Experimental-Control Comparisons
(1) Transtheoretical-based tailored Anti-bullying programme [Evers et al., 2007]
(2) Norwegian Anti-bullying programme [Galloway & Roland, 2004]
(3) SAVE [Ortega & Del Rey, 1999; Ortega et al., 2004]
(4) Kia Kaha [Raskauskas, 2007]

Age-Cohort Designs
(1) Respect [Ertesvag & Vaaland, 2007]
Olweus Bullying Prevention Programme [OBPP]:
(2) First Bergen Project against Bullying; 'Bergen 1' [1983-1985]
(3) First Oslo Project against Bullying; 'Oslo 1' [November 1999-November 2000]
(4) New National Initiative Against Bullying in Norway; 'New National' [2001-2007]
(5) Five-year Follow-up in Oslo; 'Oslo 2' [2001-2006]
[Olweus, 1991; 1992; 1993b; 1994a; 1994b; 1994c; 1995; 1996a; 1996b; 1996c; 1997a; 1997b; 1997c; 2004a; 2004b; 2005a; 2005b; Olweus & Alsaker, 1991]
(6) Finnish Anti-bullying programme [Salmivalli et al., 2004; 2005]

Table 5. Key Features of Evaluations.[3]

Randomized Experiments			
Project	**Components of the programme**	**Participants**	**Research Design**
Baldry & Farrington (2004)	Kit of 3 videos and a booklet divided into 3 parts; used in active methods such as role-playing, group discussions and focus groups	239 students aged 10–16 in 13 schools: • 131 in the experimental group • 106 in the control group • experimental and control students from the same schools but from 10 different classes; classes randomly assigned	Intervention and control groups, random assignment, pretest and post-test measures
Cross et al. (2004)	Targeting 3 levels: a) the whole-school community ('whole-school planning and strategy manual') b) students' families (home activities linked to each classroom-learning activities; 16 skills-based newsletter items) c) grades 4-5 students along with their teachers (classroom curriculum)	2,068 students (aged 9-10 from 29 schools) of which: −1046 intervention students − 922 control students −15 intervention schools −14 control schools	Pretest and posttest data from intervention and control schools; 3-year randomised control trial
De Rosier (2004); De Rosier & Marcus (2005)	Program for children experiencing peer dislike, bullying or social anxiety; highly structured manualized intervention combining social learning and cognitive-behavioural techniques	1079 students − 50.8% boys − 49.2% girls − mean age: 8.6 years of which: − 415 eligible to participate in S.S.GRIN (664 children as non-identified)	Pre-test, post-test, experimental and control groups; 18 children in each school (11 public elementary schools; North Carolina) randomly assigned to the treatment group and the remainder of the list assigned to no-treatment control group
Fekkes et al. (2006)	An anti-bullying school program including anti-bullying training for teachers, a whole-school anti-bullying policy, an anti-bullying curriculum	3816 students aged 9 to 12 years (50% of the sample girls)	Two-year follow-up randomized intervention-group/control-group design; schools randomly assigned

---

[3] All dates in the tables specify the year of publication of the report [not the year the programmes took place] with the exception of the Olweus evaluations; for the latter, the period the programme took place is shown. Not all published reports of a specific programme are presented in this table, only the most relevant ones.

Frey et al. (2005)	Training manual for staff (staff training) including a core instructional session for all school staff and two in-depth training sessions for counsellors, administrators and teachers; classroom curriculum (10 semi-scripted skill lessons); parent engagement (take-home letters etc)	A random sub-sample (N=544) of a longitudinal study (N=1023) observed and their behaviour being coded	Pre-test, post-test, experimental and control groups, schools randomly assigned
Hunt (2007)	Information at parent and teacher meetings about the nature of bullying in schools; school staff conducted a 2-hour classroom-based discussion of bullying using activities from an anti-bullying work-book	• 44 students at T1 (155 intervention students and 289 control students) and of those 318 at T2	Pre-test, post-test, experimental and control groups; schools randomly assigned to intervention or wait-list condition
Jenson & Dieterich (2007)	Youth Matters Prevention Curriculum; series of instructional modules; 10-session module during each of the four semesters of 2 academic years	Fourth-graders from 28 schools: 456 control students and 670 experimental students	Group-randomized trial; fourth grade classrooms from 28 schools randomly assigned
Rosenbluth et al. (2004)	5 program components including classroom curriculum; staff training; policy development; parent education; support services for individual students	Fifth graders from elementary schools (929 students in intervention group and 834 in the comparison group)	Pre-test, post-test, intervention and control groups; pair of schools matched and randomly allocated to experimental or control conditions
Salmivalli et al. (2007)	Universal/whole-school intervention; Indicated intervention/work with individual students; comprehensive programme with manuals for teachers, information for parents; increased supervision; internet-virtual learning environments; web-based discussions forum for teachers; peer support for bullies and victims of bullying	All Finnish comprehensive schools invited to volunteer; of the 300 schools who were willing to participate, a representative sample of 78 schools was chosen; programme still running/ no final results yet	An age-cohort design and a randomized experiment 'nested' in the same programme; only results for the latter available

Before-After, Experimental-Control Comparisons			
Project	**Components of the programme**	**Participants**	**Research Design**
Andreou et al. (2007)	Set of curricular activities to create classroom opportunities for a) awareness raising, b) self-reflection and c) problem-solving situations relevant to bullying	454 pupils: • 206 control: 123 boys and 83 girls • 248 experimental: 126 boys and 122 girls Sample size by grade: 145 fourth grade 162 fifth grade 147 sixth grade	An experimental pre-test, posttest design with a control group. Classes assigned to the experimental and control groups on the basis of teachers' willingness to be involved in the intervention
Bauer et al. (2007)	The Olweus Bullying Prevention Program; Components targeting school-, classroom-, individual- and community-level interventions	• 4959 intervention students of which: – 2522 females –1672 sixth graders –1629 seventh graders –1588 eighth graders • 1559 control students of which: –782 females – 570 sixth graders – 515 seventh graders – 449 eighth graders	A nonrandomized controlled trial with 10 public middle schools (7 intervention – implementing the Olweus Bullying Prevention Program – and 3 control)
Ciucci & Smorti (1998)	Three levels: school (first two years) to promote an anti bullying policy; class and individual level (third year): Quality Circles & Role Playing to promote cooperative and problem-solving skills	167 students participated in the treatment group. 140 students part of the control group All children from one secondary school.	Experimental pretest, posttest control group design
Melton et al. (1998)	Inspired by the OBPP; school-wide, classroom, individual and community interventions based on the OBPP	Fourth through eighth grade students from six non-metropolitan school districts. Districts organised into matched pairs: Group A schools: implemented the project for 2 years Group B schools: served as a comparison group for the first year of the project and received the intervention the second year. Baseline: 6389 students [grade 4-6] Time 1: 6263 students [grade 5-7] Time 2: 4928 students [grade 6-8]	Before-after, experimental-control comparison with 3 measurements: baseline [March 1995], T1 [March 1996] and T3 [May 1997]

Menard et al. (2008)	Comprehensive, school-based intervention; classroom curriculum (7 core sessions and 2 optional);	All students in each of the third-through fifth-grade classrooms in 7 elementary schools [3497 students] and all students in sixth-through eighth-grade classrooms in 3 middle schools [1627]	Multiple non-equivalent control group pre-test posttest design with ex ante selection of treatment and comparison groups; matched treatment and comparison groups at baseline
Olweus: Bergen 2 [1997–1998]	School level [e.g. Staff discussion groups; Bullying Prevention Coordinating Committee]; Classroom level [e.g. classroom rules]; individual level [e.g. supervision of students]; and community level components	Approximately 2,400 students in grades 5, 6, and 7 [OBPP had been in place for only 6 months when the second measurement took place]	An experimental pre-test, posttest design with a control group; eleven intervention and eleven comparison schools
Pepler et al. (2004)	Systemic school-based program; 3 similar elements of intervention across the 3 schools: staff training; codes of behaviour; improved playground intervention	Pupils from 3 schools (aged 5 to 11); 2 classes from each grades 1-6 (12 classes in all) from each school were randomly selected to participate; 319 children from school A and 300 children from school B the first year of the program; 325, 240 and 303 children from schools A, B and C accordingly during the second year; 306, 163 and 289 children from school A, B and C accordingly in the second year of the program	Quasi experimental with two waiting-list controls. In year 1, school A started the program and school B served as a waiting-list control. In year 2, school A continued the program, school B formally started the program while school C served as a waiting list control. In year 3, schools A and B continued the program, while school C began its formal involvement in the anti-bullying program
Rahey & Craig (2002)	12-week program based on the Bully Proofing Your School Program; psycho-educational program within the classroom; a peer mediation program; groups for children referred for involvement in bullying/victimisation	Students from one intervention (114 boys and 126 girls) and one comparison school (123 boys and 128 girls); children in grades one through eight	An experimental pre-test, posttest design with a control group [one experimental school and one control school]
Rican et al. (1996)	Programme inspired by the OBPP; components of the OBPP – e.g. Olweus videocassette – used along with other methods (e.g. 'class charter'	8 fourth grade elementary school classes used [half in each condition] −100 students in experimental condition − 98 students in control condition	Pretest posttest experimental-control comparison

26

Stevens et al. (2000)	Training sessions for teachers; manual with video; three modules; booster sessions	1,104 students aged 10-16 from 18 schools: – 151 primary and 284 secondary students in Treatment with Support – 149 primary and 277 secondary students in Treatment without Support – 92 primary and 151 secondary students in the Control Group	Experimental pretest/posttest comparison including a control group [2 experimental groups – Treatment with Support and Treatment without Support – and one control group]
Whitney et al. (1994)	Whole school approach; curriculum classroom strategies; the Heartstone Odyssey; quality circles; 'Only playing Miss' theatrical play; peer counselling; bully courts; changes to playgrounds and lunch breaks	27 schools in total in this second survey, 8309 students aged 8-16 from 16 primary and 7 secondary (intervention) schools; 4 control schools; 1 primary (99 pupils) and 3 secondary (1742 pupils)	Pretest and posttest 18 months later. 3 year follow-up in 4 intervention schools

Other Experimental-Control Comparisons			
Project	**Components of the programme**	**Participants**	**Research Design**
Evers et al. (2007)	The Build Respect, Stop Bullying™ Program was offered; a multi-component intervention package	12 middle schools and 13 high schools in the USA (1237 middle and 1215 high school students) : – 483 middle and 309 high school students in control group – 488 middle and 375 high school students in Treatment 1 – 266 middle and 531 high school students in Treatment 2	3X2 experimental design crossing 3 experimental groups with 2 treatment groups; pretest and post-test measures; schools matched on key variables (type of community, region of country and % of students eligible for free lunches)
Galloway & Roland (2004)	Professional development program for teachers; 4 inservice days over a 9-month period; 15 2-hour peer supervision sessions; handouts for teachers	9 intervention schools and 6 control groups: – comparison sample 1 – experimental sample 1 – experimental sample 2 – comparison sample 2 300-350 pupils in each sample apart from comparison sample 2 [151 students]	Longitudinal design with two experimental and two comparison samples of first graders – primary schools – only in a two-year period [1992–1994]
Ortega et al. (2004)	Educational intervention model; democratic management of interpersonal relationships; co-operative group work; education of feelings and values; direct intervention with high-risk students	In the 5 intervention schools: – 731 intervention pupils at pretest and 901 intervention students at posttest In the 4 control schools – 440 control pupils	5 intervention schools [3 primary; 2 secondary] had pretest and posttest measures, compared to 4 control schools with only posttest measures. Follow-up after 4 years

27

| Raskauskas (2007) | A whole-school approach | 49 schools – excluding 4 schools that intended to implement the program [31 intervention schools that implemented Kia Kaha for a 3-year period with 22 control schools all together] | Comparison of intervention schools with matched-comparison groups |

Age-Cohort Designs			
Project	**Components of the programme**	**Participants**	**Research Design**
Ertesvag & Vaaland (2007)	Teachers and school management staff participate in series of seminars; a 2-day seminar for the school management personnel and school representatives was also run in advance of the implementation period	• Pupils from 3 primary and 1 secondary school in Norway • Pupils in grades 5–6 (aged 11–13) at the primary schools and grades 8-10 (aged 14–16 years) at the secondary school • Number of pupils completing the survey at T1–T4 was: 745, 769, 798 and 792 respectively	'Age-longitudinal design with adjacent or consecutive cohorts' with four measurement points
Olweus: Bergen 1 [1983–1985]	School level [e.g. Staff discussion groups; Bullying Prevention Coordinating Committee]; Classroom level [e.g. classroom rules]; individual level [e.g. supervision of students]; and community level components	Students from 112 grade 4–7 classes in 42 primary and junior high schools Each of the 4 age cohorts consisted of 600-700 subjects with roughly equal distribution of boys and girls	Extended selection cohorts design with 3 measurements; May 1983; May 1984 and May 1985
Olweus: Oslo 1 [1999–2000]	School level [e.g. Staff discussion groups; Bullying Prevention Coordinating Committee]; Classroom level [e.g. classroom rules]; individual level [e.g. supervision of students]; and community level components	Approximately 900 students [at both time points] in grades 5 through 7	Extended selection cohorts design with 2 measurements; 1999 and 2000
Olweus: New National [2001–2007]	School level [e.g. Staff discussion groups; Bullying Prevention Coordinating Committee]; Classroom level [e.g. classroom rules]; individual level [e.g. supervision of students]; and community level components	Students in grades 4 through 7 from only 3 – out of 5 – different cohorts of schools are provided	Extended selection cohorts design; data provided for 3 measurements: October 2001, October 2002 and October 2003

| Olweus: Oslo 2 [2001–2006] | School level [e.g. Staff discussion groups; Bullying Prevention Coordinating Committee]; Classroom level [e.g. classroom rules]; individual level [e.g. supervision of students]; and community level components | Data for assessments for the 14 out of 19 Oslo schools from the first cohort are provided

Students in grades 4–7 followed from 2001 until 2005. Students in grades 8–10 followed from 2001 until 2003 | Extended selection cohorts design; data provided for 5 measurements for students in grades 4 through 7; data provided for 3 measurements for students in grades 8 through 10 |
|---|---|---|---|
| Salmivalli et al. (2004); Salmivalli et al. (2005) | Intervention training for teachers; class-level interventions; school-level interventions [whole-school anti-bullying policy]; individual-level interventions | 8 schools from Helsinki and 8 schools from 4 towns near Turku 1,220 students aged 9-12 in 16 schools [600 girls] | Age-longitudinal design with adjacent cohorts |

4. Descriptions of Included Programmes

Next we describe the elements of the included anti-bullying programmes. These descriptions are based on the best available data and on the information provided in reports evaluating the intervention (category 6), rather than in reports describing the programme (category 3). The rationale underlying this decision refers to the fact that the way in which a programme was designed and the way it was implemented in the school may be two different procedures that do not necessarily have everything in common. For all programmes we have attempted to contact the evaluators of the programme.

4.1 Randomized Experiments

Bulli and Pupe

'Bulli and Pupe' was an intervention programme concerned with bullying and family violence. The programme, developed by Baldry (2001), was 'directed towards the individual and peer group, and aimed to enhance awareness about violence and its negative effects' (Baldry & Farrington, 2004, p. 3). The intervention package consisted of three videos and a booklet divided into three parts; each video was linked to one part of the booklet. Each part of the booklet was meant to take the form of an interactive lesson where professionals, experienced in school and juvenile processes, discussed three issues according to the structure of the manual.

The first part of the booklet, entitled 'Bullying among peers', emphasized teen violence among peers. The booklet presented vignettes and graphics that reported research findings on bullying in an attempt to raise students' awareness of this issue. The corresponding video showed teenagers talking about bullying based on their own experiences and judgments. The second part of the booklet, entitled 'Children witnessing domestic violence', analysed the effects of domestic violence on children and the repercussions for school achievement and peer relations. In the accompanying video, children in a shelter for battered women were presented, talking about their personal experiences and emotions. Finally, the third part of the booklet, entitled 'Cycle of violence', dealt with the long-term effects of violence on adults who were victims of violence in their childhood. The corresponding video consisted of an interview conducted with a 19-year old boy who had a violent father.

The programme was in the first place delivered in three days by experts who, together with teachers, discussed about bullying, read the booklet and analysed its content. The programme was taken over by teachers who once a week created a facilitation group and allowed chil-

dren to discuss any problems they encountered with their peers. The programme was more effective with secondary students because it required its participants to have good interpersonal and cognitive skills (Baldry & Farrington, 2004, p. 4).

Friendly Schools

'Friendly Schools' was a theoretically grounded programme. Its educational techniques (e.g. role modelling, drama activities, skills training, etc.) were based on notions derived from Social Cognitive theory, the Health Belief Model and Problem Behaviour theory (Cross et al., 2004, p. 191). An interesting aspect of this programme is that it was based on the results of a systematic review (Cross et al., 2004, p. 187), which provided a set of key elements to be included in the final intervention strategy. The programme targeted bullying at three levels: a) the whole-school community; b) the students' families; and c) the fourth and fifth grade students and their teachers.

With regard to the whole-school intervention component, in each school, a Friendly Schools Committee was organised with key individuals (e.g. a parent representative, a school psychologist, a school nurse, teaching staff) who could co-ordinate and successfully sustain the anti-bullying initiative. Each committee was provided with a four-hour training, designed to build members' capacity to address bullying. Each member was provided with a specific strategy manual. The manual was a step-by-step guide on how to implement the anti-bullying initiative. It included among others the Pikas 'Method of Shared Concern' and the 'No Blame' approach.

With regard to the family intervention component, this included home activities linked to each classroom-learning activities. Parents were also provided with 16 skills-based newsletter items (eight for each year of the intervention) that aimed to provide research information on bullying as well as advice to parents on what to do if their child was a perpetrator or a victim of bullying behaviour.

Moving on to the Grade 4 and 5 classroom curriculum, the Friendly Schools curriculum consisted of nine learning activities per year. The curriculum was offered by trained teachers in three blocks of three 60-minute lessons, over a three-school-term period. The learning activities aimed to promote awareness of what was bullying behaviour; to help students to become assertive and talk about bullying with teachers and parents; and to promote peer and adult discouragement of bullying behaviour.

Finally, the Friendly Schools programme offered manuals to teachers. The teacher manuals were designed to be entirely self-contained so as to maximize the likelihood of teacher implementation. Friendly Schools project staff also provided teacher training (a six-hour course) for all intervention teachers.

S.S. GRIN

The Social Skills Group Intervention (S.S.GRIN) was a school-based programme that aimed to help children enhance their social skills. S.S.GRIN was designed as a social-skills training intervention for peer-rejected, victimized and socially anxious children. It could be applied to an array of problems that are social in nature (e.g. aggression, low self-esteem, depression, social anxiety, social withdrawal) not just bullying (De Rosier & Marcus, 2005, p. 140). The authors argued that the programme went beyond the most common social-skills training (De Rosier & Marcus, 2005, p. 141) by emphasizing the cognitive aspects of relations and emotions. That is, children were not only taught pro-social skills, but they were also taught, on the cognitive level, how to identify negative perceptions and behaviours in an effort to help children to regulate their own emotions as well as enhance their coping skills.

Overall, the programme was a combination of social-learning and cognitive-behavioural techniques, used to help children build social skills and positive relationships with peers. It was a highly structured, manualized program (De Rosier, 2004, p. 197) with a number of sessions containing scripts and activities to undertake. Each session included didactic instruction combined with active practice such as role-playing, modelling and hands-on activities (De Rosier, 2004, p. 197). The children participated in group sessions for eight consecutive weeks. Each session lasted approximately an hour. The groups were led by each school's counsellor and an intern, who were trained and supervised by one of the programme instructors (De Rosier & Marcus, 2005, p. 143).

Dutch Anti-Bullying Programme

The anti-bullying initiative in the Netherlands was inspired by the Olweus programme (Fekkes et al., 2006, p. 639). The programme was specifically designed to tackle bullying behaviour by involving teachers, parents and students. It offered a two-day training session for teachers in order to inform them about bullying behaviour and to instruct them about how to deal with bullying incidents in schools. During the intervention period, teachers had access to the training staff for additional advice. Intervention schools were supported by an external organization named KPC, which specialized in training school staff and in assisting schools in setting up new curricula and guidelines. The core intervention programme included: a) anti-bullying training for teachers; b) a bullying survey; c) anti-bullying rules and a written anti-bullying school policy; d) increased intensity of surveillance; and f) information meetings for parents.

During the intervention, there was careful dissemination of the anti-bullying programme to intervention schools. Also, the researchers provided information about the number of intervention and control schools, which have used the above-mentioned elements of intervention. Finally, intervention schools were supplied with the booklet 'Bullying in

schools: how to deal with it' and with a 'Bullying Test', a computerized questionnaire that children could complete anonymously in the classroom.

Steps to Respect

The Step to Respect program aimed to tackle bullying by a) increasing staff awareness, b) fostering socially responsible beliefs, and c) teaching social-emotional skills so as to promote healthy relationships (Frey et al., 2005, p. 481). The programme included staff and family training manuals, a programme guide and lesson-based curricula for third-through sixth-grade classrooms (Hirschstein et al., 2007, p. 7).

Components at a whole school level consisted of an anti-bullying policy and procedures, staff training and parents meetings, all aiming at sharing understanding of bullying and its consequences and increasing adult awareness, monitoring and involvement. At the classroom level, the proposed activities consisted of teaching friendship skills, emotion regulation skills, identifying types of bullying, teaching prevention strategies and peer group discussion. The aim was to improve peer relations and reduce the risk of victimisation, assess level of safety and recognize, report and refuse bullying. At the individual level, students involved in bullying were approached and coached based on the 'Four-A Responses': affirm behaviour, ask questions, assess immediate safety and act.

The S to R training manual consisted of an instructional session for all school staff and two in-depth training sessions for counsellors, administrators and teachers. There were also videos accompanying the programme. With regard to staff training, there were two levels of training: all school staff received an overview of the programme goals and principal aspects of the programme (programme guide). Teachers, counsellors and administrators received additional training in how to coach students involved in bullying, based on behavioural skills training, co-operative learning and role-playing.

The student curriculum comprised skills and literature-based lessons delivered by third- through sixth-grade teachers during a 12–14 week period. The intervention consisted of 10 semi-scripted skills lessons with topics such as joining groups, distinguishing reporting from tattling and being a responsible bystander.

Finally, with regard to the parent intervention, administrators informed parents about the programme and the school's anti-bullying policy and procedures. Parents could also benefit from other resources such as letters provided to them and newsletters describing whole-school anti-bullying activities undertaken at school.

Anti-Bullying Intervention in Australian Secondary Schools

This anti-bullying intervention consisted of several activities that aimed to increase awareness and identification of bullying, to promote empa-

thy for targets of bullying and to provide students with strategies to cope with bullying (Hunt, 2007, p. 22). The intervention was based on an educational anti-bullying programme, which was delivered by teachers. There was no specific training for teachers. Information about bullying was provided at parent and teacher meetings. Teacher meetings were held in conjunction with regular staff meetings whilst parent meetings were held after hours. A summary of the information covered at parent meetings was also published in the school newsletter in an attempt to target the wider parent population. Finally, the programme included a two-hour classroom-based discussion of bullying (offered by teachers) using activities from an anti-bullying workbook written by Murphy and Lewers (2000).

Youth Matters

The Youth Matters programme used 'a curricular and a modified systemic approach to bullying prevention' (Jenson & Dieterich, 2007, p. 287). The aim of the curriculum was to strengthen peer and school norms against antisocial behaviours by addressing critical issues (issue modules) such as the difference between teasing and bullying, building empathy, risks and norms surrounding aggression and so on. The curriculum also aimed to promote skills (skill modules; structured skills training sessions) that students could use in order to stay safe at school, cope with bullying, enhance their social skills and improve their peer relationships. To address systemic issues associated with bullying, curriculum modules terminated with the development of classroom or school-wide projects, which placed emphasis on the negative consequences of bullying for students.

The curriculum consisted of ten-session modules. Each module included a 30–40 page story, the content of which was directly linked to the structured skills training sessions. When looking at the implementation of the programme, all curriculum materials were 'language sensitive': translated into Spanish for use in the three Spanish-speaking classrooms included in the evaluation. Youth Matters curriculum modules were offered to fourth and fifth graders. According to Jenson and Dieterich (2007, p. 287), grades 4 and 5 were selected 'based on an appropriate fit between developmental ability and curricula'.

The Youth Matters programme was based on a theoretically grounded curriculum. The curriculum was based on theoretical constructs derived from the Social Development Model. The latter integrated perspectives from three theories (i.e. social control theory, social learning theory and differential association theory) and proposed that four factors inhibit the development of anti-social development in children. These were: a) bonding or attachment to family, schools and positive peers; b) belief in the shared values or norms of the above-mentioned social units; c) external constraints or consistent standards against antisocial behaviour; and d) social, cognitive and emotional skills that can

34

be seen as protective tools for children to solve problems and perform adequately in social situations. The Youth Matters curriculum addressed each of these four core areas.

Expect Respect

Expect Respect was a school-based programme that aimed to promote awareness and effective responses to bullying and sexual harassment. The project was developed by Safe Place, the sole provider of comprehensive sexual and domestic violence prevention and intervention services in Austin, Texas (Rosenbluth et al., 2004, p. 211). The programme targeted the involvement of all members of the school community in recognising and responding to bullying and sexual harassment. The overall project design was inspired by the work of Olweus (Rosenbluth et al., 2004, p. 212). Expect Respect consisted of five core programme components, namely a classroom curriculum, staff training, policy development, parent education and support services.

The classroom curriculum was based on 12 weekly sessions adapted from a specific manual called 'Bullyproof: a teachers' guide on teasing and bullying for use with fourth and fifth grade students' (Whitaker et al., 2004, p. 330). The Bullyproof curriculum was designed to be taught in conjunction with literature typically read by fourth and fifth graders. Although the anti-bullying curriculum was designed to be implemented by teachers, within the framework of the Expect Respect programme, it was jointly led by Safe Place Staff and teachers or school counsellors (Whitaker et al., 2004, p. 331). The curriculum aimed to increase the ability and willingness of bystanders to intervene in bullying situations, thus reducing the social acceptability of bullying and sexual harassment. The Bullyproof lessons included writing assignments, role-plays of how to intervene in bullying situations, class discussions and so on.

With regard to the staff training, a six-hour training was provided to project staff, counsellors, and fifth grade teachers. The training was given by the author of the specific manual and aimed to prepare school personnel to respond effectively to bullying incidents. In addition, three-hour training sessions were provided once per semester for all personnel, including bus drivers, cafeteria workers, hall monitors and office staff. The training presentation included research on bullying and sexual harassment; strategies to enhance mutual respect among students; practice in using lessons from the curriculum; and methods for integrating the lessons into other subject areas including language arts and health.

School administrators were encouraged to develop an anti-bullying policy (policy development) in their school to ensure consistent responses by all staff members to incidents of bullying and sexual harassment. Principals were expected to present the policy to school staff, students and parents. In order to facilitate the overall procedure of policy development, Expect Respect staff provided an initial policy template to school administrators (Whitaker et al., 2004, p. 332) and each school

was encouraged to expand this initial policy in accordance with the specific needs of their unit.

The Expect Respect programme also included parent training. Educational presentations were offered to parents twice a year, providing information about the project. The information given to parents through these meetings (as well as through parent newsletters sent home) was aimed at enhancing parents' strategies to help children involved in bullying as bullies, victims, bully-victims or bystanders.

Further support services were provided such as continuous assistance of school counsellors by Safe Place staff. School counsellors were given a specialized session on how to deal with students who were repeatedly involved in bullying as either perpetrators or victims. They were also provided with a comprehensive resource manual containing reading and resource materials on bullying, sexual harassment and domestic violence.

KiVa

The name of the specific project is an acronym of the expression 'Kiusaamista Vastaan' which means 'against bullying'. The word 'kiva' in Finnish means 'nice' and this is why this acronym was chosen for the specific anti-bullying initiative in Finland. Regarding the overall perspective of the programme, the KiVa project included a universal and an indicated intervention. The universal intervention referred to efforts made to influence the group norms whilst the indicated intervention referred to the way in which specific cases were handled in schools through individual and group discussions between the teacher and the students involved (Salmivalli et al., 2007, p. 6).

The KiVa programme included a large variety of concrete materials for students, teachers and parents. It also utilized the Internet and virtual learning environments (e.g. computer games against bullying) aiming in this way to enhance students' attitudes against bullying. Also, students received their own personal user ID, which they could use as a password before the completion of each web-based questionnaire on bullying. KiVa included 20-hour student lessons, which were carried out by student teachers. The lessons involved discussions, group work, short films about bullying, and role-playing exercises. After each lesson, a class rule was adopted, based on the central theme of the lesson.

A unique feature of the KiVa programme was the use of an anti-bullying computer game. The game involved five levels and the teacher always activated the next level of the game after the relevant lesson was completed. Students were able to begin using the game after the third lesson; the second level of the programme was played after the fifth lesson, and so on until the end of the school year. Each level of the computer game included three components that were named as 'I know', 'I can' and 'I do'. In the first component, students were informed about basic facts on bullying. In the second component, the 'I can'-component, students moved around in the virtual school and faced different chal-

lenging bullying incidents. Finally, the third component was used to encourage students to make use of their knowledge and skills in real life situations.

Another important element of the KiVa project was the teacher training. Teachers were also provided with vests that they could use during playtime while supervising the school yard. This simple technique aimed to enhance teachers' visibility in the schoolyard and to signal that bullying was taken seriously in the school. Also, all teachers carrying out the KiVa programme could seek advice from a web-based discussion forum, where they could share experiences and ideas about bullying with other colleagues.

Within the school framework, the programme also facilitated the use of a peer support group for victims of bullying. The classroom teacher was expected to arrange a group with 2–4 classmates – those who were pro-social and had high status in the class – who were expected to provide support to victimized students, thus sustaining healthy peer relationships. An interesting element in the KiVa programme is that it incorporated both punitive and non-blame approaches when dealing with perpetrators of bullying. Half of the school teams were instructed to use more punitive approaches (e.g. 'what you have done is wrong and it has to stop right now) whilst the rest of the school teams were instructed to use no-blame approaches in their discussions with children (e.g. 'your classmate is also having a hard time and this is why he behaves like that; what could we do to help him?'). There was also co-operative group work among experts when dealing with children involved in bullying.

Finally, the KiVa programme involved parents. A parents' guide was sent to the home and provided information about bullying and advice on how parents could be involved to reduce this problem. Information nights for parents were also organised and provided.

4.2 Before-After/Experimental-Control Comparisons

Greek Anti-Bullying Programme

The Greek anti-bullying initiative was a four-week intervention programme that aimed to minimize both bullying and victimization. The conceptual framework of the Greek anti-bullying programme was based on the theoretical model proposed by Salmivalli in 1999 (Andreou et al., 2007, p. 696), according to which changing an individual's behaviour (e.g. the bully's behaviour) entailed motivating not only the particular person but also the rest of the group members (participant roles' approach).

The programme was embedded within the wider curriculum of the fourth-, fifth- and sixth-grade classrooms and consisted of eight instructional hours, each hour corresponding to one curricular activity. The

curricular activities were presented to students by their classroom teachers who received training beforehand. The teacher training consisted of five 4-hour meetings and aimed to increase awareness of the bullying problem and its seriousness as well as to raise teachers' self-efficacy in implementing the programme (Andreou et al., 2007, p. 697).

The Greek anti-bullying curriculum was divided into three parts in accordance with the three main theoretical axes proposed by Salmivalli in 1999, namely: a) awareness-raising; b) self-reflection; and c) commitment to new behaviours (Andreou et al., 2007, pp. 697–698).

In line with the first axis (awareness-raising), small-group and whole-class discussions were conducted (over three instructional hours) that aimed to increase students' awareness of the bullying problem. Corresponding materials included a real snap-shot from the playground, a story entitled 'A new friend' and students' own drawings. In line with the second theoretical axis (self-reflection), two instructional hours involving classroom discussions were conducted. These discussions placed emphasis on the participant roles that students took in the bullying process. Corresponding materials involved each students' completion of open-ended sentences. Through this activity students were intended to reflect on critical issues around the causes, benefits, feelings, and consequences of adopting different roles. In line with the final axis (commitment to new behaviours), three instructional hours of small-group and whole-class discussions were conducted concerning different ways of approaching or solving the peer-conflict situation and the formulation of class rules. Corresponding materials involved an open-ended comic-strip for group completion to find a solution to the bullying situation presented in the relevant story.

Seattle Trial of the Olweus Programme

The Olweus Bullying Prevention Programme (OBPP) was implemented and evaluated in a non-randomized controlled trial in a cohort of ten Seattle middle schools (Bauer et al., 2007, p. 267). The overall programme was in absolute concordance with the Olweus programme and aimed at improving peer relations and promoting a safe and positive school environment by addressing and tackling the problem of bullying.

Intervention schools received consultation by district trainers prior to implementation. The programme components corresponded to several levels of intervention such as the whole-school level, the classroom level, the individual level and the community level. At the school level, the programme started with an 'official start date' during which a school assembly took place aiming to present the overall programme to students, introduce the basic concepts and raise enthusiasm among students. The core components of the programme at the school level also included a coordinating committee, the members of which were responsible for the initial planning and oversight of the implementation of the intervention. Regular staff discussions were also organised with the goal

of fostering collaboration in implementation efforts. School anti-bullying rules were presented to students that set clear guidelines about the students' behaviour that was expected within the school. School surveillance was a crucial element of the anti-bullying programme. Tracking and identifying 'hot spots' of bullying was crucial in reducing the percentage of bullying incidents whilst continuous surveillance on behalf of the teachers involved constant reminders that bullying was an unacceptable form of behaviour in the school. Teachers in the intervention schools received teacher training.

The programme aimed to raise awareness of the problem of bullying among the parents and the overall community as well. Involving parents and the overall community was an important element of the programme since students' behaviour could not be seen as fragmented: socially acceptable forms of behaviour should be positively reinforced within and outside the school community.

Progetto Pontassieve

The programme was delivered in a period of three years, and it consisted of two main parts. During the first two years it was delivered more at the school level whereas the third year was more at the class and individual level (Ciucci & Smorti, 1998). During the first year a training course for teachers took place addressing psychosocial risks for children and bully-victim problems. At the end of the training, a study was conducted to reveal how serious was the problem of bullying and what were its characteristics. The second year of the intervention included a counselling service for each individual who was affected by bullying.

The intervention took place in the third year and was based on the use of two different methods: Quality Circles, where pupils had to co-operate to find practical solutions to their problems, with the use of the Interpersonal Process Recall which consisted of the recording of one Quality Circle and discussion about it. The other method used was Role Playing conducted in small groups with subsequent class discussions, which helped students to examine possible strategies to face and over-take bullying problems. The aims of both of these methods were to make students aware that they could intervene in an efficient way to reduce bullying.

South Carolina Programme

This programme involved the implementation of the OBPP in South Carolina schools. It was a comprehensive school-based anti-bullying programme essentially inspired by the Norwegian model (Melton et al., 1998, p. 72; p. 74) and aimed to target bullying at the school, classroom, individual and community levels.

In accordance with the OBPP, the South Carolina programme included a school-wide intervention. In each school, coordinating commit-

tees planned and guided the school's anti-bullying initiative throughout the various phases of the project. The committees consisted of school psychologists or counsellors and representative teachers, students and parents. In each school, a survey was conducted prior to the implementation of the programme, which aimed to assess the nature and extent of bullying problems in the school. The survey results were presented during a school conference day that aimed to increase students' awareness about this problem. There were school-wide events to launch the programme. Another element of the programme at the school level included teacher surveillance in order to track down 'hot-spots' of bullying.

At the classroom level, core elements of the programme included the formulation of clear anti-bullying rules, the use of consistent sanctions for violating the rules, the use of consistent praise of pro-social behaviour by teachers and the scheduling of regular classroom meetings or discussions during which teachers and peers discussed issues related to bullying in their school. Teachers had a wide variety of materials that they could use in the classroom such as videos and classroom materials, a teachers' guide, and programme newsletters that they could consult ('Bully-Free Times').

At the individual level, interventions included discussions with bullies and their parents and the development of safety plans for chronic victims of bullying. Informational newsletters for parents were also provided. At the community level, an effort was made to involve community members in the anti-bullying initiative by a) making the programme known among a wide range of residents in the local community, b) engaging community members in the school's anti-bullying activities and c) engaging community members, students and school personnel in anti-bullying efforts within the community (e.g. by introducing programme elements into summer church school classes).

Other elements of the programme included the involvement of school-based mental health professionals to assist the development of individual interventions with children who were frequently involved in bullying as perpetrators or victims, the development of American versions of several materials used in the OBPP and the development of additional materials for teachers and other school staff such as teachers' guide books and teachers' newsletters.

Bully-Proofing Your School

'Bully-Proofing Your School' (BPYS) was a comprehensive, school-based intervention programme for the prevention of bullying, with three major components: a) heightened awareness of the problem of bullying, involving a questionnaire to measure the extent of bullying and the creation of classroom rules related to zero tolerance for bullying; b) teaching students protective skills for dealing with bullying, resistance to victimization and providing assistance to potential victims by teaching assertiveness skills; and c) creation of a positive school climate where stu-

dents were encouraged to work as positive and supportive bystanders (Menard et al., 2008, p. 7). The primary targets of BPYS were elementary and middle school students. School staff were involved as both secondary targets of intervention (since changes in their behaviour was a requirement for the construction of a positive anti-bullying school environment) and as agents delivering the intervention to students. Teachers were given information and strategies that they could use while delivering the intervention.

The intervention in the classes consisted of a classroom curriculum, which included seven sessions of approximately 30–40 minutes. Each session was delivered by a teacher or by mental health staff. After completion of the classroom curriculum materials, teachers were encouraged to hold weekly classroom meetings during which students could be helped to reflect on their behaviours. Parents were offered information through newsletters. Individual parents of students involved in bullying as either perpetrators or victims were given consultation. The complete BPYS programme ran over a period of three years. The first year was devoted to implementing the full curriculum and the following two years were intended to reinforce all the activities delivered in the first year.

Toronto Anti-bullying Programme

The Toronto anti-bullying programme was inspired by the OBPP (Pepler et al., 2004, p. 125). It was based on the understanding that bullying is a problem that extends far beyond the individual children; it involved the peer group and the teachers, as well as the parents of children (Pepler et al., 2004, p. 127). The programme included several preventive elements implemented at the school, parent and classroom levels, as well as additional work with specific students involved in bullying as perpetrators or victims.

The level of implementation of the programme varied across the intervention schools. However, in all intervention schools three critical elements were found: staff training, codes of behaviour and improved playground supervision. At the school level an emphasis was placed on developing a positive code of behaviour among students, engaging teachers and promoting positive playground interactions. At the parent level, information nights were held during which parents were informed about the problem of bullying in their school. Also, information about the programme and its objectives was sent home. At the classroom level, children were involved in developing classroom rules against bullying. Further classroom activities aimed to change students' attitudes and to promote healthy relationships among peers. At the individual level, children involved in bullying as perpetrators or victims received specialized intervention through consultation and though engaging their parents. Follow-up monitoring of these cases helped school authorities to establish that bullying incidents were terminated or discontinued.

Ecological Anti-bullying Programme

The Ecological Anti-bullying programme examined peer group and school environment processes 'utilizing a systemic interactional model with evaluations at each level of intervention' (Rahey & Craig, 2002, p. 283). The overall aim of the programme was the creation of a supportive and safe school environment in which firm limits against bullying were established. The specific goals of the programme included raising awareness of the problem of bullying, increasing empathy, encouraging peers to speak against bullying and formulating clear rules against bullying.

The 12-week programme was based on the 'Bully Proofing Your School' (BPYS) programme which was designed to increase the understanding of bullying and decrease the incidence of bullying (Rahey & Craig, 2002, p. 285). The programme elements included a psycho-educational component implemented within each classroom, a peer mediation component and specialized groups for children involved in bullying.

At the school-wide level, the psycho-educational programme was implemented by psychology students who received training sessions and manuals prior to intervention. Prior to the programme, at a school assembly the programme was introduced to students. The assembly signalled the formal beginning of the intervention. The classroom programmes involved interactive educational approaches such as role playing and puppet techniques. The topics addressed were bullying and victimization, conflict resolution, empathy, listening skills and individual differences (Rahey & Craig, 2002, p. 286).

Individual programmes for children involved in bullying were also part of the intervention. The relevant sessions consisted of social skills, listening, empathy training and supportive counselling. Each weekly session lasted 45 minutes. The programme also included intervention at the teacher level. Teacher programmes consisted of meetings with teachers to discuss bullying, intervention approaches, and student support for those directly involved in bullying. During the intervention, the programme coordinators met with principals and teachers to offer support.

Short Intensive Intervention in Czechoslovakia

The anti-bullying intervention in Czechoslovakia was inspired by the OBPP and borrowed elements from it, such as the Olweus videocassette on bullying (Rican et al., 1996, p. 399). The Olweus bullying questionnaire was used to measure several aspects of bullying within the schools. A peer nomination technique was also used to identify bully and victim scores. The relevant results from both measurement scales were presented to teachers in the intervention schools to increase awareness of the problem of bullying. The programme researchers discussed with the teachers 'possibilities of an individual approach to the bullies as well as to the victims' (Rican et al., 1996, p. 399).

As another intervention element, teachers were instructed to introduce relevant ethical aspects into the curriculum where possible: the ideal of knighthood was suggested for history classes and the ideal of consideration for the weak was introduced in sentences used for dictation and analysis (Rican et al., 1996, p. 400). Another element of the intervention involved the use of a method called 'class charter'. Specifically, children were asked to indicate how they would like their teachers and other classmates to behave towards them as well as how students should behave towards teachers and among themselves. The final aim of this classroom activity was the construction of a set of rules and principles, which was then signed by all pupils in the classroom and placed there in a visible position. Finally, the Olweus videocassette on bullying was shown to children and was used as a means of promoting the anti-bullying idea in the school.

Flemish Anti-bullying Programme

The Flemish anti-bullying intervention was inspired by the OBPP (Stevens et al., 2000, p. 198). Following the Norwegian approach, this programme introduced a school-based strategy that included several activities: a) for individual students involved in bullying (individual-level intervention); b) for parents and teachers (school-level intervention); and c) for the peer group (class-level intervention). Intervention schools (in the condition of Treatment with Support) received specific training sessions on the anti-bullying programme and individualized feedback on intervention strategies during the implementation process (Stevens et al., 2000, p. 200). Those in the 'Treatment without Support' condition received no additional help from the researchers.

The anti-bullying programme consisted of three modules. A manual with a video (entitled 'How was your day?') was also provided. The first module, targeting the school level, involved the formulation of an anti-bullying policy on behalf of the school staff. During the first module, information sessions were held for teachers, non-teaching staff and parents that aimed to increase awareness of bully/victim problems. The second module, targeting the class level, involved curriculum-based activities for the peer group. Active teaching methods such as modelling techniques and role-playing were used as a means of enhancing students' behavioural changes. The curriculum activities also involved social skills training to help students learn the appropriate way to intervene in bullying situations. Clear classroom rules against bullying were also set. The third module, targeting the individual level, consisted of direct work with children involved in bullying as either perpetrators or victims. This module was based on social learning theory (Stevens et al., 2000, p. 199).

Sheffield Anti-Bullying programme

The Sheffield anti-bullying initiative offered a marked variety of materials that teachers could use to address the problem of bullying. These materials were based on existing knowledge and ideas, but not on a systematic evaluation of the effects and relative success of different interventions (Smith, 1997, p. 194). The core intervention was based on a whole school policy on bullying (Smith, 1997, p. 195). Schools were given a choice of additional interventions covering: a) curriculum work (e.g. video, drama, literature, quality circles); b) playground interventions (e.g. surveillance, training lunchtime supervisors in recognising bullying, improving the playground environment); c) working with individuals and small groups (e.g. peer counselling, assertiveness training for victims, the Pikas method).

Curriculum-based strategies included a variety of materials and activities that aimed to increase children's awareness of the problem of bullying. A video entitled 'Sticks and Stones' could be used by teachers. The film showed interviews with students, a scenario depicting bullying episodes and clips from the operation of a bully-court (Smith et al., 2004, p. 102). The video came with a manual containing ideas on how to start a discussion, use drama and engage students in creative writing activities. To deal with racial issues another video was available, entitled 'White Lies', which specifically addressed issues of racial bullying. A drama, entitled 'Only playing Miss' aimed to address issues related to harassing behaviours. A novel, entitled 'The Heartstone Odyssey', gave teachers the chance to address through literature the issue of bullying. This was a story for primary students, which tackled the issues of racial harassment and bullying. The use of quality circles was also part of the curriculum-based anti-bullying strategies. They consisted of a group of students who met together to identify and address problems related to bullying, to find effective solutions that they then presented to the class teacher or senior management team (Smith et al., 2004, p. 103).

Other components of the Sheffield anti-bullying initiative involved individual work with children directly involved in bullying, peer counselling and increased playground surveillance. Peer counselling involved a 'listening line' for other students (Smith et al., 2004, p. 104): students formed small teams comprising two or three counsellors and one receptionist. Each team was directed by a supervising teacher; students never intervened in bullying situations themselves. Direct work with children involved in bullying as perpetrators was carried out though a method developed by Anatol Pikas, entitled 'Shared Concern', which was based on a structured script that could guide teachers' discussions with students involved in bullying. Making changes to playgrounds and training of lunchtime supervisors were also part of the intervention strategies.

The intervention programme did not indicate which and how many of these methods had to be used in order for the project to be successful. The interested reader can find however in several places the extent to

44

which the elements of the programme were implemented within each school (e.g. Eslea & Smith, 1998, p. 208; Smith et al., 2004, p. 101).

4.3 Other Experimental-Control Comparisons

Transtheoretical-based Tailored Anti-bullying Programme

This anti-bullying initiative involved 'transtheoretical-based tailored programmes that provided individualized and interactive computer interventions to populations of middle and high school students involved in bullying as bullies, victims and/or passive bystanders' (Evers et al., 2007, p. 398). The intervention involved only three 30-minute computer sessions during the school year for the students and a 10-page manual for staff and parents with optional activities. According to the programme designers, the transtheoretical model is 'a theory of behaviour change that applies particular change processes like decision-making and reinforcement to help individuals progress at particular stages of change' (Evers et al., 2007, p. 398).

Intervention materials included the 'Build Respect, Stop Bullying' programme, which is a multi-component, internet-based computer system (Evers et al., 2007, p. 402). Students initiated the programme by running a multimedia CD which brought them to the programme website. Students could use the programme by creating a login name based on personal information and a password. Once the students registered for the programme, logged in and consented to be involved in the intervention study, they were given instructions on how to proceed. This multi-media programme also included short movies (videos) of students giving testimonials about bullying (Evers et al., 2007, p. 403).

Other elements of the programme included: a) a 10-page family guide, sent to children's homes, which provided brief information about the multi-media programme and its relation to the anti-bullying initiative; and b) a 10-page staff guide, which included general information about bullying and how to support student change, classroom activities and information on how to work with parents. Teachers were not provided with any training.

Norwegian Anti-bullying Programme

This anti-bullying initiative was based on a pilot study conducted in primary schools in a town in the South of Norway. Based on the theoretical perspective of the programme, teachers' professional development is a crucial factor affecting the quality of school life for both school staff and students. Teachers are constantly called to deal with child problem behaviour. Thus, it was argued that 'investing' in teachers' professional development and helping teachers enhance their coping skills and tactics, could be very productive in reducing children's antisocial behaviour, including bullying. As Galloway and Roland (2004, p. 45) put it 'the implications for the argument presented above are that

attempts to reduce bullying can, and should, form an integral part of wider ranging attempts to improve the quality of teaching and learning. Teachers should perceive an anti-bullying initiative as assisting them in their core work, from which they derive their job satisfaction and for which they are rightly held accountable'.

The core element of the intervention within this programme was teacher training, which consisted of four in-service days over a nine-month period. A handout summarizing the content of the course was distributed to teachers in each meeting. In addition, the programme included 15 two-hour peer supervision sessions, the aim of which was to give teachers the opportunity to discuss the practical implications of the theoretical concepts introduced in the in-service days.

SAVE

The SAVE anti-bullying programme was based on an educational model which placed emphasis on an ecological approach to analyzing bullying and violence in general (Ortega et al., 2004, p. 169). The model proposed the design of an educational project regarding interpersonal relationships based on the dimension of convivencia (coexistence) and on the dimension of activity. The theoretical notion of convivencia signalled the spirit of solidarity, fraternity, cooperation, harmony and a desire for mutual understanding, the desire to get on well with others and the resolution of conflict through dialogue and other non-violent ways (Ortega et al., 2004, p. 169).

Three processes were relevant to the design of the SAVE programme, namely: a) management of the social environment and of the ways in which children interact; b) the specific method of instructive action; and c) activities that were geared towards feelings and values of education (Ortega et al., 2004, p. 170).

The programme was based on the principle of democratic management of interpersonal relationships in which teachers, without losing their authority, gave students the opportunity to have an active and participative role in decision-making. Co-operative group work was another element of the intervention. The programme included direct intervention work with students at risk or involved in bullying. For these children a variety of additional preventive measures were offered such as quality circles, conflict mediation, peer support, the Pikas Method, assertiveness and empathy training (Ortega et al., 2004, p. 172). Finally, the programme included training sessions for teachers and work with families but the extent to which these were implemented varied across schools (Ortega et al., 2004, p. 176).

Kia Kaha

Kia Kaha was designed as an anti-bullying programme, but it also met the requirements of two essential areas within the curriculum framework: social sciences and health/physical well-being (Raskauskas, 2007,

p. 10). The programme involved a whole-school approach to tackling bullying and victimization. In the Maori language the word 'kia kaha' means to stand strong, which is why this name was used 'to represent the need for the whole-school community to stand strong to prevent bullying' (Raskauskas, 2007, p. 9). The programme covered issues such as peer relationships, identifying and dealing with bullying, making personal choices, developing feelings of self-worth, respecting differences and working co-operatively to build a safe classroom environment.

The Kia Kaha curriculum used several resources, including a teachers' guide, with an overview of the programme, instructions on how to plan and implement the lessons, a video and information to be sent home to parents. The video included five bullying situations that provided the basis for discussing both on what was happening and what could be done. Students were taught to take steps to defuse bullying situations: Stop, Think, Consider Options, Act, Follow-up. The student and teacher components were delivered through the regular classroom curriculum.

Police Education Officers (PEOs) are trained as educators and are involved in youth education in New Zealand. PEOs visited schools and introduced the programmes offered by the police, including Kia Kaha. PEOs introduced and tried to convince principals to use the whole-school approach in their schools. They also trained the teachers in the programme, hosted parent nights and taught up to four lessons of the curriculum.

4.4 Age-Cohort Designs

Respect Programme

Respect, previously running under the name Connect, was a programme that aimed to tackle different types of child problem behaviour, such as disobedience, off-task behaviour, bullying and aggression. The programme was implemented in both primary and secondary schools. The Respect programme worked on the system level by including all school personnel, pupils and parents in an attempt 'to improve the quality of the school at the individual, at the class and at the school levels' (Ertesvag & Vaaland, 2007, p. 714). The programme was based on four basic principles (Ertesvag & Vaaland, 2007, p. 716): a) adults were expected to act as sources of authority. This involved an authoritative approach that aimed to create a warm and caring environment; b) the programme was broad-based involving all persons in the school and intervening at all levels (individual, classroom and school level); c) adults should act consistently in order to ensure that they made an impact on student behaviour; and d) the programme was based on the notion of continuity, which implied a long-term commitment to the previous three principles.

Within the programme framework, teachers and school management staff participated in series of seminars. The staff training sessions introduced the basic principles of the programme and practical approaches to

the prevention of child problem behaviour along with some illustrative examples. A two-day seminar for schools' management and other key school personnel was run in advance of the implementation period. Within each school, a one-day workshop took place with the main goal of ensuring that the school staff understood their own school's implementation process. Other short-term training sessions took place during the intervention period (Ertesvag & Vaaland, 2007, p. 718). Within each school, a project group shared day-to-day responsibility for implementing the programme. Among the different intervention schools, a network was established with the aim of discussing knowledge, experiences and challenges related to programme implementation.

Finally, there were four main strategies in the implementation of the programme, namely a) having a whole school approach to the problem of bullying; b) using an authoritative approach to classroom leadership; c) choosing the right timing of the intervention and, finally, d) commitment to the principles of the programme.

Olweus Bullying Prevention Programme (OBPP)

The OBPP was a multi-level programme aiming at targeting the individual, the school, the classroom and the community level. Apart from marked mass-media publicity, the program started with a one-day school conference during which the problem of bullying was addressed between school staff, students and parents. This signalled the formal commencement of the intervention. Two different types of materials were produced: a handbook or manual for teachers (entitled 'Olweus' core programme against bullying and antisocial behaviour') and a folder with information for parents and families. The programme also included: a) CD-programme that was used for assessing and analysing the data obtained at the pretest period, so that school-specific interventions could then be implemented; b) a video on bullying; c) the Revised Olweus Bully/Victim Questionnaire and d) the book 'Bullying at school: what we know and what we can do'.

The anti-bullying measures mainly targeted three different levels of intervention: the school, the classroom and the individual. At the school level, the intervention included:

- Meetings among teachers to discuss ways of improving peer-relations; staff discussion groups.
- Parent/teacher meetings to discuss the issue of bullying.
- Increased supervision during recess and lunchtime.
- Improvement of playground facilities so that children have better places to play during recess time.
- A questionnaire survey.
- The formation of a coordinating group.

At the classroom level the intervention included:
- Students were given information about the issue of bullying and were actively involved in devising class rules against bullying.
- Classroom activities for students included role-playing situations that could help students learn how to deal better with bullying.
- Class rules against bullying.
- Class meetings with students.
- Meetings with the parents of the class.

At the individual level the intervention included:
- Talks with bullies and their parents and enforcement of non-hostile, non-physical sanctions.
- Talks with victims, providing support and providing assertiveness skills training to help them learn how to successfully deal with bullying; also, talks with the parents of victims.
- Talks with non-involved children to make them become effective helpers.
- Development of individual intervention plans.

An interesting feature of the OBPP is that it offered guided information about what schools should do at both the intervention and the maintenance period. 'The Olweus program demands significant commitment from the school during the "introductory period" which covers a period of about 18 months. Later the methodology acquired by the staff and the routines decided by the school may be maintained using less resources … Yet, even for the maintenance period, the programme offers a point by point description of what the school should do to continue its work against bullying in accordance with Olweus methodology (Olweus, 2004c, p. 1). Also, at the school level training was offered to the whole school staff, with additional training provided to the coordinators and key personnel. These were responsible for coordinating the overall anti-bullying initiative in their school. The programme also included cooperation among experts and teachers(e.g. psychologists) who worked with children involved in bullying.

Finnish Anti-Bullying programme
The Finnish anti-bullying programme used a participant role approach to bullying (Salmivalli et al., 2005, p. 467). In agreement with this approach to bullying, three steps in curriculum-based preventive work involved: a) raising awareness of the issue of bullying; b) encouraging students' self-reflection on their own behaviour; and c) commitment to anti-bullying behaviours (Salmivalli et al., 2007, pp. 467–468).

The core element of the intervention involved a one-year teacher training. This training was provided in four sessions/meetings carried out throughout the school year. During the training teachers were given feedback about the situation in their own classes (based on the results of

the pre-intervention data) and information about alternative methods of intervening to prevent bullying at the individual, class and school level. Also, teachers were offered advice about individual cases that they found difficult to deal with. During the training, teachers were provided with anti-bullying materials that they could use along with the formal curriculum activities or materials. These materials involved, for example, overhead transparencies and suggestions for discussions as well as role-playing exercises developed by a group of drama teachers, 'Theatre in Education'. For interventions at the individual level, teachers were presented with several methods that they could use individually with specific children involved in bullying, such as the method of 'Shared Concern', the 'No Blame' approach and the Farsta method (Salmivalli et al., 2007, p. 471). Regardless of the method used, the role of systematic follow-ups after the initial work was strongly emphasized. At the school level, teachers were encouraged to take the anti-bullying message to their school and to promote the process of developing a whole-school anti-bullying policy.

5. Analysis of Included Evaluations

5.1 Key Results

Table 6. Key Results of Evaluations.

Project Information	Bullying	Victimization
Randomized Experiments		
Baldry & Farrington (2004)	**Younger** EB: M 1.69 (2.15) 58 EA: M 2.69 (3.31) 26 CB: M 1.54 (2.20) 57 CA: M 1.57 (2.20) 72 **Older** EB: M 2.54 (3.59) 63 EA: M 2.31 (3.07) 99 CB: M 2.11 (2.44) 46 CA: M 3.39 (3.99) 36	**Younger** EB: M 3.66 (4.36) 59 EA: M 2.24 (3.50) 29 CB: M 3.25 (3.50) 56 CA: M 1.85 (2.62) 71 **Older** EB: M 3.64 (4.89) 64 EA: M 2.31 (3.89) 99 CB: M 1.84 (2.35) 44 CA: M 2.79 (2.48) 38
Cross et al. (2004)	EB: 13.0% (1038) EA1: 16.4% (992) CB: 15.1% (919) CA1: 15.2% (875)	EB: 16.2% (982) EA1: 13.2% (990) EA2: 14.7% (869) CB: 15.7% (860) CA1: 13.9% (880) CA2: 14.6% (792)
De Rosier (2004); De Rosier & Marcus (2005)	EB: M .09 (1.08) 187 EA1: M .15 (1.22) 187 EA2: M .15 (1.32) 134 CB: M .13 (1.18) 194 CA1: M .07 (1.13) 194 CA2: M .14 (1.05) 140	EB: M .31 (1.10) 187 EA1: M .38 (1.16) 187 EA2: M .31 (1.12) 134 CB: M .27 (1.06) 194 CA1: M .26 (1.12) 194 CA2: M .42 (1.22) 140
Fekkes et al. (2006)	EB: 5.1% (1101) EA1: 7.9% (1098) EA2: 6.6% (686) CB: 5.1% (1110) CA1: 8.9% (1108) CA2: 7.3% (895)	EB: 17.7% (1106) EA1: 15.5% (1104) EA2: 14.0% (688) CB: 14.6% (1115) CA1: 17.3% (1112) CA2: 11.9% (897)
Frey et al. (2005)	**Direct** EB: M .46 (.59) 563 EA: M .48 (.62) 457? CB: M .56 (.66) 563 CA: M .62 (.71) 457? **Indirect** EB: M .88 (.72) 563 EA: M .90 (.74) 457? CB: M .94 (.73) 563 CA: M .96 (.83) 457?	EB: M 1.01 (.79) 563 EA: M .90 (.82) 457? CB: M 1.07 (.82) 563 CA: M 1.01 (.83) 457?

Notes: E = Experimental, C = Control, B = Before, A = After (A1, A2, A3, A4 = post tests 1, 2, 3, 4). M = Mean (SD in parentheses, followed by N). LOR = Logarithm of odds ratio, SE = Standard error. E1, E2, E3, C1, C2, C3 = 3 schools in experimental or control conditions. F1, F2, F3 = Fall in 3 years. S1, S2, S3 = Spring in 3 years. L, H = Low, high implementation. NA = Not available. ET = Treatment with support. EW = Treatment without support.

Hunt (2007)	**Bullying Alone** EB: M 1.30 (.60) 152 EA: M 1.17 (.46) 111 CB: M 1.30 (.66) 248 CA: M 1.31 (.64) 207 **Bullying in Group** EB: M 1.47 (.70) 152 EA: M 1.39 (.72) 111 CB: M 1.36 (.75) 248 CA: M 1.41 (.76) 207	EB: M 1.86 (1.21) 152 EA: M 1.53 (1.12) 111 CB: M 1.71 (1.05) 248 CA: M 1.52 (1.08) 207
Jenson & Dieterich (2007)	LOR = .161, SE = .280 (N = 667)	LOR = .491, SE = .286 (N = 668)
Rosenbluth et al. (2004)	EB: 10.6% (929) EA: 17.0% (741?) CB: 11.2% (834) CA: 17.8% (665?)	EB: 40.8% (929) EA: 36.7% (741?) CB: 47.5% (834) CA: 34.7% (665?)
Salmivalli et al. (2007)	EB: 5.8% (3894) EA: 3.4 % (3404) CB: 6.0 % (3406) CA: 5.1 % (2499)	EB: 16.1 % (3898) EA: 9.0 % (3408) CB: 16.4 % (3403) CA: 14.4 % (2501)
Before-After, Experimental-Control Comparisons		
Andreou et al. (2007)	EB: M 10.43 (3.40) 248 EA1: M 10.06 (3.80) 246 EA2: M 10.45 (4.09) 234 CB: M 9.87 (3.65) 206 CA1: M 10.85 (3.72) 207 CA2: M 10.81 (3.94) 203	EB: M 10.74 (3.61) 248 EA1: M 10.63 (3.90) 248 EA2: M 10.21 (3.49) 235 CB: M 10.62 (3.78) 206 CA1: M 11.17 (3.68) 206 CA2: M 11.03 (3.89) 201
Bauer et al. (2007)	NA	**Physical** EB: 13.8% (4531) EA: 14.6% (4419) CB: 16.3% (1373) CA: 17.5% (1448) **Relational** EB: 24.8% (4607) EA: 24.7% (4480) CB: 30.4% (1408) CA: 30.2% (1456)
Ciucci & Smorti (1998)	EB 46.7% (167) EA 49.7% (169) CB 43.9% (140) CA 51.4% (141)	EB 44.9%(167) EA 50.3% (169) CB 37.4% (140) CA 47.4% (141)
Melton et al (1998)	EB 24% (3904) EA 20% (3827) CB 19% (2485) CA 22% (2436)	EB 25% (3904) EA 19% (3827) CB 24% (2485) CA 19% (2436)

Notes: E = Experimental, C = Control, B = Before, A = After (A1, A2, A3, A4 = post tests 1, 2, 3, 4). M = Mean (SD in parentheses, followed by N). LOR = Logarithm of odds ratio, SE = Standard error. E1, E2, E3, C1, C2, C3 = 3 schools in experimental or control conditions. F1, F2, F3 = Fall in 3 years. S1, S2, S3 = Spring in 3 years. L, H = Low, high implementation. NA = Not available. ET = Treatment with support. EW = Treatment without support.

Menard et al. (2008)	**Elementary School** **Physical** B: r = -.063 (708) A1 r = .044 (636) A2: r = .102 (708) A3: r = .116 (735) A4: r = .047 (710) **Relational** B: r = -.103 (708) A1: r = -.066 (636) A2: r = .080 (708) A3: r = .134 (735) A4: r = .052 (710) **Middle School** **Physical** B: r = .040 (280) A1: r = -.128 (306) A2: r = .009 (339) A3: r = .080 (354) A4: r = .049 (348) **Relational** B: r = .019 (280) A1: r = -.009 (306) A2: r = .092 (339) A3: r = .094 (354) A4: r = .092 (348)	**Elementary School** **Physical** B: r = .005 (708) A1: r = -.009 (636) A2: r = .052 (708) A3: r = .109 (735) A4: r = .101 (710) **Relational** B: r = -.027 (708) A1: r = -.028 (636) A2: r = .109 (708) A3: r = .051 (735) A4: r = .067 (710) **Middle School** **Physical** B: r = .060 (280) A1: r = .032 (306) A2: r = -.022 (339) A3: r = -.031 (354) A4: r = .040 (348) **Relational** B: r = .014 (280) A1: r = .036 (306) A2: r = -.053 (339) A3: r = -.027 (354) A4: r = -.003 (348)
Olweus/Bergen 2	EB 5.6% (1278) EA 4.4% (1296) CB 4.1% (1111) CA 5.6% (1168)	EB 12.7% (1297) EA 9.7% (1320) CB 10.6% (1117) CA 11.1% (1179)
Pepler et al. (2004)	E2S1: 32% (300) E2F2: 27% (240) E2F1: 26% (300) E2S2: 20% (240) E2S1: 32% (300) E2F3: 16% (163) E2F1: 26% (300) E2S3: 14% (163) C3F2: 23% (303) C3S2: 23% (303) E3F2: 23% (303) E3S3: 14% (289) E3S2: 23% (303) E3F3: 13% (289) C2F1: 26% (300) C2S1: 32% (300)	E2S1: 42% (300) E2F2: 57% (240) E2F1: 52% (300) E2S2: 48% (240) E2S1: 42% (300) E2F3: 41% (163) E2F1: 52% (300) E2S3: 38% (163) C3F2: 41% (303) C3S2: 39% (303) E3F2: 41% (303) E3S3: 28% (289) E3S2: 39% (303) E3F3: 28% (289) C2F1: 52% (300) C2S1: 42% (300)

Notes: E = Experimental, C = Control, B = Before, A = After (A1, A2, A3, A4 = post tests 1, 2, 3, 4). M = Mean (SD in parentheses, followed by N). LOR = Logarithm of odds ratio, SE = Standard error. E1, E2, E3, C1, C2, C3 = 3 schools in experimental or control conditions. F1, F2, F3 = Fall in 3 years. S1, S2, S3 = Spring in 3 years. L, H = Low, high implementation. NA = Not available. ET = Treatment with support. EW = Treatment without support.

Rahey & Craig (2002)	**Junior Children** EB: M .206 (.570) 125 EA: M .254 (.779) 125 CB: M .105 (.526) 67 CA: M .224 (.487) 67 **Senior Children** EB: M .425 (.895) 138 EA: M .521 (.916) 138 CB: M .264 (.503) 176 CA: M .391 (.714) 176	**Junior Children** EB: M 1.22 (1.34) 125 EA: M .783 (1.19) 125 CB: M 1.09 (1.29) 67 CA: M .881 (1.33) 67 **Senior Children** EB: M .440 (.863) 138 EA: M .890 (1.29) 138 CB: M .563 (1.03) 176 CA: M .685 (1.11) 176
Rican et al. (1996)	EB: 19.0% (100) EA: 7.1% (98) CB: 13.3% (98) CA: 11.2% (98)	EB: 18.0% (100) EA: 7.1% (98) CB: 16.3% (98) CA: 14.3% (98)
Stevens et al. (2000)	**Primary School** ETB: M 1.02 (.11) 120 ETA1: M 1.02 (.11) 120 ETA2: M 1.02 (.10) 120 EWB: M 1.12 (.14) 100 EWA1: M 1.12 (.13) 100 EWA2: M 1.10 (.14) 100 CB: M 1.05 (.12) 70 CA1: M 1.07 (.12) 70 CA2: M 1.10 (.15) 70 **Secondary School** ETB: M 0.99 (.09) 226 ETA1: M 1.02 (.11) 226 ETA2:M 1.00 (.10) 226 EWB: M 1.02 (.11) 185 EWA1: M 1.02 (.10) 185 AWA2: M 1.01 (.12) 185 CB: M 1.02 (.12) 115 CA1: M 1.03 (.13) 115 CA2: M 1.02 (.12) 115	**Primary School** ETB: M 1.09 (.13) 120 ETA1: M 1.08 (.15) 120 ETA2: M 1.06 (.13) 120 EWB: M 1.16 (.15) 100 EWA1: M 1.16 (.14) 100 EWA2: M 1.10 (.15) 100 CB: M 1.14 (.15) 70 CA1: M 1.14 (.13) 70 CA2: M 1.11 (.15) 70 **Secondary School** ETB: M 1.02 (.11) 226 ETA1: M 1.05 (.14) 226 ETA2: M 1.02 (.12) 226 EWB: M 1.04 (.12) 185 EWA1: M 1.03 (.11) 185 EWA2: M 1.03 (.13) 185 CB: M 1.03 (.12) 115 CA1: M 1.04 (.13) 115 CA2: M 1.02 (.13) 115
Whitney et al. (1994)	**Primary** EB: M 1.50 (.83) 2519 EA: M 1.45 (.77) 2370 CB: M 1.57 (.92) 91 CA: M 1.84 (.97) 99	**Primary** EB: M 1.95 (1.13) 2523 EA: M 1.90 (1.09) 2380 CB: M 1.88 (1.10) 91 CA: M 1.97 (1.11) 99
Other Experimental-Control Comparisons		
Evers et al. (2007)	**Middle School** E: 68% B (202) C: 81% B (377) **High School** E: 62% B (359) C: 79% B (221)	**Middle School** E: 65% (202) C: 83% (377) **High School** E: 63% (359) C: 78% (221)
Galloway & Roland (2004)	E: M .34 (.60?) 672 C: M .40 (.60?) 475	E: M .87 (.78?) 675 C: M 1.07 (.78?) 475

Notes: E = Experimental, C = Control, B = Before, A = After (A1, A2, A3, A4 = post tests 1, 2, 3, 4). M = Mean (SD in parentheses, followed by N). LOR = Logarithm of odds ratio, SE = Standard error. E1, E2, E3, C1, C2, C3 = 3 schools in experimental or control conditions. F1, F2, F3 = Fall in 3 years. S1, S2, S3 = Spring in 3 years. L, H = Low, high implementation. NA = Not available. ET = Treatment with support. EW = Treatment without support.

Ortega et al. (2004)	E: 4.1% (910) C: 6.5% (440)	E: 4.2% (910) C: 8.5% (440)
Raskauskas (2007)	E: M .45 (.75) 1539? C: M .53 (.85) 1542?	E: M .84 (1.10) 1554 C: M 1.03 (1.18) 1557
Age-Cohort Designs		
Ertesvag & Vaaland (2007)	**Grade 5** B: M .29 (.32) 118 A1: M .31 (.43) 126 A2: M .21 (.33) 151 A3: M .17 (.38) 143 **Grade 6** B: M .36 (.38) 152 A1: M .28 (.43) 129 A2: M .17 (.25) 130 A3: M .21 (.30) 140 **Grade 7** B: M .31 (.32) 147 A1: M .32 (.39) 160 A2: M .30 (.40) 134 A3: M .15 (.28) 140 **Grade 8** B: M .32 (.49) 123 A1: M .25 (.33) 128 A2: M .41 (.60) 112 A3: M .25 (.49) 123 **Grade 9** B: M .34 (.55) 95 A1: M .32 (.48) 128 A2: M .35 (.59) 112 A3: M .33 (.49) 122 **Grade 10** B: M .35 (.49) 112 A1: M .41 (.55) 99 A2: M .38 (.60) 149 A3: M .31 (.56) 124	**Grade 5** B: M .54 (.49) 118 A1: M .53 (.53) 126 A2: M .43 (.48) 151 A3: M .44 (.54) 143 **Grade 6** B: M .46 (.46) 152 A1: M .50 (.57) 129 A2: M .38 (.47) 130 A3: M .39 (.46) 140 **Grade 7** B: M .44 (.51) 147 A1: M .39 (.52) 160 A2: M .44 (.52) 134 A3: M .24 (.46) 140 **Grade 8** B: M .30 (.57) 123 A1: M .21 (.34) 128 A2: M .57 (.74) 112 A3: M .32 (.40) 123 **Grade 9** B: M .26 (.39) 95 A1: M .26 (.46) 128 A2: M .36 (.55) 112 A3: M .44 (.55) 122 **Grade 10** B: M .35 (.60) 112 A1: M .27 (.34) 99 A2: M .24 (.40) 149 A3: M .24 (.34) 124
Olweus/Bergen 1	**Grades 5-7** B 7.28% (1689) A1 5.02% (1663) **Grades 6-7** B 7.35% (1294) A2 3.60% (1103)	**Grades 5-7** B 9.98% (1874) A1 3.78% (1691) **Grades 6-7** B 9.92% (1297) A2 3.55% (1115)
Olweus/Oslo 1	**Grades 5-7** B 6.4% (874) A 3.1% (983)	**Grades 5-7** B 14.4% (882) A 8.5% (986)
Olweus/New National	**Grades 5-7** B 5.7% (8370) A1 3.6% (8295) **Grades 6-7** B 5.1% (8222) A2 2.6% (8473)	**Grades 5-7** B 15.2% (8387) A1 10.2% (8299) **Grades 6-7** B 13.2% (8238) A2 8.7% (8483)

Notes: E = Experimental, C = Control, B = Before, A = After (A1, A2, A3, A4 = post tests 1, 2, 3, 4). M = Mean (SD in parentheses, followed by N). LOR = Logarithm of odds ratio, SE = Standard error. E1, E2, E3, C1, C2, C3 = 3 schools in experimental or control conditions. F1, F2, F3 = Fall in 3 years. S1, S2, S3 = Spring in 3 years. L, H = Low, high implementation. NA = Not available. ET = Treatment with support. EW = Treatment without support.

| Olweus/Oslo 2 | Grades 4-7
B 5.5% (2682)
A1 2.8% (3077)
A2 2.3% (3022)
A3 2.8% (2535)
A4 2.7% (2834)
Grades 8-10
B 6.2% (1445)
A1 5.7% (1449)
A2 4.1% (1526) | Grades 4-7
B 14% (2695)
A1 9.8% (3077)
A2 8.8% (3026)
A3 8% (2538)
A4 8.4% (2967)
Grades 8-10
B 7.1% (1452)
A1 6.8% (1462)
A2 5.2% (1532) |
| Salmivalli et al.
(2005) | Grade 4
B: M .15 (.36) 389
L: M .08 (.26) 247
H: M .03 (.18) 125
Grade 5
B: M .11 (.32) 417
L: M .12 (.32) 258
H: M .07 (.25) 131 | Grade 4
B: M .14 (.34) 389
L: M .10 (.29) 247
H: M .06 (.24) 125
Grade 5
B: M .13 (.33) 417
L: M .11 (.32) 258
H: M .07 (.26) 131 |

Notes: E = Experimental, C = Control, B = Before, A = After (A1, A2, A3, A4 = post tests 1, 2, 3, 4). M = Mean (SD in parentheses, followed by N). LOR = Logarithm of odds ratio, SE = Standard error. E1, E2, E3, C1, C2, C3 = 3 schools in experimental or control conditions. F1, F2, F3 = Fall in 3 years. S1, S2, S3 = Spring in 3 years. L, H = Low, high implementation. NA = Not available. ET = Treatment with support. EW = Treatment without support.

Table 6 summarizes key results of the included evaluations from the 30 different programmes. Wherever possible, this table shows either (a) prevalence (of bullies or victims) and the number on which this is based, or (b) mean score (on bullying or victimization scales) and the associated standard deviation and number on which this is based. Where the desired information was not reported, we requested it from the researchers, but they sometimes did not reply. In the rare cases where both prevalence and means were provided, we chose to show prevalence. The only exception was Raskauskas (2007), who provided prevalence only for victimization but mean scores for both bullying and victimization. In this case, in the interests of showing comparable data on bullying and victimization, we reported the mean scores.

In most cases, we had no choice of what prevalence figure to report. Very few researchers showed several categories of bullying or victimization (e.g. never, a few times, about once a fortnight, almost once a week, more than once a week; see Raskauskas, 2007, p. 20). If they had, we could perhaps have used the area under the ROC curve as our effect size measure (see e.g. Farrington, Jolliffe & Johnstone, 2008).

Where we could choose which prevalence figure to report, we chose the prevalence of bullying (or victimization) more than once or twice, because the definition of bullying specifies repeated acts. The criterion recommended by Olweus (1991) was "2–3 times a month or more". However, we did not set the criterion high if this produced a low prevalence, because it would then be difficult to detect any effect.

For example, Cross et al. (2004, p. 202) showed figures for "almost every day", "once every 2–3 weeks", "once or twice" and "not at all".

For victimization, our criterion was "once every 2–3 weeks or more often". For bullying, we used "ever bullied" because the criterion of "once every 2–3 weeks or more often" yielded prevalences no greater than 5%. However, we did not show prevalences of bullying for the second follow-up (EA2, CA2 in Table 6) because the published figures seemed clearly incorrect.

We followed the researchers in the way they split up their results for analysis. Baldry and Farrington (2004) presented results separately for younger (age 11–12) and older (age 13-14) children; Frey et al. (2007) presented results separately for direct and indirect bullying; Evers et al. (2007), Menard et al. (2008) and Stevens et al. (2006) presented results separately for different categories of schools; Menard et al. (2008) also presented results separately for physical and relational bullying; Ertesvag and Vaaland (2007) and Salmivalli et al. (2005) presented results separately for different grades; and Salmivalli et al. (2005) and Stevens et al. (2006) presented results separately for different implementation conditions.

As far as possible, we show prevalence (or means) for the experimental condition before and after the intervention (EB, EA) and the control condition before and after the intervention (CB, CA). Where there are several posttests, we show results obtained in all of these. Jenson and Dieterich (2007) did not report prevalence or means but reported coefficients (logarithms of odds ratios) in logistic regression models. Menard et al. (2008) reported phi correlations between experimental/control and bully/nonbully (or victim/nonvictim). Where question marks are shown after numbers, we have estimated them based on information provided by the researchers.

The most problematic numbers in Table 6 are for the Pepler et al. (2004) evaluation. This had a complex design. In year 1 (1992–93), school 1 received the anti-bullying programme and school 2 served as a control. In year 2 (1993–94), school 1 continued to receive the programme, school 2 also received the programme, and school 3 served as a control. In year 3 (1994–95), all three schools received the programme. Self-report measures of bullying and victimization (in the previous two months) were taken in the fall and spring of each year.

In analysing the data, we wanted to take advantage of both the experimental-control comparison and the before and after measures, because the combination of these designs is stronger than either alone. We could do this by the judicious choice of comparison schools and assessment times. For example, for school 2, fall of year 1 was before and spring of year 2 was after the intervention. An appropriate comparison would be fall of year 2 and spring of year 2 for school 3, both of which were before any intervention. Therefore, school 3 could be regarded as a control while school 2 was regarded as an experimental school for this comparison. In Table 6, spring and fall in an experimental school (before and after the intervention) are always compared with spring and

fall in a control school (with no intervention). We should, however, point out that Pepler et al. (2004, pp. 129–130) stated that:

"Even though no official interventions were implemented, the process of change appears to have started in School B [2] and School C [3] during the assessment-only phase. Therefore, our data analyses were conducted within school rather than between the intervention and control schools".

In light of this, our effect size estimates for this programme may be conservative.

For Rosenbluth et al. (2004), we only show one follow-up period (at the end of the semester, immediately after the programme) because only three of the six schools provided later follow-up data. For Hunt (2007), our figures are based on correspondence with Caroline Hunt where she indicated that her published victimization figures (p.24) were scored in the opposite direction. We have reversed the direction of scoring in Table 6. For Salmivalli et al. (2007), we only show the second follow-up, because this was done at the same time of the year as the before measure. We are concerned to minimize seasonal effects on bullying and victimization. We are very grateful to Christina Salmivalli for giving us preliminary results from this evaluation.

For Rahey and Craig (2002), we used questions about bullying in the previous week, based on correspondence with Leila Rahey. The Stevens et al. (2000) data appear implausible; bullying and victimization were measured using 8-point scales, but all the means are remarkably close to 1. For the present, until we can obtain clarification, we have not included their data in our meta-analysis. For Whitney et al. (1994), Mike Eslea kindly provided raw data, but unfortunately the data about the control secondary schools was missing. Therefore, we show data only for the primary schools. Evers et al. (2007) was a before-after, experimental-control design, but unfortunately they only reported data on how many of the bullies (or victims) at the pretest continued to be bullies (or victims) at the posttest. We have therefore classified this among the other experimental-control comparisons. We have asked the authors to provide the prevalence of bullying and victimization before and after in experimental and control conditions, but (as of mid-July 2008) we have not yet received this information.

5.2 Analysis of Effect Sizes

Table 7 shows the analysis of effect sizes for bullying. The measure of effect size is the weighted mean odds ratio (OR) with its associated 95% confidence interval (CI). Where the CI includes the chance value of 1.0, the OR is not statistically significant. The Z-value (based on a unit normal distribution) measures the statistical significance more accurately; Z-values greater than 1.96 or less than -1.96 are statistically sig-

nificant. The calculation of the OR and its associated CI are explained in the Technical Appendix.

Table 7. Effect Sizes for Bullying.

Project	OR	CI	Z	P
Randomized Experiments				
Baldry & Farrington (2004)	1.14	0.51 – 2.58	0.32	ns
Cross et al. (2004)	0.77	0.51 -- 1.15	-1.28	ns
De Rosier (2004)	0.87	0.63 – 1.21	-0.82	ns
Fekkes et al. (2006)	1.12	0.74 – 1.69	0.53	ns
Frey et al. (2005)	1.04	0.81 – 1.34	0.31	ns
Hunt (2007)	1.46	0.93 – 2.28	1.66	.097
Jenson & Dieterich (2007)	1.17	0.57 – 2.41	0.44	ns
Rosenbluth et al. (2004)	0.99	0.63 – 1.58	-0.03	ns
Salmivalli et al. (2007)	1.47	1.02 – 2.13	2.07	.038
Weighted mean (FE)	1.07	0.93 – 1.22	0.94	Ns
Before-After, Experimental-Control				
Andreou et al. (2007)	1.75	1.20 – 2.57	2.87	.004
Ciucci & Smorti (1998)	1.20	0.58 – 2.47	0.49	ns
Melton et al. (1998)	1.52	1.24 – 1.85	4.10	.0001
Menard et al. (2008)	1.64	1.37 – 1.96	5.32	.0001
Olweus/Bergen 2	1.79	0.98 – 3.26	1.90	.057
Pepler et al. (2004)	1.69	1.22 – 2.35	3.12	.002
Rahey & Craig (2002)	1.19	0.70 – 1.99	0.64	ns
Rican et al. (1996)	2.52	0.60 – 10.52	1.27	ns
Whitney et al. (1994)	2.12	1.17 – 3.87	2.48	.013
Weighted mean (FE)	1.60	1.44 – 1.79	8.43	.0001
Other Experimental-control				
Evers et al. (2007)	2.15	1.50 – 3.09	4.15	.001
Galloway & Roland (2004)	1.20	0.91 – 1.59	1.27	ns
Ortega et al. (2004)	1.63	0.84 – 3.14	1.45	ns
Raskauskas (2007)	1.20	1.01 – 1.42	2.11	.035
Weighted mean (FE)	1.31	1.15 – 1.49	4.04	.0001
Weighted mean (RE)	1.43	1.09 – 1.88	2.55	.011
Age-Cohort Designs				
Ertesvag & Vaaland (2004)	1.34	1.13 – 1.58	3.35	.0008
Olweus/Bergen 1	1.69	1.25 – 2.28	3.43	.0006
Olweus/Oslo1	2.14	1.18 – 3.87	2.51	.012
Olweus/New National	1.78	1.54 – 2.06	7.81	.0001
Olweus/Oslo2	1.75	1.35 – 2.26	4.27	.0001
Salmivalli et al. (2005)	1.31	1.07 – 1.61	2.56	.010
Weighted mean (FE)	1.57	1.44 – 1.71	10.19	.0001
Weighted mean (RE)	1.56	1.34 – 1.81	5.75	.0001
Weighted mean (FE)	1.43	1.35 – 1.51	12.76	.0001
Weighted mean (RE)	1.41	1.28 – 1.55	7.08	.0001

Note: OR = Odds Ratio; CI = Confidence Interval; FE = Fixed Effects; RE = Random Effects.

In general, results obtained with shorter follow-up periods were combined with results obtained with longer follow-up periods to produce the OR and the CI. However, in the case of Olweus/ Oslo 2, where there were four follow-up assessments for grades 4–7 but only two follow-up assessments for grades 8–10, the OR was based on only the two

common follow-up assessments. With age-cohort designs, the before assessment was regarded as the control condition and the after assessment was regarded as the experimental condition. In general, only one short and one long follow-up assessment was analysed. For Ertesvag and Vaaland (2007), the shortest (A1) and longest (A3) follow-up assessments were analysed, and results obtained in all six grades were combined. For Pepler et al. (2004), the first four experimental comparisons (e.g. E2S1-E2F2) were each compared with the first control comparison (C3F2–C3S2) because it was considered that these were the most valid comparisons. As in all other cases, all four ORs were combined into a single OR.

Only one of the nine randomized experiments (Salmivalli et al., 2007) found a significant effect of the programme on bullying, although one other evaluation (Hunt, 2007) reported a near-significant effect. Overall, the nine randomized experiments yielded a weighted mean OR of 1.07, indicating a very small and non-significant effect of these programmes on bullying. In contrast, five of the nine evaluations with before-after/experimental-control designs found a significant effect, and one other (Olweus/Bergen 2) reported a near-significant result. Overall, these nine studies yielded a large weighted mean OR of 1.60 (p < .0001). Fixed effects (FE) models were used in both cases because the effect sizes were not significantly heterogeneous.

Two of the four other experimental-control comparisons found significant effects on bullying, and the weighted mean OR for all four studies varied between 1.31 (FE) and 1.43 (Random Effects model or RE). All six age-cohort designs yielded significant effects, with weighted mean ORs of 1.57 (FE) and 1.56 (RE). Over all 28 studies, the weighted mean ORs were 1.43 (FE) and 1.41 (RE), indicating a substantial effect of these programmes on bullying. To give a concrete example, if there were 20 bullies and 80 non-bullies in the experimental condition and 26 bullies and 74 non-bullies in the control condition, the OR would be 1.41. Hence, OR = 1.41 can correspond to 30% more bullies in the control condition (or conversely 23% fewer bullies in the experimental condition).

Table 8. Effect Sizes for Victimization.

Project	OR	CI	z	P
Randomized Experiments				
Baldry & Farrington (2004)	1.69	0.76 – 3.78	1.29	ns
Cross et al. (2004)	1.07	0.79 – 1.43	0.42	ns
De Rosier (2004)	1.04	0.75 – 1.45	0.24	ns
Fekkes et al. (2006)	1.25	0.95 – 1.65	1.61	ns
Frey et al. (2005)	1.09	0.76 – 1.56	0.44	ns
Hunt (2007)	1.26	0.67 – 2.36	0.71	ns
Jenson & Dieterich (2007)	1.63	0.78 – 3.41	1.31	ns
Rosenbluth et al. (2004)	0.70	0.50 – 0.97	-2.14	.032
Salmivalli et al. (2007)	1.66	1.39 – 1.99	5.64	.0001
Weighted mean (FE)	1.19	1.06 – 1.33	2.96	.003
Weighted mean (RE)	1.18	0.96 – 1.44	1.59	Ns
Before-After, Experimental-Control				
Andreou et al. (2007)	1.48	1.01 – 2.16	1.99	.047
Bauer et al. (2007)	1.01	0.85 – 1.18	0.06	ns
Ciucci & Smorti (1998)	1.21	0.70 – 2.12	0.69	ns
Melton et al. (1998)	1.06	0.91 – 1.23	0.70	ns
Menard et al. (2008)	1.22	1.02 – 1.46	2.14	.032
Olweus/Bergen 2	1.43	1.04 – 1.95	2.23	.026
Pepler et al. (2004)	0.94	0.71 – 1.24	-0.42	ns
Rahey & Craig (2002)	0.79	0.47 – 1.33	-0.87	ns
Rican et al. (1996)	2.43	0.62 – 9.73	1.28	ns
Whitney et al. (1994)	1.26	0.80 – 1.98	0.98	ns
Weighted mean (FE)	1.10	1.01 – 1.20	2.12	.041
Other Experimental-Control				
Evers et al. (2007)	2.33	1.62 – 3.35	4.59	.0001
Galloway & Roland (2004)	1.59	1.20 – 2.11	3.26	.001
Ortega et al. (2004)	2.12	1.15 – 3.91	2.40	.016
Raskauskas (2007)	1.35	1.14 – 1.60	3.54	.0004
Weighted mean (FE)	1.54	1.35 – 1.75	6.45	.0001
Weighted mean (RE)	1.70	1.31 – 2.21	3.98	.0001
Age-Cohort Designs				
Ertesvag & Vaaland (2004)	1.18	0.99 – 1.39	1.88	.060
Olweus/Bergen 1	2.89	2.14 – 3.90	6.93	.0001
Olweus/Oslo 1	1.81	1.23 – 2.66	3.03	.002
Olweus/New National	1.59	1.45 – 1.73	10.18	.0001
Olweus/Oslo 2	1.48	1.25 – 1.77	4.44	.0001
Salmivalli et al. (2005)	1.30	1.06 – 1.60	2.47	.014
Weighted mean (FE)	1.52	1.43 – 1.63	12.65	.0001
Weighted mean (RE)	1.60	1.34 – 1.91	5.16	.0001
Weighted mean (FE)	1.35	1.29 – 1.41	13.14	.0001
Weighted mean (RE)	1.23	1.12 – 1.34	4.62	.0001

Note: OR = Odds Ratio; CI = Confidence Interval; FE = Fixed Effects; RE = Random Effects.

Table 8 shows the analysis of effect sizes for victimization (being bullied). Only two of the randomized experiments found significant effects of the programme on victimization. However, the FE model indicated a significant effect of these programmes (OR = 1.19). The RE model was not significant, because of the increase in the variance of each OR. Three of the 10 studies with before-after/experimental-control designs

yielded significant results, and the weighted mean OR (1.10) was just statistically significant. The FE model was used because the effect sizes were not significantly heterogeneous.

All four studies with other experimental-control designs found significant results, with weighted mean ORs of 1.54 (FE) and 1.70 (RE). Five of the six age-cohort designs yielded significant results, and the other one (Ertesvag & Vaaland, 2004) was nearly significant. The weighted mean ORs were 1.52 (FE) and 1.60 (RE). Over all 29 studies, the weighted mean ORs were 1.35 (FE) and 1.23 (RE), indicating significant effects of these programmes on victimization. To give a concrete example, if there were 20 victims and 80 non-victims in the experimental condition, and 25 victims and 75 non-victims in the control condition, then OR = 1.33. If there were 24 victims and 76 non-victims in the control condition, then OR = 1.26. Hence, these values of the OR correspond to 20%-25% more victims in the control condition (or conversely, 17-20% fewer victims in the experimental condition).

We conclude that the following 12 anti-bullying programmes were clearly effective in reducing bullying and victimization: Andreou et al. (2007), Ertesvag and Vaaland (2007), Evers et al. (2007), Melton et al. (1998), Olweus/Bergen2, Olweus/Bergen 1, Olweus/Oslo1, Olweus/New National, Olweus/Oslo2, Raskauskas (2007), Salmivalli et al. (2005) and Salmivalli et al. (2007). The following 10 programmes were probably effective, as judged by their effect sizes or by their significance in relation to either bullying or victimization or by other results presented by the authors: Baldry and Farrington (2004), Fekkes et al. (2006), Hunt (2007), Jenson and Dieterich (2007), Menard et al. (2008), Pepler et al. (2004), Rican et al. (1996), Whitney et al. (1994), Galloway and Roland (2004) and Ortega et al. (2004). The remaining 7 programmes had little effect on bullying or victimization: Bauer et al. (2007), Ciucci and Smorti (1998), Cross et al. (2004), De Rosier (2004), Frey et al. (2005), Rahey and Craig (2002) and Rosenbluth et al. (2004). In fact, there were some indications that the Rosenbluth et al. (2004) programme had harmful effects. Why were some programmes effective and others ineffective? We will address this question in section 6.

5.3 Effect Size versus Research Design

Tables 7 and 8 show that the weighted mean odds ratio effect size measure varies across the four types of research design. In order to test whether this variation is statistically significant, it is necessary to calculate the heterogeneity between groups or QB (Lipsey & Wilson, 2001, pp. 135-138). For bullying, QB = 29.08 (3 df, p <.0001). For victimization, QB = 41.85 (3 df, p <.0001). Therefore, we can conclude that effect sizes varied significantly across research designs. Weisburd, Lum and Petrosino (2001) also found lower effect sizes in randomized experiments than in other designs.

As mentioned earlier, the randomized experiments and before-after/experimental-control designs might be regarded by some researchers as methodologically superior to the other experimental-control and age-cohort designs. However, all designs have advantages and problems. For example, randomized experiments can (if a sufficiently large number of units is randomly assigned) minimize many threats to internal validity. However, experiments on bullying usually randomly assign only a small number of schools (see Table 9) and are vulnerable to differential attrition. The age-cohort design, on the other hand, largely eliminates problems of differential attrition (as well as selection, aging, and regression effects) but is potentially vulnerable to history and testing effects. However, Olweus (2005a) argued convincingly that these were unlikely, especially since the effects of programmes have been investigated in many different time periods. Overall, we conclude that these are the best four designs that have been used to evaluate the effects of anti-bullying programmes, and we give credence to results obtained in all of them.

Table 9. Units of Random Allocation.

Children:
De Rosier (2004) => 18 children in each of the 11 schools allocated to the experimental group; rest to control group (N = 415)

Classes:
Baldry & Farrington (2004) => 10 classes (N = 239)

Schools:
Cross et al. (2004) => 29 schools (N = 1957)
Fekkes et al. (2006) => 50 schools (N = 2221)
Frey et al. (2005) => 6 schools (N = 1126)
Hunt (2007) => 7 schools (N = 400)
Jenson & Dieterich (2007) => 28 schools (N = 685)
Rosenbluth et al. (2004) => 12 schools (N = 1763)
Salmivalli et al. (2007) => 78 schools (N =7301)

6. Coding of Study Features

6.1 Key Features of the Evaluation

We have already discussed one feature of the evaluation, namely the research design. In order to investigate the relationship between evaluation features and effect size in a comparable way, all features were dichotomized (in order to produce roughly equal groups, as much as possible). For example, research design was dichotomized into (1) randomized experiments plus before-after/experimental-control designs (20 studies) versus (2) other experimental-control designs plus age-cohort designs (10 studies). Other features of the evaluation that were investigated were as follows:

(a) Sample size (experimental plus control conditions), dichotomized into 1,500 children or more (14) versus 1,499 children or less (16). Several meta-analyses (e.g. Farrington & Welsh, 2003) have found a negative relationship between effect size and sample size.
(b) Publication date, dichotomized into 2004 or later (19) versus 2003 or earlier (11).
(c) Average age of the children, dichotomized into 10 or less (13) versus 11 or more (13).
(d) Location in the USA (8) versus other places (22).
(e) Location in other places (23) versus Norway (7).
(f) Location in other places (13) versus Europe (17).
(g) Outcome measure, dichotomized into others (23) versus a dichotomous measure of two or more times per month (7). This latter measure was associated with larger effect sizes than mean scores or simple prevalences.

6.2 Key Elements of the Programme

Each anti-bullying programme included a variety of intervention elements. Table 10 summarizes the elements of the intervention in different programmes and their frequency. In constructing this table we consulted the evaluators of the various programmes, and sent them our coding of the elements of the intervention for their programme. By mid-July 2008, we had received feedback on 24 out of 30 programmes and relevant changes were made to the coding where appropriate. For instance, even though the 'Controlled Trial of OBPP' (Bauer et al, 2007) included an anti-bullying video, this anti-bullying method was involved in only two out of seven intervention schools, so we did not code this element as included in this programme. For similar reasons, for 'Youth Matters' (Jenson & Dieterich, 2007) we did not code the use of anti-bullying videos, even though the formal description of the programme included this method.

Element 1 (whole-school anti-bullying policy) involves the presence of a formal anti-bullying policy on behalf of the school. In many schools, as indicated by researchers, such a policy was already in effect. It was not possible for us know whether, for each programme, the same anti-bullying policy was incorporated in the intervention schools.

Element 2 (classroom rules) refers to the use of rules against bullying that students were expected to follow. In many programmes, these rules were the result of cooperative group work between the teachers and the students, usually after some extent of exposure of the students to the philosophy or messages of the anti-bullying programme. In many cases the rules were written on a notice that was displayed in a distinctive place in the classroom.

Element 3 (school conferences) refers to the organization of school assemblies during which children were informed about bullying. In many programmes, these conferences were organized after the pretest data collection and aimed to inform students about the extent of bullying behaviour in their school. This was perceived as an initial way to sensitize students about bullying and as a means of announcing the formal beginning of the intervention programme in the school.

Element 4 (curriculum materials) refers to the use of materials about bullying during classroom lessons. Some programmes were curriculum-based whereas in others teachers incorporated anti-bullying materials into the regular curriculum.

Element 5 (classroom management) refers to an emphasis on classroom management techniques in detecting and dealing with bullying behaviour.

Element 6 (cooperative group work) refers to the cooperation among different professionals (usually among teachers and some other professional groups) in working with bullies and victims of bullying.

Elements 7 and 8 (work with bullies and victims) concern individualized work (not offered at the classroom level) with children involved in bullying as victims or perpetrators. In most programmes, this service was offered by professionals, such as interns or psychologists, who collaborated with teachers in the school.

Element 9 (work with peers) refers to the formal engagement of peers in tackling bullying. This could involve the use of several strategies such as peer mediation (students working as mediators in the interactions among students involved in bullying) and peer mentoring, which was usually offered by older students. The philosophy of many anti-bullying programmes also placed emphasis on the engagement of bystanders in bullying situations in such a way that disapproval of bullying behaviour was expressed adequately while support was offered to victims.

Elements 10 and 11 (information for teachers and parents): Many programmes offered information for teachers and parents, but it was not possible for us to assess the quality of the information provided. For instance, many programmes reported the presence of a manual that teachers could consult in the implementation of the intervention, but the extent to which this manual was structured is difficult for us to assess. The same can be said about the information provided to parents. It was clear to us that programmes differed a lot in the quality of this information. In some programmes parents were provided with newsletters regarding the anti-bullying initiative in their school, while in others parents were provided with guides on how to help their child deal with bullying as well as information about the anti-bullying initiative implemented in their school. However, the overall information that we had regarding this element of the intervention did not allow us to differentiate among different levels of its implementation across programmes.

Element 12 (improved playground supervision): Some anti-bullying programmes aimed to identify 'hot-spots' or 'hot-times' of bullying (mostly during playtime or lunchtime) and provided improved playground supervision of children.

Element 13 (disciplinary methods): Some programmes emphasized punitive methods in dealing with bullying situations. One programme (KiVa; Salmivalli et al., 2007) used both punitive and non-punitive methods. In half of the 78 intervention schools teachers were encouraged to use strong disciplinary methods whilst in the rest of the intervention schools teachers were encouraged to deal with bullying situations in a non-punitive way.

Elements 14 and 15 (Non-punitive methods): Some programmes included restorative justice approaches and other non-punitive methods such as the 'Pikas method' and the 'No Blame' approach in dealing with children involved in bullying.

Element 16 (school tribunals and bully courts) was not used to any great extent in any of the present studies. Bully courts were offered as an optional element within the Sheffield programme, but no school actually established one.

Element 17 (teacher training): This was coded as present or absent. We also coded both the duration (number of meetings among experts and teachers) as well as the intensity (number of hours) of this training (see later). Again, we sent emails to the evaluators of the different programmes and asked for their advice. Some researchers were responsive and offered us adequate information on both the duration and the intensity of teacher training to the extent that we could be confident about our accuracy in coding these elements. For other programmes,

however, we could not code one or both of these features of teacher training.

Element 18 (parent training): For all programmes this refers to the organization on behalf of the school of 'information nights/educational presentations' for parents and/or 'teacher-parent meetings' during which parents were given information about the anti-bullying initiative in the school.

Elements 19 and 20 (videos and virtual reality computer games): Some programmes utilized technology in their anti-bullying materials such as the use of anti-bullying videos or virtual reality computer games to raise students' awareness regarding bullying.

We also coded other features of the intervention programmes:

(a) The number of elements included out of 20, dichotomized into 10 or less (15 programmes) versus 11 or more (15 programmes). Olweus (2005a) reported a 'dose-response' relationship between the number of components implemented in a school and the effect on bullying.
(b) The extent to which the programme was not (17) or was (13) inspired by the work of Dan Olweus.
(c) The duration of the programme for children, dichotomized into 240 days or less (12) versus 270 days or more (17).
(d) The intensity of the programme for children, dichotomized into 19 hours or less (12) versus 20 hours or more (11).
(e) The duration of the programme for teachers, dichotomized into 3 days or less (13) versus 4 days or more (11).
(f) The intensity of the programme for teachers, dichotomized into 14 hours or less (13) versus 15 hours or more (13).

Table 10. Programme Elements and Their Frequency.

Element	Frequency*
1. Whole-school anti-bullying policy	22
2. Classroom rules	23
3. School conferences/assemblies providing information about bullying to children	17
4. Curriculum materials	26
5. Classroom management	23
6. Cooperative group work among experts (e.g. teachers, counsellors and interns)	16
7. Work with bullies	17
8. Work with victims	18
9. Work with peers (e.g. peer mediation, peer mentoring, peer group pressure as bystanders)	9
10. Information for teachers	29
11. Information for parents	20
12. Improved playground supervision	10
13. Disciplinary methods	10
14. Non-punitive methods (e.g. Pikas, No Blame)	6
15. Restorative Justice approaches	1
16. School tribunals; school bully courts	0
17. Teacher training	21
18. Parent training/meetings	11
19. Videos	15
20. Virtual Reality computer games	3

* out of 30 studies

6.3 Effect Size versus Study Features

There have been few other attempts to relate effect size to programme elements (see e.g. Kaminski, Valle, Filene & Boyle, 2008). Table 11 shows the programme elements and design features that were significantly (or nearly significantly in two cases) related to effect sizes for bullying. Because of small numbers, five of the 20 programme elements could not be investigated (curriculum materials, information for teachers, restorative justice approaches, school tribunals/bully courts, and virtual reality computer games). As explained before, the significance test is based on the heterogeneity between groups QB. The weighted mean odds ratio effect sizes are also given for the different categories.

The most important programme elements that were associated with a decrease in bullying were parent training, improved playground supervision, disciplinary methods, school conferences, information for parents, classroom rules, classroom management, and videos. In addition, the total number of elements, and the duration and intensity of the programme for children and teachers, were significantly associated with a decrease in bullying. Also, programmes inspired by the work of Dan Olweus worked best. Regarding the design features, the programmes worked better with older children, in smaller-scale studies, in Norway

specifically, and in Europe more generally. Older programmes, and those in which the outcome measure was two times per month or more, also worked better. No programme element was significantly associated with an increase in bullying.

Table 11. Significant Relationships with Bullying.

	Cat (N) OR	Cat (N) OR	QB	P
Programme Elements				
Parent training	No (18) 1.29	Yes (10) 1.67	20.57	.0001
Playground supervision	No (19) 1.30	Yes (9) 1.68	19.77	.0001
Disciplinary methods	No (18) 1.30	Yes (10) 1.66	18.27	.0001
Duration for children	240- (10) 1.18	270+ (17) 1.51	14.13	.0002
Total elements	10- (15) 1.32	11+ (13) 1.60	11.77	.0006
Based on Olweus	No (17) 1.31	Yes (11) 1.59	11.15	.0008
Intensity for children	19- (11) 1.28	20+ (10) 1.65	9.23	.002
School conferences	No (13) 1.33	Yes (15) 1.57	8.64	.003
Duration for teachers	3- (12) 1.34	4+ (11) 1.61	7.42	.006
Information for parents	No (10) 1.27	Yes (18) 1.50	7.09	.008
Intensity for teachers	14- (12) 1.32	15+ (12) 1.52	5.63	.018
Classroom rules	No (7) 1.22	Yes (21) 1.46	4.55	.033
Classroom management	No (7) 1.23	Yes (21) 1.46	4.10	.043
Videos	No (14) 1.35	Yes (14) 1.50	3.17	.075
Design Features				
Age of children	10- (13) 1.18	11+ (11) 1.58	22.02	.0001
Publication Year	04+ (18) 1.31	03- (10) 1.69	18.75	.0001
Outcome measure	Other (21) 1.33	2+M (7) 1.74	18.51	.0001
In Norway	Rest (21) 1.34	Nor (7) 1.58	7.76	.005
In Europe	Rest (12) 1.32	EU (16) 1.53	6.47	.011
Sample size	1500+ (15) 1.31	1499- (13) 1.47	2.89	.089

Notes: Cat = Category of variable; OR = Weighted mean odds ratio; QB = heterogeneity between groups; Duration in days; Intensity in hours; Outcome Measure 2+M: two times per month or more (versus other measures).

Table 12 shows the programme elements and design features that were significantly related to effect sizes for victimization (being bullied). Effect sizes for bullying and victimization were significantly correlated ($r = .58$, $p < .0001$). The most important programme elements that were associated with a decrease in victimization were videos, disciplinary methods, work with peers, parent training, cooperative group work and playground supervision. In addition, the duration of the programme for children and teachers, and the intensity of the programme for teachers, were significantly associated with a decrease in victimization. Regarding the design features, the programmes worked better with older children, in Norway specifically and in Europe more generally, and they were less effective in the USA. Older programmes, those in which the outcome measure was two times per month or more, and those with other experimental-control and age-cohort designs, also worked better. No programme element was significantly associated with an increase in victimization.

Table 12: Significant Relationships with Victimization

	Cat (N) OR	Cat (N) OR	QB	P
Programme Elements				
Videos	No (15) 1.15	Yes (10) 1.51	32.96	.0001
Duration for teachers	3- (13) 1.18	4+ (11) 1.55	27.80	.0001
Disciplinary methods	No (19) 1.21	Yes (10) 1.50	21.64	.0001
Intensity for teachers	14- (13) 1.19	15+ (12) 1.50	21.38	.0001
Duration for children	240- (11) 1.13	270+ (17) 1.42	18.08	.0001
Work with peers	No (20) 1.11	Yes (9) 1.41	15.43	.0001
Parent training	No (19) 1.23	Yes (10) 1.47	15.24	.0001
Cooperative group work	No (14) 1.22	Yes (15) 1.42	9.51	.002
Playground supervision	No (19) 1.29	Yes (10) 1.41	3.87	.049
Design Features				
Outcome measure	Other (22) 1.18	2+M (7) 1.64	49.19	.0001
In Europe	Rest (13) 1.13	EU (16) 1.52	40.90	.0001
Design	12 (19) 1.13	34 (10) 1.53	40.73	.0001
In Norway	Rest (22) 1.20	Nor (7) 1.55	30.77	.0001
Not in USA	US (8) 1.10	Rest (21) 1.45	27.26	.0001
Publication year	04+ (19) 1.23	03- (10) 1.52	21.04	.0001
Age of children	10- (13) 1.18	11+ (12) 1.43	12.80	.0004

Notes: Cat = Category of variable; OR = Weighted mean odds ratio; QB = heterogeneity between groups; Design: 12 = randomized experiments + before-after/experimental-control versus 34 = other experimental-control + age-cohort designs; Duration in days; Intensity in hours; Outcome Measure 2+M: two times per month or more (versus other measures)

Variables that might help to explain differential treatment effects in meta-analysis (e.g. elements of the intervention, study features, etc.) cannot be assumed to be statistically independent. Researchers should try to disentangle the relationships among them and identify those that truly have significant independent relationships with effect sizes (Lipsey, 2003, p. 78). Multivariate techniques can be used to solve this problem in meta-analysis (Hedges, 1982). Weighted regression analyses (Lipsey & Wilson, 2001, pp. 138–140) were carried out to investigate which elements of the programmes and which features of the evaluations were independently related to bullying and victimization effect sizes (LORs). These analyses were severely limited by the small number of studies. Nevertheless, they showed that the most important elements of the programme that were related to a decrease in bullying were parent training and improved playground supervision (Table 13).

Table 13: Results of Weighted Regression Analyses

	B	SE (B)	Z	p
Bullying Effect Size				
Elements				
Parent training	.176	.066	2.68	.007
Playground supervision	.169	.067	2.53	.011
Elements + Features				
Age of children	.227	.064	3.56	.0004
Playground supervision	.202	.065	3.12	.002
Victimization Effect Size				
Elements				
Videos	.223	.052	4.26	.0001
Intensity for Teachers	.166	.052	3.20	.001
Elements + Features				
Outcome measure	.238	.053	4.46	.0001
Design	.180	.053	3.38	.0007

The most important elements of the programme that were related to a
decrease in victimization were videos and the intensity of the pro-
gramme for teachers. When the design features were added, the most
important factors that were related to a decrease in bullying were the
age of the children and improved playground supervision. The most
important factors that were related to a decrease in victimization were
the outcome measure (two or more times per month versus other meas-
ures) and the design (other experimental-control comparisons and age-
cohort designs versus other designs).

7. Conclusions

7.1 Summary of Main Findings

The present systematic review shows that school-based anti-bullying programmes are often effective, and that particular programme elements were associated with a decrease in bullying and victimization (being bullied). No programme element was significantly associated with an increase in bullying or victimization.

The most important programme elements that were associated with a decrease in bullying were parent training, improved playground supervision, disciplinary methods, school conferences, information for parents, classroom rules, classroom management, and videos. In addition, the total number of elements, and the duration and intensity of the programme for children and teachers, were significantly associated with a decrease in bullying. Also, programmes inspired by the work of Dan Olweus worked best. Regarding the design features, the programmes worked better with older children, in smaller-scale studies, in Norway specifically, and in Europe more generally. Older programmes, and those in which the outcome measure was two times per month or more, also yielded better results.

The most important programme elements that were associated with a decrease in victimization were videos, disciplinary methods, work with peers, parent training, cooperative group work and playground supervision. In addition, the duration of the programme for children and teachers, and the intensity of the programme for teachers, were significantly associated with a decrease in victimization. Regarding the design features, the programmes worked better with older children, in Norway specifically and in Europe more generally, and they were less effective in the USA. Older programmes, those in which the outcome measure was two times per month or more, and those with other experimental-control and age-cohort designs, also yielded better results.

7.2 Policy Implications

Meta-analysis remains the most systematic and credible procedure available for synthesizing the results of quantitative intervention studies in a manner that can inform practitioners, policy makers, and researchers (Lipsey, 2003, p. 80). The main policy implication of our review is that new anti-bullying programmes should be designed and tested based on our results. These could be grounded in the successful Olweus programme but should be modified in light of the key programme elements that we have found to be most effective.

Also, future evaluations of anti-bullying programmes should be designed in light of our results. In particular, programmes should be targeted on children aged 11 or older, rather than on younger children.

The outcome measure of bullying or victimization should be two times per month or more. Also, cost-benefit analyses of anti-bullying programmes should be carried out, to investigate how much money is saved for the money expended (Welsh, Farrington, & Sherman 2001). Saving money is a powerful argument to convince policy-makers and practitioners to implement intervention programs (Farrington, 2008, p. 59). Finally, anti-bullying programmes should pay more attention to theories of bullying and victimization.

7.3 Implications for Future Research

Contrary to some previous reviews (Ferguson et al., 2007, p. 410) the present systematic review shows that school-based anti-bullying programmes are effective. There are many implications of our review for future research. Several questions have been raised that should be addressed. For example:

- Why are there different effects of programme elements and design features on bullying and victimization?
- Why do results vary in different countries?
- Why do results vary by research design?
- Why do programmes work better with older children?
- Why are larger and more recent studies less effective than smaller-scale and older studies?
- Why do results vary with the outcome measure of bullying or victimization?

Future evaluations should have before and after measures of bullying and victimization in experimental and control schools. Bullying and victimization should be carefully defined and measured. Since it is difficult to randomly assign a large number of schools, it may be best to place schools in matched pairs and randomly assign one member of each pair to the experimental condition and one member to the control condition. It seems unsatisfactory to randomly assign school classes because of the danger of contamination of control children by experimental children. Only children who are tested both before and after the intervention should be analysed, in order to minimize problems of differential attrition. Research is needed on the best methods of measuring bullying, on what time periods to enquire about, and on seasonal variations.

It is important to develop methodological quality standards for evaluation research that can be used by systematic reviewers, scholars, policy makers, the mass media, and the general public in assessing the validity of conclusions about the effectiveness of interventions in reducing crime (Farrington, 2003, p. 66). Such quality standards could include guidelines to programme evaluators with regard to what elements of the intervention should be included in published reports, perhaps

under the aegis of the Campbell Collaboration Crime and Justice Group (Farrington & Petrosino, 2001; Farrington & Weisburd, 2007). If these guidelines had been in existence, they would have been very helpful in the ambitious task we have undertaken to fully code the elements of the intervention in all studies.

With a positive response from 24 out of 30 evaluators of anti-bullying programmes (up to mid-July 2008), we have been fairly successful. However, many things still need to be done. For instance, because of time limitations and lack of information, we were unable to detect the varying results of the intervention programmes according to subgroups of students – subgroups defined for example by gender, ethnicity, participant roles in bullying, developmental needs and/or capacities of children. Other researchers have also indicated the lack of specific intervention work based on the above factors (Pepler, Smith & Rigby, 2004; Smith & Ananiadou, 2003). Several of the above features were not mentioned in reports, making it difficult for us to code them. For the 20 programme elements that we did code, only one study (Fekkes et al., 2006) provided the percentage of intervention and control schools that actually implemented these elements.

Future systematic reviewers could attempt to detect the impact of anti-bullying programmes for different subgroups of students. Future reports should provide key information about features of evaluations, according to a checklist that should be developed (inspired perhaps by the CONSORT Statement for medical research: Altman et al., 2001; Moher, Schulz, & Altman, 2001). Information about key elements of programmes, and about the implementation of programmes, should be provided. Where bullying and victimization are measured on 5-point scales, the full 5 x 2 table should be presented, so that the Area Under the ROC Curve (AUC) could be used as a measure of effectiveness. This would avoid the problem of results varying with the particular cut-off points that are chosen.

Research is needed to develop and test better theories of bullying and victimization, for example using vignettes with children to ask about what factors promote or prevent bullying. The advantages and disadvantages and validity of different outcome measures (e.g. self-reports, peer ratings, teacher ratings, systematic observation) should be studied. The short-term and long-term effects of anti-bullying programmes should be investigated in prospective longitudinal studies. Effects on different types of bullying, and effects on different types of children, teachers, schools, and contexts, should be investigated.

Ideally, interventions should be based on theories of bullying and victimization (Baldry & Farrington, 2007, p. 201). These theories should guide programme development. Other researchers have emphasized on the importance of using theoretically grounded interventions as well. As Eck (2006, p. 353) puts it: '…if we are to improve our ability to give valid crime policy advice, we must begin to treat crime theory

74

more seriously. Accounting for the theoretical support for anti-crime interventions will put our generalizations on sounder epistemological foundations than the current reliance on naive induction'.

In conclusion, results obtained so far in evaluations of anti-bullying programmes are encouraging. The time is ripe to mount a new programme of research on the effectiveness of these programmes, based on our findings.

Technical Appendix: Effect Size and Meta-Analysis

In order to carry out a meta-analysis, every evaluation must have a comparable effect size. The most usual effect sizes for intervention studies are the standardized mean difference d and the odds ratio OR (Lipsey & Wilson, 2001). Where researchers reported the prevalence of bullying (or victimization), we calculated OR. Where researchers reported mean scores, we calculated d. It is easy to convert d into OR (see later), and this is what we did. We carried out our meta-analysis using the natural logarithm of OR (LOR) and converted the results back into OR for case of interpretation. We will explain this for bullying but the same methods were used for victimization.

Odds Ratio

The OR is calculated from the following table:

	Non-Bullies	Bullies
Experimental	a	b
Control	c	d

Where a, b, c, d are numbers of children

$OR = (a*d) / (b*c)$

* indicates multiplication

An OR greater than 1 indicates a desirable effect of the anti-bullying programme, while an OR less than 1 indicates an undesirable effect. The chance value of the OR is 1, indicating no effect.

For example, the figures for the first posttest of Fekkes et al. (2007) were as follows:

	Non-Bullies	Bullies	% Bullies
Experimental	1101	87	7.9
Control	1009	99	8.9

Here, $OR = (1101*99) / (1009*87) = 1.14$

The statistical significance of an OR is assessed by calculating the LOR:

$LOR = Ln (OR)$

Here, $LOR = Ln (1.14) = 0.131$

The variance of LOR, VLOR, is as follows:
VLOR = (1/a) + (1/b) + (1/c) + (1/d)
Here, VLOR = 0.0236

The standard error of LOR, SELOR, is the square root of the variance:
Here, SELOR = 0.1535
Once SELOR is known, it is easy to calculate confidence intervals for OR. The 95% confidence interval (CI) about LOR is + or − 1.96 * SELOR
Hence, the lower CI = 0.131 − 1.96 * 0.1535   = -- 0.170

The corresponding lower CI for the OR is:
ORLOCI = Exp (- 0.170) = 0.84
Where Exp indicates the exponential.
Similarly, the higher CI = 0.131 + 1.96 * 0.1535 =   0.432

The corresponding higher CI for the OR is:
ORHICI = Exp (0.432) = 1.54
The confidence intervals are symmetrical about LOR (0.131 + or - 0.301) but not about OR (1.14, CI 0.84 – 1.54).

The significance of LOR is tested as follows:
Z = LOR / SELOR
Where z is an observation from a unit normal distribution with mean = 0 and standard deviation = 1
Here, Z = 0.85

Since this is below the value (1.96) corresponding to p = .05, we conclude that the OR of 1.14 is not statistically significant, and hence that the anti-bullying programme did not cause a significant decrease in bullying.

Standardized Mean Difference d

d is calculated as follows:
        d = (MC – ME) / SP
Where MC = Mean score in control condition
        ME = Mean score in experimental condition
        SP = Pooled standard deviation (SD)

The pooled variance, VP, is as follows:
        VP = [(NC – 1)* VC + (NE – 1)* VE] / (NC + NE – 2)
Where NC = Number in control condition
        VC = Variance of control scores
        NE = Number in experimental condition
        VE = Variance of experimental scores

As an example, for bullying by older children after the intervention of Baldry and Farrington (2004):

MC = 3.39
VC = 15.92 (SD = 3.99, squared)
NC = 36
ME = 2.31
VE = 9.425 (SD = 3.07, squared)
NE = 99

VP = [(35 * 15.92) + (98 * 9.425)] / 133 = 11.134
Hence, SP = 3.337
d = (3.39 – 2.31) / 3.337 = 0.324

To a close approximation, the variance of d, Vd, is as follows:
Vd = (NC + NE) / (NC * NE)
Here, Vd = (36 + 99) / (36 * 99) = 0.03788

Hence, the standard error of d is as follows:
SEd = 0.195

The significance of d can be tested as follows:
Z = d / SEd
Here, Z = 0.324 / 0.195 = 1.66
Since this is below 1.96, this value of d is not statistically significant.

d can be converted into LOR using the following equation (Lipsey & Wilson, 2001, p.202):
LOR = d / 0.5513
Hence, LOR = 0.587

Similarly, the SE of LOR is as follows:
SELOR = SEd / 0.5513
Here SELOR = 0.354
Z = LOR / SELOR = 1.66 as before

In one case where phi correlations were reported as effect sizes (Menard et al., 2008), we use the following equation to convert r to d (Lipsey & Wilson, 2001, p.63):
d = 2 * r / sqrt (1 – r * r)

To a good approximation:
SEd = 2 * Ser

The SE of r is calculated using the transformation:
Zr = 0.5 * Ln [(1 + r) / (1 – r)]
and VAR (Zr) = 1 / (N – 3)
The analysis then proceeded as above.

Before and After Measures

Where there are before and after measures of bullying, the appropriate effect size measure is:
LOR = LORA - LORB
Where LORA = LOR after
LORB = LOR before

Fekkes et al. (2007) had a before measure of bullying, with ORB = 1.01 and
LORB = 0.010
Therefore, for Fekkes et al.,
LOR = 0.131 - 0.010 = 0.121

The variance of this LOR is as follows:
VLOR = VLORA + VLORB – 2 * COV
Where COV = Covariance

Because LORA and LORB are positively correlated, VLOR will be less than (VLORA + VLORB). However, the covariance is usually not reported. In general, VLOR will be between (VLORA + VLORB) / 2 and (VLORA + VLORB). Therefore, we estimate it as half-way between these values:
VLOR = 0.75 * (VLORA + VLORB)

For Fekkes et al. (2007):
VLOR = 0.75 * (0.0373 + 0.0236) = 0.0457
Consequently,SELOR = 0.214
OR = exp (LOR) = exp (0.121) = 1.13
The confidence intervals are 0.74 – 1.72
Z = 0.121 / 0.214 = 0.57
Again, this is less than 1.96, so this LOR is not significant.

Combining LORs Within a Study

It is common for a study to yield more than one LOR. In this case, the weighted average LOR is calculated. For example, for Baldry and Farrington (2004):
For older children, LOR after = 0.587, LOR before = - 0.247;
LOR (older) = LORO = 0.587 – (- 0.247) = 0.834
SELORO can be calculated as 0.432
For younger children, LOR after = - 0.801, LOR before = - 0.125:
LOR (younger) = LORY = - 0.801 – (- 0.125) = - 0.676
SELORY can be calculated as 0.464
Each LOR is weighted by its inverse variance (1 / VLOR).
WO = 1 / (SELORO * SELORO)
      = 1 / (0.432 * 0.432) = 5.358
WY = 1 / (SELORY * SELORY)
      = 1 / (0.464 * 0.464) = 4.651

Where WO = Weighting of LORO
WY = Weighting of LORY
LOR combined = LORC = [(LORO * WO) + (LORY * WY)] / (WO + WY) = [(0.834 * 5.358) + (-0.676 * 4.651)] / (5.358 + 4.651) = 0.133

The variance of LORC, VLORC, is:
VLORC = 1 / (WO + WY) = 1 / (5.358 + 4.651) = 0.0998
Therefore, SELORC = 0.316
ORC = exp (LORC) = exp (0.133) = 1.14
The confidence intervals are 0..62 --- 2.12
Z = LORC / SELORC = 0.133 / 0.316 = 0.42
This is not significant.

Correction for Clustering

The standard techniques assume that individuals are allocated to experimental or control conditions, so that each individual is independent of each other individual. However, in evaluations of anti-bullying programmes, it is usually the case that school classes (not individual children) are allocated to conditions. In this case, it is necessary to correct standard errors of effect sizes for the effects of clustering (Hedges, 2007).

The correction depends on an estimate of the intraclass correlation (rho). This is not usually reported. However, Murray and Blitstein (2003) carried out a systematic review of articles reporting intraclass correlations and found that, for youth studies with behavioural outcomes, rho was about 0.025. Also, Olweus (2008) informed us that: "I have made a number of such estimates on my large scale samples for being bullied and bullying others and … the intraclass correlation at the classroom level is typically in the .01 to .04 range". We therefore estimate that rho = 0.025. All the calculations assume equal sizes of clusters (school classes).

We will not correct effect sizes because the correction for clustering has a negligible impact on them. The correction for d (and, by implication, for LOR) is as follows:
Corrected d = d * sqrt [1 – [2 * (n – 1) * rho] / (N – 2)]
Where n = cluster size (school classes) and N = total sample size
For typical values of n = 30 and N = 500,
Corrected d = d * sqrt [1 – (2 * 29 * 0.025) / 498] = d * 0.998
Because this is very close to d, we do not correct effect sizes for clustering.

We need to correct standard errors of effect sizes. To a very good approximation, corrected Vd = Vd * [1 + (n – 1) * rho]
Where Vd = variance of d
Assuming n = 30 and rho = 0.025, corrected Vd = Vd * 1.725

We therefore multiply all variances of effect sizes by 1.725 and all standard errors of effect sizes by 1.313 in order to take account of the clustering of children in school classes.

For example, returning to Baldry and Farrington (2004), LORC = 0.133 and SELORC = 0.316. We multiply SELORC by 1.313 to obtain:

Corrected SELORC = 0.415

Corrected z = 0.133 / 0.415 = 0.32

Meta-Analysis

We use standard methods of meta-analysis, following Lipsey and Wilson (2001). In the simplest fixed effects model, the weighted mean effect size is as follows:

WMES = sum (Wi * ESi) / sum (Wi)

Where WMES = weighted mean effect size

ESi = effect size in the ith study

Wi = weighting in the ith study = 1 / Vi

Where Vi = variance of effect size in the study

SE (WMES) = sqrt [1 / sum (Wi)]

And Z = WMES / SE (WMES)

In order to test whether all effect sizes are randomly distributed about the weighted mean, the Q statistic is calculated:

Q = sum [Wi * (ESi – WMES) * (ESi – WMES)]

Q is distributed as chi-squared with (k-1) degrees of freedom, where k is the number of effect sizes. If Q is statistically significant, it is desirable to use a random effects model, in which a constant Vx is added to each variance Vi.

corrected Vi = Vi + Vx

Vx = [Q – k + 1] / [sum (Wi) – sum (Wi* Wi) / sum (Wi)]

The weighted mean ES and its variance are then calculated as before using the corrected Vi. In general, we report results obtained with both fixed effects (FE) and random effects (RE) models because both have advantages and disadvantages. The main disadvantage of the RE model is that it sometimes gives almost equal weight to each study, rather than giving greater weight to the larger studies in calculating the weighted mean effect size. The main disadvantage of the FE model is that it may not fit the data if studies are significantly heterogeneous. As mentioned, we use OR and LOR as the main measures of effect size in this report.

References[4]

Altman, D.G., Schulz, K.F., Moher, D., Egger, M., Davidoff, F., Elbourne, D., Gotzsche, P.C., & Lang, T. (2001). *The revised CONSORT Statement for reporting randomized trials: explanation and elaboration.* Annals of Internal Medicine, 8, 663–694.

Ananiadou, K. & Smith, P.K. (2002). Legal requirements and nationally circulated materials against school bullying in European countries. *Criminal Justice*, 2, 471–491.

Andershed, H., Kerr, M., & Stattin, H. (2001). Bullying in schools and violence on the streets: are the same people involved? *Journal of Scandinavian Studies in Criminology and Crime Prevention*, 2, 31–49.

* Andreou, E., Didaskalou, E., & Vlachou, A. (2007). Evaluating the effectiveness of a curriculum-based anti-bullying intervention program in Greek primary schools. *Educational Psychology*, 27, 693–711.

* Baldry, A.C. (2001). *Bullying in schools: Correlates and intervention strategies.* PhD Thesis, Cambridge University. Index to Theses Database, 51–8145.

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.

* Bauer, N.S., Lozano, P., & Rivara, F.P. (2007). The effectiveness of the Olweus bullying prevention program in public middle schools: A controlled trial. *Journal of Adolescent Health*, 40, 266–274.

Campbell, D.T. & Stanley, J.C. (1966). *Experimental and quasi-experimental designs for research.* Chicago: Rand McNally.

* Ciucci, E. & Smorti, A. (1998). Il fenomeno delle pretonenze nella scuola: problemi e prospettive di intervento [The phenomenon of bullying in school: problems and prospects for intervention]. *Psichiatria dell'infanzia e dell'adolescenza*, 65, 147–157.

Cook, T.D. & Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis issues for field settings.* Chicago: Rand McNally.

Cowie, H. (2000). Aggressive and bullying behavior in children and adolescents. In G. Boswell (Ed.), *Violent children and adolescents: Asking the question why* (pp. 138–150). London: Whurr.

---

[4] References with an asterisk indicate evaluation reports which have been included in the meta-analysis.

* 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.

* DeRosier, M.E. (2004). Building relationships and combating bullying: Effectiveness of a school-based social skills group intervention. *Journal of Clinical Child and Adolescent Psychology*, 33, 196–201.

* DeRosier, M.E. & Marcus, S.R. (2005). Building friendships and combating bullying: effectiveness of S.S.GRIN at one-year follow-up. *Journal of Clinical Child and Adolescent Psychology*, 34, 140–150.

Eck, J.E. (2006). When is a bologna sandwich better than sex? A defense of small-n case study evaluations. *Journal of Experimental Criminology*, 2, 345–362.

Egger, M., Smith, D.G., & O'Rourke, K. (2001). Rationale, potentials and promise of systematic reviews. In M, Egger, D.G. Smith & D.G. Altman (Eds.), *Systematic reviews in health care: Meta-analysis in context* (pp. 23–42). London: BMJ Books.

* Ertesvag, S. K. & Vaaland, G. S. (2007). Prevention and reduction of behavioural problems in school: An evaluation of the Respect program. *Educational Psychology*, 27, 713–736.

Eslea, M., & Smith, P.K. (1998). The long-term effectiveness of anti-bullying work in primary schools. *Educational Research*, 40, 203–218.

* Evers, K.E., Prochaska, J.O., Van Marter, D.F., Johnson, J.L., & Prochaska, J.M. (2007). Transtheoretical-based bullying prevention effectiveness trials in middle schools and high schools. *Educational Research, 49*, 397–414.

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. (2003). Methodological quality standards for evaluation research. *Annals of the American Academy of Political and Social Sciences, 587*, 49–68.

Farrington, D.P. (2008). Conduct disorder, aggression and delinquency. In R.M. Lerner & L. Steinberg (Eds.), *Handbook of Adolescent Psychology* (3rd ed.). Hoboken, NJ: Wiley (in press).

Farrington, D.P., Jolliffe, D., & Johnstone, L. (2008). *Assessing violence risk: A framework for practice*. Edinburgh: Risk Management Authority Scotland.

Farrington, D.P. & Petrosino, A. (2001). The Campbell Collaboration Crime and Justice Group. *Annals of the American Academy of Political and Social Science*, 578, 35–49.

Farrington, D.P. & Weisburd, D. (2007). The Campbell Collaboration Crime and Justice Group. *The Criminologist, 32* (1), 1–5.

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.

* Fekkes, M., Pijpers, F.I.M., & Verloove-Vanhorick, S.P. (2006). Effects of antibullying school program on bullying and health complaints. *Archives of Pediatrics and Adolescent Medicine*, 160, 638–644.

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

Fonagy, P., Twemlow, S.W., Vernberg, E., Sacco, F.C., & Little, T.D. (2005). Creating a peaceful school learning environment: The impact of an anti-bullying program on educational attainment in elementary schools. *Medical Science Monitor*, 11 (7), 317–325.

* Frey, K.S., Edstrom, L.V.S., & Hirschstein, M. K. (2005). The Steps to Respect program uses a multi-level approach to reduce playground bullying and destructive playground behaviours. In D.L. White, M.K. Faber, & B.C. Glenn (Eds.), *Proceedings of Persistently Safe Schools 2005*, (pp. 47–55). Washington, DC: Hamilton Fish Institute, George Washington University.

* 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.

* Galloway, D. & Roland, E. (2004). Is the direct approach to reducing bullying always the best? In P.K. Smith, D. Pepler, & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 37–53). Cambridge: Cambridge University Press.

Hawker, D.S.J. & Boulton, M.J. (2000). Twenty years research on peer victimisation and psychosocial maladjustment: A meta-analytic review of cross-sectional studies. *Journal of Child Psychology and Psychiatry*, 41, 441–455.

Hedges, L.V. (1982). Fitting continuous models to effect size data. *Journal of Educational Statistics*, 7, 245–270.

Hedges, L.V. (2007). Effect sizes in cluster-randomized designs. *Journal of Educational and Behavioral Statistics*, 32, 341–370.

Heydenberk, R.A., Heydenberk, W.R., & Tzenova, V. (2006). Conflict resolution and bully prevention: Skills for school success. *Conflict Resolution Quarterly*, 24, 55–69.

* Hirschstein, M. K., Van Schoiack Edstrom, L., Frey, K. S., Snell, J. L., & Mackenzie, E. P. (2007). Walking the talk in bullying prevention: Teacher implementation variables related to initial impact of Steps to Respect Program. *School Psychology Review*, 36, 3–21.

Howard, K.A., Flora, J., & Griffin, M. (1999). Violence-prevention programs in schools: State of science and implications for future research. *Applied and Preventive Psychology*, 8, 197–215.

* Hunt, C. (2007). The effect of an education program on attitudes and beliefs about bullying and bullying behaviour in junior secondary school students. *Child and Adolescent Mental Health*, 12 (1), 21–26.

* Jenson, J.M. & Dieterich, W.A. (2007). Effects of a skills-based prevention program on bullying and bully victimization among elementary school children. *Prevention Science*, 8, 285–296.

* Jenson, J.M., Dieterich, W.A. & Rinner, J.R. (2005a). Effects of a skills-based prevention program on bullying and bully victimization among elementary school children. Paper given at the Annual Meeting of the American Society of Criminology. Toronto, Canada (November).

* Jenson, J.M., Dieterich, W.A., & Rinner, J.R. (2005b). *The prevention of bullying and other aggressive behaviors in elementary school students: Effects of the Youth Matters curriculum*. Paper given at the Annual Meeting of the Society for Prevention Research, Washington, DC (May).

* Jenson, J.M., Dieterich, W.A., Powell, A., & Stoker, S. (2006a). *Effects of a skills-based intervention on aggression and bully victimization among elementary school children*. Paper given at the Annual Meeting of the Society for Prevention Research, San Antonio, Texas (June).

* Jenson, J.M., Dieterich, W.A., Powell, A., & Stoker, S. (2006b). *Effects of the Youth Matters prevention curriculum on bullying and other aggressive behaviors in elementary school students*. Paper given at the Annual Meeting of the Society for Social Work and Research, San Antonio, Texas (January).

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).

Kaiser-Ulrey, C. (2003). *Bullying in middle schools: A study of B.E.S.T. – Bullying Eliminated from Schools Together – an anti-bullying programme for seventh grade students*. PhD Thesis. Tallahassee: Florida State University, College of Education.

Kalliotis, P. (2000). Bullying as a special case of aggression: Procedures for cross-cultural assessment. *School Psychology International*, 21, 47–64.

Kaminski, J.W., Valle, L.A., Filene, J.H., & Boyle, C.Y. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology*, 36, 567–589.

Leary, M.R., Kowalski, R.M., Smith, L., & Phillips, S. (2003). Teasing, rejection and violence: Case studies of the school shootings. *Aggressive Behavior*, 29, 202–214.

Lipsey, M.W. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *The Annals of the American Academy of Political and Social Sciences*, 587, 69– 81.

Lipsey, M.W. (2008). *Personal communication* (January 30).

Lipsey, M.W. & Wilson, D.B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Littell, J.H., Corcoran, J., & Pillai, V. (2008). *Systematic reviews and meta-analysis*. Oxford: Oxford University Press.

Martin, F.D.F., Martinez, M. del. C.P., & Tirado, J.L.A. (2005). Design, implementation and evaluation of a bullying prevention pilot program [Spanish: Diseno, aplicacion y evaluacion de un Programa Piloto para la Prevencion del Maltrato entre companeros]. *Revista Mexicana de Psicologia*, 22, 375–384.

* Melton, G.B., Limber, S.P., Flerx, V., Nation, M., Osgood, W., Chambers, J., Henggeler, S., Cunningham, P., & Olweus, D. (1998). *Violence among rural youth*. Final report to the Office of Juvenile Justice and Delinquency Prevention, Washington DC.

* Menard, S., Grotpeter, J., Gianola, D., & O'Neal, M. (2008). *Evaluation of Bullyproofing your School: final report*. Downloaded from NCJRS: http://www.ncjrs.gov/pdffiles1/ nij/grants/221078.pdf

Meraviglia, M.G., Becker, H., Rosenbluth, B., Sanchez, E., & Robertson, T. (2003). The Expect Respect Project: Creating a positive elementary school climate. *Journal of Interpersonal Violence*, 18, 1347–1360.

Moher, D., Schulz, K.F., & Altman, D. (2001). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Journal of the American Medical Association*, 285, 1987–1991.

Murray, D.M. & Blitstein, J.L. (2003). Methods to reduce the impact of intraclass correlation in group-randomized trials. *Evaluation Review*, 27, 79–103.

Mytton, J., DiGuiseppi, 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. (1992). Bullying among school children: Intervention and prevention. In R.D. Peters, R.J. McMahon, & V.L. Quinsey (Eds.), *Aggression and violence throughout the lifespan* (pp. 100–125). London: Sage.

* Olweus, D. (1993). Bully/victim problems among schoolchildren: Long-term consequences and an effective intervention program. In S. Hodgins (Ed.), *Mental disorder and crime* (pp. 317–349). Thousand Oaks, CA: Sage.

* Olweus, D. (1994a). Bullying at school: Basic facts and effects of a school based intervention program. *Journal of Child Psychology and Psychiatry*, 35, 1171–1190.
* Olweus, D. (1994b). Bullying at school: Basic facts and an effective intervention programme. *Promotion and Education*, 1, 27–31.
* Olweus, D. (1994c). Bullying at school: long-term outcome for the victims and an effective school-based intervention program. In L.R. Huesmann (Ed.), *Aggressive behavior: Current perspectives* (pp. 97–130). New York: Plenum.
* Olweus, D. (1995). Peer abuse or bullying at school: Basic facts and a school-based intervention programme. Prospects, 25 (1), 133–139.
* Olweus, D. (1996a). Bullying or peer abuse in school: Intervention and prevention. In G. Davies, S. Lloyd-Bostock, M. McMurran, & C. Wilson (Eds.), *Psychology, law and criminal justice: International developments in research and practice* (pp. 248–267). Berlin: Walter de Gruyter.
* Olweus, D. (1996b). Bullying at school: Knowledge base and effective intervention. *Annals of the New York Academy of Sciences*, 784, 265–276.
* Olweus, D. (1996c). Bully/victim problems at school: Facts and effective intervention. *Reclaiming Children and Youth: Journal of Emotional and Behavioral Problems*, 5 (1), 15–22.
* Olweus, D. (1997a). Bully/victim problems in school: Knowledge base and an effective intervention project. *Irish Journal of Psychology*, 18, 170–190.
* Olweus, D. (1997b). Bully/victim problems in school: Facts and intervention. *European Journal of Psychology of Education*, 12, 495–510.
* Olweus, D. (1997c). Tackling peer victimization with a school-based intervention program. In D.P. Fry & K. Bjorkqvist (Eds.), *Cultural variation in conflict resolution: Alternatives to violence* (pp. 215–232). Mahwah, NJ: Erlbaum.
* Olweus, D. (2004a). The Olweus Bullying Prevention Programme: Design and implementation issues and a new national initiative in Norway. In P.K. Smith, D. Pepler, & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 13–36). Cambridge: Cambridge University Press.
* Olweus, D. (2004b). Bullying at school: Prevalence estimation, a useful evaluation design, and a new national initiative in Norway. *Association for Child Psychology and Psychiatry Occasional Papers*, 23, 5–17.
Olweus, D. (2004c). *Continuation of the Olweus programme against bullying and antisocial behaviour*. Olweus-Gruppen; Mot mobbing (October).
* Olweus, D. (2005a). A useful evaluation design, and effects of the Olweus bullying prevention program. *Psychology, Crime and Law*, 11, 389–402.

* Olweus, D. (2005b). *Bullying in schools: Facts and intervention*. Paper presented at the IX International Meeting on Biology and Sociology of Violence, Under the Honorary Presidency of H.M. Queen Sofia, Valencia, Spain (October).

Olweus, D. (2008). *Personal communication* (June 5).

* 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.

* Ortega, R. & Del Rey, R. (1999). *The use of peer support in the S.A.V.E. project*. Paper presented at the IXth European Conference on Developmental Psychology, Spetses, Greece (September).

* Ortega, R., Del-Rey, R., & Mora-Mercan, J. A. (2004). SAVE Model: An anti-bullying 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.

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., Smith, P.K., & Rigby, K. (2004). Looking back and looking forward: Implications for making interventions work effectively. In P.K. Smith, D. Pepler, & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 307–324). Cambridge: Cambridge University Press.

Petticrew, M. & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*. Malden: Blackwell.

* Pintabona, Y.C. (2006). *Frequently bullied students: Outcomes of a universal school-based bullying preventive intervention on peer victimization and psychological health*. PhD Thesis, Curtin University of Technology, Bentley, Western Australia.

* Rahey, L. & Craig, W.M. (2002). Evaluation of an ecological program to reduce bullying in schools. *Canadian Journal of Counselling*, 36, 281–295.

* Raskauskas, J. (2007). *Evaluation of the Kia Kaha anti-bullying programme for students in years 5–8*. Wellington: New Zealand Police.

* Rican, P., Ondrova, K., & Svatos, J. (1996). The effect of a short, intensive intervention upon bullying in four classes in a Czech Town. *Annals of the New York Academy of Sciences*, 794, 399–400.

Rigby, K. (2002). *A meta-evaluation of methods and approaches to reducing bullying in preschools and early primary school in Austra-*

lia. Canberra: Attorney General's Department, Crime Prevention Branch.

Roland, E. (1989). Bullying: The Scandinavian tradition. In D. P. Tattum & D. A. Lane (Eds.), *Bullying in schools* (pp. 21–32). Stoke-on-Trent: Trentham Books.

Roland, E., & Idsoe, T. (2001). Aggression and bullying. *Aggressive Behavior*, 27, 446–462.

* Rosenbluth, B. Whitaker, D.J., Sanchez, E., & Valle, L.A. (2004). The Expect Respect Project: Preventing bullying and sexual harassment in US elementary schools. In P.K. Smith, D. Pepler & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 211–233). Cambridge: Cambridge University Press.

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. (2008). *Personal communication* (June 18).

* Salmivalli, C., Karna, A., & Poskiparta, E. (2007). From peer put-downs to peer support: A theoretical model and how it translated into a national anti-bullying program. In S.R. Jimerson, S.M. Swearer, & D.L. Espelage (Eds.), *The International Handbook of School Bullying*. Mahwah, New Jersey: Lawrence Erlbaum Associates (chapter submitted for review for publication).

* Salmivalli, C., Kaukiainen, A., Voeten, M., & Sinisammal, M. (2004). Targeting the group as a whole: The Finnish anti-bullying intervention. In P.K. Smith, D. Pepler & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 251–275). Cambridge: Cambridge University Press.

* Salmivalli, C., Kaukiainen, A., Voeten, M. (2005). Anti-bullying intervention: Implementation and outcome. *British Journal of Educational Psychology*, 75, 465–487.

Salmivalli, C., & Nieminen, E., (2002). Proactive and reactive aggression among school bullies, victims and bully-victims. *Aggressive Behavior*, 28, 30–44.

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton-Mifflin.

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. (1997). Bullying in schools: The UK experience and the Sheffield Anti-bullying Project. *Irish Journal of School Psychology*, 18, 191–201.

Smith, P.K., & Ananiadou, K. (2003). The nature of school bullying and the effectiveness of school-based interventions. *Journal of Applied Psychoanalytic Studies*, 5, 189–209.

Smith, P.K., Ananiadou, K., & Cowie, H. (2003). Interventions to reduce school bullying. *Canadian Journal of Psychiatry*, 48, 591–599.

Smith, P. K., Cowie, H., Olafsson, R. F., & Liefooghe, 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.) (2004a). *Bullying in schools: How successful can interventions be?* Cambridge: Cambridge University Press.

* Smith, P.K., Sharp, S., Eslea, M., & Thompson, D. (2004b). England: the Sheffield project. In P.K. Smith, D. Pepler & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 99–123). Cambridge: Cambridge University Press.

Spröber, N. (2006a). *Entwicklung und Evaluation eines Mehrebenen-Programms zur Prävention von „Bullying" an weiterführenden Schulen und zur Förderung der positiven Entwicklung von Schülern*. Dissertation der Fakultät für Informations- und Kognitionswissenschaften der Eberhard-Karls-Universität Tübingen zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften.

* 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.

Stevens, V., De Bourdeaudhuij, I. & Van Oost, P. (2001). Anti-bullying interventions at school: aspects of programme adaptation and critical issues for further programme development. *Health Promotion International*, 16, 155–167.

* Stevens, V., Van Oost, P., & De Bourdeaudhuij, I. (2000). The effects of an anti-bullying intervention programme on peers' attitudes and behaviour. *Journal of Adolescence*, 23, 21–34.

* Stevens, V., Van Oost, P., & De Bourdeaudhuij, I. (2001). Implementation process of the Flemish anti-bullying intervention and relation with program effectiveness. *Journal of School Psychology*, 39 (4), 303–317.

* Stevens, V., Van Oost, P., & De Bourdeaudhuij, I. (2004). Interventions against bullying in Flemish schools. In P.K. Smith, D. Pepler & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 141–165). Cambridge: Cambridge University Press.

Ttofi, M.M. & Farrington, D.P. (2008). Bullying: Short-term and long-term effects, and the importance of Defiance Theory in explanation and prevention. *Victims and Offenders*, 3, 289–312.

Twemlow, S.W., Fonagy, P., & Sacco, F.C. (2005). A developmental approach to mentalizing communities: II. The Peaceful Schools experiment. *Bulletin of the Menninger Clinic*, 69, 282–304.

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.

Warden, D., Moran, E., Gillies, J., Mayes, G., & Macleod, L. (1997). An evaluation of a children's safety training programme. *Educational Psychology*, 17, 433–448.

Weisburd, D., Lum, C.M., & Petrosino, A. (2001). Does research design affect study outcomes in criminal justice? *Annals of the American Academy of Political and Social Science*, 578, 50–70.

Welsh, B.C., Farrington, D.P., & Sherman, L.W. (Eds.) (2001). *Costs and benefits of preventing crime*. Boulder, Co: Westview Press.

* Whitaker, D.J., Rosenbluth, B., Valle, L.A., & Sachez, 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: Erlbaum.

* Whitney, I., Rivers, I., Smith, P.K., & Sharp, S. (1994). The Sheffield Project: Methodology and findings. In P.K. Smith & S. Sharp (Eds.), *School bullying: Insights and perspectives* (pp. 20–56). London: Routledge.

Wiefferink, C., Hoekstra, H., Beek, J.T., & Van Dorst A. (2006). Effects of an anti-bullying programme in elementary schools in the Netherlands. *European Journal of Public Health*, 16, 76–76.

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.

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.

Woods, T., Coyle, K., Hoglund, W., & Leadbeater, B. (2007). Changing the contexts of peer victimization: The effects of a primary prevention program on school and classroom levels of victimization. In J.E. Zins, M.J. Elias & C.A. Maher (Eds.), *Bullying, victimization, and peer harassment: A handbook of prevention and intervention* (pp. 369–388). New York: Haworth Press.

Other reports in this series

Closed-Circuit Television Surveillance and Crime Prevention
Authors: Brandon C. Welsh and David P. Farrington
A systematic review of the effects of closed circuit television surveillance based on 41 evaluations from five countries. Published in October 2007.
ISBN: 978-91-85664-79-5

Improved Street Lighting and Crime Prevention
Authors: David P. Farrington and Brandon C. Welsh
A systematic review of the effects on crime of improved street lighting based on 13 evaluations from the United States and United Kingdom. Published in October 2007.
ISBN: 978-91-85664-78-8

Effectiveness of Neigbourhood Watch In Reducing Crime
Authors: Trevor H. Bennett, Katy R. Holloway, David P. Farrington
A systematic review of the effects of neighbourhood watch based on 36 evaluations. Published in March 2008.
ISBN: 978-91-85664-91-7

The Influence of Mentoring on Reoffending
Authors: Darrick Jolliffe and David P. Farrington
A rapid evidence assessment of the effects of mentoring based on 18 evaluations. Published in March 2008.
ISBN: 978-91-85664-90-0

Effects of Early Family/Parent Training Programs on Antiscocial Behavior & Delinquency
Authors: Alex R. Piquero, David P. Farrington, Brandon C. Welsh, Richard Tremblay and Wesley G. Jennings.
A systematic review of effects of early family/parent training based on 55 evaluations. Published in June 2008.
ISBN: 978-91-85664-95-5

Effectiveness of Treatment in Reducing Drug-Related Crime
Katy R. Holloway, Trevor H. Bennett, David P Farrington
A systematic review of the effects of drug treatment programs on crime based on 75 evaluations. Published in October 2008.
ISBN: 978-91-86027-10-0

The reports can be ordered from:
Fritzes Kundservice, 106 47 Stockholm, Sweden
Phone: + 46(0)8 690 91 90, fax: + 46(0)8 690 91 91,
e-mail: order.fritzes@nj.se