Campbell Systematic Reviews
2009:6
First published: 15 December 2009
Last updated: 8 March 2010

School-Based Programs to Reduce Bullying and Victimization

David P. Farrington, Maria M. Ttofi

THE CAMPBELL COLLABORATION

Colophon

Colophon

Title School-Based Programs to Reduce Bullying and Victimization

Institution The Campbell Collaboration

Authors Farrington, David P.
Ttofi, Maria M.

DOI 10.4073/csr.2009.6

No. of pages 147

Last updated 15 December 2009

Citation Farrington, D. P., Ttofi, M. M. School-Based Programs to Reduce Bullying and Victimization.
Campbell Systematic Reviews 2009:6
10.4073/csr.2009.6

Copyright © Farrington et al.
This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Keywords

Support/Funding This systematic review was funded by the SFI Campbell, The Danish National Centre for Social Research and the National Institute of Justice (United States). The Campbell Collaboration funded the editing process.

Potential Conflicts of Interest Neither of the authors has any financial interest in any existing or planned anti-bullying programs.

Corresponding author David P. Farrington
Professor
Institute of Criminology,
Cambridge University
Sidgwick Avenue
Cambridge CB3 9DT
UK
(44) 1223-872555
dpf1@cam.ac.uk

Campbell Systematic Reviews

Editors-in-Chief	Mark W. Lipsey, Vanderbilt University, USA Arild Bjørndal, Norwegian Knowledge Centre for the Health Services & University of Oslo, Norway
Editors	
Crime and Justice	David B. Wilson, George Mason University, USA
Education	Chad Nye, University of Central Florida, USA Ralf Schlosser, Northeastern University, USA
Social Welfare	Julia Littell, Bryn Mawr College, USA Geraldine Macdonald, Queen's University, UK & Cochrane Developmental, Psychosocial and Learning Problems Group
Managing Editor	Karianne Thune Hammerstrøm, The Campbell Collaboration
Editorial Board	
Crime and Justice	David Weisburd, Hebrew University, Israel & George Mason University, USA Peter Grabosky, Australian National University, Australia
Education	Carole Torgerson, University of York, UK
Social Welfare	Aron Shlonsky, University of Toronto, Canada
Methods	Therese Pigott, Loyola University, USA Peter Tugwell, University of Ottawa, Canada

The Campbell Collaboration (C2) was founded on the principle that systematic reviews on the effects of interventions will inform and help improve policy and services. C2 offers editorial and methodological support to review authors throughout the process of producing a systematic review. A number of C2's editors, librarians, methodologists and external peer-reviewers contribute.

The Campbell Collaboration
P.O. Box 7004 St. Olavs plass
0130 Oslo, Norway
www.campbellcollaboration.org

Table of Contents

TABLE OF CONTENTS **3**

EXECUTIVE SUMMARY/ABSTRACT **5**
Background 5
Objectives 5
Search Strategy 5
Selection Criteria 6
Results 6
AUTHORS' CONCLUSIONS 6

1 INTRODUCTION **8**
1.1 Impetus for the Systematic Review 8
1.2 Definition of Bullying 9
1.3 Background 10
1.4 Objectives of the Review 11

2 METHODS **13**
2.1 Measuring the Effects of a Program 13
2.2 Criteria for Inclusion or Exclusion of Studies 15
2.3 Searching Strategies 17

3 RESULTS OF SEARCHES **19**
3.1 Studies Found 19
3.2 Included Evaluations 20

4 DESCRIPTIONS OF INCLUDED PROGRAMS **21**
4.1 Randomized Experiments 21
4.2 Before-After, Experimental-Control Comparisons 31
4.3 Other Experimental-Control Comparisons 44
4.4 Age-Cohort Designs 47

5 ANALYSIS OF INCLUDED EVALUATIONS **53**
5.1 Key Results 53
5.2 Analysis of Effect Sizes 57
5.3 Effect Size Versus Research Design 60

6 CODING OF STUDY FEATURES **62**
6.1 Key Features of the Evaluation 62

6.2 Key Elements of the Program 63
6.3 Effect Size Versus Study Features 66
6.4 Publication Bias Analyses 68

7 CONCLUSIONS 69
7.1 Summary of Main Findings 69
7.2 Policy Implications 69
7.3 Implications for Future Research 73

8 OTHER TOPICS 76
8.1 Acknowledgements 76
8.2 Conflict of Interest 76
8.3 Plans for Updating the Review 76

9 TECHNICAL APPENDIX: EFFECT SIZE AND META-ANALYSIS 78

10 REFERENCES 86

11 APPENDIX/TABLES AND FIGURES 102
11.1 Table 1: List of Databases Searched 102
11.2 Table 2: List of Journals Searched from 1983 until May 2009 103
11.3 Table 3: Categorization of Reports Based on their Relevance to the Present Review 104
11.4 Table 4: Percentage of Reports and Evaluations[a] of Programs Within Each Category 104
11.5 Table 5: 89 Reports of 53 Different Evaluations* 105
11.6 Table 6: Key Features of 53 Evaluations 108
11.7 Table 7: Key Results of 44 Program Evaluations 126
11.8 Table 8: Effect Sizes for Bullying 135
11.9 Table 9: Effect Sizes for Victimization 137
11.10 Table 10: Units of Random Allocation* 138
11.11 Table 11: Program Elements and Their Frequency 139
11.12 Table 12: Significant Relationships with Bullying 140
11.13 Table 13: Significant Relationships with Victimization 141
11.14 Table 14: Results of Weighted Regression Analyses 142
11.15 Figure 1: Number of Reports in Each Category Within Year Period 143
11.16 Figure 2: Key Features of the Evaluation 144
11.17 Figure 3: Intervention Components 145
11.18 Figure 4: Forest Graph for Bullying 146
11.19 Figure 5: Forest Graph for Victimization 147
11.20 Figure 6: Funnel Plot for Bullying 148
11.21 Figure 7: Funnel Plot for Victimization 148

Executive Summary/Abstract

BACKGROUND

School bullying has serious short-term and long-term effects on children's physical and mental health. Various anti-bullying programs have been implemented world wide and, more rarely, evaluated. Previous narrative reviews, summarizing the work done on bullying prevention, as well as previous meta-analyses of anti-bullying programs, are limited. 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.

OBJECTIVES

This report presents a systematic review and meta-analysis of the effectiveness of programs designed to reduce school bullying perpetration and victimization (i.e. being bullied). The authors indicate the pitfalls of previous reviews and explain in detail how the present systematic review and meta-analysis addresses the gaps in the existing literature on bullying prevention.

SEARCH STRATEGY

In the present report, we go beyond 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 May 2009; searching for international evaluations in 18 electronic databases and in languages other than English; and focusing only on programs that are specifically designed to reduce bullying and not aggressive behavior (i.e. the outcome variables specifically measure bullying). Leading researchers in the area of school bullying were also contacted via e-mail.

SELECTION CRITERIA

Studies were included in this review if they evaluated the effects of an anti-bullying program by comparing an experimental group who received the intervention with a control group who did not. The word 'experimental' here refers to students who received the program and does not necessarily imply randomization. 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) quasi-experimental 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. Both published and unpublished (e.g. PhD theses) reports were included. Reports concerning an evaluation of a program had to clearly indicate that bullying or victimization were included as outcome measures. Bullying and victimization could be measured using self-report questionnaires, peer ratings, teacher ratings, or observational data.

RESULTS

We found a total of 622 reports that were concerned with bullying prevention. The number of reports on anti-bullying programs and on the necessity of tackling bullying has increased considerably over time. Only 89 of these reports (describing 53 different program evaluations) could be included in our review. Of the 53 different program evaluations, only 44 provided data that permitted the calculation of an effect size for bullying or victimization. Our meta-analysis of these 44 evaluations showed that, overall, school-based anti-bullying programs are effective in reducing bullying and victimization (being bullied). On average, bullying decreased by 20% – 23% and victimization decreased by 17% – 20%. The effects were generally highest in the age-cohort designs and lowest in the randomized experiments. It was not clear, however, that the randomized experiments were methodologically superior in all cases, because sometimes a very small number of schools (between three and seven) were randomly assigned to conditions, and because of other methodological problems such as differential attrition. Various program elements and intervention components were associated with a decrease in both bullying and victimization. Work with peers was associated with an increase in victimization. We received feedback from researchers about our coding of 40 out of 44 programs. Analyses of publication bias show that the observed effect sizes (for both bullying and victimization) were based on an unbiased set of studies.

AUTHORS' CONCLUSIONS

Results obtained so far in evaluations of anti-bullying programs are encouraging. The time is ripe to mount a new long-term research strategy on the effectiveness of these programs, based on our findings. The main policy implication of our review is

that new anti-bullying programs should be designed and tested based on the key program elements and evaluation components that we have found to be most effective. We recommend that a system of accrediting anti-bullying programs should be developed, supervised by an international body such as the International Observatory on Violence in Schools.

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, 2008a) 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 programs 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 programs, 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 program more effective than another? What intervention elements can predict the success of a program in reducing school bullying? These questions have inspired our research.

Our systematic review follows 26 years of intervention research (from 1983 to the end of May 2009) and is based on extensive literature searches. Our meta-analytic approach offers a quantitative summary of effect sizes of anti-bullying programs 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 primarily involves imbalance of power and repeated acts.

Our review is also concerned with victimization (being bullied). The majority of evaluations of bullying prevention programs aimed to reduce both bullying and victimization. We report results for these outcome measures (i.e. bullying and victimization) separately. With few exceptions (e.g. Menesini et al., 2003), most evaluations did not report other outcome measures such as the prevalence of bully-victims (i.e. children who both bully and are bullied by others). Consequently, our review is restricted to the effectiveness of programs to reduce bullying and victimization only.

Bullying is a type of aggressive behavior (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 rumors spread about you, having belongings taken away, teasing and threatening (Baldry & Farrington, 1999). Cyber bullying is a recent development (Smith et al., 2008) and it may be too recent to have high quality evaluations of school-based programs that target this form of bullying. Our aim is to review programs that are specifically intended to prevent or reduce school bullying, not programs that are intended to prevent or reduce school aggression or violence. It is possible that programs designed to reduce school aggression or other problem behaviors also reduced school bullying, and vice versa; however, as much as possible, we have focused specifically on bullying.

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

The first large-scale anti-bullying program was implemented nationally in Norway in 1983. A more intensive version of the national program 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 program. Since then at least 15 other large-scale anti-bullying programs, some inspired by Olweus and some based on other principles, have been implemented and evaluated in at least 10 other countries. Baldry and Farrington (2007) reviewed sixteen major evaluations in eleven different countries, of which five involved an uncontrolled methodological design. They concluded that eight of them produced desirable results, two produced mixed results, four produced small or negligible effects, and two produced undesirable results. The present review includes many more evaluations (i.e. 53 in total) and attempts to investigate the effectiveness of program components. Special efforts were made to avoid problems arising from duplicate publications. For example, the Flemish Anti-bullying Program was evaluated once and the results were disseminated in four publications. However, in contrast to previous reviews (e.g. Merrell, Gueldner, Ross and Isava, 2008), we carefully coded it as one evaluation. As another example, study findings on the effectiveness of the Olweus Bullying Prevention Program were disseminated in 22 publications. However, the program was tested in only eight separate evaluations.

American research is generally targeted on school violence or peer victimization rather than bullying. There are a number of existing reviews of school violence programs and school-based interventions for aggressive behavior (e.g. Howard, 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 programs that are explicitly designed to reduce bullying and that explicitly measure bullying.

The most informative single source of reports of anti-bullying programs is the book edited by P.K. Smith and his colleagues (2004a), which contains descriptions of 13 programs 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.D. Smith, Schneider, Smith and Ananiadou (2004), who summarized effect sizes in 14 whole-school anti-bullying programs, and by Vreeman and Carroll (2007), who reviewed 26 school-based programs. These two prior reviews are of high quality. However, neither carried out a full meta-analysis measuring weighted mean effect sizes and correlations between study features and effect sizes.

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 and with only 15 programs specifically concerned with bullying. Another meta-analytic review was published by Ferguson, San Miguel, Kilburn and Sanchez (2007). However, this included searches in only one database (PsycINFO) for articles published between the years 1995 and 2006 (p. 406). It included outcome variables that measured 'some element of bullying behavior or aggression toward peers, including direct aggressive behavior toward children in a school setting' (p. 407). The latest meta-analytic review was completed by Merrell et al. (2008). However, this included searches in only two databases (PsycINFO and ERIC) for studies only published in English, and it included a wide range of outcome measures; there were only eight studies where the outcome was self-reported bullying and only ten studies where the outcome was self-reported victimization.

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 May 2009.
- 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 programs that are specifically designed to reduce bullying and not aggressive behavior (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 programs in reducing school bullying. Our aim is to locate and summarize all the major evaluations of programs in developed countries. Bullying has been studied in (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 to include evaluations (if available) in all these countries. We aim to measure effect sizes in each evaluation and to investigate which features (e.g. of programs and students) are related to effect sizes. We aim to make recommendations about which components of programs are most effective in which circumstances, and hence about how future anti-bullying programs might be improved. We also aim to describe in detail anti-bullying programs that have been evaluated using a controlled methodological design. We also aim to make recommendations about how the design and analysis of evaluations of anti-bullying programs might be improved in future. However, we are of course limited by the information that is available in published and unpublished reports.

2 Methods

2.1 MEASURING THE EFFECTS OF A PROGRAM

How can the effects of an anti-bullying program on bullying and victimization 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 program 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 pre-test measurement causes a change in the post-test 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 post-test 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-history 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 (the counter-factual), in order to estimate what would have happened in the absence of the intervention. If children were merely measured before and after receiving the intervention, it would be impossible to disentangle the impact of the program 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 probably have decreased because of aging effects alone.

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

In research on anti-bullying programs, schools or school classes, rather than children, are usually randomly assigned to receive the program. In some evaluations, a very small number of schools (between three and seven) were randomly assigned, threatening statistical conclusion validity. It is not true in all cases that randomized experiments on anti-bullying programs 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.

The word 'experimental' as used here is equivalent to 'treated' and does not necessarily imply randomization. It refers to students who received the program. Non-randomized comparisons of experimental and control classes with no prior measures of bullying are clearly inferior to non-randomized comparisons with prior measures. Where there are no prior measures of bullying, it is important to include some pre-test 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 after the intervention 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. Overall, the experimental-control comparisons and age-cohort designs might be regarded by some researchers as methodologically inferior to the randomized experiments and experimental-control/before-after designs, but all designs have advantages and problems. These are the best four designs that have been used to evaluate the effects of anti-bullying programs, and we will give credence to all of them in providing useful information about the effectiveness of anti-bullying programs.

2.2 CRITERIA FOR INCLUSION OR EXCLUSION OF STUDIES

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

(a) The study described an evaluation of a program designed specifically to reduce school (kindergarten to high school) bullying. Studies of aggression or violence were 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 program 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, peer ratings, teacher ratings, or observational data.

(d) The effectiveness of the program was measured by comparing students who received it (the experimental condition) with a comparison group of 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) pre-test 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 program, but Olweus compared students of age X after the program (the experimental condition) with different students of the same age X in the same schools before the program (the control condition). We include this kind of age-cohort design because arguably the experimental and control students are comparable (at least in age and in attending the same schools).

(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 programs 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 evaluations of programs involving controlled methodological designs were included in the systematic review but not in the meta-analysis because they did not provide enough data to allow us to calculate an effect size (see Table 6, page 107). Some other controlled studies are included (e.g. Salmivalli, Karna, & Poskiparta, 2009)[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.

[1] Personal communication with Christina Salmivalli via e-mail (June 18, 2008) and with Antti Karna (May 22, 2009).

In our review published by the Swedish National Council for Crime Prevention (Ttofi, Farrington, & Baldry, 2008), we set a minimum initial sample size of students (total in experimental and control conditions) of 200 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 correlations 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 pre-test and post-test) 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. In this Campbell review, we include all studies irrespective of sample size, but we distinguish the smaller studies (less than 200 students) in our tables (8 and 9) of effect size.

In the Swedish review, in the interests of maximizing comparability, we only included measures of bullying based on self-reports by students. These are the most common measures used in the evaluation of anti-bullying programs, and we believe that they are the most useful measures (see e.g. Solberg & Olweus, 2003). In this Campbell review, however, we include measures of bullying based on peer and teacher reports. In the very rare cases where more than one measure was reported (e.g. Alsaker & Valkanover, 2001), we chose, first self-reports, second, peer reports, and third, teacher reports.

2.3 SEARCHING STRATEGIES

(a) We started by searching for the names of established researchers in the area of bullying prevention (e.g. Australia, Ken Rigby; Canada, Debra Pepler; England, Peter K. Smith; Finland, Christina Salmivalli; Spain, Rosario Ortega; Norway, Dan Olweus). This searching strategy was used in different databases in order to initially obtain as many evaluations of known research programs as possible.

(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, see appendix page 101). In all databases, the same key words were used with different combinations. More specifically:

- Bully/Bullies/Anti-Bullying/Bully-Victims/Bullying
- AND: School
- AND: Intervention/Program/Outcome/Evaluation/Effect/Prevention/Tackling /Anti-bullying

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 (List of Journals Searched from 1983 until May 2009, see appendix page 102) gives a list of the journals that we have hand-searched, either online or in print, from 1983 until the end of May 2009. 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 2008, 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. Ciucci & Smorti, 1998; Gini, Benelli, & Casagrande, 2003; 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 622 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 reports were categorized based on a relevance scale that we constructed (Table 3: Categorization of Reports Based on Their Relevance to the Present Review, see appendix page 103).

Table 4 (Percentage of Reports and Evaluations of Programs Within Each Category, see appendix page 104) shows the percentage of studies within each category. It also shows the number of evaluations that were included in the meta-analysis. The vast majority of reports (40.7%) were somewhat relevant (category 2), making general suggestions about reducing bullying or, more rarely, reviewing anti-bullying programs. With regard to the reports that we were not able to obtain (16, or 2.6%), most of them were Masters or PhD theses. Moving on to the obtained reports, only 89 (14.3%) were eligible for inclusion in our Campbell review (categories 5 and 6). It is regrettable that a fair number of evaluations of anti-bullying programs were excluded from our review (category 4: 11.4%) because of their (uncontrolled) methodological design.

The number of reports concerned with anti-bullying programs has increased markedly over time, as indicated in Figure 1 (see appendix, page 142). The total time period was divided into 5-year chunks as follows: 1983-1987, 1988-1992, 1993-1997, 1998-2002 and 2003-2009.

The most obvious increase of interest in implementing and evaluating bullying prevention programs occurred in the latest period. In the last six years or so (up to the end of May 2009), 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.

Of the 89 reports (of 53 evaluations) that are eligible for inclusion in our comprehensive Campbell review, 62 reports involved 32 evaluations of programs with a sample size more than 200, and 15 reports involved 12 evaluations of

programs with a sample size less than 200. Twelve reports of nine evaluations did not provide enough data to allow the calculation of an effect size and were, therefore, not included in the meta-analysis.

3.2 INCLUDED EVALUATIONS

The 89 reports of 53 evaluations 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 5 (see appendix page 104) lists the 89 reports included in the present systematic review. For each evaluation, all relevant reports are presented so that readers can follow up according to their own interests. Within each of the four categories of research design, reports were grouped based on the program evaluation 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 program by Whitney, Rivers, Smith and Sharp (1994) was placed in category 6, because information was provided about the effectiveness of the program which was evaluated using an age-cohort design (with schools being the unit of analysis). 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 schools that received the program. 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 program), the report by Stevens, De Bourdeaudhuij and Van Oost (2001) was placed in category 2 because it reviewed several anti-bullying programs and did not present outcome data on one specific program. Table 6 (see appendix, page 107) summarizes key features of the 53 different evaluations that are included in this report. Recall that 9 evaluations did not provide enough data to allow the calculation of an effect size. These are specified in Table 5 (see appendix, page 104), which also presents the reason for exclusion of the nine evaluations.

4 Descriptions of Included Programs

Next we provide an in-depth narrative review of the programs that have been evaluated in the past and that were included in our meta-analysis. These descriptions are based on the best available data and on the information provided in reports evaluating the intervention (categories 5 and 6), rather than in reports describing the program (category 3). The rationale underlying this decision refers to the fact that the way in which a program 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 programs we have attempted to contact the evaluators of the program. We have received positive feedback from researchers regarding the way we coded 40 out of the 44 evaluations (all except: Ciucci and Smorti, 1998; Pagliocca et al., 2007; Raskauskas, 2007; Rican et al., 1996).

4.1 RANDOMIZED EXPERIMENTS

4.1.1 Bulli and Pupe (Italy)

Bulli and Pupe' was an intervention program concerned with bullying and family violence. The program, 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', analyzed 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 program was in the first place delivered in three days by experts who, together with teachers, discussed about bullying, read the booklet and analyzed its content. The program was taken over by teachers who once a week created a facilitation group and allowed children to discuss any problems they encountered with their peers. The program was more effective with secondary students because it required its participants to have good interpersonal and cognitive skills (Baldry & Farrington, 2004, p. 4).

4.1.2 Project Ploughshares Puppets for Peace (Canada)

Project Ploughshares Puppets for Peace (P4 program) was an anti-bullying program that aimed to educate elementary school students about bullying and conflict resolution (Beran & Shapiro, 2005, p. 703). The P4 program used puppets and a 30-minute script. Using three-feet, hand-and-rod puppets, two puppeteers enacted a story that involved direct and indirect bullying, as well as a successful resolution to this scenario. These behaviors occurred among two female puppets and a male puppet friend.

After watching the play, students were invited to identify the bullying behaviors. During the discussion, four main strategies –presented as '4 Footsteps'– to deal with bullying were suggested to pupils: a) ignore, b) say stop, c) walk away and d) get help. The show took approximately 45 minutes and aimed to increase children's awareness about which behaviors could be categorized as bullying and to show various strategies that children who were bullied and/or who witnessed bullying could use to discourage it (Beran & Shapiro, 2005, p. 703).

4.1.3 Short Video Intervention (England)

This anti-bullying strategy, involved a single viewing of an anti-bullying video, entitled Sticks and Stones, and aimed to examine its effects on secondary school students' views of, and involvement in, bullying. The program aimed to examine both attitudes toward bullying and the actual behavior since 'it would not be unreasonable to propose that these attitudes will influence actual behavior' (Boulton & Flemington, 1996, p. 334). The program involved only one school that had no prior anti-bullying policy.

The video presented pupils (either in groups or on their own) talking about bullying, their views about this phenomenon and their personal experiences of bullying. The

video also involved a number of bullying scenes (see Boulton & Flemington, 1996, p. 337 for examples).

4.1.4 Friendly Schools (Australia)

'Friendly Schools' was a theoretically grounded program. Its educational techniques (e.g. role modeling, drama activities, skills training, etc.) were based on notions derived from Social Cognitive theory, the Health Belief Model and Problem Behavior theory (Cross et al., 2004, p. 191). An interesting aspect of this program 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 program 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 organized 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 activity. 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 behavior.

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 behavior; to help students to become assertive and talk about bullying with teachers and parents; and to promote peer and adult discouragement of bullying behavior.

Finally, the Friendly Schools program 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.

4.1.5 S.S. GRIN (USA)

The Social Skills Group Intervention (S.S.GRIN) was a school-based program 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 program 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 behaviors in an effort to help children to regulate their own emotions as well as enhance their coping skills.

Overall, the program was a combination of social-learning and cognitive-behavioral 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, modeling 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 counselor and an intern, who were trained and supervised by one of the program instructors (De Rosier & Marcus, 2005, p. 143).

4.1.6 Dutch Anti-Bullying Program

The anti-bullying initiative in the Netherlands was inspired by the Olweus program (Fekkes et al., 2006, p. 639). The program was specifically designed to tackle bullying behavior by involving teachers, parents and students. It offered a two-day training session for teachers in order to inform them about bullying behavior 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 program 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 program 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.

4.1.7 SPC and CAPSLE Program (USA)

This evaluation compared the effects of two intervention packages with a treatment-as-usual condition (Fonagy et al., 2009). Nine schools were randomly allocated to the two experimental and one control (Treatment As Usual) conditions after a stratified allocation procedure, which was used to stratify schools based on the percentage of low-income students (indicated by students' free- and reduced-lunch status). In the experimental conditions, the full intervention was offered for two years (the efficacy phase) with a limited third year of intervention (the maintenance phase).

The first experimental condition involved a 'School Psychiatric Consultation' (SPC), a manualized protocol that aims to address mental health issues of children with disruptive behavioral problems, internalizing problems, or poor academic performance. SPC was a school-level intervention focused on individual children. Three child psychiatry residents, supervised biweekly by a senior child psychiatrist, delivered mental health consultation following the SPC manual for four hours per week. The psychiatric residents attended weekly school resource meetings and consulted directly with teachers, parents and other school personnel, through classroom observations and meetings, providing 140 consultations for 65 students in year 1 and 97 consultations for 45 students in year 2.

The second experimental condition involved the implementation of CAPSLE ('Creating a Peaceful School Learning Environment'), a manualized psychodynamic approach addressing the co-created relationship between bullies, victims and bystanders. In contrast to SPC, CAPSLE represents a whole-school intervention approach. It aimed to modify the educational and disciplinary school climate. A CAPSLE team drawn from school staff in the pilot project led implementation in the two intervention years using a training manual. In year 1, teachers received a day of group training, students received nine sessions of self-defense training, and the CAPSLE team consulted with school staff monthly. Year 2 started with a school-wide half-day refresher self-defense course, and consultation continued with counselors, teachers and adult/peer mentor programs. In year 3 (the maintenance phase), self-defense training continued as in year 2.

CAPSLE includes several anti-bullying materials that can be used by teachers such as a Teacher Discipline Manual (used in the teacher training), a Student Workbook, Buttons and Magnets and Patches (used as a way of reinforcing of desirable student behavior), Parent Warning Notes (notifying parents about specific problem behavior of the child) as well as anti-bullying videos that can be used during the physical education lessons (and videos that can be used by parents). CAPSLE also

includes the Gentle Warrior Program, a 12-week curriculum specifically designed for physical education teachers. For CAPSLE, intervention fidelity was assessed using a teacher self-report measure that required teachers to state the frequency with which various CAPSLE program components were implemented.

4.1.8 Steps to Respect (USA)

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 program included staff and family training manuals, a program 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 victimization, 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 behavior, 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 counselors, administrators and teachers. There were also videos accompanying the program. With regard to staff training, there were two levels of training: all school staff received an overview of the program goals and principal aspects of the program (program guide). Teachers, counselors and administrators received additional training in how to coach students involved in bullying, based on behavioral skills training, cooperative 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 program 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.

4.1.9 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 empathy 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 program, 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 program includes a two-hour classroom-based discussion of bullying (offered by teachers) using activities from an anti-bullying workbook written by Murphy and Lewers (2000).

4.1.10 Youth Matters (USA)

The Youth Matters program 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 behaviors 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 program, 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 program 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 anti-social behavior; and d) social, cognitive and emotional skills that can 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.

4.1.11 KiVa (Finland)

The name of this 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 program, 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 program 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 program 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 program 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 challenging 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 program 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 program 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 program 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 program 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 organized and provided.

4.1.12 Behavioral Program for Bullying Boys (South Africa)

This program targeted male youth, from a sub-economic collared suburb, involved in bullying. The program was based on the findings of an in-depth needs assessment within three schools and targeted a specific number of male students aged sixteen who (based on the results of the questionnaire that had been administered) were 'considered to be a serious threat to the harmonious functioning of everyday school life' (Meyer & Lesch, 2000, p. 59). The theoretical basis of the program could be found in the Social Interactional Model for the development of aggression (Meyer & Lesch, 2000, p. 61) and involved a behavioral approach for tackling the problem of bullying. The program was implemented by psychology students for ten non-consecutive weeks, with twenty hour-long sessions held twice weekly at the school, during school hours.

The components of the 17-session behavioral program included homework tasks, modeling, self-observation, role-plays, and a token economy system for reinforcing positive behaviors. According to the program designers 'the chief contingency for behavioral change was the token economy system, using Wonderland Games tokens, chocolates and cinema tickets as rewards for non-bullying behavior' (Meyer & Lesch, 2000, p. 62). Each participant was monitored by himself and by a 'buddy' who was selected in each session prior to the monitoring. Each session included an opportunity for feedback on the students' progress in the week, a discussion of a

relevant applied topic, role-playing, games and drawing. The program designers pointed out the limitations of the intervention strategy. As they indicate (Meyer and Lesch, 2000, p. 67) 'the program was too short and structured to address the issues that were disclosed in sessions, as the severity of the nature of the aggression in the schools and vast social problems was seriously underestimated'.

4.1.13 Expect Respect (USA)

Expect Respect was a school-based program 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 program targeted the involvement of all members of the school community in recognizing 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 program 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 program, it was jointly led by Safe Place Staff and teachers or school counselors (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, counselors, 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 program 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 counselors by Safe Place staff. School counselors 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.

4.1.14 Pro-ACT + E Program (Germany)

Pro-ACT + E was a universal, multidimensional program that aimed to prevent bullying in secondary schools (Sprober et al., 2006). It involved a cognitive-behavioral approach to the problem of bullying and victimization by building up pro-social behavior. The program was universal: it did not involve specific work with perpetrators or victims of bullying. However, it included both teacher and parent training and a two-hour classroom discussion with students about violence problems. The program offered curriculum materials that aimed to increase awareness in relation to the problem of bullying and placed emphasis on specific issues such as classroom management and classroom rules against bullying.

4.2 BEFORE-AFTER, EXPERIMENTAL-CONTROL COMPARISONS

4.2.1 Be-Prox Program (Switzerland)

The Be-Prox program was specifically designed to tackle bullying and victimization among kindergarten students. According to Alsaker and Valkanover (2001, pp. 177-178) 'the somewhat higher adult-children ratio, the interest of preschool teachers in socialization, the greater flexibility as to scheduling and teaching, and the admiration of many preschoolers for their teachers are ideal conditions for the implementation of preventive programs against bully/victim problems'. The basic principle of Be-Prox was to enhance preschool teachers' capacity to handle

bully/victim problems (Alsaker, 2004, p. 291). The program engaged teachers in an intensive focused supervision for approximately four months. Central features of Be-Prox were the emphasis on group discussions, mutual support and co-operation between consultants and teachers and between teachers and parents (Alsaker, 2004, pp. 292-293).

The teacher training was provided in six steps (Alsaker, 2004; figure 15.1, p. 292). Initially, teachers were given information about victimization (step 1) and the implications of this information were discussed (step 2). During the third step specific implementation tasks were introduced and the teachers worked in groups in preparation for the practical implementation (step 4). After this preparation, teachers implemented specific preventive elements in the classroom (step 5) for a specific period of time. After that, teachers met and discussed their experiences of the implementation of the preventive measures (step 6).

In eight meetings over a four-month period, issues related to the prevention of bullying were addressed. The main purpose of the first meeting was sensitization. Teachers were asked to describe any possible bully/victim problems in their schools and were then given information about bullying and other types of aggressive behavior. They were also presented with the main principles of the program. The importance of contact between kindergarten teachers and children's parents was also emphasized and teachers were advised to consider the possibility of organizing a meeting with parents. In the second meeting, the importance of setting limits and rules to preschool children was discussed. Teachers were invited to elaborate some behavior codes in their classroom in collaboration with the children and to be ready to present them during the third meeting. Also, as a second homework task, teachers were asked to organize a parent meeting.

During the third meeting, teachers discussed their experiences of implementing classroom rules against bullying. The main focus of this meeting was the need for consistent teacher behavior, the difference between positive and negative sanctioning and the use of basic learning principles in the classroom. The main focus of the fourth session was on the role and responsibility of children who were not involved in bullying and of bystanders in the prevention of victimization. Teachers were asked to draw some kind of personality profiles of passive and aggressive victims and of bullies and to present them to the rest of the group. After this task, teachers were presented with research findings regarding the characteristics of children who were or were not involved in bullying. As a homework task for the next meeting, teachers were asked to systematically observe non-involved children and to develop some means of involving them in the prevention of victimization.

During the fifth meeting, research-based information about motor development and body awareness among preschool children was presented to teachers. A discussion

between teachers and program researchers of children's self-perceptions of strength, of peers' perceptions of strengths of victims of bullies, and other motor characteristics of children, aimed to yield important insights. The overall discussion and exchange of information among teachers aimed to promote teachers' understanding about how to change these perceptions within the classroom setting. Specific goals to be achieved within the classroom were clearly set, such as training in empathy and body awareness among children, participation and involvement of non-involved children and talks with all the children about the situation in their kindergarten. During the sixth meeting, time was given to reflect on the goals formulated at the beginning of the prevention program. Teachers were also given time to discuss their experiences with implementing the goals of the fifth meeting within the classroom settings. The last two meetings followed a similar format, with time given for reflection on goals achieved, problems dealt with, and an overall evaluation of the program.

4.2.2 Greek Anti-Bullying Program

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

The program 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 program (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 behaviors (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 student's 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 behaviors), 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.

4.2.3 Seattle Trial of the Olweus Program (USA)

The Olweus Bullying Prevention Program (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 program was in absolute concordance with the Olweus program 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 program 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 program started with an 'official start date' during which a school assembly took place aiming to present the overall program to students, introduce the basic concepts and raise enthusiasm among students. The core components of the program 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 organized with the goal of fostering collaboration in implementation efforts. School anti-bullying rules were presented to students that set clear guidelines about the students' behavior that was expected within the school. School surveillance was a crucial element of the anti-bullying program. 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 behavior in the school. Teachers in the intervention schools received teacher training.

The program 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 program since students' behavior could not be seen as fragmented: socially acceptable forms of behavior should be positively reinforced within and outside the school community.

4.2.4 Dare to Care; Bully Proofing Your School Program (Canada)

'Dare to Care; Bully Proofing Your School' was a modification of the 'Bully Proofing Your School' program (Beran et al., 2004, p. 103), which in turn was modeled on the Olweus Program. This anti-bullying program placed emphasis on clinical support to victims and perpetrators of bullying in the form of individual and group counseling. It also enabled collaboration with community services. The essence of the program was to encourage accountability for creating solutions among all parties involved in the education system (Beran et al., 2004, p. 104).

The program included several steps. Program facilitators provided to school personnel information and training on issues related to bullying in schools (in a full-day professional development workshop). This workshop aimed to ensure that the program principles would be reflected in the overall curriculum and would be sustained over time. Information was also given to parents. Then, students, parents and school staff collaborated in the development of a school anti-bullying policy. This policy had the aim of identifying caring and aggressive behaviors and consequences of those behaviors, but with a focus on reparation rather than punishment. The anti-bullying policy was posted throughout the school. Finally, the program involved the implementation, on behalf of the teachers, of a classroom curriculum that educated children about the nature of bullying and strategies to avoid victimization. The curriculum included discussion, role-plays, artwork, books, videos and skits presented to school staff, parents and other children.

4.2.5 Progetto Pontassieve (Italy)

The program 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 the problem of bullying was and what its characteristics were. The second year of the intervention included a counseling 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 cooperate 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 overtake 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. Transtheoretical-based Tailored Anti-bullying Program (USA)

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

Other elements of the program included: a) a 10-page family guide, sent to children's homes, which provided brief information about the multi-media program 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.

4.2.6 Social Skills Training (SST) Program (England)

Social Skills Training was a program specifically designed to support 'chronic victims' of bullying (Fox and Boulton, 2003, p. 237). The general aim of the program was to help children improve their social skills, therefore reducing a child's individual risk of victimization (Fox and Boulton, 2003, p. 234). The program involved an eight-week course during which children learnt how to use both problem-solving and relaxation skills, how to think positively, how to modify their non-verbal behavior and how to use some verbal strategies such as 'fogging' and 'mirroring' (Fox and Boulton, 2003, p. 235).

During the program, victims of bullying were gathered in groups of five to ten and were exposed to the aims of the program for one hour per week. Two trainers delivered the one-hour sessions throughout the program. The first week was dedicated to children introducing each other and listening to each other's problem. The next two sessions dealt with issues of friendship and aimed to help children form strong friendships (e.g. having conversations; asking to join in), whilst the

fourth session dealt with issues of body language: teaching children how to modify their non-verbal behavior in a way that would protect them from being victimized. During the fifth session children learned how to be assertive whilst in the next two sessions children were taught how to deal with the bully. The eighth session signaled the end of the program.

4.2.7 Stare bene a scuola: Progetto di prevenzione del bullismo (Italy)

This intervention was based on the curriculum activities and the whole-school approach because it tried to involve all people in a school (Gini et al., 2003). The program was delivered to 6 schools and included several activities. Teachers were first trained in three days on 'cooperative learning' and in particular on the Jigsaw technique. Teachers then had an on-going supervision once every fifteen days. The intervention in the class lasted 4 months with two meetings a week. The intervention was directed towards the following areas: a) awareness of the body and what it feels; b) emotional awareness; and c) bullying awareness. These areas were dealt with in each of the sessions, starting from the first one. For each thematic area, several activities were conducted and several methods were used.

4.2.8 ViSC Training Program (Germany)[2]

The Viennese Social Competence Training program aimed to provide students 'with systematic theoretically-based guidance in becoming responsible and competent actors in conflict situations' (Atria and Spiel, 2007, p. 179). It was specifically designed for disadvantaged adolescents aged fifteen to nineteen who were considered at risk for future problems (Atria & Spiel, 2007, p. 179). The theoretical basis of the programs drew its main ideas from social information processing theory and from research that approached the problem of bullying as a group phenomenon (Gollwitzer et al, 2006, p. 126).

The ViSC program consisted of thirteen lessons which were divided into three phases: a) impulses and group dynamics; b) reflection; and c) action. The first phase, entitled 'impulses and group dynamics', consisted of six lessons and the main aim was to enhance students' competence in dealing with critical situations by teaching them how to look at social situations from different perspectives using vignette stories, discussions and role-plays. The second phase, reflection, involved one lesson during which pupils reflected on what had been learned in the first phase of the program. The last phase, action, consisted of six lessons during which the trainer asked students to define how they wanted to benefit from the remaining lessons. The trainer collected students' individual ideas, evaluated them and –along with the students– put them in practice in alignment with the global goal of the program: enhancing pupils' social competence. The third phase of the program was

[2] For the evaluation that is included in the meta-analysis. The program was also evaluated in Austria, but due to lack of data it was not possible for us to include the specific evaluation in the meta-analysis.

flexible and it could involve several projects suggested by pupils such as a movie production, a work of art, the organization of a party, and so on. This flexibility was allowed and was, in fact, a main feature of ViSC because organizing such projects 'involves a variety of critical situations, in which alternative, non-aggressive response options can be probed, rehearsed, and evaluated for success' (Gollwitzer et al, 2006, p. 126).

Based on the design of the program, the training of students was conducted by trainers other than their teachers. The trainers participated in instruction workshops and were also supervised during the training by the ViSC developers' team at the University of Vienna (Gollwitzer et al, 2006, p. 127). According to the principles of the program, it was essential for the trainer to avoid receiving any information about individual students offered by teachers; students' assessments should be based on standardized diagnostic measures (Atria and Spiel, 2007, p. 184). Moreover, the training was conducted during regular class time and teachers were advised to attend the lessons, so that the program was taken seriously by the students. ViSC has been implemented and evaluated three times: by Gollwitzer (2005), by Atria and Spiel (2007) and by Gollwitzer et al. (2006).

4.2.9 Granada Anti-bullying Program (Spain)

This program was a pilot anti-bullying program with the following aims: a) to establish children's involvement in bullying within different participant roles/categories; b) to reduce the number of students involved in the phenomenon as bullies, victims and bully-victims; c) to increase the number of students who are categorized as non-involved in bullying, through the enhancement of pro-social skills; and d) to identify the threats to fidelity of the program and establish the validity of the pilot program with the possibility of replicating it in future (Martin et al., 2005, p. 376). Forty-nine sixth graders from one Spanish primary school in Granada participated in the program.

The program designers gathered information about the social, educational and economic background of the school, of the students' families and the community in general. That was done during 3 meetings/seminars of three hours each. Parents, teachers and members of the educational team attended those meetings. Through these meetings, it was established that the program should target interpersonal relationships of the children. It was decided that the program would be curriculum-based as part of the normal program of the school. It was decided that the program would be implemented by one of the researchers because the teachers did not have enough qualifications to do it and because of lack of time and resources for teacher training. Parents and teachers were provided with information about bullying [a dossier/file] that they could use to discuss the problem of bullying with children. Also, teachers could attend the intervention program so that later they would be able to implement it by themselves. Parents were invited to attend some talks on

bullying that would be given by the implementation team so that the program could be continued outside the school. The program was implemented for five months at the classroom level (30 sessions; 3 sessions per week with one tutor, i.e. one of the evaluators).

During the first 5 sessions, the tutor informed the children about peer bullying. Topics covered in the first 5 sessions involved issues such as concept of bullying, types of bullying, how to identify it, individual and group differences in bullying, and classroom rules against bullying. From the 6th to the 21st sessions, the program emphasis was on the emotional and social abilities of the children. Several topics were covered such as: identification and expression of emotions during bullying situations; communication abilities; ability to pose questions; ability of children to give and receive complements and complaints; ability to say no in life; ability to ask for a change of behavior; and ability to solve interpersonal problems. From the 17th to the 21st sessions, the program placed emphasis on mediation.

From the 22nd to the 25th sessions, the program emphasis was on human rights. Several topics were covered such as: freedom and equality, respect of private life, respect for other people's belongings, and respect for others' opinions. Similarly, from the 26th to 30th sessions, the emphasis was on moral education. During the whole program (sessions 1 – 30), there was also an emphasis on the inhibition of impulsivity and enhancement of reflexivity. For the enhancement of reflexivity, the program designers used a specific program called 'Programa de Intervencion para Aumentar la Attention y la Reflixividad' [PIAAR] developed by Gargallo (2000; see Martin et al., 2005, p. 378). This focuses on cognitive techniques that aim to inhibit impulsivity and enhance self-control. The program also included role-playing, peer mediation, guided discussion, brainstorming, and drawings.

The authors acknowledge several problems with the implementation of the program such as: little involvement by parents and teachers; implementation of the program lessons during recess time or during the physical education program; lack of time to cover all the topics; no second follow-up because of difficulties of following the children; problems with the size and selection of the sample; the instrument they used; and possible contamination of results because of the way they categorized the children (Martin et al., 2005, p. 382). These pitfalls could easily be spotted. For example, the evaluators indicate that they implemented the program with the most aggressive sixth graders who had the worst interpersonal problems (Martin et al., 2005, p. 738). This made it difficult to know whether any changes in bullying in the experimental condition were attributable to the effectiveness of the program or to regression to the mean. Also, even though they distributed a self-report questionnaire, they categorized children based on those questionnaires only after teachers' suggestions.

4.2.10 South Carolina Program (USA)

This program involved the implementation of the OBPP in South Carolina schools. It was a comprehensive school-based anti-bullying program 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 program included a school-wide intervention. In each school, coordinating committees planned and guided the school's anti-bullying initiative throughout the various phases of the project. The committees consisted of school psychologists or counselors and representative teachers, students and parents. In each school, a survey was conducted prior to the implementation of the program, 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 program. Another element of the program at the school level included teacher surveillance in order to track down 'hot-spots' of bullying.

At the classroom level, core elements of the program 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 behavior 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 program 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 program 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 program elements into summer church school classes).

Other elements of the program 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.

4.2.11 Bully-Proofing Your School (USA)

'Bully-Proofing Your School' (BPYS) was a comprehensive, school-based intervention program 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 students 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 behavior 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 behaviors. Parents were offered information through newsletters. Individual parents of students involved in bullying as either perpetrators or victims were given consultation. The complete BPYS program 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.

4.2.12 Befriending Intervention Program (Italy)

Befriending intervention was an anti-bullying program that relied mainly on a peer support model. The overall aims of the program were: a) to reduce bullying episodes through developing in bullies an awareness of their own and others' behavior; b) to enhance children's capacity to offer support to the victims of bullying; c) to enhance responsibility and involvement on the part of bystanders; and d) to improve the quality of interpersonal relationships in the class group (Menesini et al, 2003, p. 1).

The anti-bullying intervention was offered in five steps (Menesini et al, 2003, p. 5). During the first phase, which targeted the class level (class intervention), several activities were offered aiming to increase children's awareness of pro-social and helping behaviors and to promote positive attitudes towards others. Through work at the class level, the school authorities sensitized and prepared the whole-school population for the new service that the school unit was about to implement. In this way, another goal was achieved, namely developing values and attitudes toward 'peer support activities' in the whole-school population.

During the second phase of the program, the 'peer supporters' were selected. Approximately three to four supporters were allocated in each classroom and were selected based on a combination of techniques, such as self- and peer-nominations. These children were then trained in special full-day sessions or in regular meetings during school time (phase three) so that they knew how to deal with other children and how to facilitate interactions amongst other children. Teachers and other professionals (psychologists and social workers) took part in these sessions as well. The overall aim of this phase of the anti-bullying program was to help peer supporters to enhance their listening and communication skills since they would be the mediators in the interactions among children.

During the fourth phase of the program, peer supporters worked in their classes with the assistance and close monitoring of their teachers. The teachers in each class organized 'circle meetings' during which the needs of specific children involved in bullying (target children) were identified. Target children were contacted and, after their consent and cooperation, were offered help by the peer supporters. Peer supporters were not only assigned to specific tasks involving the target children but were also supervised by the teachers so that they were given constant feedback on their on-going work in the class.

During the final phase of the Befriending Intervention, the leading group of peer supporters was involved in training other children in the class, so that more children could be involved in the program (in the transmission of training and passing on the roles).

4.2.13 Toronto Anti-bullying Program (Canada)

The Toronto anti-bullying program 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 program 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 program varied across the intervention schools. However, in all intervention schools three critical elements were found: staff training, codes of behavior and improved playground supervision. At the school level an emphasis was placed on developing a positive code of behavior 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 program 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.

4.2.14 Ecological Anti-bullying Program (Canada)

The Ecological Anti-bullying program 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 program was the creation of a supportive and safe school environment in which firm limits against bullying were established. The specific goals of the program 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 program was based on the 'Bully Proofing Your School' (BPYS) program which was designed to increase the understanding of bullying and decrease the incidence of bullying (Rahey & Craig, 2002, p. 285). The program 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 program was implemented by psychology students who received training sessions and manuals prior to intervention. Prior to the program, at a school assembly the program was introduced to students. The assembly signaled the formal beginning of the intervention. The classroom programs 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 programs for children involved in bullying were also part of the intervention. The relevant sessions consisted of social skills, listening, empathy training and supportive counseling. Each weekly session lasted 45 minutes. The program also included intervention at the teacher level. Teacher programs consisted of meetings with teachers to discuss bullying, intervention approaches, and student support for those directly involved in bullying. During the intervention, the program coordinators met with principals and teachers to offer support.

4.2.15 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 program 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.

4.3 OTHER EXPERIMENTAL-CONTROL COMPARISONS

4.3.1 Norwegian Anti-bullying Program

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 program, 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 behavior. 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 anti-social behavior, 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 program 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 program 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.

4.3.2 B.E.S.T-Bullying Eliminated from Schools Together (USA)

BEST was implemented in one K-12 developmental research school in Northern Florida. It is based on the Kia Kaha anti-bullying program (see later), although it is different in many ways. The evaluation of the program was based on data from 7th graders. BEST is a complex alteration of the Kia Kaha, having foundations within social cognitive theory and social competence theory (Kaiser-Ulrey, 2003, pp. 18 – 19). The program was implemented by four school teachers (as opposed to the Kia Kaha program which was primarily implemented by Police Officers). The 12-week program ran twice per week in concordance with the established curriculum, with activity sessions lasting no longer than 45 minutes per session.

BEST was divided into four modules with three activity pods per module. The program placed emphasis on social problem solving techniques, awareness raising and rules against bullying, and included teacher training and a teacher manual along with a student evaluation form. The program also included a parent evaluation form along with parent information, sent home to inform parents, accompanied by contact information for the researcher in the event that they should have any questions (Kaiser-Ulrey, 2003, p. 84). Detentions (e.g. a suspension of one day) and conferences were given to students who committed bullying-related behaviors (Kaiser-Ulrey, 2003, p. 93). In the initial Kia Kaha program, researchers could make use of anti-bullying videos that were specific to the Maori culture. Alteration of these videos, in order to make them compatible with the American culture, was unsuccessful. Instead, the teachers made use of scenarios/stories that they could incorporate in the anti-bullying sessions.

4.3.3 SAVE (Spain)

The SAVE anti-bullying program in Seville 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 signaled 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 program, 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 program 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 program 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 program 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).

4.3.4 Kia Kaha (New Zealand)

Kia Kaha was designed as an anti-bullying program, 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 program 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 program 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 program, 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, and 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 programs 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 program, hosted parent nights and taught up to four lessons of the curriculum.

4.4 AGE-COHORT DESIGNS

4.4.1 Respect Program (Norway)

Respect, previously running under the name Connect, was a program that aimed to tackle different types of child problem behavior, such as disobedience, off-task behavior, bullying and aggression. The program was implemented in both primary and secondary schools. The Respect program 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 program 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 program 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 behavior; and d) the program was based on the notion of continuity, which implied a long-term commitment to the previous three principles.

Within the program framework, teachers and school management staff participated in series of seminars. The staff training sessions introduced the basic principles of the program and practical approaches to the prevention of child problem behavior 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 program. Among the different intervention schools, a network was established with the aim of discussing knowledge, experiences and challenges related to program implementation.

Finally, there were four main strategies in the implementation of the program, 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 program.

4.4.2 Olweus Bullying Prevention Program, OBPP (Norway)

The OBPP was a multi-level program 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 signaled the formal commencement of the intervention. Two different types of materials were

produced: a handbook or manual for teachers (entitled 'Olweus' core program against bullying and antisocial behavior') and a folder with information for parents and families. The program also included: a) CD-program that was used for assessing and analyzing the data obtained at the pre-test 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 children not involved to make them become effective helpers.

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 program 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 program also included cooperation among experts and teachers (e.g. psychologists) who worked with children involved in bullying.

4.4.3 Donegal Anti-bullying Program (Ireland)

The Donegal anti-bullying program in Ireland was implemented in the county of Donegal. Of the 100 primary schools in the county, 42 were involved, but data from 22 schools were included in the evaluation of the program. The Donegal program was inspired by the Norwegian anti-bullying initiative (O'Moore & Minton, 2004, p. 277) and is based on four elements:

a) Training of a network of professionals: Eleven teachers were trained through a program of workshops, to provide further training and support for staff, students and parents in the intervention schools.

b) Teachers' resource pack: A pack containing information about bullying was given to each member of the trained network. The pack provided materials with an overall emphasis on classroom management, the development of a positive atmosphere in class and schools, staff leadership, and parent-teacher cooperation.

c) Parents' resource pack: An information leaflet was produced specifically for the purposes of the intervention, providing information to parents about prevalence, types of bullying, causes and effects, as well as advice on how to deal with bullying.

d) Work with students: An element involving a general awareness-raising campaign. Awareness-raising regarding the issue of bullying was facilitated through age-related handbooks that were given to students, through peer leadership and, at the organizational level, through emphasis on the creation of a positive school environment by teachers and school professionals in general.

4.4.4 Chula Vista Olweus Bullying Prevention Program (USA)

With funding from the California Department of Education and the Office of the Attorney General of California, three elementary schools of the city of Chula Vista implemented the OBPP (Pagliocca et al., 2007). The OBPP is described above and the description will not be repeated in detail here.

In the early stages of planning for the OBPP, each school appointed a Bullying Prevention Coordinating Committee, with several members such as: the principal; teachers; parents; students; a school counselor; a psychologist or social worker; other non-teaching school staff (e.g. a librarian); a Family Resource Coordinator; and a police department person. Each of the BPCCs was provided with a two-day

training by a certified trainer of the OBPP. A full day of training in the Olweus model for teachers and other school staff was also provided. Parents were also provided with written materials in English or Spanish as appropriate. Arrangements were also made by all three schools to provide face-to-face presentation of the OBPP information to parents. Schools' anti-bullying rules were publicized at the community level by posting of them off school grounds at local stores frequented by parents and students.

The evaluation relied on a number of sources of information related to the operation of the OBPP in the Chula Vista schools such as:

- Key Informant Survey and Interview

A total of nine Key Informants participated in this component of the evaluation of the program. Four were representatives from the school district, three were from the Family Resources Centers affiliated with the participating schools, and two were from the Chula Vista Police Department (CVPD). The Key Informants were asked to answer questions about their role in the OBPP, the training received by the school staff, the materials used by the project, and the implementation of the core components of the OBPP.

- Pre- and Post-Intervention Questionnaires

These questionnaires were administered in English or Spanish as appropriate, with Spanish versions developed by an external consultant addressing not only the literal translation but also considering cultural interpretations of specific terms and phrases. Apart from the Olweus Bully/Victim Questionnaire, the Chula Vista OBPP included a Parent Questionnaire, a Teacher Questionnaire and a Playground Supervisor Questionnaire.

- Workbooks for Bullying Prevention Coordinating Committees

As a standard part of their initial training, each Bullying Prevention Coordinating Committee began developing a 'Workbook' to describe and document its schools' plan for implementing the OBPP. In addition to their use in planning and development in the early stages of the project, the workbooks were also designed to be 'working documents' that would guide the project and reflect the fidelity with which the program was implemented. Key Informants made use of the Workbooks in the early planning stages of the project, with continuing use dropping off after the first year. The Chula Vista OBPP evaluation also relied on 'Quarterly Self-Evaluation Reports'.

4.4.5 Finnish Anti-Bullying program

The Finnish anti-bullying program in Helsinki and Turku 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 behavior; and c) commitment to anti-bullying behaviors (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.

4.4.6 Sheffield Anti-Bullying program (England)

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 recognizing bullying, improving the playground environment); c) working with individuals and small groups (e.g. peer counseling, 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 behaviors. 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 counseling and increased playground surveillance. Peer counseling involved a 'listening line' for other students (Smith et al., 2004, p. 104): students formed small teams comprising two or three counselors 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 program 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 which the elements of the program were implemented within each school (e.g. Eslea & Smith, 1998, p. 208; Smith et al., 2004, p. 101).

5 Analysis of Included Evaluations

5.1 KEY RESULTS

Table 7 (Key Results of 44 Program Evaluations, see appendix page 125) summarizes key results of the 44 program evaluations that presented effect size data. Our aim was to identify the best available effect size measures in each evaluation. 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. We have received responses concerning 40 of the 44 program evaluations.

In the rare cases where both prevalence and means were provided, we chose to show prevalence. For example, Alsaker and Valkanover (2001) provided mean scores for bullying and victimization based on scales completed by teachers, but prevalence of bullying and victimization based on peer reports. Peer reports present data on overall bullying and victimization, while teacher reports were based on different types of bullying and victimization (e.g. isolation, having belongings stolen etc). We chose to present evaluation data based on prevalence (and on peer reports rather than on teacher reports). Raskauskas (2007) 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" and this was used by other researchers inspired by

Olweus. However, we did not set the criterion high if this produced a low prevalence, because it would then have been 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. O'Moore and Minton (2004) provided prevalence figures[3] for 'not at all', 'once or twice', 'sometimes', 'once a week' and 'several times a week'. We used prevalence figures for 'sometimes', 'once a week' and 'several times a week' (combined). For the evaluation data by Pagliocca et al. (2007), we present prevalence based on 'two or three times a month', 'about once a week' and 'several times a week' (combined).

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) and Menard et al. (2008) 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; Salmivalli et al. (2005) also presented results separately for different implementation conditions; Sprober et al. (2006) presented results separately for different types of bullying (mean scores for verbal and physical bullying) and for two different experimental conditions; and Meyer and Lesch (2000) presented data separately for different schools. The methods used to combine two or more effect sizes presented in a study into a single effect size are specified in the Technical Appendix.

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 post-tests, we show results obtained in all of these unless there were specific reasons for not doing so. For example, Meyer and Lesch (2000) presented complete data for peer nominations of bullying for the pre-test and two post-test measurements, but had many missing data for the second post-test measurement of bullying based on self-reports. Moreover, it is not clear on which sample size the peer nominations were based on. Because of our preference for self-reports over peer reports and our doubts about the peer sample size, we chose to show results for self-reports for the baseline and the first follow-up only. Also, they had one experimental condition compared with

[3] E-mail correspondence with Stephen Minton dated June 10, 2009.

two control conditions (i.e. 'no treatment at all' versus a 'play-control group' in which children were supervised and played with adult mentors). We were concerned about the possible impact of the play activity on the outcome variable, so we chose to present data comparing the experimental with the 'no treatment' control group. 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/non-bully (or victim/non-victim). Where question marks are shown after numbers, we have estimated them.

The most problematic numbers in Table 7 are for the Pepler et al. (2004) evaluation. This had a complex design. In year 1 (1992-93), school 1 received the anti-bullying program and school 2 served as a control. In year 2 (1993-94), school 1 continued to receive the program, school 2 also received the program, and school 3 served as a control. In year 3 (1994-95), all three schools received the program. Self-report measures of bullying and victimization (in the previous two months) were taken in the fall and spring of each year.

In analyzing 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 program may be conservative.

For Rosenbluth et al. (2004), we only show one follow-up period (at the end of the semester, immediately after the program) because only three of the six schools provided later follow-up data. For Hunt (2007), our figures are based on e-mail 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 7. For Salmivalli et al. (2009), we only show the second follow-up, because this was carried out at the same time of the year as the before measure. Because we are concerned to minimize seasonal effects on bullying and victimization, we aimed to choose assessments carried out at the same time of

the year. We are very grateful to Christina Salmivalli and Antti Karna for giving us preliminary results from this evaluation, restricted to students who were tested both before and after. For Fonagy et al. (2009), we only show baseline data versus first follow-up data, both collected in the same month (October) of each year. The published report shows results for the randomized trial after EM multiple imputation procedure was used to estimate missing data. We are very grateful to Peter Fonagy for providing the (non-imputed) mean scores for the CAPSLE intervention along with the relevant Ns. The report includes results of a School Psychiatric Consultation (SPC) intervention as well, but we do not present data for this intervention. This is because only a fraction of the students received this intervention (Fonagy et al., 2009, p. 4) but evaluation data are presented for all students.

For Rahey and Craig (2002), we used questions about bullying in the previous week, based on e-mail correspondence with Leila Rahey. The results obtained in the Flemish Anti-Bullying program (Stevens et al., 2000) were excluded. Bullying and victimization were each measured using eight items, each measured on a five-point scale (from "it has not happened" to "several times a week"). It might be expected, therefore, that scores might range from eight to 40. And yet, the mean scores in the crucial table (8.1 in Stevens et al. 2004) were all between 0.99 and 1.16, with the vast majority between 1.00 and 1.10. This was because only logarithms of scores were reported. We requested the raw data from Veerle Stevens, but she informed us (e-mail October 3, 2008) that she no longer had access to the data. Since all the means were so close to 1.0 (making the test of the effects the program very insensitive) and since we did not know the number of students on which each mean was based, we decided to exclude this program from our analyses.

The evaluation presented by Kim (2006) was also excluded from the meta-analysis. The researcher presented before and after data for victimization, yet they were clearly implausible, yielding an odds ratio of 248.1. For the Atria and Spiel (2007) evaluation, we only had data for the two follow-up periods, but no baseline data because of many missing values. In further e-mail communications with Moira Atria and Dagmar Strohmeier, we were specifically advised not to include this evaluation of the VISC program (e-mail dated June 7, 2008). Evers et al. (2007) was a before-after, experimental-control design, but they only reported data in their published article on how many of the bullies (or victims) at the pre-test continued to be bullies (or victims) at the post-test. We were, however, able to classify this among the before-after/ experimental-control comparisons because Kerry Evers kindly provided the necessary data via e-mail[4].

We are very grateful to several researchers for the information they provided to us via e-mail correspondence which enabled us to calculate effect sizes. For instance,

[4] E-mail dated May 28, 2009.

the published report by O'Moore and Minton (2004) is based on evaluation data of pre-test and post-test measurement periods in experimental schools, but with no control schools. This was originally classified in category 4 as an uncontrolled before-after design. This evaluation could be included in the meta-analysis, however, if it was analyzed as an age-cohort design, which is what we have done. Mona O'Moore and Stephen Minton kindly provided evaluation data for students in grade 4 ('control' students) before the program and for different grade 4 students who had received the anti-bullying program for a year ('experimental' students).

As another example, the Sheffield program (Whitney et al., 1994) is based on a before-after experimental-control design, but we could not obtain the necessary control data to analyze it in this way. Because Peter K. Smith and Mike Eslea kindly provided us with raw data from the experimental schools, we were able to analyze this evaluation based on an age-cohort design. The evaluation of B.E.S.T by Kaiser-Ulrey (2003) was based on a before-after/ experimental-control comparison. However, data are presented only for the follow-up period. Because Kaiser-Ulrey presented data supporting the equivalence of individuals in the experimental and control conditions at the pre-test measurement period, we decided to categorize this evaluation under other-experimental control comparisons. Finally, other researchers (e.g. Helen Cowie, Reiner Hanewinkel, Maila Koivisto) tried to supply us with the data that we requested, but were unable to do so because they could not retrieve the data because of the passage of time. The study by Twemlow et al. (2005) was not included in the meta-analysis because the data of this evaluation were included in the later evaluation by Fonagy et al. (2009). Despite our persistent attempts via e-mail, we never managed to get any results of the evaluation by Wiefferink et al. (2006).

5.2 ANALYSIS OF EFFECT SIZES

Table 8 (see appendix. Page 134) shows the analysis of effect sizes for bullying. The measure of effect size is the 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; Z-values greater than 1.96 or less than -1.96 are statistically significant. The calculation of the OR and its associated CI are explained in the Technical Appendix. Smaller studies (N < 200 students) are indicated with an asterisk. In all cases, the effect sizes for smaller studies were non-significant. Random-effects models were used to calculate the weighted mean effect sizes. Initial values of Q (from the fixed-effect model) are shown in Table 8. Figure 4 (see appendix, page 145) shows the accompanying forest graph for bullying effect sizes. In this figure, the measure of effect size is the logarithm of OR (LOR).

In general, results obtained for different types of bullying (e.g. physical, verbal) were combined, because the aim was to produce one summary OR for each evaluation.

The methods used in combining effect sizes are specified in the Technical Appendix. A combined OR is presented, for instance, in the case of Bauer et al. (2007), Frey (2005), Hunt (2007), Sprober et al. (2006) and Menard et al. (2008). Results obtained for different schools (e.g. Evers et al., 2007; Menard et al., 2008; Meyer & Lesch, 2000) and for different ages (e.g. Baldry & Farrington, 2004; Rahey & Craig, 2002; Salmivalli et al., 2005; Whitney et al., 1994) were also combined. Also, 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 were analyzed. For Ertesvag and Vaaland (2007), the shortest (A1) and longest (A3) follow-up assessments were analyzed, 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 14 randomized experiments (Fonagy et al., 2009) found a significant effect of the program on bullying, although one other evaluation (Hunt, 2007) reported a near-significant effect. Overall, the 14 randomized experiments yielded a weighted mean OR of 1.10, indicating a very small and non-significant effect of these programs on bullying. In contrast, five of the 14 evaluations with before-after/experimental-control designs found a significant effect, and one other (Olweus/Bergen 2) reported a near-significant result. Overall, these 14 studies yielded a large weighted mean OR of 1.60 ($p < .0001$).

One of the four other experimental-control comparisons found significant effects on bullying (Ortega et al., 2004), and the weighted mean OR for all four studies was 1.20 ($p = .010$). Seven of nine age-cohort designs yielded significant effects, with an overall weighted mean OR of 1.51 ($p < .0001$). Over all 41 studies, the weighted mean OR was 1.36 ($p < .0001$), indicating a substantial effect of these programs 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. If there were 25 bullies and 75 non-bullies in the control condition, OR = 1.33. Hence, OR = 1.36 can correspond to 25% - 30% more bullies in the control condition (or conversely 20% - 23% fewer bullies in the experimental condition).

Table 9 (see appendix, page 136) shows the analysis of effect sizes for victimization. Only three of the 11 randomized experiments found significant effects of the program on victimization but the weighted mean OR of 1.17 was just significant ($p = .050$). Five of the 17 studies with before-after/experimental-control designs yielded significant results, and the weighted mean OR of 1.22 was statistically significant ($p = .007$).

Three of the four studies with other experimental-control designs found significant results, yielding a significant weighted mean OR of 1.43 ($p < .006$). Seven of the nine age-cohort designs yielded significant results, and another one (O'Moore & Minton, 2004) was nearly significant. The weighted mean OR of 1.44 was statistically significant ($p < .0001$). Over all 41 studies, the weighted mean OR was 1.29 ($p < .0001$), indicating significant effects of these programs 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, this value of the OR can correspond to 20%-25% more victims in the control condition (or conversely, 17% - 20% fewer victims in the experimental condition). Figure 5 (see appendix, page 146) shows the accompanying forest graph for victimization effect sizes. In this figure, the measure of effect size is the logarithm of OR (LOR).

Based on significant ORs, we conclude that the following 19 anti-bullying programs appeared to be effective in reducing bullying and/or victimization: Alsaker and Valkanover (2001), Andreou et al. (2007), Ertesvag and Vaaland (2007), Evers et al. (2007), Fonagy et al. (2009), Galloway and Roland (2004), Melton et al. (1998), Menard et al. (2008), Olweus/Bergen 1, Olweus/Bergen 2, Olweus/Oslo 1, Olweus/Oslo 2, Olweus/New National, Ortega et al. (2004), Pepler et al. (2004), Raskauskas (2007), Salmivalli et al. (2005), Salmivalli et al. (2009), and Whitney et al. (1994).

Based on non-significant and small ORs [i.e. less than 1.4], we conclude that the following 17 anti-bullying programs appeared to be ineffective in reducing bullying and/or victimization: Bauer et al. (2007), Beran and Shapiro (2005), Beran et al. (2004), Boulton and Flemington (1996), Ciucci and Smorti (1998), Cross et al. (2004), De Rosier (2004), Fekkes et al. (2006), Fox and Boulton (2003), Frey et al. (2005), Gini et al. (2003), Gollwitzer et al. (2006), Kaiser-Ulrey (2003), Meyer and Lesch (2000), Pagliocca et al. (2007), Rahey and Craig (2002), and Sprober et al. (2006). However, it should be noted that, in 9 out of 17 cases, the evaluation involved small N (i.e. less than 200). Also, in three cases (Fekkes et al., 2006; Gollwitzer et al., 2006; Pagliocca et al., 2007), analyses presented by the researchers suggested that the program was effective.

Based on a significant OR less than 1, one program appeared to be damaging: Rosenbluth et al. (2004). However, this might possibly have been a chance finding consequent upon the large number of statistical tests.

In the remaining seven cases, ORs were large (bigger than 1.4) but non-significant: Baldry and Farrington (2004), OR = 1.69 for victimization (ns); Hunt (2007), OR = 1.46 for bullying (p = .097); Jenson and Dieterich (2007), OR = 1.63 for victimization (ns); Martin et al. (2005), OR = 2.56 for bullying (ns) and OR = 1.97 for victimization (ns); Menesini et al. (2003), OR = 1.60 for bullying (ns) and OR = 1.42 for victimization (ns); O'Moore and Minton (2004), OR = 2.12 for bullying (ns) and OR = 1.99 for victimization (p = .059); Rican et al. (1996), OR = 2.52 for bullying (ns) and OR = 2.46 for victimization (ns). These programs may have been effective but we cannot draw firm conclusions about them. Why were some programs effective and others ineffective? We will address this question in section 6.

5.3 EFFECT SIZE VERSUS RESEARCH DESIGN

Tables 8 and 9 (see appendix) 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 = 31.88 (3 df, p <.0001). For victimization, QB = 19.85 (3 df, p = .0002). 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 are randomly assigned) minimize many threats to internal validity. However, experiments on bullying often randomly assign only a small number of schools (see Table 10, appendix page 137), therefore reducing statistical conclusion validity, and are vulnerable to differential attrition. For example, there was differential attrition in the Salmivalli et al. (2009) evaluation, with many more students lost in the control condition (27%) than in the experimental condition (13%). This differential attrition created higher effect sizes than when (as in the present report) the analysis was based only on students known before and after (OR for bullying = 1.47 in Swedish Report, 1.38 here; OR for victimization = 1.66 in Swedish Report, 1.55 here).

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 programs 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 programs, and we give credence to results obtained in all of them.

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 (31 studies) versus (2) other experimental-control designs plus age-cohort designs (13 studies). Other features of the evaluation that were investigated were as follows:

(a) Sample size (experimental plus control conditions), dichotomized into 900 children or more (22) versus 899 children or less (22). 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 (27) versus 2003 or earlier (17).

(c) Average age of the children, dichotomized into 10 or less (19) versus 11 or more (25). The calculation of average age is problematic. For example, students in grade 4 (age 10 – 11) could range from 10.000 to 11.999, and we therefore estimated their average age as 11. Researchers who calculated average ages based on integer values of age (rather than exact values to several decimal places) might have reported an average age of 10.5 in this case.

(d) Location in the USA and Canada (15) versus other places (29).

(e) Location in other places (37) versus Norway (7).

(f) Location in other places (19) versus Europe (25).

(g) Outcome measure, dichotomized into others (34) versus a dichotomous measure of two or more times per month (10). This latter measure was associated with larger effect sizes than mean scores or simple prevalences.

Figure 2 (see appendix, page 143) shows key features of the evaluation for each study.

6.2 KEY ELEMENTS OF THE PROGRAM

Each anti-bullying program included a variety of intervention elements. Table 11 (see appendix page 138) summarizes the elements of the intervention in different programs and their frequency. In constructing this table we consulted the evaluators of the various programs, and sent them our coding of the elements of the intervention for their program. We received feedback on 40 out of 44 evaluations 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 program. 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 program included this method. In other cases, intervention components were added. For example, the Befriending Intervention Program (Menesini et al., 2003) included both parent training/meetings and teacher training, though these intervention components were not presented in the published report. After our communication with the leading evaluator of the program[5], we decided to include these components. Similarly, the published paper by Fonagy et al. (2009) does not clearly specify 'work with peers', 'school conferences' and 'non-punitive methods' (especially the 'No Blame' approach), but after our communication with the leading researcher of the program, we included these components[6].

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 program, 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 programs, 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 program. 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 programs, these

[5] Personal communication via e-mail from Ersilia Menesini (June 1, 2009).
[6] Personal communication via e-mail from Peter Fonagy (June 29, 2009).

conferences were organized after the pre-test data collection and aimed to inform students about the extent of bullying behavior 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 program in the school.

Element 4 (curriculum materials) refers to the use of materials about bullying during classroom lessons. Some programs 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 behavior.

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 programs, 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 programs also placed emphasis on the engagement of bystanders in bullying situations in such a way that disapproval of bullying behavior was expressed adequately while support was offered to victims.

Elements 10 and 11 (information for teachers and parents): Many programs offered information for teachers and parents, but it was not possible for us to assess the quality of the information provided. For instance, many programs 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 programs differed a lot in the quality of this information. In some programs 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 programs.

Element 12 (improved playground supervision): Some anti-bullying programs 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 programs emphasized punitive methods in dealing with bullying situations. One program (KiVa; Salmivalli et al., 2009) 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 programs included restorative justice approaches and other non-punitive methods such as the 'Pikas method' and the 'No Blame' approach (now termed 'Support Group Method') 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 program, 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 e-mails to the evaluators of the different programs 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 programs, however, we could not code one or both of these features of teacher training.

Element 18 (parent training/meetings): For all programs 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 programs 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 programs:

(a) The number of elements included out of 20, dichotomized into 10 or less (25 programs) versus 11 or more (19 programs). Olweus (2005a) and Smith (1997, p. 198) 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 program was not (27) or was (17) inspired by the work of Dan Olweus.
(c) The duration of the program for children, dichotomized into 240 days or less (23) versus 270 days or more (20).
(d) The intensity of the program for children, dichotomized into 19 hours or less (21) versus 20 hours or more (14).
(e) The duration of the program for teachers, dichotomized into 3 days or less (21) versus 4 days or more (20). Where programs did not include teacher training, then teacher duration was coded as zero in the dataset.
(f) The intensity of the program for teachers, dichotomized into 9 hours or less (18) versus 10 hours or more (21). Where programs did not include teacher training, then teacher intensity was coded as zero in the dataset.

Figure 3 (see appendix, page 144) shows the intervention components utilized in each evaluation. Figures 2 and 3 show our coding system in detail.

6.3 EFFECT SIZE VERSUS STUDY FEATURES

There have been few other attempts to relate effect size to program elements (see e.g. Kaminski, Valle, Filene & Boyle, 2008). Table 12 (see appendix, page 139) shows the program elements and design features that were significantly (or nearly significantly in two cases) related to effect sizes for bullying. Because of small numbers in one category, four of the 20 program elements could not be investigated (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 and heterogeneity (Q) are also given for the different categories.

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

Table 13 (see appendix. page 140) shows the program elements and design features that were significantly related to effect sizes for victimization (being bullied). Weighted effect sizes for bullying and victimization were significantly correlated ($r = .51$, $p < .0001$). The most important program elements that were associated with a

decrease in victimization were disciplinary methods, parent training/meetings, videos and cooperative group work. In addition, the duration and intensity of the program for children and teachers were significantly associated with a decrease in victimization. Work with peers was associated with an increase in victimization, in agreement with other research showing that programs targeting delinquent peers tend to cause an increase in offending (e.g. Dishion et al., 1999; Dodge et al., 2006). Work with peers was also associated with an increase in bullying, but not significantly so (OR = 1.42 for no work with peers, OR = 1.35 for work with peers). Regarding the design features, the programs worked better with older children, in Norway specifically and in Europe more generally, and they were less effective in the USA and Canada. Older programs, 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.

Our finding that anti-bullying programs work better with older children (age 11 or older) conflicts with the arguments of Peter Smith (forthcoming). Therefore, we examined this finding in more detail, by dividing the average age into four categories: 6-9 (12 programs), 10 (7 programs), 11-12 (14 programs), and 13-14 (11 programs). The weighted mean OR for bullying steadily increased with age: 1.21 (6-9), 1.23 (10), 1.44 (11-12) and 1.53 (13-14); QB = 15.65, 3 df, p = .001. Similarly, the weighted mean OR for victimization steadily increased with age: 1.17 (6-9), 1.25 (10), 1.26 (11-12) and 1.37 (13-14); QB =7.24, 3 df, p = .065. These results confirm our conclusion that the effectiveness of programs increases with the age of the children.

Variables that might help to explain differential treatment effects in meta-analysis (e.g. elements of the intervention) 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 programs 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 program that were related to a decrease in bullying were parent training/meetings and disciplinary methods (Table 14, see appendix page 141). When all the intensity and duration factors from Table 12 were added, the most important program elements were intensity for children and parent training/meetings.

The most important elements of the program that were associated with a decrease in victimization were videos and disciplinary methods. Work with peers was

associated with an increase in victimization. When all the intensity and duration factors from table 13 were added, the most important elements were work with peers (negatively related), the duration of the program for children, and videos.

6.4 PUBLICATION BIAS ANALYSES

While a meta-analysis will yield a mathematically accurate synthesis of the studies included in the analysis, if these studies are a biased sample of all relevant studies, then the mean effect computed by the meta-analysis will reflect this bias (Borenstein et al., 2009, p. 277). To assess publication bias, we used the Duval and Tweedie's trim-and-fill procedure and visually inspected the resulting funnel plot. Analyses were conducted separately for bullying and victimization, based on the LOR.

Figure 6 (see appendix, page 147) shows that, for bullying, the observed studies are symmetrically distributed around the vertical line (indicating the LOR point estimate) that divides the funnel plot in half. This symmetry suggests the absence of publication bias. This was confirmed by the results of the Trim and Fill analysis. No imputed values were added and the OR and confidence intervals were not changed.

The same procedure was followed for victimization. The results of the funnel plot (Figure 7, see appendix page 147) suggest that publication bias should not be a problem for our results since, again, the studies are symmetrically distributed around the mean effect size. However, one imputed study was added to the figure. Using Trim and Fill the imputed OR was 1.28 (CI 1.17 – 1.40), compared with the original estimate of OR = 1.29 (CI 1.18 – 1.42). The key point is that the adjusted estimate is very close to the original, supporting the absence of publication bias affecting our results.

7 Conclusions

7.1 SUMMARY OF MAIN FINDINGS

The present systematic review shows that school-based anti-bullying programs are often effective, and that particular program elements were associated with a decrease in bullying and victimization. One program element (work with peers) was significantly associated with an increase in victimization.

We conclude that, on average, bullying decreased by 20% - 23% and victimization by 17% – 20%. The effects were generally highest in the age-cohort designs and lowest in the randomized experiments. It was not clear, however, that the randomized experiments were methodologically superior, because of very small numbers of schools randomized in some cases, and because of other methodological problems such as differential attrition.

The most important program elements that were associated with a decrease in both bullying and victimization were parent training/meetings, disciplinary methods, the duration of the program for children and teachers and the intensity of the program for children and teachers. Regarding the design features, the programs worked better with older children and in Norway specifically. Older programs and those in which the outcome measure was two times per month or more also yielded better results. Various other intervention components and key features of the evaluation were significantly related with the reduction of either bullying or victimization.

7.2 POLICY IMPLICATIONS

In developing new policies and practices to reduce bullying, policy-makers and practitioners should draw upon high quality evidence-based programs that have been proved to be effective. New anti-bullying initiatives should be inspired by existing successful programs but should be modified in light of the key program elements that we have found to be most effective (or ineffective). For example, it seems from our results that work with peers should not be used. It should be borne in mind, however, that we have discovered the program elements that are most highly correlated with effectiveness. This does not prove that they cause effectiveness, but this is the best evidence we have at present.

We recommend that a system of accrediting effective anti-bullying programs should be developed. In England and Wales in 1996, a system of accrediting effective programs in prison and probation was established (McGuire, 2001). For a program to be accredited, it had to meet explicit criteria based on knowledge about what worked to reduce offending. Only accredited programs can be used in England and Wales, and similar systems have been developed in other countries such as Scotland and Canada. A similar system should be developed for accrediting anti-bullying programs in schools to ensure that programs contain elements that have been proved to be effective in high quality evaluations. This accreditation system could perhaps be organized by an international body such as the International Observatory on Violence in Schools.

New anti-bullying programs should be disseminated using high quality standards of implementation in a way that ensures that the program is more likely to have an impact. The quality of a program is undoubtedly important, but so is the way in which it is implemented. Implementation procedures should be transparent in order to enable researchers to know whether effects are related to key features of the intervention or key feature of the evaluation. It is sad, for instance, that only two of the 44 evaluations included in our meta-analytic review (Fekkes et al., 2006; Smith et al., 2004) provided key information about the percentage of intervention and control schools that implemented each intervention component.

Our results show that the intensity and duration of a program is directly linked to its effectiveness, and other researchers (Olweus, 2005; Smith, 1997) also found a 'dose-response' relationship between the number of components of a program that were implemented in a school and its effect on bullying. For example, both the duration (number of days) and intensity (number of hours) of teacher training were significantly related to the reduction of bullying and victimization. Similarly, the duration (number of days) and intensity (number of hours) of the program for children were significantly related to the reduction of bullying and victimization. What these findings show is that programs need to be intensive and long-lasting to have an impact on this troubling problem. It could be that a considerable time period is needed in order to build up an appropriate school ethos that efficiently tackles bullying.

New anti-bullying initiatives should also pay attention to enhancing playground supervision. For bullying, playground supervision was one of the elements that were most strongly related to program effectiveness. It is plausible that this is effective since a lot of bullying occurs during recess time. Improving the school playground environment (e.g. through reorganization and/or identification of 'hot spots') may also be a promising and low-cost intervention component.

Disciplinary methods (i.e. firm methods for tackling bullying) were an intervention component that was significantly related to both bullying and victimization. To some extent, this finding may be attributable to the big effects of the Olweus program, which included a range of firm sanctions, including serious talks with bullies, sending them to the principal, making them stay close to the teacher during recess time, and depriving them of privileges.

The results of the KiVa project promise to provide useful answers in future about the effectiveness of disciplinary methods[7]. An interesting element of the KiVa program is that it incorporated both punitive and non-punitive approaches to deal with perpetrators of bullying. Half of the school teams, chosen at random, were instructed to use more punitive approaches (e.g. 'what you have done is wrong and it has to stop right now') while the other half 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?'). A very preliminary observation from the current unpublished data is that disciplinary methods (the punitive approach) seem to work better for younger children (grade 4), while non-punitive approaches seem to work better for older children (grade 6), and for grade 5 children there seemed to be little difference. The current results are very important in that they also suggest the necessity of developing more age-specific programs.

Contrary to the arguments of Peter Smith (forthcoming) the results of our review show that programs have a bigger impact on bullying for older children (age 11 or older). This is an age range when bullying is decreasing anyway. Peter Smith argued that programs were less effective in secondary schools because negative peer influence was more important and because secondary schools were larger and students did not spend most of their time with one teacher who could be very influential. We speculate that programs may be more effective in reducing bullying by older children because of their superior cognitive abilities, decreasing impulsiveness, and increasing likelihood of making rational decisions. Many programs are based on social learning ideas of encouraging and rewarding pro-social behavior and discouraging and punishing bullying. These programs are likely to work better, for example, in building empathy and perspective-taking skills with older students.

Perhaps surprisingly, establishing a whole-school anti-bullying policy was significantly related to effect sizes for bullying but not for victimization (being bullied). Nor was individual work with bullies or victims. We recommend that more efforts should be made to implement effective programs with individual bullies and victims, perhaps based on child skills training programs (Losel & Beelman, 2003).

---

[7] Personal communication from Christina Salmivalli (January 31, 2009).

Most current programs, with some exceptions (e.g. DeRosier, 2004; Fox & Boulton, 2003; Gollwitzer et al., 2006), are not.
New anti-bullying initiatives should go beyond the scope of the school and target wider systemic factors such as the family. Studies indicate that bullied children often do not communicate their problem to anyone while parents and teachers often do not talk to bullies about their conduct (e.g. Fekkes, Pijpers, & Verloove-Vanhorick, 2005). In our systematic review, parent training/meetings were significantly related to a decrease in both bullying and victimization. These findings suggest that efforts should be made to sensitize parents about the issue of school bullying through educational presentations and teacher-parent meetings. Future anti-bullying initiatives should also bring together experts from various disciplines and make the most of their expertise. In our review, cooperative group work among experts was significantly related to the reduction of both bullying and victimization.

Future evaluations of anti-bullying programs should be designed in light of our results. Attention should be paid not only to the quality of the program but also to the way it is implemented. The present review has shown that different features of the evaluation were significantly related to a decrease in bullying and victimization. In particular, the way bullying was measured and the age of the children were important. It would be regrettable if some evaluations of anti-bullying programs did not establish the effectiveness of the program only because of the way the outcome variable was measured. Programs 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. Programs implemented in Norway seem to work best and this could be related to the long tradition of bullying research in Scandinavian countries. Other factors are that Scandinavian schools are of high quality, with small classes and well-trained teachers, and there is a Scandinavian tradition of state intervention in matters of social welfare (J.D. Smith et al., 2004a, p. 557).

Importantly, cost-benefit analyses of anti-bullying programs 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, 2009, p. 59). There never has been a cost-benefit analysis of an anti-bullying program.

In our opinion, anti-bullying programs should be based more on theories of bullying and victimization. Most past programs have been based on general social learning ideas. Future programs should be based on newer theories such as defiance theory and restorative justice approaches (Ttofi & Farrington, 2008a; 2008b). For example, poor social relationships at school can be repaired through restorative justice approaches that involve bringing together all children (bullies, victims, and other children) 'in a participatory process that addresses wrongdoing while offering

respect to the parties involved' (Morrison, 2007, p. 198). Defiance theory is useful because it places emphasis on improving bonding to the sanctioner, shame management, and legitimate, respectful sanctioning of antisocial behavior.

7.3 IMPLICATIONS FOR FUTURE RESEARCH

Interestingly, previous reviews (Ferguson et al., 2007; Merrell et al., 2008) concluded that anti-bullying programs had little effect on school bullying. We attribute their results to the relatively limited searches done and also to the inclusion criteria (e.g. not clearly focusing on bullying; including uncontrolled evaluations) that previous researchers have set (see section 1.3). After completing our more extensive review, we believe that their conclusions are incorrect. The present systematic review shows that school-based anti-bullying programs 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 program elements and design features on bullying and victimization?
- Why do results vary in different countries?
- Why do results vary by research design?
- Why do programs 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. In order to investigate the effects of different program elements, children could be randomly assigned to receive or not receive them. 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 analyzed, 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 program 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 researchers regarding our coding for 40 out of 44 evaluations of anti-bullying programs, we have been quite successful. However, because of time limitations and lack of information, we were unable to study the varying results of the intervention programs 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 (Smith & Ananiadou, 2003; Pepler, Smith, & Rigby, 2004). Most of the above features were not mentioned in reports, making it difficult for us to code them. For the 20 program elements that we did code, only two studies (Fekkes et al., 2006; Smith et al., 2004) provided the percentage of intervention and control schools that actually implemented these elements. More studies of implementation are needed.

Future systematic reviewers could attempt to detect the impact of anti-bullying programs 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 programs, and about the implementation of programs, 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 (Farrington, Jolliffe, & Johnstone, 2008). This would avoid the problem of results varying according to 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 programs 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 (Ttofi & Farrington, 2008a, 2008b). These theories should guide program 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 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 programs are encouraging. The time is ripe to mount a new program of research on the effectiveness of these programs, based on our findings.

8 Other Topics

8.1 ACKNOWLEDGEMENTS

For providing helpful information, we are very grateful to Francoise Alsaker, Eleni Andreou, Celia Arriaga, Moira Attria, Nerissa Bauer, George Belliveau, Tanya Beran, Sara Black, Catherine Blaya, Michael Boulton, Helen Cowie, Wendy Craig, Melissa DeRosier, Sigrun Ertesvag, Mike Eslea, Kerry Evers, Peter Fonagy, Claire Fox, David Galloway, Vicente Garrido, Gianluca Gini, Mario Gollwitzer, Reiner Hanewinkel, Robin Heydenberk, Warren Heydenberk, Caroline Hunt, Jeffrey Jenson, Antti Karna, Cheryl Kaiser-Ulrey, Maila Koivisto, Elmien Lesch, Sue Limber, Francisco Martin, Ersilia Menesini, Stephen James Minton, Christof Nagele, Dan Olweus, Mona O'Moore, Rosario Ortega, Debra Pepler, Yolanda Pintabona, Leila Rahey, Ken Rigby, Elizabeth Rock, Erling Roland, Barri Rosenbluth, Christina Salmivalli, Philip Slee, J. David Smith, Peter K. Smith, Nina Sprober, Veerle Stevens, Dagmar Strohmeier, Sue Swearer, Stuart Twemlow, and Scott Watson. For statistical advice, we are very grateful to Patricia Altham, Mark Lipsey, David Murray and David Wilson. For funding, we are extremely grateful to the U.S. National Institute of Justice, the Nordic Campbell Center and the Swedish National Council for Crime Prevention.

8.2 CONFLICT OF INTEREST

Neither of the authors has any financial interest in any existing or planned anti-bullying programs.

8.3 PLANS FOR UPDATING THE REVIEW

We plan to update this systematic review every 3 years in accordance with Campbell Collaboration guidelines.

8.4 NOTE ON LAST UPDATE OF REVIEW

The changes outlined below were made to the review after its initial publication. The page numbers refer to the published review dated 15 December 2009. These corrections to date do not affect any of the conclusions.

p. 58 para 3 line 1: change "nine" to "14"
line 3: change "nine" to "14"
line 4: change "1.12" to "1.10"
line 5: delete (p = .084)

p.58 para 4 line 5: change "1.37" to "1.36"
line 9: change "1.37" to "1.36"

p. 59 line 2: insert "11" before "randomized"
line 3: change to "on victimization but the weighted mean OR of 1.17 was just significant (p = .050)."
line 5: change "1.20" to "1.22" and ".012" to ".007"

p. 59 para 2 line 2: change "1.41 (p < .0001)" to "1.43 (p = .006)"

p. 134 Table 8 last line of "Randomized Experiments":
change "1.12 0.98-1.28 1.73 .084" to "1.10 0.97-1.26 1.44 ns"

p. 135 last line of data:
change "1.37 1.27-1.48 8.04 .0001" to "1.36 1.26-1.47 7.86 .0001"

p. 136 Table 9 last line of "Randomized Experiments":
change "1.14 0.97-1.33 1.59 ns" to "1.17 1.00-1.37 1.96 .050"
last line of "Before-After, Exerimental-Control":
change "1.20 1.04-1.38 2.50 .012" to "1.22 1.06-1.40 2.72 .007"
p. 136 last line:
change "1.41 1.23-1.62 4.90 .0001" to "1.43 1.11-1.85 2.73 .006"

9 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.

9.1.1 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 program, 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 post-test of Fekkes et al. (2007) were as follows:

	Non-Bullies	Bullies	% Bullies
Experimental	1011	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 program did not cause a significant decrease in bullying.

9.1.2 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 $SELOR = 0.354$

$$Z = LOR / SELOR = 1.66 \text{ as before}$$

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):

$$d = r / \text{sqrt}\,[(1 - r * r) * p * (1-p)]$$

Where p is the proportion of the sample in the experimental condition as opposed to the control condition.

To a good approximation:

$$SEd = 2 * Ser$$

The SE of r is calculated using the transformation:

$$Zr = 0.5 * \text{Ln}\,[(1 + r) / (1 - r)]$$

$$\text{and } VAR\,(Zr) = 1 / (N - 3)$$

The analysis then proceeded as above.

9.1.3 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.

9.1.4 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.

9.1.5 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 programs, 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 behavioral 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

9.1.6 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. We always used 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. As mentioned, we use OR and LOR as the main measures of effect size in this report.

10 References[8]

*Alsaker, F.D. (2004). Bernese program against victimization in kindergarten and elementary schools. In P.K. Smith, D. Pepler & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 289-306). Cambridge: Cambridge University Press.

*Alsaker, F.D., & Valkanover, S. (2001). Early diagnosis and prevention of victimization in kindergarten. In J. Juvonen & S. Graham (Eds.), *Peer harassment in school* (pp.175-195). New York: Guilford.

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.

Atria, M. & Spiel, C. (2007). Viennese Social Competence (ViSC) Training for students: Program and evaluation. In J.E. Zins, M.J. Elias, & C.A. Maher (Eds.), *Bullying, victimization and peer harassment: A handbook of prevention and intervention* (pp. 179 – 197). New York: The Haworth Press, Inc.

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

[8] References with an asterisk indicate reports of evaluations that have been included in the meta-analysis.

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.

*Beran, T., & Shapiro, B. (2005). Evaluation of an Anti-Bullying program: Student reports of knowledge and confidence to manage bullying. *Canadian Journal of Education*, 28 (4), 700-717.

*Beran, T., Tutty, L., & Steinrath, G. (2004). An Evaluation of a Bullying Prevention Program for Elementary Schools. *Canadian Journal of School Psychology,* 19 (1/2), 99-116.

Borenstein, M., Hedges, L.V., Higgins, J.P.T., & Rothstein, H.R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.

*Boulton, M.J., & Flemington, I. (1996). The effects of a short video intervention on secondary school pupils' involvement in definitions of and attitudes towards bullying. *School Psychology International*, 17, 331-345.

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.

Cowie, H., Smith, P.K., Boulton, M., & Laver, R. (1994). *Cooperation in the multi-ethnic classroom: The impact of cooperative group work on social relationships in middle schools*. London: David Fulton.

*Cross, D., Hall, M., Hamilton, G., Pintabona, Y., & Erceg, E. (2004). Australia: The Friendly Schools project. In P.K. Smith, D. Pepler, K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 187-210). Cambridge: Cambridge University Press.

*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.

Dishion, T.J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist*, 54, 755 – 764.

Dodge, K.A., Dishion, T.J., & Lansford, J.E. (2006) (Eds.). *Deviant Peer Influences in Programs for Youth*. New York: Guildford.

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. (2009). 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, P.S. (2005). Bullying: Who does what, when and where? Involvement of children, teachers and parents in bullying behavior". *Health Education Research*, 20, 81-91.

*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.M., Nelson, J.M., Dill, E.J., Little, T.D., & Sargent, J.A. (2009). A cluster randomized controlled trial of child-focused psychiatric consultation and a school systems-focused intervention to reduce aggression. *Child Psychology and Psychiatry*, online first.

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.

*Fox, C. & Boulton, M. (2003). Evaluating the effectiveness of a social skills training (SST) programme for victims of bullying. *Educational Research*, 45, 231-247.

*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.

*Gini, G., Belli, B., Casagrande, M. (2003). Le prepotenze a scuola: una esperienza di ricerca-intervento antibullismo [Bullying at school: an experience of research-intervention against bullying], *Eta Evolutiva*, 76, 33-45.

*Gollwitzer, M., Eisenbach, K., Atria, M., Strohmeier, D., & Banse, R. (2006). Evaluation of aggression-reducing effects of the 'Viennese Social Competence Training'. *Swiss Journal of Psychology*, 65, 125-135.

Hanewinkel, R. (2004). Prevention of bullying in German schools: An evaluation of an anti-bullying approach. In P.K. Smith, D. Pepler, & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be?* (pp. 81-97). Cambridge: Cambridge University Press.

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.

*Karna, A., Voeten, M., Little, T.D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2009). A large-scale evaluation of the KiVa anti-bullying program. *Child Development*, in press.

Kim, J.U. (2006). The effect of a bullying prevention program on responsibility and victimization of bullied children in Korea. *International Journal of Reality Therapy*, 26 (1), 4-8.

Koivisto, M. (2004). A follow-up survey of anti-bullying interventions in the comprehensive schools of Kempele in 1990-1998. In P.K. Smith, D. Peppler, & K. Rigby (Eds.), *Bullying in schools: how successful can interventions be?* (pp. 235-251). Cambridge: Cambridge University Press.

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.

*Limber, S.P., Nation, M., Tracy, A. J., Melton, G. B., & Flerx, V. (2004). Implementation of the Olweus Bullying Prevention Program in the Southeastern United States. In P. K. Smith, D. Pepler, & K Rigby (Eds.), *Bullying in schools. How successful can interventions be?* (pp. 55-80). Cambridge: Cambridge University Press.

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.

Losel, F. & Beelman, A. (2003). Effects of child skills training in preventing antisocial behavior: A systematic review of randomized evaluations. *Annals of the American Academy of Political and Social Science*, 587, 84 – 109.

*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.

McGuire, J. (2001). What works in correctional intervention? Evidence and practical implications. In G.A. Bernfeld, D.P. Farrington, & A.W. Leschield (Eds.), *Offender Rehabilitation in Practice: Implementing and Evaluating Effective Programs* (pp. 3 – 19). Chichester, UK: Wiley.

*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

*Menesini, E., & Benelli, B. (1999). ☐*Enhancing children's responsibility against bullying: Evaluation of a befriending intervention in Middle School Children,* Paper presented at the IXth European Conference on Developmental Psychology, Spetses, Greece, September, 1999.

*Menesini, E., Codecasa, E., Benelli, B., & Cowie, H. (2003). Enhancing children's responsibility to take action against bullying: evaluation of a befriending intervention in Italian Middle schools. *Aggressive Behavior*, 29, 1-14.

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.

Merrell, K.W., Gueldner, B.A., Ross, S.W., & Isava, D.M. (2008). How effective are school bullying intervention programs? A meta-analysis of intervention research. *School Psychology Quarterly*, 23 (1), 26 – 42.

*Meyer, N. & Lesch, E. (2000). An analysis of the limitations of a behavioral programme for bullying boys from a sub-economic environment. *Southern African Journal of Child and Adolescent Mental Health*, 12 (1), 59-69.

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.

Morrison, B. (2007). *Restoring Safe School Communities: A Whole School Response to Bullying, Violence and Alienation*. Sidney, Australia: The Federation Press.

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.

*O'Moore, M. (2005). *Prevention programme for teachers*. Paper presented at the IX International Meeting on Biology and Sociology of Violence, Under the Honorary Presidency of H.M. Queen Sofia, October, 6-7, 2005, Valencia.

*O'Moore, A.M., & Milton, S. J. (2004). Ireland: The Donegal primary school antibullying project. In P. K. Smith, D. Pepler, & K. Rigby (Eds.), *Bullying in schools: How successful can interventions be*? (pp. 275-288). Cambridge: Cambridge University Press.

*Ortega, R. & Del Rey, R. (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.

*Pagliocca, P.M., Limber, S.P., & Hashima, P. (2007). *Evaluation report for the Chula Vista Olweus Bullying Prevention Program*. Final report prepared for the Chula Vista Police Department.

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 Australia*. 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. (2009). From peer putdowns 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.), *Handbook of bulling in schools: An international perspective* (pp. 441 – 454). New York: Guilford.

*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. forthcoming. Bullying in primary and secondary schools: Psychological and organizational comparisons". In S.R. Jimerson, S.M. Swearer, & D.L. Espelage(Eds.), *The International Handbook of School Bullying*. Mahwah, NJ: Lawrence Erlbaum.

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., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary schools. *Journal of Child Psychology and Psychiatry*, 49 (4), 376 – 385.

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.

Solberg, M.E. & Olweus, D. (2003). Prevalence estimation of school bullying wit the Olweus bully/victim questionnaire. *Aggressive Behavior*, 29, 239 – 268.

*Spröber, N. (2006). 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*.

*Sprober, N., Schlottke, P.F., & Hautzinger, M. (2006). ProACT + E: Ein Programm zur Pravention von "bullying" an Schulen und zur Forderung der positiven Entwicklung von Schulern: Evaluation eines schulbasierten, universalen, primar-praventiven Programms fur weiterfuhrende Schulen unter Einbeziehung von

Lehrern, Schulern und Eltern. [German: ProACT + E: A programme to prevent bullying in schools and to increase the positive development of students. Evaluation of a school-based, universal, primary preventive programme for secondary schools that includes teachers, students and parents]. *Zeitschrift fur Klinische Psychologie und Psychotherapie: Forschung und Praxis,* 35, 140-150.

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. (2008a). Bullying: Short-term and long-term effects, and the importance of Defiance Theory in explanation and prevention. *Victims and Offenders*, 3, 289-312.

Ttofi, M.M. & Farrington, D.P. (2008b). Reintegrative Shaming Theory, Moral Emotions and Bullying. *Aggressive Behavior*, 34, 352 – 368.

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 (4), 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.

11 Appendix/Tables and Figures

11.1 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

11.2 TABLE 2: LIST OF JOURNALS SEARCHED FROM 1983 UNTIL MAY 2009

- Archives of Pediatrics and Adolescent Medicine, 1983 [vol. 137; 1] until May 2009 [vol. 163; 5]
- Aggression and Violent Behavior, 1996 [vol. 1] until 2009 [vol. 14; 3]
- Aggressive Behavior, 1983 [vol.9; 1] until 2009 [vol. 35; 3]
- Australian Journal of Education, 2000 [vol. 44] until 2007 [vol. 51] until 2009 [vol. 53; 1]
- Australian Journal of Educational and Developmental Psychology, 2001 [vol. 1] until 2008 [vol.8]
- British Journal of Educational Psychology, 1983 [vol. 53] until 2009 [vol. 79; 2]
- 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] until 2009 [vol. 24; 1]
- Child Development, 1983 [vol. 34; 1] until 2009 [vol. 80; 2]
- Criminal Justice and Behavior, 1983 [vol. 10; 1] until 2009 [vol. 36; 6]
- Crisis-The journal of Crisis Intervention and Suicide Prevention, 2001 [vol. 22] until 2009 [vol. 30; 1]
- Developmental Psychology, 1983 [vol. 19; 1] until 2009 [vol. 45; 3]
- Educational Psychology, 1983 [vol. 3; 1] until 2009 [vol. 29; 2]
- Educational Psychology in Practice, 1985 [vol. 1] until 2009 [vol. 25; 1]
- Educational Psychology Review, 1989 [vol. 1] until 2009 [vol. 21; 1]
- Educational Research, 1983 [vol. 25] until 2009 [vol. 51; 1]
- International Journal on Violence and Schools, January 2006 until 2008 [vol. 5 – 7]
- Intervention in School and Clinic, 1983 [vol. 18; 3] until 2009 [vol. 44; 5]
- Journal of Educational Psychology, 1983 [vol. 75; 1] until 2009 [vol. 101; 2]
- Journal of Emotional Abuse, 1997 [vol. 1; 1] until 2008 [vol. 8; 4]
- Journal of Experimental Criminology, 2005 [vol. 1] until 2009 [vol. 5; 2]
- Journal of Interpersonal Violence, 1986 [vol. 1] until 2009 [vol. 24; 6]
- Journal of School Health, 2005 [75; 1] until 2009 [vol. 79; 6]
- Journal of School Violence, 2001 [vol. 1; 1] until 2009 [vol. 8; 2]
- Journal of Youth and Adolescence, 1983 [vol. 12; 1] until 2009 [vol. 38; 5]
- Justice Quarterly, 1986 [vol. 1] until 2009 [vol. 26; 2]
- Pastoral Care in Education, 1983 [vol. 1] until 2009 [vol. 27; 1]
- Psychology, Crime and Law, 1994 [vol. 1] until 2009 [vol. 15; 3]
- Psychology in the Schools, 1983 [vol. 20] until 2009 [vol. 46; 5]
- Scandinavian Journal of Psychology, 1983 [vol. 24; 1] until 2009 [vol. 50; 2]
- School Psychology International, 1983 [vol. 4] until 2008 [vol. 29; 1] until 2009 [vol. 30; 2]
- School Psychology Review, 1983 [vol. 12; 1] until 2008 [vol. 37; 1]
- Studies in Educational Evaluation, 1983 [vol. 9] until 2009 [vol. 35; 1]
- Swiss Journal of Psychology, 1999 [vol. 58; 1] until 2009 [vol. 68; 1]
- Victims and Offenders, 2006 [vol. 1] until 2009 [vol. 4; 2]
- Violence and Victims, 1986 [vol. 1; 1] until 2009 [vol. 24; 2]

11.3 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 programs [description of more than one anti-bullying program]; and/or reviews of anti-bullying programs; and/or placing emphasis on suggestions/recommendations for reducing bullying.

3: **Medium relevance**; description of a specific anti-bullying program.

4: **Strong relevance**; evaluation of an anti-bullying program, 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 program that has an experimental and control condition [N may be < 200; teacher and peer nominations may also be included as outcome measures].

6: **Also included in the Swedish review**; evaluation of an anti-bullying program that has an experimental and control condition [N > 200, self-reported bullying as outcome measure].

11.4 TABLE 4: PERCENTAGE OF REPORTS AND EVALUATIONS[A] OF PROGRAMS WITHIN EACH CATEGORY

Category	Reports [N]	Evaluations [N]	Percentage
Not Obtained	16	---	2.6 %
Category 1	100	---	16.1 %
Category 2	253	---	40.7 %
Category 3	93	---	15.0 %
Category 4	71	---	11.4 %
Category 5	18	15 [3 excluded][b]	2.9 %
Category 6	71	38 [6 excluded][c]	11.4 %

a. When applicable
b. 3 evaluations presented in 3 reports were excluded from the meta-analysis (see table 5 for relevant references)
c. 6 evaluations presented in 9 reports were excluded from the meta-analysis (see table 5 for relevant references)

11.5 TABLE 5: 89 REPORTS OF 53 DIFFERENT EVALUATIONS*

Randomized Experiments

(1) ViSC Training Program [Atria & Spiel, 2007]; category 5 => excluded due to many missing values

(2) Bulli & Pupe [Baldry, 2001; Baldry & Farrington, 2004]; category 6

(3) Project Ploughshares Puppets for Peace [Beran & Shapiro, 2005]; category 5

(4) Short Video Intervention [Boulton & Flemington, 1996]; category 5

(5) Friendly Schools [Cross et al., 2004; Pintabona, 2006]; category 6

(6) S.S.GRIN [De Rosier, 2004; De Rosier & Marcus, 2005]; category 6

(7) Dutch Anti-bullying Program [Fekkes et al., 2006]; category 6

(8) SPC and CAPSLE Program [Fonagy et al., 2009]; category 6

(9) Steps to Respect [Frey, Edstrom & Hirschstein, 2005; Frey et al., 2005; Hirschstein et al., 2007]; category 6

(10) Anti-bullying Intervention in Australian Secondary Schools [Hunt, 2007]; category 6

(11) Youth Matters [Jenson & Dieterich, 2007; Jenson et al., 2005a; 2005b; 2006a; 2006b]; category 6

(12) Kiva [Karna et al., 2009 Salmivalli et al., 2009]; category 6

(13) Korean Anti-Bullying Program [Kim, 2006]; category 5 => excluded; data produced implausible effect size

(14) Behavioral Program for Bullying Boys [Meyer & Lesch, 2000]; category 5

(15) Expect Respect [Rosenbluth et al., 2004; Whitaker et al., 2004]; category 6

(16) Pro-ACT+E [Sprober, 2006; Sprober et al., 2006]; category 5

(17) The Peaceful Schools Experiment [Twemlow et al., 2005]; category 6 => excluded; part of a larger evaluation by Fonagy et al., 2009

Before-After, Experimental-Control Comparisons

(1) Be-Prox [Alsaker & Valkanover, 2001; Alsaker, 2004]; category 5

(2) Greek Anti-bullying program [Andreou et al., 2007]; category 6

(3) Seattle Trial of the Olweus Program [Bauer et al., 2007]; category 6

(4) Dare to Care: Bully Proofing your School Program [Beran et al., 2004]; category 5

(5) Progetto Pontassieve [Ciucci & Smorti, 1998]; category 6

(6) Cooperative Group Work Intervention [Cowie et al., 1994]; category 5 => excluded due to lack of data

(7) Transtheoretical-based tailored Anti-bullying program [Evers et al., 2007]; category 6

(8) Social Skills Training (SST) Program [Fox & Boulton, 2003]; category 5

(9) Stare bene a scuola: Progetto di prevenzione del bullismo [Gini et al., 2003]; category 5

(10) Viennese Social Competence (ViSC) Training [Gollwitzer et al., 2006]; category 5

(11) Conflict Resolution Program [Heydenberk et al., 2006]; category 6 => excluded due to lack of data

(12) Granada Anti-bullying Program [Martin et al., 2005]; category 5

(13) South Carolina Program; implementation of OBPP [Melton et al., 1998; Limber et al., 2004]; category 6

(14) 'Bullyproofing your School' Program [Menard et al., 2008]; category 6

(15) Befriending Intervention Program [Menesini & Benelli, 1999; Menesini et al., 2003]; category 5

(16) New Bergen Project against Bullying; 'Bergen 2' [1997-1998]; category 6

(17) Toronto Anti-bullying program [Pepler et al., 2004]; category 6

(18) Ecological Anti-bullying program [Rahey & Craig, 2002]; category 6

(19) Short intensive intervention in Czechoslovakia (Rican et al., 1996]; category 6

(20) Flemish Anti-bullying program [Stevens, De Bourdeaudhuij & Van Oost, 2000; Stevens, Van Oost & De Bourdeaudhuij, 2000; Stevens et al, 2001; Stevens et al., 2004]; category 6 => excluded due to nature of data

(21) Anti-bullying Intervention in the Netherlands [Wiefferink et al., 2006]; category 6 => excluded due to lack of data

Other Experimental-Control Comparisons

(1) Norwegian Anti-bullying program [Galloway & Roland, 2004]; category 6

(2) BEST [Kaiser-Ulrey, 2003]; category 5

(3) SAVE [Ortega & Del Rey, 1999; Ortega et al., 2004]; category 6

(4) Kia Kaha [Raskauskas, 2007]; category 6

Age-Cohort Designs

(1) Respect [Ertesvag & Vaaland, 2007]; category 6

(2) Anti-bullying Intervention in Schleswig-Holstein, Germany [Hanewinkel, 2004];

category 6 => excluded due to lack of data

(3) Anti-bullying Intervention in Kempele schools [Koivisto, 2004]; category 6 => excluded due to lack of data

Olweus Bullying Prevention Program [OBPP]; category 6:

(4) First Bergen Project against Bullying; 'Bergen 1' [1983-1985]; category 6

(5) First Oslo Project against Bullying; 'Oslo 1' [November 1999-November 2000]; category 6

(6) New National Initiative Against Bullying in Norway; 'New National' [2001-2007]; category 6

(7) Five-year Follow-up in Oslo; 'Oslo 2' [2001-2006]; category 6

[Olweus, 1991; 1992; 1993b; 1994a; 1994b; 1994c; 1995; 1996a; 1996b; 1996c; 1997a; 1997b; 1997c; Olweus, 2004a; 2004b; 2005a; 2005b; Olweus & Alsaker, 1991]

(8) Donegal Anti-Bullying Program [O'Moore & Minton, 2004; O'Moore, 2005]; category 6

(9) Chula Vista OBPP [Pagliocca et al., 2007]; category 6

(10) Finnish Anti-bullying program [Salmivalli et al., 2004; 2005]; category 6

(11) Sheffield Anti-bullying program [Whitney et al., 1994; Smith, P.K., 1997; Smith et al., 2004b]; category 6

* Nine evaluations [presented in 12 reports] were excluded from the meta-analysis.

11.6 TABLE 6: KEY FEATURES OF 53 EVALUATIONS[9]

Randomized Experiments			
Project	**Components of the Program**	**Participants**	**Research Design**
Atria & Spiel (2007); category 5; not included in the meta-analysis [Austria]	Program specifically designed for disadvantaged adolescents aged15 to 19; program divided in 3 phases [median: 17 years in the study]	112 students [57 boys and 55 girls; grades 9 and10] participated • 55 children in the treatment group • 57 children in the control group All children from one secondary school	Experimental pre-test post-test control group design; two classes from the same school randomly assigned to experimental, and two classes to control conditions [blind study with regard to data collection; p. 187]; 2 pre-test and 2 post-test measurements
Baldry & Farrington (2004); category 6 [Italy]	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, pre-test and post-test measures
Beran & Shapiro (2005); category 5	Program for victims of bullying and for bullying awareness; use of a 45-minute	129 elementary students [69 boys] in grades 3 and 4 from two schools	Experimental pre-test post-test control group design and a three-

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

[Canada]	puppet show; 4 Footsteps to tackle bullying	• 66 students in the experimental group • 63 students in the control group	month follow-up; children within each classroom [p. 704] randomly assigned
Boulton & Flemington (1996); category 5 [England]	The 'Sticks and Stones' video was viewed by the experimental children and was discussed in the classroom with their teacher	From only one secondary school: 82 girls and 88 boys drawn from 2 classes from Years 7, 8, 9 and 10.	Experimental pre-test post-test control group design; 1 class from 4 year groups [7, 8, 9, and 10] randomly assigned to the experimental condition and another one to control condition
Cross et al. (2004); category 6 [Australia]	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: • 1,046 intervention students • 922 control students • 15 intervention schools • 14 control schools	Pre-test and post-test data from intervention and control schools; 3-year randomized control trial
De Rosier (2004); De Rosier & Marcus (2005); category 6 [USA]	Program for children experiencing peer dislike, bullying or social anxiety; highly structured manualized intervention combining social learning and	1,079 students • 50.8% boys • 49.2% girls • mean age: 8.6 years of which: • 415 eligible to participate in S.S.GRIN	Pre-test, post-test, experimental and control groups; 18 children in each school (11 public elementary schools; North Carolina) randomly assigned to the treatment

	cognitive-behavioral techniques	(664 children as non-identified)	group and the remainder of the list assigned to no-treatment control group
Fekkes et al. (2006); category 6 [Netherlands]	An anti-bullying school program including anti-bullying training for teachers, a whole-school anti-bullying policy, an anti-bullying curriculum	3,816 students aged 9 to 12 years (50% of the sample girls)	Two-year follow-up randomized intervention group control-group design; schools randomly assigned
Fonagy et al (2009); category 6 [USA]	Implementation -& comparison- of two manualized programs: SPC and CAPSLE; two years of active intervention and one year of minimal input maintenance intervention	Children from nine elementary school children (3rd and 5th graders) • 3 schools randomly allocated to CAPSLE experimental condition (188 children per school) • 3 schools randomly allocated to SPC condition (131 children per school) • 3 schools randomly allocated to TAU/control condition (120 children per school)	Cluster-level randomized controlled trial with stratified restricted allocation; schools randomly assigned
Frey et al. (2005); category 6 [USA]	Training manual for staff (staff training) including a core instructional session for all school staff and two in-depth training sessions for counselors, administrators and teachers; classroom	A random sub-sample (N= 544) of a longitudinal study (N=1023) observed and their behavior being coded.	Pre-test, post-test, experimental and control groups, schools randomly assigned

	curriculum (10 semi-scripted skill lessons); parent engagement (take-home letters etc)		
Hunt (2007); category 6 [Australia]	Information at parent and teacher meetings about the nature of bullying in schools; school staff conducted a 2-hour classroom-based discussion of bullyi8ng using activities from an anti-bullying work-book	• 444 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); category 6 [USA]	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
Karna et al. (2009); category 6 [Finland]	Universal/whole-school intervention; Indicated intervention/work with individual students; compre-hensive program with manuals for teachers, information for parents; increased supervision; internet-virtual learning environments; web-based discussions forum for teachers; peer support for	All Finnish comprehensive schools invited to volunteer; of the 300 schools who were willing to participate, a representative sample of 78 schools was chosen; program still running/ no final results yet	An age-cohort design and a randomized experiment 'nested' in the same program; only results for the latter available

	bullies and victims of bullying		
Kim, J.U. (2006); category 5; not included in the meta-analysis [Korea]	Program for victims of bullying derived from reality therapy and choice theory; 2 sessions per week for 5 consecutive weeks; summer counseling program	16 children [10 boys; 6 girls] randomly assigned to control [8] and treatment [8] conditions • fifth and sixth graders • children highly recommended as participants by their teachers	Experimental pre-test post-test control group design; children randomly assigned
Meyer & Lesch (2000); category 5 [South Africa]	Program designed for bullying boys; work with psychologists; a 17-session behavioral program implemented for 10 non-consecutive weeks, with 20 hour-long sessions held twice a week	54 boys in total from 3 primary schools=> Within each school 18 boys were matched according to level of bullying and randomly allocated in 3 conditions as follows: • 6 boys in experimental group • 6 boys in play control group • 6 boys in no-supervision control group Also: peer reports on bullying based on 50 boys who were randomly selected from grades six and seven	Before and after experimental groups design with matched participants [3 measurement times]; children randomly assigned
Rosenbluth et al. (2004); category 6 [USA]	5 program components including classroom curriculum; staff training; policy development; parent education; support	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

	services for individual students		control conditions
Sprober et al. (2006); category 5 [Germany]	Universal, multi-dimensional program for secondary pupils; cognitive-behavioral oriented program	145 secondary school students; 65 females and 80 males schools randomly assigned to 3 conditions: • proACT: class and teacher curriculum • proACT+E: class, teacher and parent curriculum • control group: unspecified intervention	Experimental pre-test post-test control group design; schools randomly assigned
Twemlow et al. (2005); category 6; not included in the meta-analysis [USA]	The Peaceful Schools Experiment; Mentalization-based approach: peer and adult mentorship; the Gentle Warriors physical education program; reflection time; classroom management/discipline plans; positive climate campaigns	•Randomized controlled trial in 9 elementary schools in the Midwest • Approximately 3,600 students exposed to the program • 3rd to 5th graders • 2 years of active intervention and 1 year of minimal input maintenance intervention	Randomized controlled trial
Before-After, Experimental-Control Comparisons			
Project	**Components of the Program**	**Participants**	**Research Design**
Alsaker & Valkanover (2001); Alsaker (2004); category 5 [Switzerland]	Program specific for kindergarten children aiming to enhance teachers' capacity to intervene in bullying situations; intensive focused supervision of	Children from 8 experimental and 8 control kindergartens: • 152 [50% girls] intervention children • 167 [50% girls) control children	Experimental pre-test post-test design with a waiting list control

	teachers for 4 months [8 meetings in total]		
Andreou et al. (2007); category 6 [Greece]	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, post-test 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); category 6 [USA]	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 non-randomized controlled trial with 10 public middle schools (7 intervention – implementing the Olweus Bullying Prevention Program– and 3 control)
Beran et al. (2004); category 5 [Canada]	Program that places emphasis on clinical support to victims and bullies in the form of individual and group counseling and in collaboration with community services	Initial Screening Sample: 197 children [120 girls] from two elementary schools • 25 children in the experimental [3-month follow-up] group; 77 children in the control group • grades 4 to 6 • control and comparison school significantly different at pre-test [table 1]	Pre-test and 3-month post-test experimental-control condition

Ciucci & Smorti (1998); category 6 [Italy]	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 are part of the control group All children are from one secondary school.	Experimental pre-test, post-test control group design
Cowie et al. (1994); category 5; not included in the meta-analysis [England]	A two-year project [summer 1990 to summer 1992] Experimental classes implementing a Cooperative Group Work training [CGW classes] Control classes implementing a Normal Curriculum program [NC classes] CGW training includes trust-building exercises, problem-solving tasks, role-play activities and discussion groups	• Final sample of 149 middle school students: CGW = 103 NC = 46 Ages: 7 to 12; 56% males • Within 2 schools, the experimental classes were matched with control classes [one of the two schools implemented the program for 1 year, but the other school increased the N of experimental classes to counterbalance] • A third school with only experimental classes for 2 years	Before-after experimental-control comparison with 4 measurement points [experimental classes compared with matched pairs]
Evers et al. (2007); category 6 [USA]	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	3X2 experimental design crossing 3 experimental groups with 2 treatment groups; pre-test and post-test measures; schools matched on key variables (type of community,

		in Treatment 1 --266 middle and 531 high school students in Treatment 2	region of country and % of students eligible for free lunches)
Fox & Boulton (2003); category 5 [England]	Specifically designed for victims of bullying; an eight-week social skills training program offered by two trainers [one hour per course]	From a screening sample of 505 children, 28 children were chosen based on peer nominations: • 15 experimental children [12 girls] • 13 waiting list control children [9 girls] • 4 schools participated [2 groups – experimental and control– from each school; 4 groups in total]	Pre-test and post-test experimental and control [waiting list] condition; [before - T1, after-T2, and 'follow-up'-T3 in the experimental group; before and after in the control group which received the program after T2 data collection]
Gini et al. (2003); category 5 [Italy]	Twice a week for a 4-month period students from the experimental group were treated in their classes. It addresses 3 educational areas: acknowledgment of the physical part of own body, working on own emotions and recognition of own bullying. The program also involves teachers through a 2-day training course	104 students from 6 classes of one school served as the experimental group, 76 students from another 6 classes from another school served as a control group	Pre-test and 5-month post test experimental-control comparisons
Gollwitzer et al. (2006); category 5 [Germany]	ViSC consists of 13 lessons divided into 3 phases: • Impulse phase	184 students from 2 German secondary schools:	Before-after experimental-control comparison; two

	[units 1-6] • Reflection phase [unit 7] • Action Phase [unit 8-13] Training conducted over 13 consecutive weeks	4 experimental classes [N = 109] 3 control classes [N = 75] Only 149 children retained for analyses	post-measurements: immediately after the end of the program [short-term follow up] and 4 months after the training [long-term follow up].
Heydenberk et al. (2006); category 6; not included in the meta-analysis [USA]	Conflict resolution program comprising seven 1-hour sessions; designed to increase affective vocabulary, emotional awareness and empathy, self-regulation and conflict-resolution skills	• 2 schools in Philadelphia participated • Treatment groups: 3rd and 4th grade students • Control groups: 3rd and 5th students Year 1: pilot study of 437 students [post-test only student evaluation] Year 2: pre-test/ post-test comparison group design with 236 treatment students and 41 comparison group students	Before-after experimental-control comparison with 2 measurement points [study 2; year 2]; only one post-test measurement in study 1 [year 1]
Martin et al. (2005); category 5 [Spain]	5-month program given by the authors and endorsed by the teachers; 30 sessions; role playing and reinforcement of social skills/enhancement of self-control; cognitive therapy approach	Students from 1 school [grade 6] Experimental group: 25 students [13 boys] Control group: 24 students [13 boys] Age range: 10 – 12 10 years: 8.16% 11 years: 85.71% 12 years: 6.13%	Before-after experimental-control comparison; non-equivalent control group; a five-month [30 sessions] intervention

Melton et al. (1998); Limber et al. (2004); category 6 [USA]	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 organized 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); category 6 [USA]	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 post-test design with ex ante selection of treatment and comparison groups; matched treatment and comparison groups at baseline
Menesini et al. (2003); category 5 [Italy]	Befriending Intervention Program; 5 phases of program implementation; emphasis on 'peer supporters'	Children from 2 secondary schools: • 9 experimental classes [94 boys and 84 girls] • 5 control classes [63 boys and 52 girls] • age range: 11 – 14	Pre-test post-test experimental-control comparison

Olweus: Bergen 2 [1997-1998]; category 6 [Norway]	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, post-test design with a control group; eleven intervention and eleven comparison schools
Pepler et al. (2004); category 6 [Canada]	Systemic school-based program; 3 similar elements of intervention across the 3 schools: staff training; codes of behavior; 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); category 6 [Canada]	12-week program based on the Bully Proofing Your School Program; psycho-educational program within the classroom; a peer mediation program; groups for	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, post-test design with a control group [one experimental school and one control school]

	children referred for involvement in bullying/ victimization		
Rican et al. (1996); category 6 [Czechoslovakia]	Program 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	Pre-test post-test experimental-control comparison
Stevens et al. (2000); category 6; not included in the meta-analysis [Belgium]	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 pre-test/post-test comparison including a control group [2 experimental groups –Treatment with Support and Treatment without Support- and one control group]
Wiefferink et al. (2006); category 6; not included in the meta-analysis [Netherlands]	No information on the one-year intervention is given	• 50 elementary schools in the Netherlands with approximately 4,000 students [aged 9 to 12] participated • 25 experimental schools • pre-test measures at the beginning of 2005/06 school year; post-test measures at school year's end	Before-after/ experimental-control design

Other Experimental –Control Comparisons			
Project	**Components of the Program**	**Participants**	**Research Design**
Galloway & Roland (2004); category 6 [Norway]	Professional development program for teachers; 4 in-service days over a 9-month period; 15 2-hour peer supervision sessions; hand outs 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– in a two-year period [1992-1994]
Kaiser-Ulrey (2003); category 5 [USA]	Based on the Kia-Kaha anti-bullying program. BEST is a complex alteration of the Kia-Kaha, having foundations within social cognitive theory and social competence theory. A 12-week intervention comprising 24 sessions of 45 minutes each. Emphasis on social problem solving techniques; awareness raising; teacher manual and teacher training; anti-bullying classroom rules	• 7th graders from one K-12 developmental research school in Northern Florida participated in the program • 58 students in the intervention group • 67 students in the control group	Intended to be before-after/ experimental-control comparison, but no data given at the pre-test; only results of analyses indicating prior equivalence of individuals within experimental and control conditions; thus, evaluation treated as 'other experimental-control design'

Ortega et al. (2004); category 6 [Spain]	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 pre-test and 901 intervention students at post-test In the 4 control schools • 440 control pupils	5 intervention schools [3 primary; 2 secondary] had pre-test and post-test measures, compared to 4 control schools with only post-test measures. Follow-up after 4 years
Raskauskas (2007); category 6 [New Zealand]	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]	Intervention schools compared with matched-comparison groups
Age-Cohort Designs			
Project	**Components of the Program**	**Participants**	**Research Design**
Ertesvag & Vaaland (2007); category 6 [Norway]	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 • 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

Hanewinkel (2004); category 6; not included in the meta-analysis [Germany]	Program based on the ideas of OBPP; 2-year intervention. • School level: questionnaire survey, playground supervision, staff meetings, teacher-parent meetings • Classroom level: classroom anti-bullying rules • Individual level: talks with bullies and victims, serious talks with parents of involved children	• In April 1994, 47 schools applied for participation in the program; a total number of 14, 788 students at the pre test measurement stage. • Schools reassured that they would not be obliged to follow-up during the implementation stage; 10 schools dropped out • 37 schools implemented the program: 6 primary, 14 Hauptschule, 8 Mittelschule, 6 Gymnasia and 3 Gesamtschule	Age-cohort design The study was initially designed as a quasi-experimental, pre-test/ post-test design. [Data assessment: not a within-individual repeated measurement; only students of same grades were compared; Hanewinkel, 2004: 86]
Koivisto (2004); category 6; not included in the meta-analysis [Finland]	• Intervention components varied from school to school and over the years * Intervention included parent-teacher meetings, anti-bullying rules, anti-bullying curriculum material, firm monitoring during recess time and a pupil-welfare group comprising the head teacher, a representative of teaching staff, the school psychologist, the school doctor and nurse	• A total number of 2729 students in grades 4, 6 and 7 from Kempele comprehensive schools • Initial survey in 1990 and follow-up assessments every two years for a total period of eight years	Age-cohort design with follow-up evaluations every 2 years

Olweus: Bergen 1 [1983-1985]; category 6 [Norway]	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]; category 6 [Norway]	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]; category 6 [Norway]	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	Data for assessments for the 14 out of 19 Oslo schools from the	Extended selection cohorts design; data provided for 5

category 6 [Norway]	Prevention Coordinating Committee]; Classroom level [e.g. classroom rules]; individual level [e.g. supervision of students]; and community level components	first cohort are provided. Students in grades 4-7 followed from 2001 until 2005. Students in grades 8-10 followed from 2001 until 2003	measurements for students in grades 4 through 7; data provided for 3 measurements for students in grades 8 through 10.
O'Moore and Minton (2004); category 6 [Ireland]	A whole-school approach to tackle bullying. A program including teacher training, information for parents, a teacher's resource pack, individual work with children involved , age-related anti-bullying handbooks	• 42 of the 100 primary schools in the county of Donegal involved in the program • evaluation of the program based on the data from 22 schools • age range of students: 6 – 11 years	Age-cohort design
Pagliocca et al (2007); category 6 [USA]	Implementation of the OBPP in Chula-Vista district schools. 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	3 primary schools participated in the program due to their higher crime rates than the state average. Over a 2-year period (2003 – 2005), a total of 3378 students in grades 3 through 6 received the program with a roughly equal distribution of boys and girls	Age-Cohort Design 3 time points; baseline (Spring 2003/T1), T2 one year later (Spring of the first year of the intervention) and T3, Spring of the second year of the intervention
Salmivalli et al. (2004); Salmivalli et al. (2005);	Intervention training for teachers; class-level interventions; school-level	8 schools from Helsinki and 8 schools from 4 towns near Turku	Age-longitudinal design with adjacent cohorts

category 6 [Finland]	interventions [whole-school anti-bullying policy]; individual-level interventions	1,220 students aged 9-12 in 16 schools [600 girls]	
Whitney et al. (1994); category 6 [England]	Whole-school approach; curriculum classroom strategies; the Heartstone Odyssey; quality circles; 'Only playing Miss' theatrical play; peer counseling; 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)	Age-cohort design

11.7 TABLE 7: KEY RESULTS OF 44 PROGRAM EVALUATIONS

Randomized Experiments		
Project Information	**Bullying**	**Victimization**
Baldry & Farrington (2004) [category 6]	**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
Beran & Shapiro (2005) [category 5]	EB M 10.41 (4.27) 66 EA M 9.68 (3.68) 66? CB M 8.91 (3.49) 63 CA M 8.61 (3.21) 63?	NA
Boulton & Flemington (1996) [category 5]	EB M 9.0 (2.1) 84 EA M 9.3 (2.2) 84 CB M 14.8 (5.3) 80 CA M 14.8 (5.1) 80	NA

Cross et al. (2004) [category 6]	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) [category 6]	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) [category 6]	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)
Fonagy et al. (2007) [category 6]	*(CAPSLE)* EB M 100.4 (9.72) 563 EA M 98.9 (9.02) 457 *(TAU)* CB M 98.2 (8.99) 360 CA M 99.3 (8.18) 274	*(CAPSLE)* EB M 100.64 (9.49) 563 EA M 99.0 (9.63) 457 *(TAU)* CB M 99.7 (9.77) 360 CA M 99.8 (9.20) 274
Frey et al. (2005) [category 6]	**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?
Hunt (2007) [category 6]	**Bullying Alone** EB: M 1.30 (0.60) 152 EA: M 1.17 (0.46) 111 CB: M 1.30 (0.66) 248 CA: M 1.31 (0.64) 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

	Bullying in Group EB: M 1.47 (0.70) 152 EA: M 1.39 (0.72) 111 CB: M 1.36 (0.75) 248 CA: M 1.41 (0.76) 207	
Jenson & Dieterich (2007) [category 6]	LOR = .161, SE = .280 (N = 667)	LOR = .491, SE = .286 (N = 668)
Karna et al. (2009) [category 6]	EB: 5.19% (3336) EA: 3.42 % (3336) CB: 5.60 % (2305) CA: 5.03 % (2305)	EB: 15.07 % (3345) EA: 9.03 % (3345) CB: 16.09 % (2306) CA: 14.27 % (2306)
Meyer & Lesch (2000) [category 5]	**School 1** E1B M 104.16 (26.24) 6 E1A M 119.5 (16.57) 6 C1B M 75.2 (34.09) 6 C1A M 74.0 (41.07) 6 **School 2** E2B M 82.0 (28.50) 6 E2A M 62.8 (20.91) 6 C2B M 86.4 (49.03) 6 C2A M 54.2 (13.92) 6 **School 3** E3B M 86.0 (17.81) 6 E3A M 75.5 (21.52) 6 C3B M 93.6 (21.83) 6 C3A M 109.4 (53.26) 6	NA
Rosenbluth et al. (2004) [category 6]	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?)
Sprober et al. (2006) [category 5]	**Verbal Bullying** E1B M 22.95 (5.64) 48? E1A1 M 23.46 (6.79) 48? E1A2 M 21.73 (4.70) 42? E2B M 22.94 (6.27) 48? E2A1 M 21.39 (3.98) 48? E2A2 M 21.38 (3.57) 42?	E1B M 20.02 (5.75) 48? E1A1 M 18.39 (5.20) 48? E1A2 M 17.71 (4.70) 42?

		E2B M 19.76 (4.26) 48?
	CB M 26.79 (6.80) 48?	E2A1 M 18.06 (3.29) 48?
	CA1 M 25.50 (5.56) 48?	E2A2 M 17.84 (3.46) 42?
	CA2 M 26.85 (7.79) 42?	
		CB M 20.38 (5.79) 48?
	Physical Bullying	CA1 M 18.82 (8.45) 48?
	E1B M 26.78 (2.37) 48?	CA2 M 19.32 (7.42) 42?
	E1A1 M 26.27 (3.51) 48?	
	E1A2 M 26.67 (3.53) 42?	
	E2B M 26.72 (4.05) 48?	
	E2A1 M 25.26 (2.43) 48?	
	E2A2 M 25.68 (2.17) 42?	
	CB M 29.08 (4.50) 48?	
	CA1 M 26.89 (3.79) 48?	
	CA2 M 28.89 (6.85) 42?	

Before-After, Experimental-Control Comparisons

Project Information	Bullying	Victimization
Alsaker & Valkanover (2001); Alsaker (2004) [category 5]	EB PR 41.4% (150) EA PR 40.1% (152) CB PR 31.7% (161) CA PR 33.5% (165)	EB PR 57.9% (150) EA PR 49.3% (152) CB PR 32.9% (161) CA PR 52.1% (164)
Andreou et al. (2007) [category 6]	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) [category 6]	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)
Beran et al. (2004)	NA	EB M 5.77 (6.1) 25

[category 5]		EA M 5.36 (5.5) 25 CB M 3.60 (3.5) 77 CA M 3.41 (3.4) 77
Ciucci & Smorti (1998) [category 6]	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)
Evers et al. (2007) [category 6]	**Middle School** EB 75.9% (266) EA 61.7% (266) CB 78.1% (483) CA 73.7% (483) **High School** EB 67.6% (531) EA 49.2 (531) CB 71.5 % (309) CA 67.0 % (309)	**Middle School** EB 82.0% (266) EA 60.2% (266) CB 80.3 % (483) CA 75.4% (483) **High School** EB 68.4% (531) EA 50.7% (531) CB 75.4% (309) CA 68.6% (309)
Fox & Boulton (2003) [category 5]	NA	EB PR M 29.47 (8.16) 15 EA PR M 34.29 (16.01) 15 CB PR M 31.54 (18.93) 13 CA PR M 33.56 (20.15) 13
Gini et al. (2003) [category 5]	EB: 11.1% (63) EA: 17.5% (63?) CB: 19.1% (47) CA: 23.4% (47?)	EB: 36.5% (63) EA: 41.3% (63?) CB: 51.1% (47) CA: 34.0% (47?)
Gollwitzer et al. (2006) [category 5]	EB M 1.56 (0.51) 89 EA1 M 1.58 (0.63) 89? EA2 M 1.46 (0.45) 89? CB M 1.54 (0.53) 60 CA1 M 1.55 (0.53) 60? CA2 M 1.57 (0.65) 60?	EB M 1.64 (0.65) 89 EA1 M 1.51 (0.60) 89? EA2 M 1.48 (0.55) 89? CB M 1.63 (0.49) 60 CA1 M 1.62 (0.60) 60? CA2 M 1.56 (0.60) 60?
Martin et al. (2005) [category 5]	EB 44% (25) EA 28% (25?) CB 20.83% (24) CA 25% (24?)	EB 28% (25) EA 20% (25?) CB 20.83% (24) CA 25% (24?)
Melton et al (1998) [category 6]	EB 24% (3904) EA 20% (3827) CB 19% (2485) CA 22% (2436)	EB 25% (3904) EA 19% (3827) CB 24% (2485) CA 19% (2436)

Menard et al. (2008) [category 6]	**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)
Menesini et al. (2003) [category 5]	EB PR M 2.24 (4.89) 178 EA PR M 2.06 (4.31) 178? CB PR M 2.04 (3.72) 115 CA PR M 3.02 (4.78) 115?	EB PR M 3.53 (6.19) 178 EA PR M 3.68 (6.68) 178? CB PR M 3.06 (5.54) 115 CA PR M 4.45 (6.90) 115?
Olweus / Bergen 2 [category 6]	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) [category 6]	E2S1: 32% (300) E2F2: 27% (240) E2F1: 26% (300) E2S2: 20% (240) E2S1: 32% (300) E2F3: 16% (163) E2F1: 26% (300) E2S3: 14% (163)	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: 23% (303) C3S2: 23% (303) E3F2: 23% (303) E3S3: 14% (289) E3S2: 23% (303) E3F3: 13% (289) C2F1: 26% (300) C2S1: 32% (300)	C3F2: 41% (303) C3S2: 39% (303) E3F2: 41% (303) E3S3: 28% (289) E3S2: 39% (303) E3F3: 28% (289) C2F1: 52% (300) C2S1: 42% (300)
Rahey & Craig (2002) [category 6]	**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) [category 6]	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)
Other Experimental-Control Comparisons		
Project Information	**Bullying**	**Victimization**
Galloway & Roland (2004) [category 6]	E: M .34 (.60?) 672 C: M .40 (.60?) 475	E: M .87 (.78?) 675 C: M 1.07 (.78?) 475
Kaiser-Ulrey (2003) [category 5]	E: M 1.51 (1.17) 58 C: M 1.36 (.83) 67	E: M 1.79 (1.31) 58 C: M 1.50 (1.12) 67
Ortega et al. (2004) [category 6]	E: 4.1% (910) C: 6.5% (440)	E: 4.2% (910) C: 8.5% (440)
Raskauskas (2007) [category 6]	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		
Project Information	**Bullying**	**Victimization**
Ertesvag & Vaaland (2007) [category 6]	**Grade 5**	**Grade 5**
	B: M .29 (.32) 118	B: M .54 (.49) 118
	A1: M .31 (.43) 126	A1: M .53 (.53) 126
	A2: M .21 (.33) 151	A2: M .43 (.48) 151
	A3: M .17 (.38) 143	A3: M .44 (.54) 143
	Grade 6	**Grade 6**
	B: M .36 (.38) 152	B: M .46 (.46) 152
	A1: M .28 (.43) 129	A1: M .50 (.57) 129
	A2: M .17 (.25) 130	A2: M .38 (.47) 130
	A3: M .21 (.30) 140	A3: M .39 (.46) 140
	Grade 7	**Grade 7**
	B: M .31 (.32) 147	B: M .44 (.51) 147
	A1: M .32 (.39) 160	A1: M .39 (.52) 160
	A2: M .30 (.40) 134	A2: M .44 (.52) 134
	A3: M .15 (.28) 140	A3: M .24 (.46) 140
	Grade 8	**Grade 8**
	B: M .32 (.49) 123	B: M .30 (.57) 123
	A1: M .25 (.33) 128	A1: M .21 (.34) 128
	A2: M .41 (.60) 112	A2: M .57 (.74) 112
	A3: M .25 (.49) 123	A3: M .32 (.40) 123
	Grade 9	**Grade 9**
	B: M .34 (.55) 95	B: M .26 (.39) 95
	A1: M .32 (.48) 128	A1: M .26 (.46) 128
	A2: M .35 (.59) 112	A2: M .36 (.55) 112
	A3: M .33 (.49) 122	A3: M .44 (.55) 122
	Grade 10	**Grade 10**
	B: M .35 (.49) 112	B: M .35 (.60) 112
	A1: M .41 (.55) 99	A1: M .27 (.34) 99
	A2: M .38 (.60) 149	A2: M .24 (.40) 149
	A3: M .31 (.56) 124	A3: M .24 (.34) 124
Olweus / Bergen 1 [category 6]	**Grades 5-7**	**Grades 5-7**
	B 7.28% (1689)	B 9.98% (1874)
	A1 5.02% (1663)	A1 3.78% (1691)
	Grades 6-7	**Grades 6-7**
	B 7.35% (1294)	B 9.92% (1297)
	A2 3.60% (1103)	A2 3.55% (1115)
Olweus/ Oslo 1 [category 6]	**Grades 5-7**	**Grades 5-7**
	B 6.4% (874)	B 14.4% (882)
	A 3.1% (983)	A 8.5% (986)

Olweus / New National [category 6]	**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)
Olweus/ Oslo 2 [category 6]	**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)
O'Moore and Minton (2004) [category] 6	**Grade 4** B 10.49 % (181) A 5.24 % (248)	**Grade 4** B 19.23 % (182) A 10.67 % (253)
Pagliocca et al. (2007) [category 6]	**Grades 3 - 6** B 27.86 % (1177) A1: 22.88 % (1088) A2: 24.33 % (1126)	**Grades 3 - 6** B 12.91 % (1177) A1: 10.84 % (1088) A2: 10.39 % (1126)
Salmivalli et al. (2005) [category 6]	**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
Whitney et al. (1994) [category 6]	**Primary** B: 10.0% (2519) A: 8.4% (2370) **Secondary** B: 6.2% (4103) A: 4.3% (4612)	**Primary** B: 26.0% (2523) A: 23.1% (2380) **Secondary** B: 10.0% (4116) A: 9.2% (4620)

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 Category 6 = Evaluations with a sample size above 200 and with self-reports as outcome measures of bullying Category 5 = Evaluations with a sample size less than 200 and/or with other outcome measures of bullying; ? = estimate by the authors.

11.8 TABLE 8: EFFECT SIZES FOR BULLYING

Project	OR	CI	Z	P
Randomized Experiments				
Baldry & Farrington (2004)	1.14	0.51 – 2.58	0.32	ns
Beran & Shapiro (2005)*	1.14	0.53 – 2.46	0.34	ns
Boulton & Flemington (1996)*	0.93	0.38 – 2.27	- 0.16	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
Fonagy et al. (2009)	1.66	1.10 – 2.50	2.41	.016
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
Karna et al. (2009)	1.38	0.92 – 2.06	1.58	ns
Meyer & Lesch (2000)*	0.68	0.16 – 2.90	- 0.52	ns
Rosenbluth et al. (2004)	0.99	0.63 – 1.58	-0.03	ns
Sprober et al. (2006)*	0.95	0.63 – 1.45	- 0.23	ns
Weighted mean (Q = 15.83, ns)	**1.10**	**0.97 – 1.26**	**1.44**	**ns**
Before-After, Experimental-Control				
Alsaker & Valkanover (2001)	1.15	0.55 – 2.40	0.36	ns
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
Evers et al. (2007)	1.65	1.15 – 2.36	2.72	.007
Gini et al. (2003)*	0.76	0.15 – 3.84	- 0.32	ns
Gollwitzer et al. (2006)*	1.23	0.63 – 2.42	0.61	ns
Martin et al. (2005)*	2.56	0.33 – 19.63	0.90	ns
Melton et al. (1998)	1.52	1.24 – 1.85	4.10	.0001
Menard et al. (2008)	1.74	1.45 – 2.09	5.98	.0001
Menesini et al. (2003)	1.60	0.81 – 3.16	1.34	ns
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
Weighted mean (Q = 6.24, ns)	**1.60**	**1.45 – 1.77**	**9.07**	**.0001**
Other Experimental-Control				
Galloway & Roland (2004)	1.20	0.91 – 1.59	1.27	ns
Kaiser-Ulrey (2003)*	0.76	0.33 – 1.76	- 0.64	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 (Q = 1.95, ns)	**1.20**	**1.04 – 1.38**	**2.57**	**.010**

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
O'Moore & Minton (2004)	2.12	0.81 – 5.55	1.53	ns
Pagliocca et al. (2007)	1.30	0.93 – 1.83	1.54	ns
Salmivalli et al. (2005)	1.31	1.07 – 1.61	2.56	.010
Whitney et al. (1994)	1.33	1.12 – 1.60	3.17	.002
Weighted mean (Q = 14.99, p = .059)	**1.51**	**1.35 – 1.70**	**7.10**	**.0001**
Weighted mean (Q = 70.89, p = 0001)	**1.36**	**1.26 – 1.47**	**7.86**	**.0001**

Note: OR = Odds Ratio; CI = Confidence Interval; * Initial N < 200

11.9 TABLE 9: 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
Fonagy et al. (2009)	1.39	1.02 – 1.91	2.07	.038
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
Karna et al. (2009)	1.55	1.28 – 1.88	4.49	.0001
Rosenbluth et al. (2004)	0.70	0.50 – 0.97	-2.14	.032
Sprober et al. (2006)*	1.15	0.64 – 2.09	0.47	ns
Weighted mean (Q = 17.94, p = .056)	**1.17**	**1.00 – 1.37**	**1.96**	**.050**
Before-After, Experimental-Control				
Alsaker & Valkanover (2001)	3.14	1.52 – 6.49	3.09	.002
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
Beran et al. (2004)*	1.04	0.28 – 3.88	0.06	ns
Ciucci & Smorti (1998)	1.21	0.70 – 2.12	0.69	ns
Evers et al. (2007)	1.79	1.23 – 2.60	3.06	.002
Fox & Boulton (2003)*	0.71	0.14 – 3.71	- 0.40	ns
Gini et al. (2003)*	0.40	0.12 – 1.40	- 1.43	ns
Gollwitzer et al. (2006)*	1.38	0.70 – 2.72	0.94	ns
Martin et al. (2005)*	1.97	0.23 – 16.78	0.62	ns
Melton et al. (1998)	1.06	0.91 – 1.23	0.70	ns
Menard et al. (2008)	1.26	1.05 – 1.51	2.48	.013
Menesini et al. (2003)	1.42	0.84 – 2.39	1.32	ns
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)	2.46	0.62 – 9.73	1.28	ns
Weighted mean (Q = 29.02, p = .024)	**1.22**	**1.06 – 1.37**	**2.72**	**.007**
Other Experimental-Control				
Galloway & Roland (2004)	1.59	1.20 – 2.11	3.26	.001
Kaiser-Ulrey (2003)*	0.65	0.28 – 1.50	- 1.02	ns
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 (Q = 5.98, ns)	**1.43**	**1.11 – 1.85**	**2.73**	**.006**

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
O'Moore & Minton (2004)	1.99	0.98 – 4.07	1.89	.059
Pagliocca et al. (2007)	0.92	0.71 – 1.21	- 0.57	ns
Salmivalli et al. (2005)	1.30	1.06 – 1.60	2.47	.014
Whitney et al. (1994)	1.14	1.00 – 1.29	2.01	.044
Weighted mean (Q = 57.04, p < .0001)	**1.44**	**1.21 – 1.72**	**4.15**	**.0001**
Weighted mean (Q = 129.82, p < .0001)	**1.29**	**1.18 – 1.42**	**5.61**	**.0001**

Note: OR = Odds Ratio; CI = Confidence Interval; * Initial N < 200

11.10 TABLE 10: UNITS OF RANDOM ALLOCATION*

Children:

De Rosier (2004) => 18 experimental students from each of 11 schools (N = 381)
Beran & Shapiro (2005) => 66 experimental students from 2 schools (N = 129)
Boulton & Flemington (1996) => 84 experimental students from 1 school (N = 164)
Meyer & Lesch (2000) => 18 experimental students from 3 schools (N = 36)

Classes:

Baldry & Farrington (2004) => 10 classes (N = 224)

Schools:

Cross et al. (2004) => 29 schools (N = 1957)
Fekkes et al. (2006) => 50 schools (N = 2221)
Fonagy et al. (2009) => 3 schools in experimental 1 condition; 3 schools in the experimental 2 condition (N = 923 in experimental 1 condition and control condition)
Frey et al. (2005) => 6 schools (N = 1126)
Hunt (2007) => 7 schools (N = 400)
Jenson & Dieterich (2007) => 28 schools (N = 668)
Karna et al. (2009) => 78 schools (N =5641)
Rosenbluth et al. (2004) => 12 schools (N = 1763)
Sprober et al. (2006) => 3 schools (N = 144)

* N showing total sample size of students in experimental and control conditions

11.11 TABLE 11: PROGRAM ELEMENTS AND THEIR FREQUENCY

Element	Frequency*
1. Whole-school anti-bullying policy	26
2. Classroom rules	31
3. School conferences/assemblies providing information about bullying to children	21
4. Curriculum materials	34
5. Classroom management	29
6. Cooperative group work among experts (e.g. teachers, counselors and interns)	24
7. Work with bullies	25
8. Work with victims	25
9. Work with peers (e.g. peer mediation, peer mentoring, peer group pressure as bystanders)	16
10. Information for teachers	39
11. Information for parents	30
12. Improved playground supervision	12
13. Disciplinary methods	13
14. Non-punitive methods (e.g. Pikas, No Blame)	11
15. Restorative Justice approaches	0
16. School tribunals; school bully courts	0
17. Teacher training	30
18. Parent training/meetings	17
19. Videos	21
20. Virtual Reality computer games	3

* Out of 44 evaluations

11.12 TABLE 12: SIGNIFICANT RELATIONSHIPS WITH BULLYING

	Cat (N) OR (Q)	Cat (N) OR (Q)	QB	P
Program Elements				
Parent training/meetings	No (24) 1.25 (ns)	Yes (17) 1.57 (ns)	19.61	.0001
Playground supervision	No (30) 1.29 (.038)	Yes (11) 1.53 (ns)	18.65	.0001
Intensity for children	19- (19) 1.25 (ns)	20+ (13) 1.62 (.0002)	18.19	.0001
Intensity for teachers	9- (16) 1.19 (ns)	10+ (20) 1.52 (.015)	17.68	.0001
Duration for children	240- (20) 1.17 (ns)	270+ (20) 1.49 (.017)	16.59	.0001
Disciplinary methods	No (28) 1.31 (.058)	Yes (13) 1.59 (ns)	13.18	.0003
Duration for teachers	3- (19) 1.22(ns)	4+ (19) 1.50 (ns)	12.73	.0004
Classroom management	No (13) 1.15 (ns)	Yes (28) 1.44 (.001)	7.91	.005
Teacher training	No (13) 1.24 (ns)	Yes (28) 1.46 (.002)	7.43	.006
Classroom rules	No (11) 1.15 (ns)	Yes (30) 1.44 (.003)	7.41	.006
Whole-school policy	No (17) 1.19 (ns)	Yes (24) 1.44 (.002)	7.12	.008
School conferences	No (21) 1.30 (.044)	Yes (20) 1.49 (.032)	6.98	.008
Total elements	10- (23) 1.30 (ns)	11+ (18) 1.48 (.016)	6.79	.009
Based on Olweus	No (25) 1.31 (.037)	Yes (16) 1.50 (.031)	6.45	.011
Information for parents	No (13) 1.21 (ns)	Yes (28) 1.44 (.001)	6.20	.013
Cooperative group work	No (19) 1.31 (ns)	Yes (22) 1.48 (.0004)	5.54	.019
Design Features				
Age of children	10- (18) 1.22 (ns)	11+ (23) 1.50 (ns)	14.95	.0001
Outcome measure	Other (31) 1.32 (.036)	2+ M (10) 1.64 (ns)	13.92	.0002
Publication date	04+ (25) 1.31 (.015)	03- (16) 1.56 (ns)	11.07	.0009
In Norway	Rest (34) 1.33 (.039)	Nor (7) 1.58 (.058)	10.15	.001
In Europe	Rest (17) 1.33 (.001)	EU (24) 1.46 (ns)	3.41	.065
Sample size	899- (21) 1.26 (.0004)	900+ (20) 1.43 (ns)	3.29	.070

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); (Q) shows significance of Q statistic for each category; ns = not significant.

11.13 TABLE 13: SIGNIFICANT RELATIONSHIPS WITH VICTIMIZATION

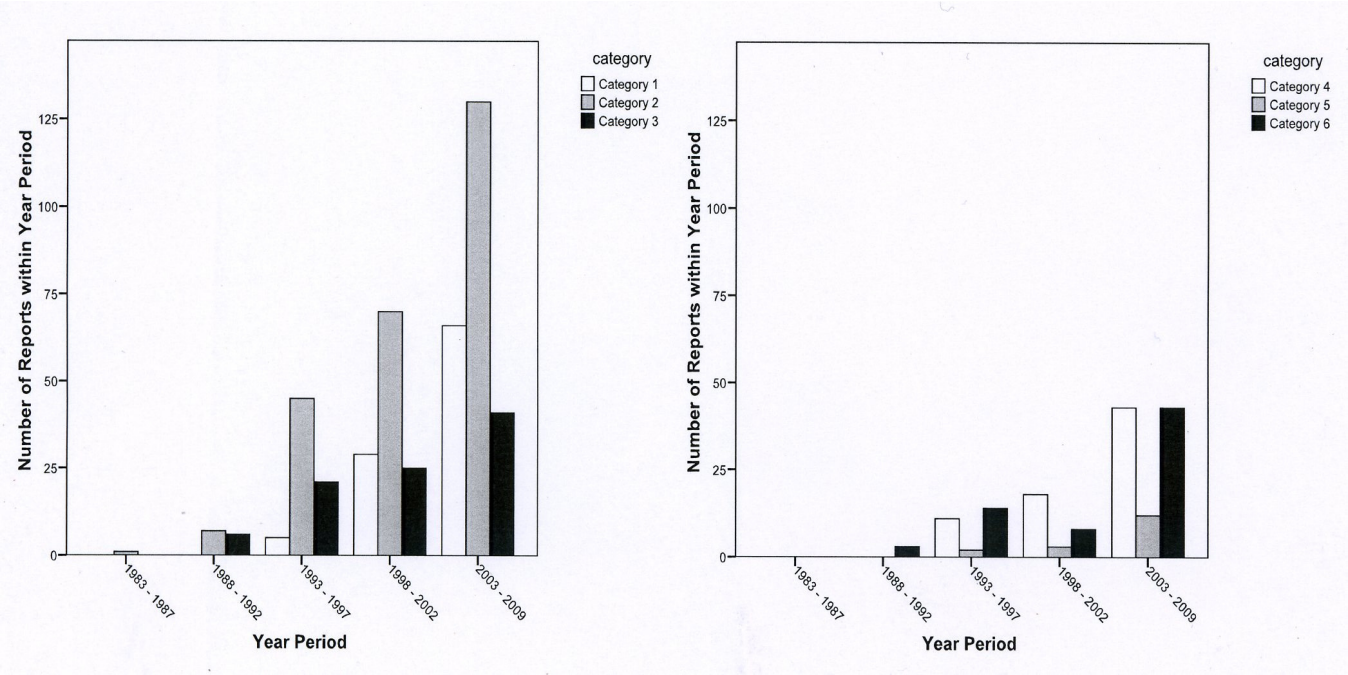
	Cat (N) OR (Q)	Cat (N) OR (Q)	QB	P
Program Elements				
Work with peers	No (25) 1.39 (.0001)	Yes (16) 1.13 (.016)	19.34	.0001
Disciplinary methods	No (28) 1.21 (.005)	Yes (13) 1.44 (.0001)	18.41	.0001
Parent training/meetings	No (24) 1.20 (ns)	Yes (17) 1.41 (.0001)	14.75	.0001
Duration for teachers	3- (18) 1.18 (ns)	4+ (20) 1.41 (.0001)	12.84	.0003
Videos	No (22) 1.17 (.002)	Yes (19) 1.38 (.0001)	12.36	.0004
Cooperative group work	No (18) 1.20 (.028)	Yes (23) 1.38 (.0001)	10.82	.001
Duration for children	240- (20) 1.15 (.007)	270+ (20) 1.35 (.0001)	10.81	.001
Intensity for children	19- (18) 1.21 (ns)	20+ (14) 1.42 (.0008)	9.40	.002
Intensity for teachers	9- (15) 1.22 (ns)	10+ (21) 1.37 (.0001)	4.83	.028
Design Features				
Outcome measure	Other (31) 1.18 (.008)	2+ M (10) 1.57 (.0001)	43.58	.0001
In Norway	Rest (34) 1.18 (.002)	Nor (7) 1.55 (.0001)	40.97	.0001
Not in US or Canada	US/Can (14) 1.06 (.024)	Rest (27) 1.42 (.0001)	39.21	.0001
In Europe	Rest (17) 1.11 (.011)	EU (24) 1.44 (.0001)	36.83	.0001
Design	12 (28) 1.16 (.010)	34 (13) 1.41 (.0001)	19.80	.0001
Publication date	04+ (26) 1.21 (.005)	03- (15) 1.42 (.0001)	15.07	.0001
Age of Children	10 – (18) 1.22 (.016)	11+ (23) 1.34 (.0001)	3.93	.047

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); (Q) shows significance of Q statistic for each category; ns = not significant

11.14 TABLE 14: RESULTS OF WEIGHTED REGRESSION ANALYSES

	B	SE (B)	Z	P
Bullying Effect Size				
(a) 20 Elements only				
Parent training/meetings	.1808	.0557	3.25	.001
Disciplinary methods	.1178	.0582	2.02	.043
(b) All Elements				
Intensity for children	.1726	.0675	2.56	.010
Parent training/meetings	.1594	.0635	2.51	.012
Victimization Effect Size				
(a) 20 Elements only				
Work with peers	- .2017	.0478	4.22	.0001
Videos	.1285	.0505	2.55	.011
Disciplinary methods	.1102	.0469	2.35	.019
(b) All Elements				
Work with peers	- .2362	.0480	4.93	.0001
Duration for children	.1498	.0536	2.79	.005
Videos	.1338	.0491	2.73	.006

11.15 FIGURE 1: NUMBER OF REPORTS IN EACH CATEGORY WITHIN YEAR PERIOD



11.16 FIGURE 2: KEY FEATURES OF THE EVALUATION

STUDY: ↓ ELEMENTS: →	N.C.	T.O.	D.C.	I.C.	D.T.	I.T.	O.M.	S.S.	P.D.	A.A.	I.L.	M.D.
Randomized Experiments												
Baldry & Farrington, 2004	A	D	E	G	✖	✖	M	O	R	T	V	Y
Beran & Shapiro, 2005	A	D	E	G	✖	✖	M	O	R	S	W2	Y
Boulton & Flemington, 1996	A	D	E	G	✖	✖	M	O	Q	T	V	Y
Cross et al., 2004	B	D	F	G	I	K	M	P	R	S	W2	Y
De Rosier, 2004	A	D	E	G	✖	✖	M	O	R	S	W1	Y
Fekkes et al., 2006	A	C	F	H	I	■	M	P	R	S	V	Y
Fonagy et al, 2009	B	D	F	■	I	■	M	P	R	S	W1	Y
Frey et al., 2005	A	D	E	G	■	K	M	P	R	S	W1	Y
Hunt, 2007	A	D	E	G	✖	✖	M	O	R	T	W2	Y
Jenson & Dieterich, 2007	A	D	F	H	✖	✖	M	O	R	S	W1	Y
Meyer & Lesch, 2000	A	D	E	H	✖	✖	M	O	Q	T	W2	Y
Rosenbluth et al., 2004	B	C	E	G	I	L	M	P	R	T	W1	Y
Salmivalli et al. 2009	B	D	F	H	J	L	N	P	R	T	V	Y
Sprober et al., 2006	B	C	E	G	J	K	N	O	R	T	V	Y
Before-After Experimental-Control Comparisons												
Alsaker & Valkanover, 2001	A	D	E	G	J	L	M	O	Q	S	V	Y
Andreou et al., 2007	A	D	E	G	J	L	M	O	R	S	V	Y
Bauer et al., 2007	B	C	E	G	J	K	M	P	R	T	W1	Y
Beran et al., 2004	A	D	E	G	I	K	M	O	R	T	W2	Y
Ciucci & Smorti, 1998	A	D	F	■	I	■	M	O	Q	S	V	Y
Evers et al., 2007	A	D	■	G	✖	✖	M	P	R	T	W1	Y
Fox & Boulton, 2003	A	D	E	G	✖	✖	M	O	Q	S	V	Y
Gini et al., 2003	A	C	E	G	J	L	M	O	Q	S	V	Y
Gollwitzer et al., 2006	A	D	F	G	✖	✖	M	O	R	T	V	Y
Martin et al., 2005	A	D	E	H	✖	✖	M	O	R	T	V	Y
Melton et al., 1998	B	C	F	H	■	■	M	P	Q	T	W1	Y
Menard et al., 2008	A	D	F	G	■	L	M	P	R	T	W1	Y
Menesini et al., 2003	B	D	E	H	J	L	M	O	Q	T	V	Y
Bergen 2 [1997-1998]	B	C	E	H	J	L	N	P	Q	T	U	Y
Pepler et al., 2004	B	C	F	■	J	K	M	O	R	S	W2	Y
Rahey & Craig, 2002	B	D	E	G	I	■	M	O	Q	S	W2	Y
Rican et al, 1996	A	C	E	■	✖	✖	M	O	Q	S	V	Y
Other Experimental-Control Comparisons												
Galloway & Roland, 2004	A	D	F	■	J	L	M	P	R	S	U	Z
Kaiser-Ulrey, 2003	B	C	E	H	J	K	M	O	Q	T	W1	Z
Ortega et al., 2004	B	D	F	H	J	L	N	P	R	T	V	Z
Raskauskas, 2007	A	D	E	G	✖	✖	M	P	R	S	W2	Z
Age-Cohort Designs												
Ertesvag & Vaaland, 2007	B	D	F	■	J	L	M	P	R	T	U	Z
Bergen 1 [1983-1985]	B	C	F	H	J	L	N	P	Q	T	U	Z
Oslo 1 [1999-2000]	B	C	F	H	J	L	N	P	Q	T	U	Z
National Norway [2001-2007]	B	C	F	H	J	L	N	P	Q	T	U	Z
Oslo 2 [2001-2006]	B	C	F	H	J	L	N	P	Q	T	U	Z
O' Moore & Minton, 2004	A	D	E	■	J	L	N	O	R	S	V	Z
Paglioca et al, 2007	B	C	F	H	J	L	N	P	R	S	W1	Z
Salmivalli et al., 2004	A	C	F	■	J	L	M	P	R	T	V	Z
Whitney et al., 1994	B	C	F	■	✖	✖	M	P	Q	T	V	Z

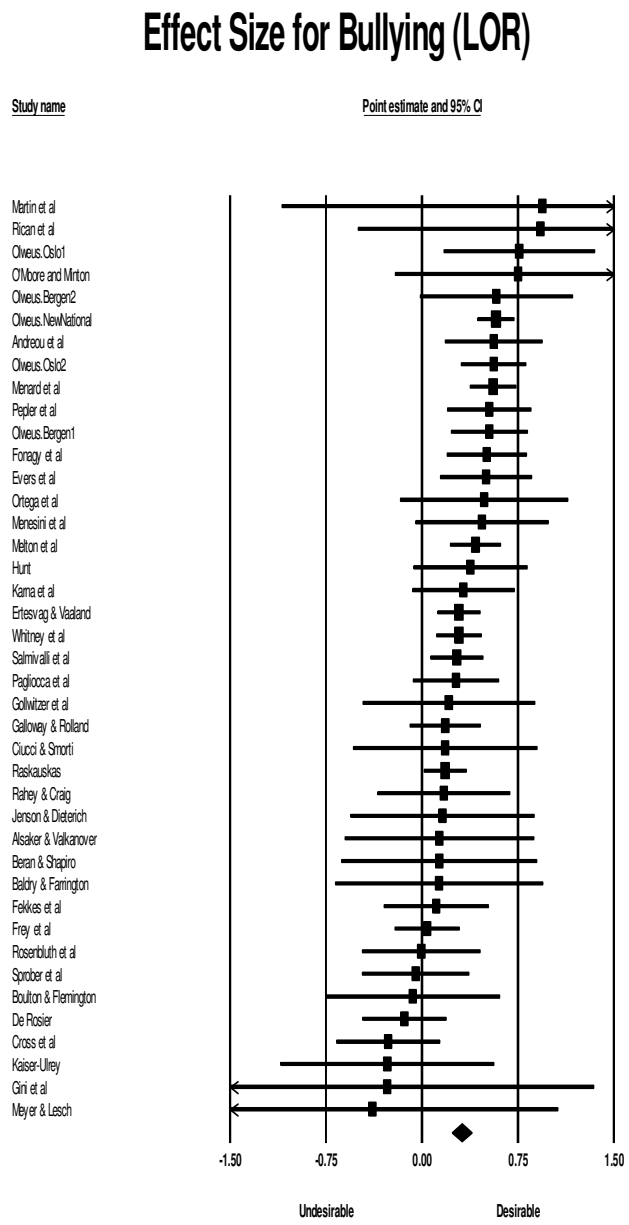
Note: N.C. = Number of Intervention Components [A = 10 or less; B = 11 or more]; T.O. = Theoretical Orientation [C = based/ inspired by Olweus; D = different from Olweus]; D.C. = Duration of Intervention for Children [E = 240 days or less; F = 270 days or more]; I.C. = Intensity of Intervention for Children [G = 19 hours or less; H = 20 hours or more]; D.T. = Duration of Intervention for Teachers [I = 3 day meetings or less; J = 4 day meetings or more]; I.T. = Intensity of Intervention for Teachers [K = 9 hours or less; L = 10 hours or more]; O.M. = Outcome Measure [M = means, prevalence, other measures; N = 2 or more times per month]; S.S. = Sample Size [O = 899 or less; P = 900 or more]; P.D. = Publication Date [Q = 2003 or before; R = 2004 or later]; A.A. = Average Age [S = 10 or less; T = 11 or more]; I.L. = Location of Intervention [U = in Norway; V = elsewhere in Europe; W1 = in the USA; W2 = other than Europe and the USA]; M.D. = Methodological Design [Y = randomized experiment or before-after experimental-control comparison; Z = other experimental-control comparison or an age-cohort design] ✖= not an intervention element; ■ = missing value

11.17 FIGURE 3: INTERVENTION COMPONENTS

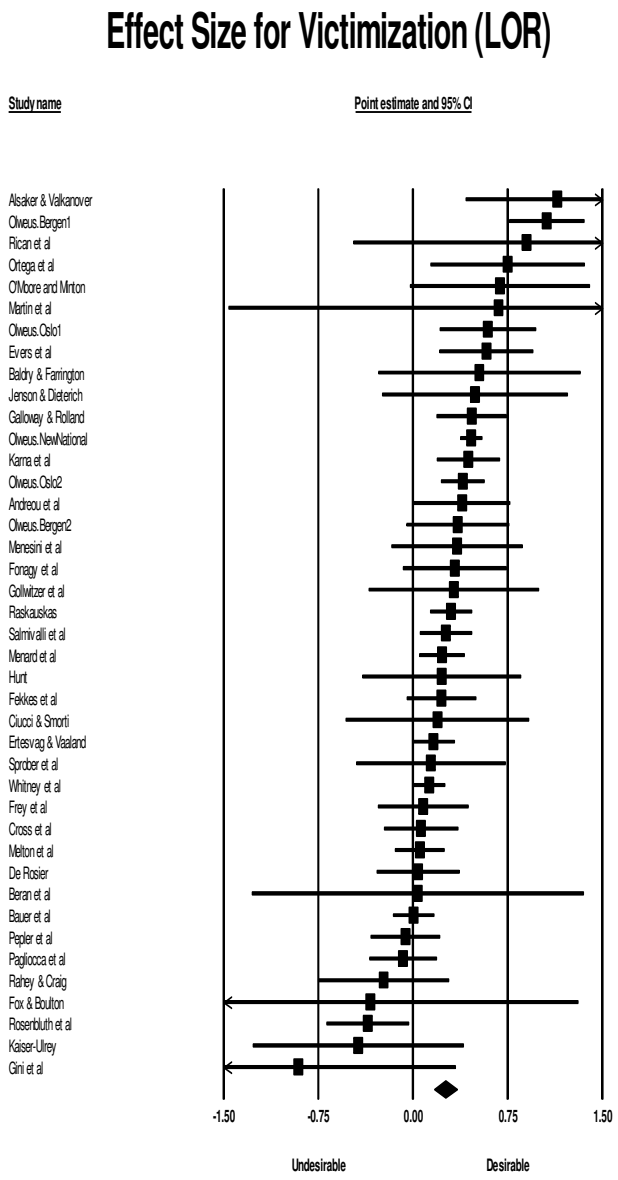
STUDY: ↓ ELEMENTS: →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Randomized Experiments																				
Baldry & Farrington, 2004	✖	✖	✔	✔	✖	✖	✖	✖	✖	✔	✖	✖	✖	✖	✖	✖	✖	✖	✔	✖
Beran & Shapiro, 2005	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✔	✖
Boulton & Flemington, 1996	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✔	✖
Cross et al., 2004	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✖	✖	✔	✖	✖	✔	✖	✔	✖
De Rosier, 2004	✖	✖	✖	✔	✖	✔	✔	✔	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Fekkes et al., 2006	✔	✔	✔	✔	✖	✖	✖	✖	✖	✔	✔	✔	✔	✖	✖	✖	✔	✖	✖	✔
Fonagy et al., 2009	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✖	✖	✔	✖	✖	✔	✔	✔	✖
Frey et al., 2005	✔	✔	✖	✔	✔	✖	✔	✔	✖	✔	✔	✖	✖	✖	✖	✖	✔	✖	✔	✖
Hunt, 2007	✔	✔	✔	✔	✔	✖	✖	✖	✖	✔	✔	✖	✖	✖	✖	✖	✖	✖	✖	✖
Jenson & Dieterich, 2007	✖	✖	✔	✔	✖	✖	✖	✖	✖	✔	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Meyer & Lesch, 2000	✖	✖	✖	✖	✖	✔	✔	✖	✖	✖	✖	✖	✖	✔	✖	✖	✖	✖	✖	✖
Rosenbluth et al., 2004	✔	✖	✔	✔	✔	✔	✖	✔	✔	✔	✔	✖	✖	✖	✖	✖	✔	✔	✖	✖
Salmivalli et al. 2009	✔	✔	✖	✔	✔	✔	✔	✔	✔	✔	✔	✖	✔	✔	✖	✖	✔	✔	✔	✔
Sprober et al., 2006	✖	✔	✖	✔	✔	✖	✔	✔	✔	✔	✔	✖	✔	✔	✖	✖	✔	✔	✔	✖
Before-After Experimental-Control Comparisons																				
Alsaker & Valkanover, 2001	✖	✔	✖	✖	✔	✖	✖	✖	✖	✔	✔	✖	✖	✔	✖	✖	✔	✔	✖	✖
Andreou et al., 2007	✖	✔	✖	✔	✔	✖	✖	✖	✖	✔	✖	✖	✖	✖	✖	✖	✔	✖	✖	✖
Bauer et al., 2007	✔	✔	✔	✔	✔	✖	✔	✔	✖	✔	✔	✔	✖	✖	✖	✖	✔	✖	✖	✖
Beran et al., 2004	✔	✖	✖	✔	✖	✔	✔	✔	✖	✔	✔	✖	✖	✔	✖	✖	✔	✖	✔	✖
Ciucci & Smorti, 1998	✖	✖	✔	✖	✖	✔	✖	✖	✔	✔	✖	✖	✖	✖	✖	✖	✔	✖	✖	✖
Evers et al., 2007	✖	✖	✖	✖	✖	✖	✖	✖	✖	✔	✔	✖	✖	✖	✖	✖	✖	✖	✔	✔
Fox & Boulton, 2003	✖	✖	✖	✖	✖	✔	✖	✔	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖	✖
Gini et al., 2003	✖	✔	✔	✔	✖	✔	✖	✖	✖	✔	✖	✖	✖	✖	✖	✖	✔	✖	✖	✖
Gollwitzer et al., 2006	✖	✔	✖	✔	✖	✖	✖	✖	✔	✔	✔	✖	✖	✖	✖	✖	✖	✖	✖	✖
Martin et al., 2005	✖	✔	✖	✔	✖	✔	✔	✔	✔	✔	✔	✖	✖	✖	✖	✖	✖	✖	✖	✖
Melton et al., 1998	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✔	✖	✖	✖	✔	✖	✔	✖
Menard et al., 2008	✔	✔	✖	✔	✔	✔	✖	✖	✖	✔	✔	✖	✖	✖	✖	✖	✔	✔	✖	✖
Menesini et al., 2003	✔	✔	✖	✖	✔	✔	✔	✔	✔	✔	✔	✖	✖	✖	✖	✖	✔	✔	✖	✖
Bergen 2 [1997-1998]	✔	✔	✔	✔	✔	✔	✔	✔	✖	✔	✔	✔	✔	✖	✖	✖	✔	✔	✔	✖
Pepler et al., 2004	✔	✔	✔	✔	✔	✖	✔	✔	✔	✔	✔	✔	✔	✖	✖	✖	✔	✔	✖	✖
Rahey & Craig, 2002	✔	✔	✖	✔	✔	✔	✔	✔	✔	✔	✖	✖	✖	✖	✖	✖	✔	✔	✖	✖
Rican et al, 1996	✖	✔	✖	✔	✔	✖	✖	✖	✖	✔	✖	✖	✔	✖	✖	✖	✖	✖	✔	✖
Other Experimental-Control Comparisons																				
Galloway & Roland, 2004	✖	✖	✖	✖	✖	✖	✖	✖	✖	✔	✖	✖	✖	✖	✖	✖	✔	✖	✖	✖
Kaiser-Ulrey, 2003	✖	✔	✖	✔	✔	✔	✔	✖	✔	✔	✔	✖	✔	✖	✖	✖	✔	✖	✖	✖
Ortega et al., 2004	✔	✔	✖	✔	✔	✔	✔	✔	✔	✔	✔	✖	✖	✔	✖	✖	✔	✖	✖	✖
Raskauskas, 2007	✔	✔	✖	✔	✔	✔	✖	✖	✖	✔	✔	✖	✖	✖	✖	✖	✖	✖	✔	✖
Age-Cohort Designs																				
Ertesvag & Vaaland, 2007	✔	✔	✔	✖	✔	✖	✔	✔	✖	✔	✔	✔	✖	✖	✖	✖	✔	✔	✔	✖
Bergen 1 [1983-1985]	✔	✔	✔	✔	✔	✔	✔	✔	✖	✔	✔	✔	✔	✖	✖	✖	✔	✔	✔	✖
Oslo 1 [1999-2000]	✔	✔	✔	✔	✔	✔	✔	✔	✖	✔	✔	✔	✔	✖	✖	✖	✔	✔	✔	✖
National Norway [2001-2007]	✔	✔	✔	✔	✔	✔	✔	✔	✖	✔	✔	✔	✔	✖	✖	✖	✔	✔	✔	✖
Oslo 2 [2001-2006]	✔	✔	✔	✔	✔	✔	✔	✔	✖	✔	✔	✔	✔	✖	✖	✖	✔	✔	✔	✖
O' Moore & Minton, 2004	✔	✖	✔	✔	✔	✖	✖	✖	✔	✔	✔	✖	✖	✔	✖	✖	✔	✔	✖	✖
Pagliocca et al., 2007	✔	✔	✔	✔	✔	✔	✔	✔	✖	✔	✔	✔	✔	✖	✖	✖	✔	✔	✔	✖
Salmivalli et al., 2004	✔	✔	✖	✔	✔	✖	✔	✔	✖	✔	✖	✖	✖	✔	✖	✖	✔	✖	✖	✖
Whitney et al., 1994	✔	✔	✔	✔	✔	✖	✔	✔	✔	✔	✔	✔	✖	✔	✖	✖	✖	✖	✔	✖

Note: 1 =whole-school anti-bullying policy; 2 =classroom rules; 3 = school conferences providing information about bullying to pupils; 4 = curriculum materials; 5 = classroom management; 6 = cooperative group work among experts [e.g. among teachers, counselors and interns]; 7 = work with bullies; 8 = work with victims; 9 = work with peers [e.g. peer mediation; peer mentoring; peer group pressure as bystanders]; 10 = information for teachers; 11 = information for parents; 12 = increased playground supervision; 13 = disciplinary methods; 14 = non-punitive methods [e.g. 'Pikas' or 'No Blame Approach']; 15 = restorative justice approaches; 16 = school tribunals/ school bully courts; 17 = teacher training; 18 = parent training; 19 = videos; 20 = virtual reality environments/ computer games

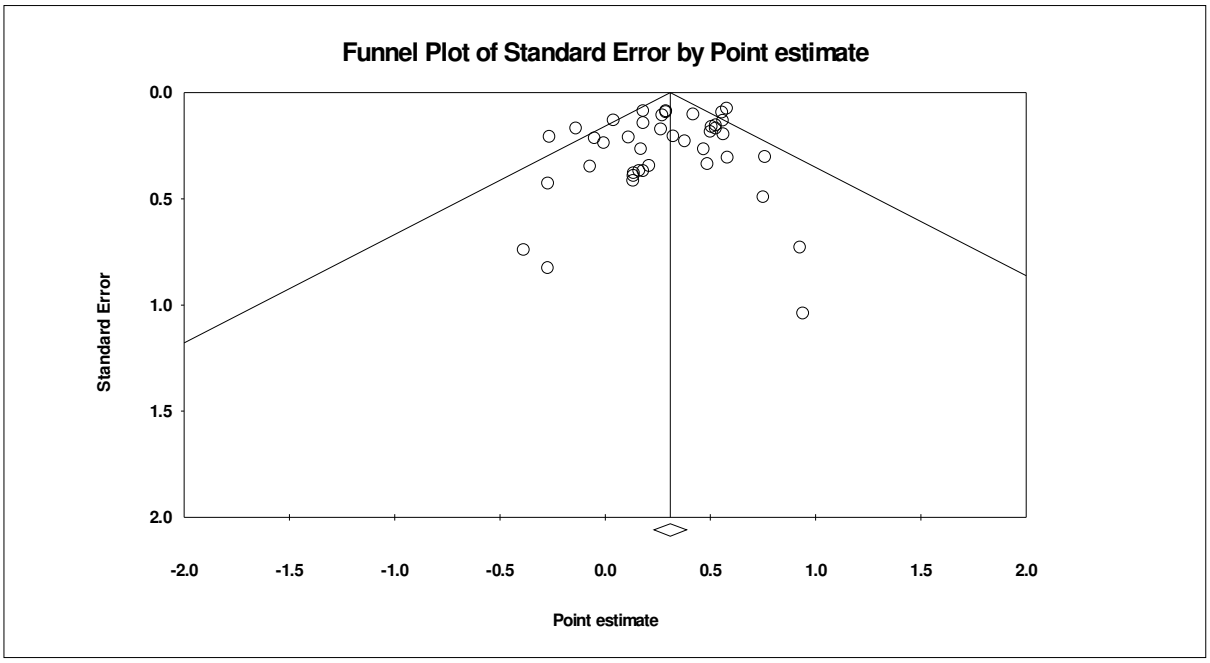
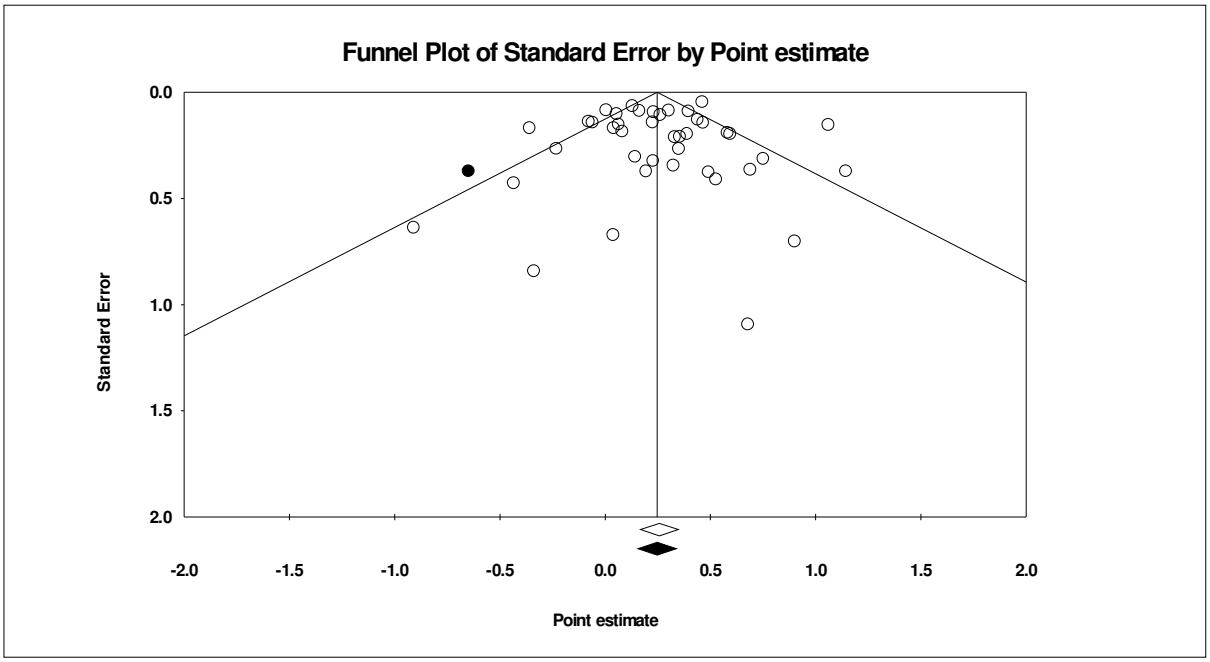
11.18 FIGURE 4: FOREST GRAPH FOR BULLYING



11.19 FIGURE 5: FOREST GRAPH FOR VICTIMIZATION



11.20 FIGURE 6: FUNNEL PLOT FOR BULLYING



11.21 FIGURE 7: FUNNEL PLOT FOR VICTIMIZATION