UNIVERSITY OF CAMBRIDGE

Evaluating the effectiveness of school- and cyber-bullying intervention and prevention programmes: Two systematic and meta-analytical reviews

Hannah Gaffney

Newnham College

February 2020

This thesis is submitted for the degree of Doctor of Philosophy

University of Cambridge

Preface

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text.

This thesis is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text.

This thesis, including footnotes, does not exceed the permitted length. The word count exclusive of bibliography and preliminary matters is 86,756.

This thesis is presented in accordance with the American Psychological Association publication manual version 7 and written in British English.

For John A. Hassett

Who always encouraged me to question everything

Rest in peace Grumpy Grandad

Acknowledgements

This thesis would not have come to fruition without the help and support of numerous different people; who must be thanked. Firstly, this dissertation would not have been possible without two individuals in particular, my doctoral supervisors, Dr. Maria Ttofi and Professor David Farrington. Maria and David took a chance on me in October 2015 and offered me a job working alongside them, even though my interview was truly horrible! They have been my backbone throughout my doctoral degree, and I am so thankful for the opportunities they have opened up for me both in the past, and in the future.

My friends and family are also deserving of a massive thank you. In particular my Mum, for always supporting my education, all twenty-five years of it, and for always taking my phone calls when I was stressed or on the verge of breaking down. I would also not be where I was today without my Dad, Rachel and extended family. My friends, both near and afar, also deserve so many thanks, for constantly being a shoulder to lean on and a constant source of wisdom and guidance. The year of 2019 was the most difficult I've ever faced, from the loss of my dear grandfather, cherished colleagues, and personal struggles, let alone trying to finish a PhD, and I truly believe I have come through this difficult stage because of my Newnham gals, my Criminology friends, and my oldest and dearest friends Sarah and Lisa.

A thank you also to my Spencer, the main man in my life. Thank you for all the cuddles and unconditional love when I would come home to escape the Cambridge bubble. Someone else who deserves a huge thank you wasn't there in the beginning but has stuck by me in the final stages of finishing the PhD. André Simoes Dias Vieira deserves all the thanks in the world. He put up with my craziness, made sure I was always well fed and for simply took such good care of me. I love you so much and I am eternally gratefully for your caring nature.

Table of Contents

List of Tables

List of Figures

Part One: Introduction

1. General Introduction

1.1 Overview

In this introductory chapter, several necessary concepts are defined and presented. The research questions are outlined, a declaration of published works is made and an outline of the thesis is provided. Furthermore, the previous related research, including the previous meta-analyses that have also examined the effectiveness of school-based anti-bullying programmes, is outlined and discussed. As such, the impetus for the present research is given. The current chapter also provides a response to critiques of previous research and defines specific terms that are used interchangeably throughout the dissertation.

1.2 Dissertation outline

This dissertation is structured as a traditional doctoral dissertation and it is based on multiple connected and published articles. The dissertation is structured in four related parts: (1) Introduction; (2) School-bullying; (3) Cyberbullying; and (4) Discussion.

The first part sets the scene for the research project and includes two chapters, including the current chapter that gives an overarching introduction to the field of research and a chapter describing the methods utilised. Chapter 2 gives a detailed outline of the methods used in this dissertation, namely, the systematic review approach and the technical methods used to conduct meta-analyses. A general overview of the steps required to undertake a systematic review is given as this is a fundamental aspect of the present research.

Chapter 2 also gives a thorough examination of all stages involved in conducting a meta-analytical review. The process to estimate raw effect sizes from primary studies that yield data as dichotomous and continuous variables is provided, along with the method of estimating the mean effect size in a meta-analysis (section 2.4). Heterogeneity and homogeneity of effect sizes are important concepts in any meta-analysis and these are described in Chapter 2 (section 2.5). Moreover, the different computational models of meta-analysis are described in detail as the model under which the results are presented is an important factor for consideration in the present research (section 2.6). Moderator analyses analogous to the ANOVA and meta-regression techniques are also explained in detail (sections 2.7 and 2.8, respectively) as these analyses are fundamental to the intricate examination of '*what works'* in effective anti-bullying programmes. Finally, Chapter 2 also provides details of the manual calculations required for the present research and the necessary corrections and adjustments needed.

Part two of this dissertation is concerned with school-bullying, and the meta-analysis to examine the effectiveness of school-based anti-bullying programmes in reducing both school-bullying perpetration and victimisation. The problem of school-bullying, as well as an explanation of the behaviours included in a definition of school bullying, is provided (Chapter 3, sections 3.2 and 3.3 respectively). A detailed synopsis of the existing literature on school-bullying is provided in Chapter 3, with a particular emphasis on the many negative and undesirable outcomes associated with experiencing bullying in school (section 3.4). This chapter highlights the importance of understanding what works in anti-bullying programmes and the potential risk and protective factors are also discussed (section 3.5). Awareness of these factors is an important element of understanding the development and application of anti-bullying programmes. Finally, a brief insight into current efforts to prevent school-bullying is provided (section 3.6) but as this is the primary focus of the current research, this section is not extensive.

Chapter 3 reviews the literature on various aspects of school-bullying behaviours, with the following chapters focusing on the systematic review and meta-analysis of primary evaluations of school-based anti-bullying programmes. Chapter 4 provides the reader with a detailed overview of the systematic review undertaken to identify eligible primary evaluations for the school-bullying meta-analysis. The inclusion and exclusion criteria (section 4.1) and detailed information regarding the searches and screening (sections 4.2 and 4.3, respectively) are outlined. Detailed information about the studies excluded for various reasons is also given (section 4.4), as well as information about the primary evaluations included in subsequent analysis (section 4.5). Chapter 5 describes the process of extracting data from these primary evaluations. Sections 5.2 and 5.3 describe the information that was extracted in relation to descriptive aspects of the primary studies and the evaluation methodologies, respectively. A brief comparative review of the strengths and limitations of the evaluation methodologies utilised by primary evaluations is also given. The majority of Chapter 5 outlines the process of coding the intervention components (i.e., specific activities and features of the anti-bullying programmes) included by the interventions. Section 5.4.1 provides detailed information about each of the components coded according to a socio-ecological framework. This chapter also describes the application of the following potential moderators and mediators in the present research: (1) conflict of interest; (2) programme specificity; (3) outcome data; and (4) possible risk of bias.

Chapter 6 presents the results of the systematic review (section 6.1) and meta-analysis (section 6.2) to examine the effectiveness of school-based anti-bullying programmes. The systematic review section outlines the breakdown of the presence and absence of several aspects of the included interventions, for example, the locations of the interventions and the types of measurement instruments used (sections 6.1.1 through 6.1.5). The numbers of studies that included, or did not include, the various intervention components are also given (section 6.1.3). The remainder of Chapter 6 presents the results of the meta-analysis and the moderator analyses to examine the potential reasons for observed differences between primary evaluations (section 6.2).

Part Three of this dissertation is concerned with the systematic and meta-analytical review of school-based intervention programmes to reduce and prevent cyberbullying behaviours. This part is structured almost identically to Part Two, in that, first a review of the cyberbullying literature is provided, followed by information regarding the systematic review methods, data extraction process and results. Given that there is currently less understanding of what works in cyberbullying intervention and prevention, the literature review (Chapter 7) review primarily focuses on the risk and protective factors associated with these relatively new forms of youth aggression. Cyberbullying has gained a lot of research attention in the decade or so since the phenomenon was first discussed, and there have been many international reviews of many aspects of these behaviours. Therefore, this literature review focuses on a narrative review of cyberbullying in the United Kingdom and Ireland, as this constitutes the most appropriate contribution to the wider research field.

Chapters 8 and 9 outline the systematic review methods and process of extracting raw data from primary evaluations of cyberbullying intervention and prevention programmes. These chapters mirror the corresponding chapters on school-bullying. The inclusion and exclusion criteria used to guide searches are presented (section 8.1), along with detailed information about the searches undertaken and the screening process (sections 8.2 and 8.3 respectively). Chapter 8 outlines both the studies excluded from the systematic review and meta-analysis and the included studies. Chapter 9 outlines the process of extracting data from these included studies, within the following sub-headings: (1) Descriptive; (2) Design; (3) Programme; and (4) Outcome.

Chapter 10 is the final component in Part Three of this dissertation and presents the results of the systematic and meta-analytical review of cyberbullying intervention and prevention programmes. The structure of this chapter is similar to that of the corresponding school-bullying chapter but given that fewer primary evaluations are included, less detail is presented. This is discussed in further detail in Chapter 10. Section 10.1 presents the systematic review of the included primary evaluations and section 10.2 outlines the results of the meta-analysis.

Whilst parts two and three of the dissertation present the methods and results of the systematic reviews for school-bullying and cyberbullying separately, part four is concerned with the overall discussion of results and post-hoc reflections. By including the interpretation of results in one chapter, rather than immediately following presentation of the results, it is hoped that similarities and contrasts between the results could be emphasised. Chapter 11 provides a detailed summary of the findings of both meta-analyses, along with an interpretation of the results (section 11.1). The implications of the results of the school-bullying meta-analysis (section 11.2), is provided including a detailed discussion of the results of moderator analyses (sections 11.2.1 through 11.2.4). Results of the cyberbullying meta-analysis are discussed in section 11.3, along with an overview of the subgroup analyses that were conducted, specifically in relation to the sample size, age of participants, evaluation methodology and the online/offline overlap in interventions (sections 11.3.1 through 11.3.3). The limitations of the research and the avenues for future research (section 11.4) are also presented in this chapter. Finally, this chapter concludes with post-hoc reflections and considerations (section 11.5). During this research project, several issues came to light, specifically the overlap between school- and cyber-bullying. Therefore, this chapter provides a detailed insight into these issues and additional avenues for future research.

1.3 Declaration of published works

Most of the work presented in this dissertation has been published in high-impact peer-reviewed academic journals. The data from both meta-analyses were published in a special issue of the review journal, Aggression and Violent Behaviour. The cyberbullying meta-analysis (i.e., Gaffney, Farrington, Espelage, & Ttofi, 2019a) and the school-bullying meta-analysis (i.e., Gaffney, Ttofi, & Farrington, 2019c) were both published in this special issue, as independent articles. Furthermore, the results of subgroup and moderator analyses in the school-bullying meta-analysis are presented in a paper, titled "*What works in anti-bullying programmes? Analysis of effective intervention components*", accepted for publication in the Journal of School Psychology (Gaffney, Ttofi, & Farrington, 2020). Comparisons of the effectiveness of intervention programmes in different countries and between repeatedly-evaluated programmes are reported in the first issue of the new highly specialized International Journal of Bullying Prevention (i.e., Gaffney, Farrington, & Ttofi, 2019b). Finally, in relation to cyberbullying, a chapter published on the prevalence of, and risk factors for, cyberbullying in the United Kingdom and Ireland also contributed to this dissertation (i.e., Gaffney & Farrington, 2018), specifically, the literature review of cyberbullying research. Finally, section 1.5 of the present chapter is largely based on a chapter submitted for publication in an upcoming Wiley-Blackwell Handbook on Bullying (Gaffney & Farrington, 2020, *in preparation*).

Each of these publications was co-authored by Dr. Ttofi and Professor Farrington, Institute of Criminology, Cambridge University, while Dr. Espelage, University of Florida contributed to the meta-analysis on cyberbullying intervention programmes. These individuals contributed invaluable edits to each publication. However, this research was undertaken with the doctoral candidate as the primary and lead researcher, under the normal supervision and guidance of appointed PhD supervisors. Where specific published papers contributed to specific chapters this is clearly specified in the text.

1.4 Research Questions

The overarching question driving this research is 'what works' in school-based bullying and cyberbullying intervention and prevention programmes. Asking 'what works' is a frequent undertaking in criminological research, and is one of the prevailing questions in evaluating intervention and prevention efforts. It is essential that we have an understanding of what is effective in prevention and intervention when implementing policies, and also in developing programmes to target specific problem behaviours.

In other words, good evaluation research involves evaluating individual programmes, but also evaluating the evaluations. Not only do we need to evaluate the effectiveness of specific programmes and initiatives that aim to reduce and/or prevent the wide array of problems that face our society today, but we also need to understand the bigger picture of *'what works'* where, and with whom. There have been numerous attempts to evaluate the effectiveness of intervention programmes, but individual evaluation studies rarely provide sufficient evidence for policy and practice recommendations (Tanner-Smith, Tipton, & Polanin, 2016).

Thus, in the context of bullying intervention and prevention, several research questions prompted the current research. The specific research questions were:

(1) Are school-based anti-bullying programmes effective?

(2) Are school-based cyberbullying intervention programmes effective?

(3) What are the specific moderator and mediator variables that are associated with the effectiveness of school-bullying intervention programmes?

(4) What are the specific moderator and mediator variables that are associated with the effectiveness of cyber-bullying intervention programmes?

One way in which researchers can construct a picture of the overall effectiveness of multiple different interventions in multiple different locations and samples is to conduct a meta-analytical review. Using this approach, meta-analysts can not only improve our understanding of how effective existing programmes are, but they can also compare and contrast primary evaluations on numerous different factors, for example, those relating to the evaluation methodology and the specific intervention activities.

In anti-bullying research, for example, there are a multitude of primary evaluations of many different programmes to reduce bullying behaviours, but the implications of the results are often limited. Primary evaluation research traditionally relies on statistical significance to indicate the effectiveness of a specific intervention, but this is impacted by methodological factors such as the sample size. For example, evaluations with small sample sizes may give a large effect that does not reach statistical significance, and evaluations with large sample sizes can often detect even a very small effect and that is statistically significant.

In evaluation research, it is becoming increasingly more important to measure effect size, particularly, the magnitude and direction of an effect size instead of relying on statistical significance (Cumming, 2014). For example, the same intervention programme may be evaluated multiple times in different samples, and researchers may find that the intervention is significantly effective in reducing an outcome in one sample, but not in another. Thus, if we rely solely on primary evaluations the overall picture can be muddied with conflicting results. How can we know if a programme is truly effective when one evaluation gives a significant effect, but another shows insignificant effects? This is where meta-analysis is very useful. Using meta-analysis, the magnitude and direction of effect sizes can be compared and contrasted.

The present research is comprised of two separate systematic and meta-analytical reviews to address the research questions. The school-bullying meta-analysis was originally intended to be an update of a previous meta-analysis (i.e., Farrington & Ttofi, 2009; Ttofi & Farrington, 2011) and the initial stages of this project were funded by a grant from the Jacobs Foundation awarded to Dr. Maria Ttofi in 2015. The purpose of this project was to update the previous meta-analysis and submit a review to the Campbell Collaboration. Whilst the review is currently under pre-publication revisions by the Campbell Collaboration, several additional analyses have been conducted for the current dissertation. For example, given the greater number of primary evaluations of anti-bullying programmes included in the present research, a greater level of detail could be extracted and used for comparative analyses. Thus, the ‘updated’ meta-analysis on school-bullying intervention programmes provides a great contribution to the existing literature. A summary of the previous research on the effectiveness of anti-bullying programmes is given in section 1.5 of the present chapter.

The cyberbullying systematic review and meta-analysis presents a novel contribution to the research literature. At the time of conducting searches for this review, there were no existing meta-analytical reviews of cyberbullying intervention programmes published in peer-reviewed journals, despite extensive research highlighting the need for such reviews. Therefore, given the great need and impetus for a meta-analytical review of cyberbullying interventions, the decision was made to conduct the cyberbullying and school-bullying reviews independently. Reflections on the applicability of this decision for future research are presented in Chapter 11.

1.5 Summary of previous research

There have been previous attempts to review the effectiveness of anti-bullying programmes in the past, the majority of which have focused on school-bullying, but little is known about the consistency of results in meta-analyses of the effectiveness of anti-bullying programmes. Therefore, the aim of the present section of this dissertation is to ‘review the reviews’. Existing systematic reviews and meta-analyses that included primary evaluations of the effects of school-based anti-bullying programmes on either school-bullying, cyberbullying, or both outcomes were identified. In total, 27 previous reviews were identified, 17 of which were systematic reviews only and 10 were systematic reviews followed by a meta-analysis of primary effect sizes.

Table 1 outlines the various aspects of the identified systematic reviews of anti-bullying programmes. The methods used by these previous reviews were fairly consistent in terms of the search terms used and databases searched. The keywords and inclusion/exclusion criteria were also comparable. The strength of these systematic reviews is that they can apply more flexibility and fluidity when evaluating primary evaluations in comparison to meta-analyses, but more structure and objectivity than narrative reviews. Therefore, systematic reviews can provide a detailed summary of the effects of programmes across a range of samples, locations, outcomes and different methodologies. Meta-analyses, however, have to be more precise and restrictive on which primary evaluations are included to ensure estimated effect sizes represent conceptually identical concepts, something that is explained in greater detail in Chapter 2.

Table 2 outlines the previous meta-analytical reviews of the quantitative effectiveness of included anti-bullying programmes. By using an online calculator to convert all effect sizes to odds ratios, and then applying a transformation described in Chapter 2 (see section 2.9) to convert the odds ratios to percentages, an estimate of the mean effectiveness is provided for each meta-analysis. This allows the reader to easily compare the findings from multiple meta-analyses. By far the greatest mean effect size, was identified by Verseveld and colleagues (2019). This meta-analysis found that interventions that targeted teacher intervention in bullying situations increased teacher interventions in bullying by approximately 45%. There was variation in meta-analyses in effectiveness of interventions to reduce school bullying perpetration and victimisation behaviours, with mean effect sizes ranging from 1% (Yeager et al., 2015, grades 8 to 13) to 23% (Ttofi & Farrington, 2011) for bullying others and from 8% (Jiménez-Barbero et al., 2016) to 24% (Merrell et al., 2008) for being bullied. The one existing cyberbullying meta-analyses suggested that included interventions were effective in reducing cyberbullying perpetration by approximately 6% and reducing cyberbullying victimisation by approximately 12% (i.e., Cleemput et al., 2014).

Admittedly, some reviews have yielded more pessimistic conclusions about the effectiveness of anti-bullying programmes (e.g., Ferguson et al., 2007; Merrell et al., 2008; Yeager et al., 2015). However, these reviews are often of poorer methodological quality than the present review, as the systematic point-by-point comparison of reviews by Ttofi, Eisner, and Bradshaw (2014) shows. For example, Merrell et al. (2008) only searched two databases and their effect size for bullying perpetration was based on only 8 studies. Yeager et al. (2015) only reviewed studies that compared the effects of programmes on different age groups, which greatly limited the number of studies included in the meta-analysis. For example, the mean effect size for bullying perpetration was only based on 16 studies. Finally, Ferguson et al. (2007) reported an effect size of r = .12 for bullying perpetration based on 23 studies and the authors describe this as a small effect. Yet, using a transformation described by Farrington and Loeber (1989), it is estimated that this mean effect size relates to an approximate 24% reduction in bullying perpetration, which does not in fact constitute a *small* change.

The majority of meta-analyses included in our review reported significant heterogeneity between primary evaluations and estimated mean effect sizes using the random effects model of meta-analysis. Many of the meta-analyses also computed moderator analysis to examine how different features and characteristics of primary evaluations may influence the overall effect size. For example, Jiménez-Barbero et al. (2016) found that higher effect sizes for bullying perpetration were associated with interventions that were implemented for less than 1 year and that were implemented with children under the age of 10 years old. Similarly, this meta-analysis found that larger effect sizes for bullying victimisation were associated with evaluations conducted after the year 2007.

Other meta-analyses examined specific intervention components (e.g., Lee, Kim, & Kim, 2015; Ttofi & Farrington, 2011). Lee et al. (2015) found that larger effect sizes were estimated for subgroups of studies that included: (1) curriculum-based intervention programmes; (2) programmes conducted with secondary school students; (3) training in emotional control; (4) peer counselling; (5) implementation of school anti-bullying policies; and (6) social skills training. Furthermore, Yeager et al. (2015) used advanced meta-analytical methods to better assess the relationship between age and the effectiveness of school-based anti-bullying programmes and found that programmes were most effective with participants aged 13 years old and younger.

Table 1

Previous reviews and meta-analyses of anti-bullying programmes

Author(s)	***Searches***	***Designs***	***Named ABPs***	***Included evaluations***
Cantone et al. (2015) (*p.* 59 - 60)	Online databases PubMed, Medline, EBSCO, were searched between 2000 and 2013 using keywords relating to bullying, cyberbullying, school, education, and mental health.	RCT designs only that were conducted with primary and secondary school students. Control group types included treatment with support and treatment without support.	Friendly schools; Steps to Respect; Peaceful Schools; CAPSLE; PATHS + Triple-P; OBPP; S.S. GRIN; Positive Action; CBT; SWPBIS; KiVa	17 evaluations of universal and focused school-based programmes to reduce either school-bullying or cyberbullying.
Chalamandaris & Piette (2015) (*p.* 133)	MEDLINE; PscyhINFO; ERIC were searched for articles published up to January 2008 using "bullying" as a keyword.	Experimental designs not specified. Participants were aged 5 to 18 years old. Types of control groups included waitlist controls and no treatment controls.	Bulli & Pupe; Project Ploughshares for Peace; Dare to Care; Bully Proofing Your School; SMART Talk; Respect; Sheffield ABP; CAPSLE; Steps to Respect; ZERO; Bully Busters; No Bullying Allowed Here; Gentle Warriors; PEACE Pack; Kidscape; Good Behaviour Game	62 articles reporting the effectiveness of school-based anti-bullying and health promotion interventions.
Cleemput et al. (2014) (*p.* 11 – 21)	Searches were conducted on databases such as ERIC, Medline, PscyhINFO, Web of Science, Social Services Abstracts, and Sociological Abstracts and Communication Abstracts for studies published from January 2003 to September 1st 2014. Keywords included terms such as, cyberbullying, electronic bullying, online bullying, cybervictim*, cyber-victim*, intervention, prevention, programme, internet safety, online safety, digital literacy, young*, teens, and young people. Academic journals	Case study and single group with pre/post-test measures evaluations (not included in meta-analysis) and quasi-experimental and RCT evaluations of participants aged 9 to 20 years old were included.	i-Safe; HAHASO; IHOP; BOC and CST programme; Quality circle approach; Cybersmart!; KiVa; ConRed; NoTrap!; Australian cyberbullying intervention; WebQuest cyberbullying; Surf-Fair; Cyberprogram 2.0; Arizona Attorney General's Prevention presentation.	19 publications of evaluations of 15 cyberbullying intervention programmes were included in the systematic review. 8 publications of 6 cyberbullying intervention programmes were included in the meta-analysis.

	were also searches and bullying research networks were contacted.			
Cox et al. (2016)	EBSCO Host was searched up to December 2013 using combinations of keywords relating to bullying, intervention, public health, adolescence/youth, and aggression/delinquency.	RCTs, quasi-experimental designs with pre- and post-test measures of bullying. Studies included a control group or comparison group, but review also included studies without control groups and were assessed on the Maryland Scale of Scientific Methods (Sherman et al., 1998).	Gatehouse Project; Whole-school anti-bullying interventions; Confident Programme	4 evaluations of anti-bullying interventions implemented in Australia.
Della Cioppa, O'Neil, & Craig (2015) (*p.* 62)	Online databases such as PsychINFO and Google Scholar were searched up to October 2014 using keywords such as "cyber bullying intervention" and/or "prevent school bullying"	Experimental and quasi-experimental studies with pre-post measurement of outcomes. Studies were assessed on a scale of scientific merit.	Survivors!; i-Safe; Missing Internet Safety programme; Cyber friendly schools project; ThinkUKnow Internet Safety programme; NoTrap!; ConRed; HAHASO; Beatbullying cybermentors; KiVa; Media Heroes	Formal evaluations of 12 cyberbullying programmes and 8 programmes that had not been formally evaluated
Earnshaw et al. (2018) (*p.* 183)	Searches conducted on databases such as PubMed, Google Scholar, EBSCO for the time period January 1st 2000 to December 31st 2015 using combinations of keywords such as bully, stigma, bias, intervention, LGBT etc.	Included studies used RCT, pretest-posttest, posttest only, and other evaluation methodologies and used the Cochrane Collaboration tool for measuring risk of bias. Participants were largely middle and secondary school students.	Names of specific interventions not provided. Categorized interventions according to the NASEM bullying report (2016) as being either: universal preventive, selective preventive, and indicated preventive.	22 articles that described 21 interventions addressing stigma-based bullying were included in the systematic review.
Evans et al. (2014) (*p.* 534 - 535)	Databases searched were: the Campbell Collaboration; Cochrane Library; ERIC; PscyhINFO; PubMed; Social Sciences Citation Index; Social Services Abstracts;	Methods included RCTs and quasi-experimental designs such as age-equivalent time-lagged contrasts and matched-group designs with post-test measures.	Bully Prevention Challenge Course; Bully Proofing Your Scool; Cool Kids Programme; Drama Programme; Empathy Training programme; FearNot!; Friendly schools; Friendly Schools, Friendly	32 articles that presented evaluation data for 24 distinct anti-bullying programmes

	Social Work Abstracts; Dissertation Abstracts; Google Scholar; Index to Thesis database; and Sociological Abstracts. Time period was June 2009 - April 2013 and keywords such as bully, victim, school, anti-bullying were used.	Participants were elementary and middle school students only.	Families; KiVa; Lunch Buddies; OBPP; Ophelia Project; Playworks; Positive Action; Restorative Whole School Approach; SWPBIS; Second Step; Social Norms Project; Steps to Respect; Take a Stand; Take the LEAD; WITs; Youth Matters; Zero programme	
Ferguson et al. (2007)	PsychINFO was searched for studies published between 1995 and 2006 using various combinations of keywords, including, school, intervention, prevention, violence, aggression, bully, externalizing.	RCT designs that randomized individuals, classrooms, or schools to control/contrast and treatment conditions. Participants were elementary to high school students.	Names of ABPs were not provided.	42 evaluations that reported 45 independent effect sizes were included in the meta-analysis.
Foody et al. (2017) (Table 1 *p.* 538–540)	PsychARTICLES, ERIC, PscyhINFO, and Education Research Complete databases were searched for articles published between January 1997 and April 2016 using combinations of bullying and cyber-bullying related keywords.	Cross-sectional and longitudinal designs were included, and the review was not specific to evaluation studies only. Participants were aged 4 to 18 years old.	Names of ABPs were not provided.	7 evaluations of intervention programmes are included in the systematic review of bullying and cyberbullying studies conducted in Ireland.
Goodman et al. (2013) (*p.* 5-6)	Searches of databases Academic Search Complete and ProQuest Education for studies published in 2005 – 2012 using keywords such as, antibully*, bully*, and intervention	Evaluations that measured bullying pre and post intervention with elementary school children.	Project Ploughshares Puppets for Peace (P4) programme; Step to Respect; Take a Stand, Lend a Hand; Stop Bullying now; WITS programme; OBPP; KiVa; Project ACHIEVE Social Skills Programme; Positive Behaviour Support programme.	10 studies that evaluated 7 intervention programmes were included in the systematic review.
Houchins et al. (2016)	PsycINFO was searched for studies published between 1980 and 2015 using combinations of keywords such	Experimental, quasi-experimental and single-case designs were included and coded for	Take a Stand, Lend a Hand, Stop Bullying Now; PS: Stories of Us; Bully Prevention	6 evaluations of anti-bullying programmes for students with disabilities were included.

(*p.* 261)	as bully*, disability*; prevention; intervention; special education. Hand searches of specialist journals were also conducted.	methodological rigor. Participants were from elementary and middle schools and with disability status.	in Positive Behaviour Support; STORIES; and Skillstreaming.	
Hutson et al. (2017) (*p.* 73-74)	Searches were conducted in MEDLINE; CINAHL; PubMed; Communication and Mass Media Complete; ERIC; and PscyhINFO for studies published up to October 2016. Keywords included cyberbullying, intervention, treatment, therapy, or programme. Hand searches of the Journal of Aggressive Behaviour were also conducted.	Included RCTs, quasi-experimental, and post-test only evaluation designs. Methodological quality was assessed.	Named interventions included: ConRed; KiVa; ViSC; Media Heroes; Cyberprogram 2.0	23 publications of 17 unique intervention programmes were included in the systematic review.
Jiménez-Barbero et al. (2016) (*p.* 166, 169)	MEDLINE, Trip Database, Cochrane Academy Search Premier, PsychINFO, ERIC and PsycARTICLES were searched between January 2000 and the end of May 2015 using keywords such as bullying, school violence and intervention or prevention programme.	Only RCT designs were included. Participants were aged 7 to 16 years old and methodological quality was assessed. Control groups included treatment as usual and waitlist groups.	STORIES; S.S. GRIN; Steps to Respect; Positive Action programme; Confident Kids; SPC and CAPSLE; Steps to Respect; KiVa; SWPBIS.	14 evaluations of school-based anti-bullying programmes that were conducted using experimental designs
Jiménez-Barbero et al. (2012) (*p.* 1647)	MEDLINE, Trip Database, Cochrane, Academy Search Premier, PsycINFO, ERIC and PsycARTICLES were searched from January 2000 using keywords such as bullying, school violence, attitudes towards violence, intervention or prevention programme and self-esteem or empathy.	RCTs, quasi-experimental, and cohort studies. Also included meta-analyses of effectiveness in the systematic review. Methodological quality was rated.	RIPP; STORIES; S.S. GRIN; Steps to Respect; Youth Matters; Positive Action; Confident Kids; SPC + CAPSLE; Adaptations of the OBPP; Bully Proofing your School; Befriending Intervention; CAPSLE; Ecological ABP; OBPP; FearNot!; Dare to Care.	54 studies were included that were a combination of meta-analysis and systematic reviews of anti-bullying programmes (n = 5) and primary evaluations of interventions.

Langford et al. (2015) (Table 3, *p.* 7)	Many databases were searched, such as: ASSIA; Australian Education Index; British Education Index; BiblioMap; Campbell Library; CENTRAL, CINAHL; EMBASE; ERIC; Global Health Database; International Bibliography of Social Sciences; Index to Theses in Great Britain and Ireland; MEDLINE; PsycINFO; Social Science Citation Index; Sociological abstracts. Searches were completed in 2011 and 2013. Keywords were not provided.	Cluster RCTs where the cluster was at the school, district, or other geographical level were included. Participants were in primary, middle, and secondary schools. Cochrane tool was used to assess risk of bias.	Friendly Schools; Friendly Schools, Friendly Families; Steps to Respect; KiVa.	6 evaluations of anti-bullying programmes were included in meta-analysis.
Lee et al. (2015) (*p.* 138-141)	Online databases MEDLINE, PsycINFO, PubMed, ERIC, and Cochrane database were searched for articles published between January 1990 and January 2010. Keywords included bully, antibullying, anti-bullying, programme, evaluat*, intervention and school.	Various designs were included, such as, longitudinal cohort designs and RCTs, non-randomized controlled trials and post-test only RCT.	Peer counselling; RWsA; Confident Kids programme; FearNot!; SPC + CAPSLE; OBPP; Steps to Respect; S.S. GRIN.	13 studies were included in the meta-analysis of the effectiveness of anti-bullying programmes in reducing bullying victimisation.
Merrell et al. (2008) (*p.* 28-30)	PsycINFO and ERIC were searched for articles published between 1980 and 2004 using keywords such as, bullying, intervention, schools, peer victimisation, and programme.	Experimental and quasi-experimental designs were included and participants were from kindergarten through to secondary schools.	Social skills programme; BEST programme; WITS; Peer support; OBPP; Bully-Proofing Your School; Expect-Respect.	16 studies included in final meta-analysis.
Nocentini et al. (2015) (*p.* 54)	Online databases PsycINFO, Scopus and PubMed were searched for evaluations published since 1996. Keywords related to the environment	Evaluation methodologies were not specified. Participants were preschool, school age, and	FearNot!; SMART Talk; Mii-School; Quest for the Golden Rule; NoTrap!; KiVa; Online Pestkoppenstoppen; Friendly ATTAC; PEACE Pack; ConRed;	32 publications relating to 13 intervention programmes were included in the systematic review.

	of the behaviour, e.g., computer game, digital game, virtual, online, the phenomenon, e.g., bullying, cyberbullying, anti-bullying, and the intervention, e.g., intervention, education, prevention, and social skill learning.	adolescence to young adulthood age.	WebQuest; The Layrinth; Empathic virtual buddy; MISAAC; User interaction paradigms.	
Polanin et al. (2012) (*p.* 50-51)	Searched five online databases, Dissertation Abstracts International, ERIC, PsycINFO, Medline, Science Direct for articles published between 1980 and 2010. Keywords included terms such as, bully or victim, bystander or participant or defender, school programme or programme, prevention or intervention, aggression. Higher education and cyber-bullying were excluded.	Studies used RCTs, non-random quasi-experimental designs, and non-random quasi-experimental matched-group designs. Participants were in (US) 3rd to 12th grades.	Curriculum-based ABP; Build Respect; CAPSLE; Steps to Respect; KiVa; Effective Bully Prevention; Befriending intervention; 5 W's approach to bullying; Expect Respect.	11 publications included in the meta-analysis of results from evaluations in 12 samples. Outcome of interest was interventions' effect on bystanders to intervene in bullying.
Rawlings et al. (2019) (*p.* 760 – 762)	PsycINFO, EMBASE, ERIC, Physical Education Index, MEDLINE, JAMA, Dissertation Abstracts, and SAGE databased were searched for articles published in English and up to 13th February 2013.	RCTs, quasi-experimental with pre- and post-test quantitative measures and single group designs evaluations conducted with North American elementary schools students.	Expect Respect, OBPP; Gentle Warrior; Positive Action; Steps to Respect; Youth Matters; WITS; Bully Busters; PEGs.	Evaluations of 10 different intervention programmes included in the systematic review.
Silva et al. (2016) (*p.* 2331)	Searches were conducted on databases such as Lilacs; PsycINFO; Web of Science; and SciELO using keywords such as bullying, school, intervention, antibullying, and	Experimental designs with no treatment control groups with participants aged 7 – 15 years old were included.	Categorised studies as either, multi-component interventions, whole-school programmes, social skills training programmes and programmes where bullying prevention activities were	18 evaluations of anti-bullying programmes were included in the systematic review.

	programme. Searches conducted in February 2015 for articles published in English, Portuguese or Spanish.		integrated into the normal school curriculum.	
Sivaraman, Nye, & Bowes (2019) (*p.* 156)	Conducted searches on databases like EMBASE, Social Sciences Citation Index; PsychINFO; ERIC; Global Health; CINAHL; PAIS Index; Education abstracts; Humanities index; ASSIA; OVID Medline; NCRJS; IBSS; PILOTS; PRISMA; Social Service Abstracts; Sociological Abstracts; British Education Index; Science Direct; Web of Science; Scopus; Cochrane; Campbell; PubMed; and Proquest. Searches were conducted between January 1987 – 30th June 2016 using keywords related to country classification, study method, school type, participant age, and bullying. 3 academic journals were also searched and bullying organizations were conducted.	Experimental and quasi-experimental designs that included a no-treatment or waitlist control group. Participants were residents of low- and middle-income countries aged between 10 and 19 years of age. Risk of bias was assessed.	REBE and ViSC; Behavioural programme for Bullying Boys; OBPP.	3 studies included in systematic review of anti-bullying programmes implemented in low- and middle-income countries.

Tanrikulu (2018) (*p.* 77, 79-83)	Online databases Academic Search Complete, Ebscohost, Google Scholar, National Thesis Databases of Council of Higher Education in Turkey, Science Direct, and Ulakbim were searched in June and July 2016 for studies published to August 2016. Keywords included terms such as cyber-bullying, cyber victimisation, prevention, intervention, cyber bully, cyber victim.	Various designs were included such as, pre-test-post-test quasi-experimental, longitudinal group-randomized, retrospective causal-comparative study, pre-test-post-test qualitative study, and RCTs. Participants were aged 11 to 19 years old.	Media Heroes; Cyber Friendly schools; Cyberprogramme 2.0; ViSC; NoTrap!; ConRed; KiVa.	17 evaluations were systematically reviewed.
Ttofi & Farrington (2011)	Using keywords such as bully, bullies, bully-victims, anti-bullying, school, intervention, prevention, programme, outcome, evaluation, and effect on databases such as the Australian Education Index, the British Education Index, Criminal Justice Abstracts, DARE, ERIC, EMBASE, Google Scholar, Index to Theses Database, MEDLINE, PsychINFO, Sociological Abstracts and Web of Science searches were conducted for articles published up to May 2009. Many journals were also searched.	Evaluations included were conducted using RCT, quasi-experimental with pre- and post-test measures, post-test only, and age cohort designs. Participants were from elementary, primary, middle and secondary schools.	ViSC; Bulli & Pupe; Project Ploughshares Puppets for Peace; Friendly Schools; S.S. GRIN; SPC + CAPSLE; Steps to Repect; Youth Matters; KiVa; Expect Respect; Be-Prox; OBPP; Progetto Pontassieve; Social Skills Training programme; Bully Proofing Your School; BEST; SAVE; Donegal ABP; Sheffield ABP	44 evaluations were included in the meta-analysis.

Verseveld et al. (2019) (*p.* 4)	Cochrane, ERIC, PsycINFO, PubMed, and Web of Science were searched up to September 2018. Keywords were based on four categories: bullying or peer victimisation, school or education, teacher or school professional, intervention or programme, and quasi-experimental design or randomized controlled trial.	RCTs and quasi-experimental designs were included and a methodological quality instrument was used.	KiVa; Bully Busters; OBPP; I DECIDE; Steps to Respect; ViSC; Sheffield Project; Expect Respect.	13 studies were included in the meta-analysis of anti-bullying programmes effect on teacher intervention in bullying situations.
Vreeman & Carroll (2007) (*p.* 79, 80-82)	MEDLINE (January 1966 to August 23rd 2004), PsycINFO, EMBASE, ERIC, Physical Education Index, Sociology, SAGE Full-text collection, and Cochrane Clinical Trials registry were searched for articles published up to August 23rd 2004. The keywords bullying or bully were used in searches.	RCTs with pre-test and post-test measures, randomized matched-pair designs; pre-test, post-test, control group design, pre-test, post-test, time-lagged comparison, and quasi-experimental with time-lagged age cohorts designs were used to evaluate interventions with participants aged between 7 and 16 years old.	Interventions were classified as curriculum interventions, multidisciplinary or whole-school interventions, social and behavioural skills group training programmes, and ‘other’ interventions.	26 studies were included in systematic review of anti-bullying programmes.

Yeager et al. (2015) (*p.* 43-45)	Online databases such as PsycINFO, ERIC, Proquest Dissertations and Theses, Google Scholar, Social Science Citation Index, EBSCO, ASSIA, PubMed, Sociological Abstracts, GALE, Academic Search Complete, MedLine, Campbell Collaboration, and Cochrane Collaboration were searched for articles published after 2009 and before September 2012. Keywords were derived from previous meta-analyses.	Experimental with pre-/post-test, single-group with pre-/post-test, and cohort-longitudinal designs were included. Participants were from kindergarten through to high school were included.	Names of ABPs were not provided.	23 reports with sufficient information to compute age-related trend were included in meta-analysis of anti-bullying programmes.

Note: RCT = randomized controlled trial; ABP = anti-bullying policy; CINAHL = Cumulative Index of Nursing and Allied Health Literature; ERIC = Education Information Resource Center; IBSS = International Bibliography of Social Sciences; NCJRS = National Criminal Justice Reference Service; LMIC = low- and middle-income countries; ASSIA = Applied Social Sciences index and abstracts

(a) A subset of named intervention programmes are included here, as this review included 65 different anti-bullying programmes.

Intervention acronyms: CAPSLE = Creating a Peaceful School Learning Environment; OBPP = Olweus Bullying Prevention Programme; PATHS = Promoting Alternative Thinking Strategies; S.S. GRIN = Social Skills Group Intervention; SWPBIS = School Wide Positive Behavioural Interventions and Supports; HAHASO = Help, Assert yourself, humour, avoid, self-talk, own it; WITS = Walk away, Ignore, Talk it out, Seek help; PEGs = Psychosocial Educational Groups for students; STORIES = The Structure/Themes/Open Communication/Reflection/Individuality/Experiential Learning/Social Problem-Solving programme; RIPP = Responding in Peaceful and Positive Ways; RWsA = Restorative Whole-school Approach

Table 2

Results of previous meta-analyses

Study	***~ N participants***	***Outcome***	***N effect sizes***	***Result***	***approximate % change***
Cleemput et al. (2014)[c]	6,373	Cyberbullying perpetration	6	g = .065, 95% CI .019, .112, p < .001	6% reduction
	9,453	Cyberbullying victimisation	6	g = .135, 95% CI .079, .190, p < .001	12% reduction
Ferguson et al. (2007)[c]	14,597	Nonviolent Bullying	23	RE r_+ = 0.12, 95% CI .08 - .17	22% reduction
Jiménez-Barbero et al. (2016)[d]	30,934	Bullying or school violence perpetration	14	d = -.12, 95% CI -.17, -.06	11% reduction
		Victimisation	8	d = -.09, 95% CI -.18, .01, n.s.	8% reduction
Langford et al. (2015)[b]	26,176	Bullying others	6	RE OR = 0.90, 95% CI 0.78, 1.04, n.s.	5% reduction
	26,256	Being bullied	6	RE OR = 0.83, 95% CI 0.72, 0.96, p < .05	9% reduction
Lee et al. (2015)[d]	19,619	Victimisation	13	RE d = -0.151, SE = .025, p < .001	14% reduction
Merrell et al. (2008)[c]	15,386	Bullying others	11	'average effect size' d = .04	4% reduction
		Being bullied	14	'average effect size' d = .27	24% reduction
		Intervened in bullying	10	'average effect size' d = .17	16% reduction
Polanin et al. (2012)[c]	12,874	Bullying bystander intervention	12	Hedge's g = .20, 95% CI	18% reduction
Ttofi & Farrington (2011)[a]	78,369	School-bullying perpetration	41	Mean OR = 1.37, 95% CI 1.27 – 1.48, p = .001	20-23% reduction
		School-bullying victimisation	41	Mean OR = 1.29, 95% CI 1.18 – 1.42, p = .001	17-20% reduction
Verseveld et al. (2019)[c]	948	Teacher intervention	13	g = 0.531, SE = 0.142, p = .013	45% change
Yeager et al. (2015)[c]	NA	Bullying others	72	*Grades 1 – 7* d = .13, z = 4.48, p < .001 *Grades 8 – 13* d = .01, z = .22, p = .83	*Grades 1 – 7* 12% reduction *Grades 8 – 13* 1% reduction

Note. RE = random effects model; MVA = Multiplicative Variance Adjustment model; r_+ = pooled correlation coefficient; CI = confidence interval; d = standardized mean difference/Cohen's d; g = Hedge's g, which is a correction for sampling variance for the standardized mean difference (d) effect size;

a. Odds ratios in these meta-analyses that are *greater* than 1 represent a *desirable* effect of the interventions, i.e., a reduction in bullying perpetration or victimisation. Similarly, odds ratios that are less than 1 represent an *undesirable* effect of the intervention, and odds ratios equal to 1 represent a null intervention effect.
b. Odds ratios in these meta-analyses that are *greater* than 1 represent a *desirable* effect of the intervention, whilst odds ratios greater than 1 represent an *undesirable* effect of the intervention.
c. In these studies positive mean effect sizes represent a desirable intervention effect.
d. In these studies negative mean effect sizes represent a desirable intervention effect.

1.5.1 Implications for the present research

The present research, specifically the school-bullying meta-analysis, is considered an update of the meta-analysis conducted by Ttofi and Farrington (2011), as has been discussed. As outlined in Table 2 this previous review found that included interventions were effective in reducing school-bullying perpetration by approximately 20 – 23% and school-bullying victimisation by approximately 17 – 20%. This previous review has also been noted as being of higher scientific methodological quality than similar meta-analyses (Ttofi et al., 2014). Therefore, this previous meta-analysis is used as a model for the present research.

However, following publication of Ttofi and Farrington's (2011) previous meta-analysis, some key bullying researchers were critical of a few of the policy recommendations suggested by the review. Namely, Smith, Salmivalli, and Cowie (2012) were sceptical of the recommendations made to implement interventions with older children (given the finding that intervention programmes were more effective with children aged 11 years or older) and that the intervention component 'work with peers' should not be used (given the finding that this component was associated with an *increase* in bullying victimisation).

Smith et al. (2012) presented their criticisms in four categories: (1) analytical procedure; (2) definitional issues; (3) historical issues; and (4) recent empirical data. In relation to the latter two categories, it is proposed that the present updated meta-analysis clearly addresses these criticisms.

Briefly, Smith et al. (2012) were concerned that Ttofi and Farrington's (2011) meta-analysis included a significant number of 'out-dated' or 'old-fashioned' intervention components (e.g., they included the impact of videos but could not evaluate the use of virtual reality games) and that therefore their conclusion did not reflect current educational or bullying prevention practice. Secondly, Smith et al. (2012) point to research which at the time of publication was unavailable to Ttofi and Farrington (2011) when conducting their analysis.

This observation is also a reflection of anti-bullying research in general. Bullying behaviours in schools continue to develop and change rapidly and researchers struggle to keep up. However, as the present research aims to not only include evaluations included by Ttofi and Farrington (2011), additional searches were conducted and so it was expected that, with the inclusion of more primary evaluations, this issue would be adequately addressed. Furthermore, planned analyses included a comparison of effect sizes in relation to the year of publication.

For example, prior to the current research project, there was only been one attempt to evaluate the effectiveness of cyberbullying intervention and prevention programmes (i.e., van Cleemput et al., 2014). Even though research on bullying online and through information communication technologies (ICTs) first emerged in the early 2000s, the phenomenon was definitively defined in 2008 (Smith et al., 2008). Moreover, as illustrated in later chapters (see Chapter 8, section 8.5) the first evaluation[1] of an intervention programme for cyberbullying behaviours was published in 2012. As Smith et al. (2008) suggest, researchers are struggling to keep pace with behavioural changes.

Smith et al. (2012) were also concerned that the Ttofi and Farrington (2011) method of dichotomizing intervention component variables meant that their analysis was correlational, and thus, subject to the flaws of correlational statistics. They suggest that instead of between-programme comparisons, comparisons of 'within-programme' factors would be a better approach. This specifically related to the finding that anti-bullying programmes were more effective when implemented with children aged 11 years or older compared to programmes implemented with children aged 10 years or younger. Smith et al. (2012) suggested that comparing effectiveness between age groups should only be done

[1] Of evaluations that were eligible for inclusion in the cyberbullying meta-analysis.

within programme, or in other words, when the same intervention programme is implemented with children of different age groups.

This is still difficult to achieve. As the present research demonstrates, anti-bullying interventions are implemented across the globe (see section 6.1.1.1). There is a large number of different intervention programmes that aim to reduce school bullying perpetration and/or victimisation, but few intervention programmes are evaluated repeatedly (see section 6.1.4.1). Therefore, comparisons of the same programme in different samples, or age groups, are possible for very few intervention programmes, but some primary studies do already publish results independently for separate age groups (see Chapter 6 for examples). Since one of the main purposes of meta-analysis is to synthesize results from *multiple* studies in order to reduce bias and reliance on statistical significance, 'within-programme' comparisons of the effectiveness of intervention components would be unhelpful in the broader picture.

Finally, Smith et al. (2012) criticise some of the defining criteria used by Ttofi and Farrington (2011) to categorize particular elements of intervention programmes; primarily, their definitions of punitive disciplinary measures and the 'work with peers' component. Smith et al.'s (2012) concerns were related to the variety of different activities that could be included under the label 'work with peers'. This is a justified observation, and in order to better understand the relationship between specific intervention activities and effect sizes, additional information on a number of intervention components was recorded. For example, additional codes to better specify the inclusion of peers in intervention activities were created. Namely, on the peer-level, components referred to the following: (1) informal engagement of peers (e.g., through group/class discussions); (2) engaging or encouraging bystanders to intervene in bullying situations; and (3) formal engagement of peers (i.e., the intervention programme was 'peer-led' or included peer mentoring techniques). Therefore, planned subgroup analyses in the present thesis are able to inform a greater understanding of the relationship between peer-involvement and the effectiveness of interventions.

1.6 Defining terms

As clearly illustrated, the present research is concerned with two forms of bullying behaviours: school-bullying and cyber-bullying. Whilst specific definitions are provided in subsequent chapters (school-bullying in Chapter 3, section 3.2; cyberbullying in Chapter 7, section 7.2) it is noteworthy to comment on the interchangeability of terms in this dissertation.

Where bullying behaviours are concerned, cyberbullying is used to refer to instances of bullying that occur online or via different ICT and social media platforms (e.g., mobile phones, email, Facebook, Instagram or Snapchat). ‘Online bullying’ is also used in this dissertation to refer to these instances of bullying. However, when referring to bullying that occurs in schools, several different terms are used interchangeably. For example, offline bullying, school-bullying, traditional bullying and bullying are used to describe instances of bullying that happen within the school environment. Additionally, in some instances ‘being bullied’ is used to refer to school-bullying victimisation and ‘bullying others’ is used to refer to school-bullying perpetration. Similarly, ‘being bullied online’ relates to cyberbullying victimisation and ‘bullying others online’ relates to cyberbullying perpetration. This use of terminology is consistent with existing bullying research.

2. Methods of systematic review and meta-analysis

2.1 Overview

The present research used two main methods to address the research questions, primarily, systematic review and meta-analysis. The following sections will provide a detailed account of these methods and later sections (e.g., see Chapters 4 and Chapter 8) will describe how these methods were applied to explore the effectiveness of school-based programmes in reducing online and offline bullying.

2.2 Systematic review

The first step in conducting any meta-analysis is to employ systematic searches of the literature in order to identify all existing includable studies (Littell, Corcoran, & Pillai, 2008). There exists many guidelines and instructions for completing a systematic review. For example, the Cochrane Collaboration provides a detailed handbook for conducting systematic reviews of interventions (Higgins & Green, 2011; version 5.1 available online at https://training.cochrane.org/handbook). Moreover, there have been several books published on the subject of systematic reviews (e.g., Gough, Oliver, & Thomas, 2017; Littell et al., 2008; Petticrew & Roberts, 2006). Furthermore, systematic reviews often cite the PRISMA guidelines for conducting searches (Preferred Reporting Items for Systematic Reviews and Meta-Analyses; Moher et al., 2009).

Overall, a systematic review involves using predetermined keywords and strict inclusion/exclusion criteria to identify, screen, appraise, and synthesize all relevant empirical studies (Zych et al., 2017). In this way, bias is minimized. Additionally, if sufficient evidence is obtained to conduct a meta-analysis (see section 2.4 for the outline of the information required) a systematic review ensures that included studies are comparable in terms of methodologies used and outcomes reported. This is essential to ensure that computed effect sizes represent the same underlying outcomes.

2.3 Meta-analysis

A meta-analysis synthesizes the observed effect sizes from the primary studies obtained in the systematic review and computes one summary mean effect size per outcome of interest, plus its variance. The strength of a meta-analysis lies in its provision of an objective synthesis of primary research. It moves away from reliance on statistical significance testing, which has been widely criticized for being influenced by many different biases (e.g., the "file drawer problem" or *p*-hacking).

Primary evaluation research has traditionally relied on statistical significance to indicate the effectiveness of an intervention programmes. However, statistical significance depends on various exogenous factors such as the sample size. Nowadays, it is considered more important to measure effect size, particularly, the *magnitude* and *direction* of an effect size, instead of focusing solely on statistical significance (Cumming, 2014). For example, the same intervention programmes may be evaluated multiple times with different samples, in different contexts and circumstances, finding that the intervention is significantly effective in reducing an outcome in one sample, but not in another. Thus, if we rely solely on primary evaluations the overall picture can be muddied with conflicting results. How can we know if a programme is truly effective when one evaluation gives a significant effect, but another shows non-significant effects? This is where meta-analysis is very useful.

Meta-analyses can be very useful for both future research and evidence-based policy. However, this approach does have limitations so researchers should take a conservative approach to estimating summary effect sizes. Heterogeneity of effect sizes is one reason the meta-analyst must be cautious when computing a summary mean effect. It is rare, particularly in social and behavioural sciences, that we can be certain that a primary study is estimating the *true* effect, without other sources of error influencing the estimate. In addition to the treatment of between-study heterogeneity, the quality of included studies can impact the weighted mean effect size produced in a meta-analysis. Meta-analysts often use measures of methodological quality to compare studies based on methodological factors and can use subgroup analyses to establish how the mean effect size may be influenced by these factors.

2.4 Effect sizes

A meta-analysis aims to estimate comparable effect sizes from multiple primary studies. The choice of effect size depends on how statistical information is reported by primary studies (Borenstein et al., 2009). In meta-analyses of intervention research, the data from primary studies are largely presented in continuous (e.g., means, standard deviations, sample sizes) or dichotomous (e.g., prevalence or percentages) forms (Wilson, 2010). Thus, the effect sizes are estimated as Cohen's *d* and Odds Ratios.

2.4.1. Dichotomous data

For primary studies that presented results as percentages of participants identifying as either bullies or victims, the odds ratio (OR) effect size was estimated. The ORs for before and after intervention time-points were calculated independently. The Comprehensive Meta-Analyses (CMA)™ software that we used to analyse effect sizes in the present report does not allow raw data for before and after time-points for primary studies that reported dichotomous outcomes to be entered separately. Thus, we were unable to use this software to calculate pre-post intervention estimates for these studies. Hence, these calculations were carried out manually, as described in section 2.8.2.

2.4.2. Continuous data

Cohen's *d* was estimated for primary studies when results were reported in the form of continuous data. Cohen's *d* is estimated as the difference between experimental and control means divided by the pooled standard deviation (Wilson, 2010), and is a common effect size used to quantify the difference between group means. The CMA software requires the meta-analyst to assign a direction for Cohen's *d* effect sizes. In the present research, effects were assigned a positive direction in cases where: (1) bullying outcomes were less in the experimental group compared to the control group; or (2) the reduction in bullying outcomes was larger in the experimental group in comparison to the change in the control group. Following this logic, a negative effect was found when there was: (1) a larger reduction in the control group compared to the experimental group; or (2) there was no change or an increase in bullying perpetration/victimisation in the experimental group but a reduction (or smaller increase in the control group). The CMA software was used to estimate the pre-post intervention effect size for these studies.

2.4.3 Mean effect sizes

In the present research, one effect size for each *independent sample* included in primary studies was estimated. Therefore, where studies reported results separately for male and female participants, or primary and secondary school students, one effect size was calculated for each group.

For comparability, all primary effect sizes were converted to odds ratios. Summary mean effects for bullying perpetration, bullying victimisation, and for each of the moderator subgroups are thus reported as odds ratios. In the present review, odds ratios greater than one represent a desirable, intervention effect; namely, a reduction of bullying in the experimental group, that is comparably larger than the change in bullying in the control group. Therefore, the change is assumed to have occurred because of the intervention programme. Similarly, odds ratios less than one represent an undesirable, intervention effect and odds ratios that equal one represents a null effect.

2.5 Heterogeneity and homogeneity

In a meta-analysis, homogeneity is the assumption that observed effect sizes (i.e., those computed from primary studies) are distributed around the summary mean effect size (i.e., that are estimated by the meta-analysis) in a manner which is no greater than what would be expected due to the (random) sampling error of primary studies (Lipsey & Wilson, 2001). Based on the chi-square distributed *Q* statistic with $k - 1$ degrees of freedom, where *k* equals the number of observed effect sizes (Hedges & Olkin, 1985), heterogeneity is calculated using estimates produced by a meta-analysis. The formula for calculating *Q*[2], the test for heterogeneity in a meta-analysis of *i* studies, is as follows:

$$Q = \sum_{i=1}^{k} W_i Y_i^2 - \frac{\left(\sum_{i=1}^{k} W_i Y_i\right)^2}{\sum_{i=1}^{k} W_i}$$

Where W_i represents the weight assigned to each study and Y_i is the observed effect size (Borenstein et al., 2009). If *Q* is statistically significant, the meta-analyst can conclude that the included effect sizes are not homogeneous (Hedges, 1982b; Rosenthal & Rubin, 1982).

In other words, the included primary studies in a meta-analysis are *heterogeneous.* This is highly probable in social and behavioural sciences research even when strict inclusion criteria are used in meta-analyses. For example, in the present research, strict inclusion criteria are used to ensure that only similar anti-bullying interventions and behavioural measures of bullying/cyber-bullying outcomes are included. Yet, some potential sources of heterogeneity in meta-analyses could include: (1) different methodologies of evaluation; (2) different implementation fidelity of intervention programmes; or (3) targeting different mechanisms of change. These limitations of the present research will be further explained in later sections (see Chapter 10). However, it is worth mentioning that, because social and behavioural sciences research aims to explain human behaviours, there are undeniably sources of error beyond our control.

The solution for dealing with the between-study variance is a key difference between meta-analytical models. The following sections will present three computational models of

[2] All mathematical formulae and terminology used are in line with those used by Borenstein et al. (2009).

meta-analysis, the fixed effect (FE) model, the random effect (RE) model, and the Multiplicative Variance Adjustment (MVA) model. The strengths and limitations of each will be discussed and the proposition that the MVA model is the most appropriate model for the present research will be argued.

2.6 Computational models

There are currently two widespread models of meta-analysis applied in social and behavioural sciences, i.e., the fixed effects and the random effects models. The present research utilises these common models of meta-analysis but also includes the MVA model (Farrington & Welsh, 2013). All three computational models are utilised in this dissertation and the rationale behind each model is explained in greater detail in this section. The main difference between these models is how between-study variance (i.e. heterogeneity) is handled, and the implications for the summary mean effect size and its variance.

To reiterate, the purpose of a meta-analysis is to synthesize observed effect sizes from primary studies and compute one summary mean effect size per outcome of interest. The primary argument between deciding how best to compute this effect size is largely based around heterogeneity. Lipsey and Wilson (2001) state that all observed effect sizes should not be treated equally, as factors such as sample size will greatly influence the precision of effect sizes. Therefore, in meta-analysis weights should be assigned to observed effect sizes to reflect these assumed differences (e.g., to give greater weight to larger studies). The method for assigning weights to observed effect sizes is the main difference between computational models of meta-analysis.

2.6.1 Fixed effects model (FE)

The FE model of meta-analysis assumes that each primary study measures an underlying *true effect* and any observed variance occurs as a result of sampling error alone. Borenstein et al. (2009) note that another term that appropriately describes the fixed-effects model is the ‘common-effect’ model, as ‘true effect’ is also used to denote the population effect. Under a FE model of meta-analysis, the study weight ($FE\ W_i$) is estimated as the inverse of the study variance (V_{Yi}).

$$FE\ W_i = \frac{1}{V_{Yi}}$$

The summary mean effect size under the FE model ($FE\ M$) is estimated as the sum of weighted effect sizes divided by the sum of weights. The variance of the summary mean effect ($FE\ V_M$) is calculated as the inverse of the sum of weights, as per the following formulae;

$$FE\ M = \frac{\sum_{i=1}^{k} W_i Y_i}{\sum_{i=1}^{k} W_i}$$

$$FE\ V_M = \frac{1}{\sum_{i=1}^{k} W_i}$$

A strength of this approach is that larger studies are appropriately assigned greater weights. Lipsey and Wilson (2001) state that a meta-analyst needs to assign weights to primary studies before estimating a summary mean effect, because not all observed effects are equal. There is a large body of literature (e.g., Cumming, 2014) in primary empirical research that strongly emphasizes the importance of power calculations and ‘large enough’ sample sizes in order to increase the external validity of a result.

It is not that studies with small samples are less valuable; they too contribute to the wider literature. However, it is widely accepted that a result obtained from a study with a larger sample is more likely to represent the true population effect. Statistical significance in empirical research is also greatly influenced by sample size, and often very small effects may be statistically significant if enough participants are tested. This increases the probability of a type I error, but in meta-analysis we avoid this problem. As meta-analysts focus on the magnitude and direction of an observed effect and not on statistical significance, it can be argued that observed effects from studies with large samples should be given more weight.

However, the FE model is largely criticized because of the assumption that the observed effects represent a shared underlying ‘true effect’ that remains consistent between evaluations. In intervention research with human participants conducted in real-world settings, such as in school-based anti-bullying programmes, this is highly unlikely to be the case. Therefore, the underlying assumption of the FE model is often incorrect and limited in the way in which it estimates the variance of the summary mean effect size.

2.6.2 Random effects model (RE)

To account for this limitation of the FE model, meta-analysts suggest that the RE model is a more appropriate way of assigning weights to observed effect sizes. This computational model accounts for sources of variance beyond sampling error. Under a RE model, we assume that the true effects of each study are normally distributed around the summary ‘true effect’ and assign weights to primary studies to account for this between-study variance. The RE model estimates weights based on the sum of variance (i.e., V_{Yi}) and between-study variance (τ^2 or tau-squared). The between-study variance is estimated using a method of moments, or the DerSimonian and Laird method (Borenstein et al., 2009).

$$\tau^2 = \frac{Q - df}{C}$$

Where C is estimated as:

$$C = \sum W_i - \frac{\sum {W_i}^2}{\sum W_i}$$

Borenstein et al. (2009) describe the use of tau-squared to estimate what the distribution of observed effect sizes around the true mean effect would be if we knew the value of the true effect. Thus, weights assigned to primary studies under a random effects model of meta-analysis are estimated as:

$$RE\ W_i = \frac{1}{V_{Yi} + \tau^2}$$

The summary mean effect therefore under a random effects model is calculated as follows:

$$RE\ M = \frac{\sum_{i=1}^{k}(RE\ W_i)(Y_i)}{\sum_{i=1}^{k} RE\ W_i}$$

Subsequently the variance of the summary mean effect is:

$$RE\ V_M = \frac{1}{\sum_{i=1}^{k} RE\ W_i}$$

However, as Borenstein et al. (2009) note, the RE model is just one way in which a meta-analyst can adjust for the heterogeneity likely to occur between primary studies. Researchers have noted various problems with this computational model. Most relevant to the present research is that, if the overall heterogeneity in a meta-analysis is high, all observed effect sizes will be assigned very similar weights using this method for estimating between-study variance. Furthermore, the inclusion of multi-site studies will greatly impact the overall result. This issue will be further outlined in later sections (see Chapter 10) using data from the meta-analysis of school-based anti-bullying programmes.

2.6.3 Multiplicative variance adjustment (MVA) model

Therefore, both the FE and RE models have limitations. Therefore, the present research also computed weighted mean effect sizes using the MVA model. This approach combines both the strengths of the FE model (i.e., larger studies = larger weights) and the RE model (i.e., adjusting for highly probable between-study variance). Weights in the MVA model are estimated as:

$$MVA\ W_i = \frac{1}{V_{Yi}}$$

The summary mean effect size is calculated as:

$$MVA\ M = \frac{\sum_{i=1}^{k} W_i Y_i}{\sum_{i=1}^{k} W_i}$$

These formulae are identical to those used for a FE model but under the MVA model the variance of the weighted summary effect size is estimated as:

$$MVA\ V_M = \left(\frac{1}{\sum_{i=1}^{k} W_i}\right)\left(\frac{Q}{df}\right)$$

Hence, as this formula suggests, under the MVA model the extra variance we assume to be present between observed effect sizes is estimated as the overall heterogeneity (*Q*) divided by the degrees of freedom (i.e., *df*; *n* of observed effect sizes – 1). The MVA model therefore accounts for heterogeneity and adjusts the variance of the weighted summary effect size by multiplying (rather than adding as per the RE model) the two values. Moreover, the *Q/df* adjustment means that the meta-analytical model fits the data more appropriately. The MVA model assigns weights to primary studies in direct proportion to the study level sampling error, as with the fixed effects model, but adjusts the standard error and confidence intervals of the mean summary effect size for between-study heterogeneity. Data from the school-bullying meta-analysis will now be used to highlight the need for this alternative approach and provide support for the MVA model as a computational model in meta-analysis.. A worked example of how the MVA model is computed is outlined in Appendix 1.

2.7 Support for the MVA model

As outlined in multiple chapters of the present dissertation, the computational model chosen to assign weights to primary studies in a meta-analysis can significantly impact the overall results. This reflection will first review these models and suggest that alternative approaches are needed.

2.7.1 Reviewing meta-analytical models

There are currently two main models of meta-analysis utilised, namely, the fixed effects model and the random effects model. The current thesis highlights the limits of both of these approaches and proposes an alternative meta-analytical computational model; the multiplicative variance adjustment (MVA; Farrington & Welsh, 2013). To demonstrate the arguments put forward in this chapter, data are drawn from the meta-analysis of 100 evaluations of school-bullying intervention programmes.

The first approach to estimating the summary mean effect in a meta-analysis is to apply the fixed-effects model but this model is arguably becoming less and less popular. This computational model assumes that each primary study is measuring an underlying *true effect* and any observed variance occurs as a result of sampling error alone. Borenstein et al. (2009) note that another term that appropriately describes the fixed-effects model is the 'common-effect' model, as 'true effect' is also used to denote the population effect. This computational model is largely criticized because between-study variance is not accounted for, and as previously discussed, studies are likely to be heterogenous in social and behavioural sciences.

To account for this limitation of the fixed effects model, prominent researchers in the field of meta-analysis suggest that the random effects model is a more appropriate way of assigning weights to observed effect sizes. This computational model accounts for sources of variance beyond sampling error, i.e., between-study variance. Under a random effects model, it is assumed that the true effects of each study are normally distributed around the summary 'true effect', and weights are assigned to primary studies to account for this between-study variance.

However, the random effects model is just one way in which a meta-analyst can adjust for the heterogeneity likely to occur between primary studies, and often this approach is *not* the most appropriate and may actually lead to less conservative summary effect sizes. There are several limitations of the random effects model. Firstly, adding additional variance in order to reduce heterogeneity is not the most ideal method of calculating a weighted mean effect size. Moreover, if heterogeneity between primary studies is high, all observed effect sizes are assigned very similar weights using this method for estimating between-study variance. There is also a significant difference in how multi-site studies are analysed (see Farrington & Welsh, 2013). This observation is consistent with previous educational studies, that have argued in experimental research the random effects model and the fixed effects model should be seen as being complementary with one another (Clarke et al., 2015). Neither approach is sufficient in meta-analyses and therefore the current thesis proposes that future research should investigate alternative computational models.

2.7.2 An alternative approach

Following meta-analyses of the overall effectiveness of school-based anti-bullying programmes and prior to conducting planned subgroup analysis, a more in-depth look at meta-analytical computational models was undertaken. As previously discussed, there are three specific models of meta-analysis referred to in this dissertation: (1) fixed effects model; (2) random effects model; and (3) the MVA model. The fixed effects and random effects models are the most common approaches, and the main issues with these computational models can be summarized as follows: the fixed effects model does not account for between-study variance and the random effects model fails to give larger studies adequate weight.

The MVA model therefore is proposed to address these issues, as previously discussed in greater detail. It is argued widely in the social and behavioural sciences literature that the fixed effects model is not appropriate to estimate weighted mean effect sizes as heterogeneity between effect sizes is not accounted for by the model. When measuring human behaviour, it is rarely applicable that 'one true effect' is measured by all studies, and there is a wide range of confounding sources of error. However, as previously argued, when a large number of evaluations are included in a meta-analysis, the procedure for accounting for heterogeneity using the random effects model fails to give appropriate weight to studies with larger samples. Therefore, the application of the MVA model appears to be the most appropriate. The MVA allows larger studies to contribute more weight to the mean summary effect size and also account appropriately for between-study variance. Using the data for the effectiveness of school-bullying anti-bullying programmes a further exploration of these issues was undertaken.

2.7.2.1 Distribution of weight. Table 3 shows the studies with the largest weights under the fixed effects model for school-bullying perpetration and victimisation outcomes. Under the fixed effects model, one can clearly see that Roland et al. (2010) contributed the largest amount of weight (w) to the overall effect size for bullying perpetration ($w = 31.85$). The next largest weight was 14.91 (Limber et al., 2018). These two studies were allocated the largest weights under a fixed effects model for bullying victimisation outcomes too, although the order was inverted and the two studies contributed similarly to the model (Limber et al., 2018 $w = 25.63$, and Roland et al., 2010 $w = 25.48$).

This means that under both the fixed effects and MVA models of meta-analysis two studies contributed approximately 47% of the weight to the mean summary effect size for school-bullying perpetration outcomes and approximately 51% of the weight for school-bullying victimisation outcomes. Moreover, an additional 28% of the relative weight for perpetration was contributed by only six studies, and 24% of relative weight for victimisation from five studies. Therefore, the weighted mean effect sizes for school-bullying perpetration (OR = 1.324) and school-bullying victimisation (OR = 1.248) largely reflect the effectiveness of programmes evaluated by only a handful of studies. In fact, these weighted effect sizes reflect roughly 8 – 9% of the included evaluations of anti-bullying programmes.

Indeed, when the weighted mean effect sizes of only these studies were computed for both school-bullying outcomes, under the fixed effects model, the results did not vary much from the overall model. For school-bullying perpetration outcomes the weighted mean effect sizes for the eight studies contributing the most amount of weight was OR = 1.359, and for school-bullying victimisation outcomes the weighted mean effect sizes for the seven studies contributing the majority of weight was OR = 1.262. The mean effect sizes were the same under the MVA model, but the confidence intervals were wider.

Table 3

Studies contributing the greatest relative weight under a fixed-effects model

	LOR	*SE*	*FE Relative Weight*	*LOR*	*SE*	*FE Relative Weight*
Study	*School-bullying perpetration*			*School-bullying victimisation*		
Roland et al. (2010)	1.417	0.018	31.85	0.304	0.018	25.48
Limber et al. (2018)	1.502	0.026	14.91	0.173	0.018	25.63
Kärnä et al. (2011a); AC	1.181	0.039	6.75	0.191	0.032	8.32
Kärnä et al. (2013), 8 – 9	1.075	0.044	5.43	-0.065	0.044	4.39
Waasdorp et al. (2012)	1.281	0.045	5.02	NA	NA	NA
Kärnä et al. (2011b), 4 – 6	1.101	0.049	4.29	0.241	0.049	3.46
Olweus New National	1.744	0.052	3.82	0.427	0.032	8.25
Kärnä et al. (2013), 2 – 3	1.165	0.004	2.29	-0.138	0.056	2.62

Table 4

Analysis of weight distribution under fixed effects and random effects models in school-bullying meta-analysis

	School-bullying perpetration		*School-bullying victimisation*	
Statistic	Fixed Effects	Random Effects	Fixed Effects	Random Effects
Mean weight	1.111	1.111	1.075	1.075
Median weight	0.165	0.965	0.15	0.95
Weight range	0.01 – 31.85	0.09 - 2.35	0.01 - 25.63	0.07 - 2.55
25th percentile	0.08	0.6	0.05	0.45
75th percentile	0.52	1.6	0.41	1.52
Sum	100	100	100	100

Since the school-bullying meta-analysis includes 100 evaluations and 103 independent effect sizes (90 for school-bullying perpetration and 93 for school-bullying victimisation) it does not seem appropriate that, so few studies contribute so significantly to estimations of the effectiveness of anti-bullying programmes. Moreover, the earlier discussion regarding the limits of the random effects model method in assigning weight appears to apply in the context of the present research. Under the random effects model of meta-analysis, the range of relative weights assigned to primary evaluations was quite limited and the interquartile range for bullying victimisation was only 1.00 and for bullying perpetration it was 1.07 (see Table 4). Such a small range of values demonstrates that very little difference exists in the relative weighting of primary evaluations. This undermines the purpose of meta-analysis. Lipsey and Wilson (2001) state that the purpose of a meta-analysis is to combine primary study effect sizes, but in such a manner that treats effect sizes with respect to factors that can influence the precision. Additionally, the median weight assigned under the random effects model was 0.95/0.965, suggesting that relative weights were distributed approximately around 1 for victimisation and perpetration outcomes. This means that there was little variation in assigned weights, again demonstrating that the random effects model does not adequately distinguish between observed effect sizes, by assigning studies with larger studies larger weights. Consequentially, it appears that in the present research large studies potentially contribute *too much* weight to a fixed effect, or MVA, mean summary effect size, and *too little* to a random effect mean summary effect size.

2.7.2.2 Further investigation. At first look this issue does not appear to impact the overall summary effect size, at least in a mathematical sense. The weighted mean effect sizes for school-bullying perpetration and victimisation did not vary greatly to the overall mean effect sizes when applying the random effects model. The overall mean effect sizes under a random effects model for perpetration and victimisation were OR = 1.308 and OR = 1.242, respectively. Therefore, the seemingly unequal distribution of weight under the fixed effects model and the comparatively more equal distribution of weight under the random effects model does not impact the summary effect size.

However, upon further examination several issues were identified. Arguably these may not impact the overall estimation of the effectiveness of anti-bullying programmes, but it will be argued that they could impact the analysis of effectiveness in relation to specific intervention components. Table 5 highlights some of the factors that may impact the subgroup analysis, i.e., the country, sample size, methodological design of the evaluation study, and the anti-bullying programme evaluated.

Table 5

Characteristics of studies given the greatest relative weight

Study	*Country*	*Sample size*	*Design*	*Programme*
Kärnä et al. (2011a)	Finland	200,000	Age Cohort	KiVa
Kärnä et al. (2011b)	Finland	8,237	RCT	KiVa
Kärnä et al. (2013); Grades 2 – 3	Finland	6,927	RCT	KiVa
Kärnä et al. (2013); Grades 8 – 9	Finland	16,503	RCT	KiVa
Limber et al. (2018)	U.S.A.	70,998	Age Cohort	OBPP
Olweus New National	Norway	16,145	Age Cohort	OBPP
Roland et al. (2010)	Norway	20,446	Age Cohort	Zero programme
Waasdorp et al. (2012)*	U.S.A.	12,334	RCT	SWPBIS

Note: * = only included bullying perpetration outcomes; RCT = randomised controlled trial; OBPP = Olweus Bullying Prevention Programmes; SWPBIS = School-wide behavioural interventions and supports.

Examining this group of studies, it is clear that they collectively provide us with little information regarding the differences between studies implemented in different countries and indeed different anti-bullying programmes. All of these studies were evaluated in either the U.S.A. or a Scandinavian country, i.e., Finland or Norway. Furthermore, the majority of these studies evaluated KiVa or the Olweus Bullying Prevention Programme (OBPP). Moreover, half of these studies utilised either a randomised controlled design or an age cohort design and the majority used large samples to evaluate the effectiveness of these programmes. In relation to the over mean effect size, it is appropriate, for many reasons, that these studies should contribute to the model more than other included evaluations. For example, as previously discussed, larger studies should be given more weight in a meta-analysis due to the increased external validity.

It is also conceivable that evaluations of KiVa and the OBPP influence the overall model because, as discussed in (Chapter 6, see section 6.3.1), these two programmes are currently the most frequently evaluated anti-bullying programmes. Indeed, repetition and consistently desirable outcomes in independent samples are strong indicators of the validity and reliability of evaluation results (Farrington et al., 2002). Whilst these studies appear to meet all of the criteria for valid and reliable scientific research (i.e., large samples, replicated findings, and randomised methodological designs), their inclusion may influence planned subgroup analysis, specifically in relation to analysis of the differences between studies including, or excluding, specific intervention components. It will be argued that it is more meaningful for future research and real-world implementation of anti-bullying programmes to exclude these eight studies from intervention component subgroup analysis. The justifications for omitting these eight studies from the subgroup analysis will now be outlined.

2.7.2.3 Large sample sizes. In primary evaluation studies, sample size and statistical power are very much one and the same. Researchers are instructed to ensure they have an appropriately large sample size in order to be able to obtain enough statistical power to detect the targeted effect. It is widely known that empirical research must avoid both Type I (false positive) and Type II (false negative) errors. While a full discussion of statistical power and type I and II errors is beyond the scope of the present research, the relationship between these concepts and sample size is still relevant.

A good methodological design is said to be one in which the relative risk of incurring a type I and type II error is balanced, and statistical power is said to be the probability that a study will correctly identify an effect (Ellis, 2010). In order to ensure statistical power, a study must use a large enough sample; if not, a study is said to be ‘under-powered’. Thus, it is appropriate that the largest studies included in the meta-analysis are inevitably those that are assigned the largest weight. However, the opposite can also be true. Ellis (2010) outlines that ‘over-powered’ studies are also inefficient. Over-powered studies are those in which the sample is too large, and the probability of even a small effect being statistically significant, typically indicated by $p < 0.05$, is increased. Therefore, if a study is over-powered the results may be as meaningless as a study that in under-powered.

It has been discussed that a meta-analysis appropriately moves the focus from statistical significance, and instead is concerned with the magnitude and direction of effects, which are arguably less influenced by sample size. However, large samples and over-powered studies are still relevant to the current argument. In order to adequately evaluate the relationship between intervention components and effect size subgroups of studies where these components are present or absent were created. Therefore, smaller numbers of studies are compared. When evaluating overall effectiveness, the power of these eight studies does not seem to influence the results too much, and the remaining evaluations contribute sufficiently to the model (i.e., the lack of difference in mean effects under fixed, MVA and random effects models). Yet, in subgroup analysis these studies with large sample sizes may over-power the results.

In the present research, one must consider how large is too large. By increasing sample sizes primary researchers strive to increase the external validity of a study, i.e., how accurately the result can be applied to the general population. But does this come at a price? In the 100 evaluations included in the school-bullying meta-analysis, there was a lot of variation in sample sizes, but there appears to be a significant jump between the numbers of participants included in these eight studies and the remainder of the included evaluations. Increasing sample size exponentially may well increase the external validity of a study, but it also increases the likelihood of confounding factors that threaten internal validity.

Thus, when conducting a meta-analysis of multiple primary studies, we are attributing greater weight to studies with potentially serious confounding variables. For example, when you increase the number of schools included in a primary study from 10 – 20 to 70 – 80 you are increasing the likelihood that these schools are fundamentally different. Previous research shows that various aspects of a school, for example school climate, school leadership, school ethos (Allen, Grigsby, & Peters, 2015; Cohen & Freiberg, 2013; Gage, Prykanowski, & Larson, 2014;) have a great impact on a range of outcomes, including the effectiveness of interventions (e.g., Wang, Berry, & Swearer, 2013). Moreover, while the greater number of schools included means that a greater number of student outcomes are collected, this also increases the number of teachers and/or trainers required to implement the intervention. Doesn't this also increase the problem of implementation fidelity? The counter argument is that random allocation accounts for these confounds, as the sources of error beyond the evaluators control are randomly allocated. However, in real-world research, random allocation may not be sufficiently well implemented in order to account for these confounds. Particularly in school-based research, it is well documented, that whilst true randomisation is the goal, random allocation processes rarely meet this idealistic standard.

Studies are meant to be comparable in a meta-analysis but comparison with respect to sample size are rare in this field. Arguably, studies with 1,000 participants *should* be given more weight than studies with say 50 participants, therefore the MVA model is most appropriate. Similarly, perhaps studies with 2,000 participants should be given more weight than studies with 1,000 studies. But when the incremental increase of sample size becomes too extreme, i.e., Brown et al. (2011) used 4,735 participants and the next largest study, Kärnä et al. (2011b) included just slightly less than double than, with 8,237 participants in Grades 4 – 6 included in the evaluation. Moreover, the increments increase greatly, with the next largest studies being Waasdorp et al. (2012) using approximately 3 times that sample size, 12,334 participants. Kärnä et al. (2013) employed 6,927 participants in Grades 1 – 3 and 16,503 participants in Grades 7 – 9.

It was decided to also exclude Kärnä et al. 2013, grades 1 - 3, n = 6,927, because even though this is closer to the study with the next largest sample size (Brown et al., 2011) there were other issues with these evaluations. Namely, the authors confirmed that it may be the case that some of the participants included in the analysis of each KiVa evaluation, even though they were independent evaluations, may have actually overlapped. In other words, it is possible that some participants in the original nationwide evaluation of KiVa using an age cohort design and published in 2011 were also included in later evaluations. Therefore, the decision was made to exclude all large KiVa evaluations that surpassed the overall mean sample size (n = 4,810 for school-bullying perpetration; n = 4,603 for school-bullying victimisation).

These large studies could have large rates of attrition and missing cases in their data, so that the actual number of participants included in the analysis could be less. However, the same problem arises in face smaller studies. If the rate of attrition is 10% in a study of 12,000 participants, 10,800 participants would be included in the analysis, which is still far larger than 3,600 participants analysed in an evaluation initially including 4,000 participants with a similar rate of attrition.

2.7.3 Additional support

The need for an alternative estimation of the summary effect size whilst also accounting for between-study variance has been frequently discussed in two main areas of research: (1) medical sciences; and (2) criminological research. Whilst the latter is primarily research conducted by the author of this paper from the University of Cambridge; the former is predominantly research produced in the University of California, Berkeley.

For example, in criminology, Farrington and Welsh (2013) compared six models of meta-analysis in order to evaluate the effectiveness of CCTV systems on rates of crime. The authors note that five of the models produced equivalent summary odds ratios, but the random effects model was the only model of meta-analysis to produce a different weighted summary odds ratio. Secondly, when heterogeneity in a meta-analysis is high, all observed effect sizes are assigned very similar weights using the RE method for estimating between-study variance, as has been demonstrated using Gaffney et al. (2019c) meta-analyses on the effectiveness of anti-bullying programmes.

Similar issues are documented in meta-analyses from the medical sciences field. Referred to as the Shore correction (Shore, Gardner, & Pannett, 1993), a number of studies have outlined the need to better account for between-study variance when assessing the relationship between several medicinal compounds and disease-related outcomes. First proposed by Armitage (1985), a simple adjustment to the confidence intervals of the summary mean effect computed under a fixed effects model (i.e., inverse weighting method) has been adopted to address the limitations of the random effects model.

In research for the present paper, a quick Google Scholar search for studies that had referenced the Shore et al. (1993) article was conducted. Ninety-four studies were found to reference this particular paper and many of these were meta-analyses that used an alternative method to calculate the between-study variance (n = 32). These studies are summarized in Table 6 along with a brief description of the justification provided to adjust the variance of the summary mean effect. Additionally, 10 studies (with some overlap) were identified that referenced the Farrington and Welsh (2013) paper on a similar alternative formula for the between-study variance adjustment: Besemer et al. (2017); Cooke & Farrington (2016); Portnoy & Farrington (2015); Gaffney et al. (2019c); Gaffney et al., (2019b); Ttofi et al. (2016); Zych, Baldry, (2019); Zych et al. (2019).

In response to the commonly identified problems with the random effects model, the alternative approach is to adjust for between-study variance *after* estimating the summary mean effect. Thus, the FE variance of the weighted mean effect size is multiplied by a constant (i.e., Q/df). In this way, the summary effect size is still estimated from weighted studies, where the assigned weights reflect the precision of the study itself (i.e., the fixed effects model), but the variance of this pooled estimate is increased to take account of the heterogeneity between observed effects.

Table 6

Description of studies that have used alternative adjustments for heterogeneity in meta-analyses.

Reference	Justification
Ayieko et al. (2014)	*"...still weighs by precision, while also taking heterogeneity into account"* p. 2
Besemer et al. (2017)	*"...takes into account heterogeneity and weights based on precision but does not increase the weights of smaller samples disproportionally"* p. 167
Bhatia, Lopipero, & Smith (1998)	*"The variance.... is adjusted by multiplying it by the ratio of the heterogeneity to its degrees of freedom"* p. 85
Carlos-Wallace et al. (2016)	*"...random-effects model does not weight studies directly on precision; it assigns smaller, less precise studies greater weight... To weight studies directly on precision while still incorporating between-study variance..."* p. 3
Chaffee & King (2012)	Adjustment was applied "*.... whenever this adjustment resulted in more conservative (wider) intervals"* p. 120
Chaffee & Weston (2010)	Shore et al. correction was used *"... whenever this adjustment resulted in more conservative (wider) CIs"* p. 4
Dorjee et al. (2018)	*"... the random effects model does not weight studies directly on precision; it assigns smaller, less precise studies greater relative weight... Fixed-effects model* [with Shore correction] *weights studies directly on precision while still incorporative of between-study variance"* p. 542
Duong et al. (2011)	The RE model *"... weighs studies based on a highly complex and non-intuitive mix of study precision,* [effect size], *and meta-analysis size...* [Shore method] *directly weighs individual studies by their precision, while between-study heterogeneity is* ***only incorporated into the summary*** **[effect size]"** p. 10 (emphasis added)

Erren et al. (2009)	When heterogeneity present *"... we increased the CI around the FES* [summary effect size from FE] *to take account of the between-study variability"* p. 998
Fahimi, Singh, & Frazee (2015)	Reported RE model and FE model summary effect size *"…with the 95% CI recalculated using the adjustment by Shore et al., where between-study heterogeneity is incorporated into calculations of variance"* p. 423
Fink & Bates (2005)	The FE model was used to weight studies and *"... Shore method of adjusting the variance and confidence interval"* p. 702
Gaertner & Thériault (2002)	*"... 95% CIs were adjusted so the* [variance] *was increased by the chi-square statistic divided by its degrees of freedom"* p. 656
Gaffney et al. (2019a)	*"... when heterogeneity is high the resulting between-study variance is large and results in approximately equal weightings for each effect size"* p. 141
Henry & Reingold (2012)	*"The Shore method, which directly weighs individual studies on their precision, was used to calculate the CI"* p. 263
Hickenbotham et al. (2012)	*"... the variance of the log of the* [effect size] *was multiplied by the ratio of the heterogeneity statistics to its degrees of freedom*" p. 3217
Kwan et al. (2004)	*"When evidence of heterogeneity was present, the 95% CI of the summary OR was readjusted… that incorporated between-study heterogeneity"* p. 529
Liu et al. (2011)	*"The Shore correction incorporates between-study heterogeneity and is usually more conservative than the fixed-effects model in estimating variance"* p. 3
Moore & Enquobahrie (2011)	*"... the random effects method can potentially lead to less conservative confidence intervals than the fixed effects inverse variance weighting methods due to a greater relative weight given to smaller studies"* p. 1532
Portnoy & Farrington (2015)	The MVA model *"... yields the same weighted mean effect size as the fixed effects model… but the variance adjustment model… exactly adjusts for the heterogeneity of the effect sizes, giving more weight to larger studies"* p. 36
Schwilk et al. (2010)	*"... the random effects model weights studies based on a highly complex and nonintuitive mix of study precision,* [effect size], *and meta-analysis size. As a*

	consequence, greater weight to smaller studies… may actually be less conservative" p. 879
Setia et al. (2006)	*"… adjusted variance was calculated by multiplying the ratio of the* [heterogeneity] *statistic to its degrees of freedom"* p. 164
Steinmaus, Nuñez, & Smith (2000)	*"… variance of the log pooled* [effect size] *was multiplied by the ratio of the heterogeneity chi-square statistic to its degrees of freedom. This adjusted variance was then used to adjust the 95 percent confidence interval"* p. 695
Steinmaus et al. (2008)	*"… study weighting is not directly proportional to study precision… can lead to summary results that are less conservative than those produced using the fixed effects model"* p. 4
Ssekitoleko, Kamya, & Reingold (2013)	*"… method of Shore et al. in order to account for between-study variance"* p. 5
Vinnikov, Blanc, & Steinmaus (2016)	*"In order to weight studies on precision, while still incorporating between-study variance, we used the fixed effects model to calculate summary* [effect size], *then adjusted their 95% CIs for heterogeneity using the method of Shore et al."* p. 1511
Welling et al. (2015)	*"…the random effects model gives relatively greater weight to smaller, less precise studies… can sometimes lead to summary results that are less conservative than those produced using the fixed effects model"* p. 153
Woolf-King et al. (2013)	*"…both the random effects model was preformed, and the fixed effects 95% CI was adjusted using the Shore method to account for between-study variance"* p. 102
Zhang et al. (2009)	One issue with the RE model "*… is that study weighting is not directly proportional to study precision and greater relative weight is given to smaller studies… can lead to summary results that are actually less conservative"* p. 158
Zhou, Smith, & Steinmaus (2004)	*".. CIs in the fixed effects model were adjusted to account for between-study variance using the method presented by Shore et al"* p. 772
Zych et al. (2019)	*"The random effects method adjusts for heterogeneity, but all studies, large and small, can have similar weights"* p. 3/4

Note. RE = random effects; FE = fixed effects; [ ] denote re-wording for purpose of present chapter and conciseness

2.8Moderator analysis

In traditional empirical research, when one wishes to compare two mean values to evaluate the difference between two participants, or two groups of participants, a *t*-test is the standard statistical test. In meta-analysis, we want to compare sub-groups of studies rather than sub-groups of individuals, so the analysis is slightly different. Guidelines provided by noted meta-analysts for this type of analysis were followed (Borenstein et al., 2009; Lipsey & Wilson, 2001).

The approach involved two steps: (1) computing the mean effect and variance for each subgroup; and (2) comparing the mean effects between subgroups (Borenstein et al., 2009). This approach has been used previously by researchers to conduct similar analyses (e.g., Kaminski et al., 2008; Ttofi & Farrington, 2011). Comparing the mean effect sizes for subgroups involves a method that is analogous to a one-way ANOVA in primary research (Hedges, 1982a; Lipsey & Wilson, 2001).

The between-studies heterogeneity is the value used to evaluate whether the difference between subgroups is statistically significant (i.e., whether the difference in weighted mean effect sizes for subgroups is, at least partially, explained by the relevant intervention component). To compare subgroups of studies in the current thesis, Borenstein et al. (2009)'s approach of using a Q-test based on analysis of variance was used. This method partitions the total variance into within-group variance (Q_W) and between-groups variance (Q_B).

When comparing two groups of studies, for example group X and group Y, the between-group variance is estimated as:

$$Q_B = Q - Q_W$$

where Q_W is the sum of the Q values for groups X and Y. A worked example of this calculation is provided in Appendix 5. This dissertation reports fixed effects values for Q_B.

Subgroup analyses in meta-analysis are observational in nature and not based on randomised comparisons and therefore false negative results may be more likely (Higgins & Green, 2011). As a result, the results of the post hoc subgroup analyses presented in the current dissertation are interpreted with caution. Where individual evaluations may not have sufficient power to test a difference between subgroups, pooling the data in a meta-analytical model the statistical power is increased. However, relatively equal numbers of studies are required in each group being compared (Oxman & Guyatt, 1992; Yusuf et al., 1991). Therefore, where established subgroups did not have relatively equal numbers of studies, subgroup analyses were not conducted.

2.9 Meta-regression

Comprehensive Meta-Analysis™ version 3 software was used to conduct meta-regression analysis to explore the relationship between continuous moderator variables and perpetration and victimisation outcomes. Weighted regression analysis (Lipsey & Wilson, 2001) was used to explore the relationship between continuous variables and school bullying perpetration and victimisation outcomes.

Meta-regression analyses were computed under the RE model and the MVA model. To apply the MVA model in meta-regression, the standard error of regression coefficients of FE models were adjusted using the *Q* value and its *df* for the mean summary effect sizes for subgroups were used to adjust the standard error to reflect between-study variance.

2.10 Manual calculations

Before estimating a weighted mean effect, a number of adjustments were made, due to the nature of the evaluations included in the sample and the limitations of available statistical software. Firstly, due to the inclusion of clusters (i.e., groups of participants) in primary evaluations, corrections for clustering were applied. This procedure is described in section 2.9.1. Secondly, because the CMA software does not have the appropriate algorithms to estimate the effectiveness of an intervention programmes that report results dichotomously, these calculations were done manually. A worked example of this calculation is provided in Appendix 2. Finally, additional manual calculations were completed to transform mean effect sizes expressed as ORs to percentage change in outcomes. This was done to help disseminate the results of the research to a non-expert audience, and an example is provided in Appendix 3 using hypothetical raw data in a 2x2 frequency table.

2.10.1 Corrections for clustering

As the present review aims to evaluate the effectiveness of school-based anti-bullying programmes, cluster-randomized trials were included. Clustering is a common phenomenon in educational evaluations (Donner & Klar, 2002), and occurs when 'clusters', not individuals, are randomly assigned to experimental conditions (Higgins, Deeks, & Altman, 2011). In other words, primary studies sometimes assigned classes or schools to intervention and control conditions, rather than individual students.

Often this approach is utilised in evaluation studies to reduce treatment contamination and increase administrative convenience (Donner, Piaggio, & Villar, 2001). However, one of the main issues with incorporating cluster-randomized trials in a meta-analysis is that participants within one cluster are likely to be more homogeneous than participants in another cluster (Higgins et al., 2011).

Thus, the variance of estimates of treatment effectiveness will be under-estimated (Donner & Klar, 2002, p. 2974). Clustering could occur for several reasons in studies included in the present report. For example: (1) classes of children, not individual children, were randomized to intervention or control condition; (2) the intervention was implemented at the classroom level (i.e., to a class or group of children at one time); or (3) the intervention was targeted at teachers, who were trained to implement the intervention in their respective classrooms.

Therefore, effect sizes in the present meta-analysis were corrected for the inclusion of clusters in primary studies. This is achieved by estimating a *design effect:*

$$1 + (M - 1) \times \text{ICC}$$

where M represents the mean cluster size in each study (e.g., the mean number of students per classroom[3]) and the ICC is the intraclass correlation coefficient.

The ICC is rarely reported by primary studies (Higgins et al., 2011; Valdebenito et al., 2018). Based on Murray and Blitstein (2003), and subsequently the strategy followed by Farrington and Ttofi (2009), an ICC of 0.025 was assumed in the current meta-analysis. The variances of effect sizes were then multiplied by this design effect estimated for each study. In the meta-analysis of school-bullying interventions there were only four studies where corrections for clustering were not required. Three studies (i.e. Berry & Hunt, 2009; Knowler & Frederickson, 2013; Meyer & Lesch, 2000) randomly assigned participants to experimental conditions, and Elledge et al. (2010) described an intervention that was not implemented in a classroom (i.e., the intervention occurred in one-on-one sessions with victims of bullying). Only two studies (i.e., Athanasiades, et al., 2015; Garaigordobil & Martínez-Valderrey, 2016) in the meta-analysis of interventions to reduce cyber-bullying randomized individuals to experimental conditions and thus were not corrected for clustering.

[3] Calculated as: total number of students / number of classrooms

Part Two: School-bullying

3. Literature Review: School-bullying

3.1 Overview

The present chapter of this thesis presents a brief but comprehensive literature review of school-bullying research. Bullying first emerged as an important topic of research in western hemisphere in the 1980s, following the tragic suicides of young boys in Norway, the reason for which was attributed to bullying victimisation (Olweus, 1993). Interestingly, around the same time in Japan, researchers defined the term *'ijime'*, which describes behaviours similar to bullying, as a cause for public concern (Morita, 1996; Morita et al., 1999; Smith, et al., 2002; Smith, Kwak, & Toda, 2016; Toda, 2016).

Since then a wealth of research on the many forms of bullying has been conducted and the importance of reducing experiences of victimisation whilst in school has been stressed repeatedly, by researchers, school stakeholders, global non-profit organizations, and educational policy makers alike. For example, in the United Kingdom, there are many organizations that work tirelessly to reduce bullying and improve the school experience for all students. These organizations include, but are not limited to, the Anti-Bullying Alliance[4], Bullying UK[5], Childline[6], The Diana Award[7], and Kidscape[8]. The following sections of this chapter present an overview of the prevalence, risk and protective factors, and associated outcomes of school-bullying. A discussion of the behaviours that constitute bullying is also provided.

3.2 Bullying definition

In order to adequately determine which interventions will effectively reduce bullying behaviours, it is important that researchers and educators start by accurately assessing the

---

[4] https://www.anti-bullyingalliance.org.uk
[5] https://www.bullying.co.uk
[6] https://www.childline.org.uk/Explore/Bullying/Pages/Bullyinginfo.aspx
[7] https://diana-award.org.uk
[8] https://www.kidscape.org.uk

prevalence of involvement in school bullying (Swearer et al., 2010). There remains some degree of disagreement in relation to definitive cut-off points for involvement in bullying (Solberg & Olweus, 2003; Swearer et al., 2010) and methods utilised for the assessment of bullying (Smith et al., 2002). However, there is better agreement in regard to the defining criteria for school bullying.

Prominent researchers in the field have defined bullying as any aggressive behaviour that incorporates three core elements, namely: (1) an intention to harm; (2) repetitive in nature; and (3) a clear power imbalance between perpetrator and victim (Center for Disease Control and Prevention, 2014; Farrington, 1993; Olweus, 1991). In other words, bullies are individuals who intend to cause harm to their victims through their actions, over a long period of time. Furthermore, victims of bullying are typically less powerful than bullies, or groups of bullies, and feel that they cannot easily defend themselves. This may be due to a physical and/or social power imbalance.

Bullying behaviours can occur in many contexts, for example, in schools, in the workplace, between siblings, and most recently, online. The present chapter is concerned only with school-bullying, namely, bullying that occurs in schools between individuals, usually aged between 4 and 18 years old. In the school context, bullying is a complex social phenomenon, that often does not happen between the bully and victim in isolation (Salmivalli, 2010). For example, individuals can be involved in bullying, not only as bullies, victims, or bully-victims (i.e., those who report bullying others and experiencing bullying victimisation themselves), but also as bystanders, defenders, or reinforcers (Zych et al., 2017).

3.3 Theoretical explanations

There are many attempts in the literature to explain why bullying occurs, and whilst a full review of each theory is beyond the scope and remit of the present research, this section will aim to provide an overview of the dominant theories and supporting experimental evidence.

Bullying is sometimes dichotomized as a form of either reactive or proactive aggression (e.g., Schwartz et al., 1998). Reactive aggression typically describes bullying behaviours that may occur in *response* to an event, such as provocation or heightened emotional state. Proactive aggression on the other hand is used to describe behaviour that is goal-directed and is normally unprovoked. A more recent adaptation of this theory is the Quadripartite Violence Typology (QVT) model that incorporates additional dimensions of the motivations that drive bullying perpetration and self-control (Howard, 2011). Specifically, the QVT approach posits that there are four distinct motivational types that explain why bullying occurs: impulsive-aversive, controlled-aversive, controlled-appetitive and impulsive-appetitive; also referred to as Rage, Revenge, Reward and Recreation (Bjornebekk & Howard, 2012; Runions, Bak, & Shaw, 2017).

In testing this theory, Runions and colleagues (2018) state that the aversive dimensions in this model are those which seek to reduce or avoid an unpleasant event or emotion, while appetitive dimensions describe those which are aimed towards seeking reward, such as pleasant experiences or emotional states. Data from nearly 2,000 Australian adolescents (aged 13 – 15 years old) was used to explore these dimensions with respect to different bullying roles. The results suggest that, when compared to uninvolved students, being involved in bullying in any role was significantly associated with higher scores on all four motives; rage, revenge, reward and recreation (Runions et al., 2018). This study suggested that bully-victims were the most motivated by all four motives in comparison to pure-bullies and pure-victims. When within-group comparisons were conducted, bully-victims reported being motivated most by 'recreation' purposes (i.e., engaging in bullying others for the purpose of enjoyment). Pure bullies were motivated most by 'reward' reasons, or in other words, reported engaging in bullying others in order to achieve a goal or positive reinforcement from peers. Evidently, pure victims reported that they would be most motivated by rage to bully others.

These findings suggest a dynamic and multi-faceted theoretical explanation for school-bullying. This theory is similar to other theories of bullying that emphasize the importance of the peer group in these incidences of aggression, particularly given the implications that those who bully may be motivated to do so in order to be ‘rewarded’, possibly by positive reinforcement from peers. One such theory that emphasizes the role of peers and the peer group in bullying is suggested by Salmivalli and colleagues in Finland. This approach has had a major impact on research, specifically in the development of the KiVa anti-bullying programme, about which more information is provided in subsequent chapters. Briefly, this theory highlights the many roles involved in bullying incidences within the school environment, particularly the importance of those beyond the bully, victim, and bully-victim roles. Salmivalli and colleagues (1996) first tested this theory and labelled bullying roles as: (1) victim; (2) bully; (3) reinforcer of the bully; (4) assistant of the bully; (5) defender of the victim; and (6) outsider. Interesting gender differences were observed, with female participants more likely to be in the role of defender or outsider and male participants most frequently in the role of bully, reinforcer, or assistant. This theory expands the role of the ‘bystander’ into active and passive participants in bullying but emphasizes the importance of understanding the complex social structures of adolescent peer groups when trying to understand bullying (Salmivalli, 2010).

Another common theoretical framework applied to school-bullying research is the ecological model, first proposed by Bronfenbrenner (1979). As previously discussed, this is the theoretical framework applied in the current research and it will be discussed in more detail throughout this dissertation. Briefly, this systems-based approach proposes that bullying can be explained by factors at many different levels of an ecological model (Baldry et al., 2015; Hong & Espelage, 2012). Factors may interact on the individual, peer, school, parent, teacher, and community levels to offer an explanation for bullying perpetration and victimisation. This has important implications for the structure and content of anti-bullying programmes.

3.4 Prevalence

Establishing the true prevalence of school-bullying is a difficult endeavour, as the results will largely depend on the measurement instrument used by the researchers. Additionally, research studies are increasingly reporting the prevalence of offline and online bullying victimisation and perpetration, and the co-occurrence of these forms of aggression (e.g., Baldry et al., 2017). The current chapter, however, prioritises offline-only bullying and issues relating to the overlap of school- and cyber-bullying will be discussed in later chapters (see Chapters 7 and 11).

The United Nations Educational, Scientific, and Cultural Organization (UNESCO, 2019) highlighted ending school violence and bullying as a major priority of their Global Education 2030 Agenda. Their report suggests that nearly one in three students indicate being bullied by their peers on at least one occasion in the previous month. The types of bullying experienced by students varied based on the location, with students in Europe and North America reporting psychological bullying as the most common form of victimisation and students in regions other than Europe and North America reporting physical bullying as the most common. Their report also highlights that overall reports of bullying are decreasing worldwide. The findings from a total of 71 countries suggest that, between 2002 and 2017, 35 countries have seen an overall decrease in reports of bullying, 23 have observed no change in the prevalence of bullying and 13 indicate an increase in bullying reports.

These findings are overall optimistic and could be attributed to the growing attention and awareness about bullying globally. A comprehensive meta-analytical review supports the findings in the UNESCO report, with a mean prevalence rate of 35% for school-bullying perpetration (*n* = 52 studies) and 36% for school-bullying victimisation (Modecki et al., 2014). Thus, effective anti-bullying programmes are still a top priority for schools and educational services around the world, even if reports suggest the percentage of children involved in bullying is falling. Furthermore, given the serious impact bullying can have on the lives of those involved, there remains a strong imperative for anti-bullying research.

3.5 Outcomes and impact

There is a wealth of research exploring the outcomes associated with school bullying, perpetration and victimisation, both in the short-term and throughout the lifespan. A recent comprehensive review of systematic reviews found that the impact of school-bullying can occur not only concurrently with perpetration and/or victimisation but also later in life (Zych, Ortega-Ruiz, & Del Rey, 2015). The current section will provide a brief overview of cross-sectional and longitudinal research that assesses the short- and long-term outcomes associated with school-bullying.

Cross-sectional research has shown that a variety of undesirable short-term outcomes are associated with school bullying. The existing research has largely focused on the mental health and behavioural problems that occur comorbidly with bullying victimisation and perpetration. These outcomes are of great concern and reinforce the need for research on effective bullying intervention and prevention. Longitudinal research has also explored the impact of childhood bullying on experiences and behaviours in adulthood.

3.5.1 Psychological outcomes

One meta-analysis of cross-sectional studies (Birkeland Nielsen et al., 2015) found a significant mean effect between school-bullying and Post-Traumatic Stress Disorder (PTSD;

American Psychiatric Association, 2000). The summary mean effect estimated from 29[9] primary studies that measured school-bullying and workplace bullying suggested that bullying victimisation was significantly correlated with a higher overall PTSD symptom score (*r* = .39, 95% CI .24 - .52). Similarly, empirical research has found a significant relationship between bullying victimisation and post-traumatic symptoms (e.g., Baldry, Sorrentino & Farrington, 2019). Not only did this study identify a significant relationship between both school- and cyber-bullying victimisation and post-traumatic stress symptoms, but there were also significant gender differences. School-bully/victims reported significantly higher levels of post-traumatic symptoms (M = 2.35, SD = 2.47, $F = 196.12$, $p < .001$) than non-involved students (M = 1.86, SD = 2.34) and female bully/victims reported higher levels of post-traumatic symptoms (M = 2.42, SD = 2.56) in comparison to male bully/victims (M = 2.27, SD = 2.38, $F = 66.617$, $p < .001$).

Several studies have also found a concerning relationship between school-bullying victimisation and suicidal ideation. Holt and colleagues (2015) conducted a large-scale meta-analysis ($n = 47$ studies) of cross-sectional studies to examine the relationship between school-bullying experiences and both suicidal ideation and suicidal behaviour. The statistically significant effects suggested that victims of bullying were 2 times more likely to report suicidal ideation (OR = 2.34, 95% CI 2.03 – 2.69) and almost 3 times more likely to report suicidal behaviour (OR = 2.94, 95% CI 2.36 – 3.67). This meta-analytical review also found that bully-victims were at an increased risk; bully-victims were nearly 4 times more likely to report suicidal ideation (OR = 3.81, 95% CI 2.13 – 6.80) and suicidal behaviour (OR = 4.02, 95% CI 2.39 – 6.76). School bullies too were not immune from the impact of their aggressive behaviour. Holt and colleagues observed that bullies were significantly more

[9] Seven studies reported the relationship between school-bullying and PTSD, but the mean effect is not given independently from workplace bullying outcomes.

likely to report suicidal ideation (OR = 2.12, 95% CI 1.67 – 2.69) and suicidal behaviours (OR = 2.62, 95% CI 1.51 – 4.55).

Using data from the Global School-based Student Health Survey, Liu, Huang, and Liu (2019) found that a history of bullying victimisation was significantly associated with suicide attempts amongst adolescents in low- and middle-income countries. This finding has especially important implications in light of recent UNESCO (2019) findings that reports of school-bullying are highest in regions such as the Middle East (41.1%), North Africa (42.7%), and sub-Saharan Africa (48.2%). Furthermore, not only is bullying victimisation associated with suicidal outcomes whilst the victimisation in ongoing, but there is also a significant impact on suicidal outcomes in adulthood. Castellví and colleagues (2016) found that bullying was significantly related to suicide attempts and incidences of suicide in youth and young adults aged 12 to 26 years of age. In a meta-analysis of longitudinal studies, five of which examined the relationship between bullying and suicide, it was found that victims of bullying were at a higher risk of attempted suicide later in life (OR = 2.39, 95% CI 1.89 – 3.01; Castellví et al., 2016).

Qualitative research with university students further demonstrates the severe impact that bullying can have on victims' lives. Participants in one study reported that they attributed many psychological problems, such as low self-esteem, body image problems, eating disorders, anxiety and depression, to experiences of bullying during school (deLara, 2019). These findings are further supported by empirical quantitative research. For example, U.S. research has found a significant relationship between bullying victimisation and higher levels of depression and lower levels of school belonging amongst students aged 11 to 15 years old (Davis et al., 2019).

These adverse outcomes can also continue into adulthood. A prospective study found that experiences of bullying (in all roles) were associated with greater risk of several adverse adulthood mental health outcomes (Copeland et al., 2013). Victims of school-bullying were more likely to report higher levels of: anxiety disorders (OR = 4.30, 95% CI 2.10 – 8.60); panic disorders (OR = 3.10, 95% CI 1.50 – 6.50); and agoraphobia (OR = 4.60, 95% CI 1.70 – 12.50), in comparison to non-involved participants, even when controlling for childhood psychiatric problems. Bully-victims were more likely to report adulthood depressive disorders (OR = 4.80, 95% CI 1.20 – 19.40) and panic disorders (OR = 14.5, 95% CI 5.70 – 36.60). There were further gender differences with respect to these associations, namely that female bully-victims were at an increased risk for agoraphobia in adulthood, but male bully-victims were not, and male bully-victims were at an increased risk for suicidality in adulthood, but female bully-victims were not.

There is a wealth of research to support the link between bullying victimisation and perpetration and a range of psychological outcomes, of which a full review is beyond the scope and remit of the present research. It is noteworthy however to discuss the extent and seriousness of these outcomes and the subsequent impetus for effective anti-bullying programmes in schools worldwide.

3.5.2 Behavioural and social outcomes

Beyond the psychological and mental health outcomes associated with school-bullying, previous research has identified a number of social and behavioural outcomes also. From routine daily habits, such as skipping breakfast (Sanders, 2019) and disrupted sleep (Geel, Goemans & Vedder, 2016), bullying can have an impact on almost every aspect of a person’s life.

A recent meta-analysis of cross-sectional studies examined the relationship between school-bullying and drug use (Valdebenito, Ttofi, & Eisner, 2015). This review concluded that both bullies (OR = 2.82, 95% CI 1.97 – 4.02) and victims (OR = 1.79, 95% CI 1.38 – 2.32) were more likely to report drug-use. Similarly, Priesman, Newman and Ford (2017) found a significant relationship between bullying victimisation specifically and adolescent substance use. The 2013 Youth Risk Behaviour Survey, a nationally representative U.S. study of approximately thirteen thousand students in Grades 9 to 12, assessed the relationship between online and offline victimisation and binge-drinking and marijuana use. The study found that adolescents who reported both offline and online victimisation were also more likely to report binge drinking behaviours (OR = 1.68, 95% CI 1.31 – 2.15) and marijuana use (OR = 1.40, 95% CI 1.10 – 1.78) when compared to non-involved adolescents. When the data for offline-only victims was examined independently, interestingly, participants were less likely to report marijuana use (OR = 0.72, 95% CI 0.59 – 0.88) in comparison to participants not involved in offline-bullying. There was no relationship between offline-only victimisation and binge-drinking behaviours.

Although involvement in school bullying is not necessarily a causal factor for undesirable life outcomes, research has found that there is an apparent association. It may be the case that the experience of school bullying functions as a stepping-stone towards undesirable life outcomes (Arsenault et al., 2010). Meta-analyses have also suggested that there is a significant relationship between weapon carrying and school-bullying (Valdebenito et al., 2017). Specifically, pure bullies (i.e., individuals involved only as perpetrators) were more likely to report weapon carrying (OR = 3.24, 95% CI 2.37 – 4.44) as were pure victims (i.e., those involved in bullying only as victims; OR = 1.79, 95% CI 1.03 – 3.11). Individuals involved in bullying as both a perpetrator and a victim (i.e., bully-victims) were also more likely to report weapon carrying than non-involved individuals (OR = 5.66, 95% CI 3.59 – 8.89). Subgroup analyses found that pure bullies and pure victims who carried a weapon were more likely to do so inside school in comparison to outside school.

These findings are further supported by recently published empirical research using data from the 2015 Youth Risk Behaviour Surveillance System (Semprevivo, Agnich, &

Peguero, 2020). Interestingly, this study found that the relationship between bullying victimisation and weapon carrying was not mediated by student race but was impacted by several individual-level risk factors, such as academic grades, depression, fighting and alcohol use. Additionally, school-bullying perpetration has been shown to be associated with offending (Ttofi et al., 2011b) and engaging in violent behaviours (Ttofi, Farrington, & Lösel, 2012) as adults.

Therefore, a bullying prevention programme could serve as a crime prevention programme, as well as a form of promoting better public health. Moreover, involvement in school bullying has been found to correlate with factors such as low academic achievement (Strøm et al., 2013). Such factors are common risk factors for youth offending and delinquency (Farrington & Welsh, 2008). School bullying is also associated with undesirable school-related outcomes such as truancy and other disciplinary problems in school (Gastic, 2008). Correlational analyses have suggested that victims of school-bullying are more likely to leave school early due to illness, as are the perpetrators of school-bullying (Kowalski & Limber, 2013). Furthermore, bullies were more likely to also report increased absence from school.

Finally, there is a well-established link between bullying victimisation during school and experiences of victimisation later in life. Longitudinal research has shown that peer victimisation at school and workplace victimisation in young adulthood are significantly related. A sample of 251 participants reported on peer victimisation aged 12 to 17 years and later on workplace victimisation at age 22 years (Brendgen & Poulin, 2018). The results showed that peer victimisation significantly predicted victimisation in the workplace, and the relationship was also partially mediated by increased depressive symptoms.

All of these factors could reflect the persistence of the same underlying construct, for example, an internalizing or depressive personality. Moreover, the causal or correlational nature of the relationship between school-bullying and these outcomes remains unclear due to the lack of longitudinal studies.

3.6 Risk and protective factors

School-bullying is a strong risk marker for several negative behavioural, health, social, and/or emotional problems, and there is also a lot of research on the risk factors that may predict bullying perpetration and victimisation. Studying risk and protective factors is another way in which researchers can attempt to explain the causes of bullying. The literature on risk and protective factors relating to school-bullying is extensive and as such, this section aims to provide a brief insight into just some of the factors associated with bullying.

Various personality traits have been established as risk factors for both school-bullying perpetration and victimisation. Studies have found that several facets of a psychopathic trait typology predict school bullying perpetration. Specifically, callous-unemotional, grandiose-manipulative, and impulsive-irresponsible traits predicted school-bullying perpetration behaviours (Orue & Calvete, 2019). Relatedly, Zych, Ttofi, and Farrington (2019) conducted a meta-analysis of 53 primary studies and found significant relationships between bullying roles and empathy and callous-unemotional traits. Bullying perpetration was significantly associated with lower levels of cognitive empathy (OR = 0.60, 95% CI 0.50 – 0.72), affective empathy (OR = 0.51, 95% CI 0.44 – 0.60), and higher levels of callous-unemotional traits (OR = 2.55, 95% CI 1.91 – 3.40), in comparison to non-bullies. With respect to bullying victimisation, no statistically significant relationships between being bullied and empathy (either cognitive or affective) was found. However, victims reported higher levels of callous-unemotional traits (OR = 1.66, 95% CI 1.13 – 2.45) in comparison to non-victims.

Other typologies of personality have been applied to school-bullying research, such as those measured by the Eysenck Junior Personality Inventory (e.g., Machimbarrena et al., 2019). In this study of 604 Spanish adolescents aged between 12 and 15 years old, both neuroticism (B = .06, $p = .048$) and psychoticism (B = .06, $p = .026$) predicted bullying victimisation, as did peer loneliness (B = .70, $p < .001$). The relationship between bullying and Eysenck's personality traits have been well documented in the bullying literature over the past few decades. Early studies, such as Slee and Rigby (1993), found that bullying was related to psychoticism. Additionally, Mynard and Joseph (1997) showed that, in comparison to non-involved students, bullies scored lower on the lie scale of the Junior Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975), victims scored lower on the extraversion scale and bully/victims scored higher on the neuroticism and psychoticism scales.

Specific individual physical traits may also be risk factors for bullying during school. Data from the Canadian sample of the 2001/2002 World Health Organization Health Behaviour in School-aged Children survey found an interesting link between weight-status and bullying (Janssen et al., 2004). In general, children who were classified as being overweight were at greater risk for bullying victimisation in comparison to their normal weight peers. Interestingly, specifically for 15 to 16-year-old participants, higher weight was significantly associated with higher involvement in bullying perpetration also. Using the ecological model, Barboza and colleagues (2009) identified significant relationships between several factors that significantly predicted bullying perpetration. For example, school bullies were more likely to report bullying victimisation, higher levels of watching television and lower levels of teacher support. Additionally, bullies reported higher levels of peer support and that both teachers and parents had low expectations regarding their academic achievement (Barboza et al., 2009). These findings have been replicated with several risk factors at the school, neighbourhood and family levels also (e.g., Bowes et al., 2009; Espelage, Polanin, & Low, 2014).

Some studies may not directly examine a particular risk or protective factor, but the results can be used to extrapolate which factors should be targeted by intervention programmes. For example, Brendgen and Poulin (2018) found that friendship support counteracted the negative link between victimisation and depressive symptoms. Attar-Schwartz, Mishna, and Khoury-Kassabri (2019) found similar results, in that the perception of support from classmates was not only related to a decreased likelihood of being bullied, but it may also mediate the relationship between the negative internalizing and externalizing behaviours associated with victimisation. However, not all risk factors have empirical support. For example, a meta-analytical review of the relationship between socio-economic status and bullying found weak associations at best (Tippett & Wolke, 2014).

Protective factors have also been identified at the different levels of an ecological model. In a systematic review of 18 meta-analytical studies Zych et al. (2019) found that a number of protective factors were significantly associated with school-bullying behaviours. Overall, the authors found that self-orientated personal competency was the strongest protective factor against bullying victimisation. Community and school factors, desirable academic achievement and other-orientated social competency were associated with the strongest protective impact on bullying perpetration. Conceptually, analysing the relationship between protective factors and bullying is difficult and beyond the scope of the present research. A comprehensive detailed review is provided by Zych et al. (2019; see also Zych, et al., 2017) and the reader is urged to consult these publications if interested in protective factors and protecting children against bullying.

3.7 Intervention and prevention

Bullying in schools is increasingly a public health concern. Given its long-term effects, it is imperative that effective intervention efforts are put in place in order to alleviate this troubling school phenomenon (Ttofi, 2015; Ttofi et al., 2011a). There are numerous different anti-bullying programmes (e.g., KiVa, Olweus Bullying Prevention Programme, ViSC), many of which have been evaluated in different countries with participants of different genders, ethnicities and ages. The aim of the present research is to estimate across each of these evaluations whether or not efforts to combat school bullying are effective. Previous research that aims to address this research question is described in more detail in Chapter 1 (see section 1.2.2). Given the wealth of research showing the serious impact that school-bullying can have the lives of bullies and victims, both in the present and the future, the need for effective anti-bullying programmes is clear. Furthermore, there are a number of risk and protective factors that could be targeted by intervention and prevention programmes, and research is needed to examine whether specific elements of programmes are more effective than others.

4. Systematic review: School-bullying

4.1 Inclusion and exclusion criteria

Strict inclusion criteria were employed in the current research were used to identify all potentially includable evaluations of anti-bullying programmes. Criteria were created before conducting searches and were similar to those used in the previous review of anti-bullying programmes (i.e., Farrington & Ttofi, 2009). Specifically, to be included in the review of *school-bullying* intervention programmes, primary evaluations must:

(1) Describe an evaluation of a school-based anti-bullying programme implemented with school-age participants (depending on the site of evaluation, ages may vary between 4 – 18 years of age);
(2) Utilise an operational definition of school-bullying that coincides with existing definitions (e.g. CDC, 2014; Farrington, 1993; Olweus, 1991);
(3) Measure school-bullying perpetration and/or victimisation using quantitative measures, such as, self-, peer-, or teacher-report questionnaires; and
(4) Use an experimental or quasi-experimental design, with one group receiving the intervention and another (control group) not receiving the intervention.

As a result, the present systematic review excludes studies that evaluate the effectiveness of intervention programmes targeting alternative forms of child/adolescent aggressive behaviours, general aggression (e.g., Leff et al., 2010), and school violence (e.g., Giesbrecht, Leadbeater, & MacDonald, 2011). Moreover, studies that only included cyberbullying outcomes were omitted from the systematic review of school-bullying intervention programmes. Other studies were excluded because they measured bullying-related non-behavioural outcomes, for example, ‘attitudes towards bullying’ (e.g., Earhart, 2011), or coping strategies for dealing with victimisation (e.g., Watson et al., 2010). In addition, studies conducted with special needs, delinquent, or psychiatric populations were excluded (e.g., Espelage, Rose, & Polanin, 2015), so that results could be generalizable to the wider mainstream school population. Studies using qualitative measures of effectiveness, such as participant perceptions of the effectiveness of the programme (e.g., Fletcher et al., 2015), were also excluded.

4.2 Searches

In order to identify potentially includable studies, Boolean searches were conducted using multiple combinations of the following keywords: *bully*; victim*; bully-victim; school; intervention; prevention; programme*; evaluation; effect*;* and *anti-bullying.* Searches were conducted on several online databases, including, but not limited to: Web of Science, PsychINFO, PsychINFO, EMBASE, DARE, ERIC, Google Scholar, and Scopus.

Databases of unpublished reports (e.g., ProQuest and ETHOS) were also searched to include grey literature in our review. This should help to minimize potential publication bias linked to larger or significant effect sizes (Easterbrook et al., 1991; McAuley, Tugwell, & Moher, 2000). In addition, evaluation studies included by previous systematic reviews were scanned, based on the name of each programme, for additional-updated evaluation results (i.e., Jiménez-Barbero et al., 2016; Jiménez-Barbero, Hernández, Esteban, & García, 2012; Cantone et al., 2015; Chalamandaris & Piette, 2015; Evans, Fraser, & Cotter, 2014).

Studies included in the previous review (Farrington and Ttofi, 2009; Ttofi & Farrington, 2011), were also included in the present systematic review. Searches for the present review were conducted up to the end of December 2016, for empirical studies published during and since 2009.

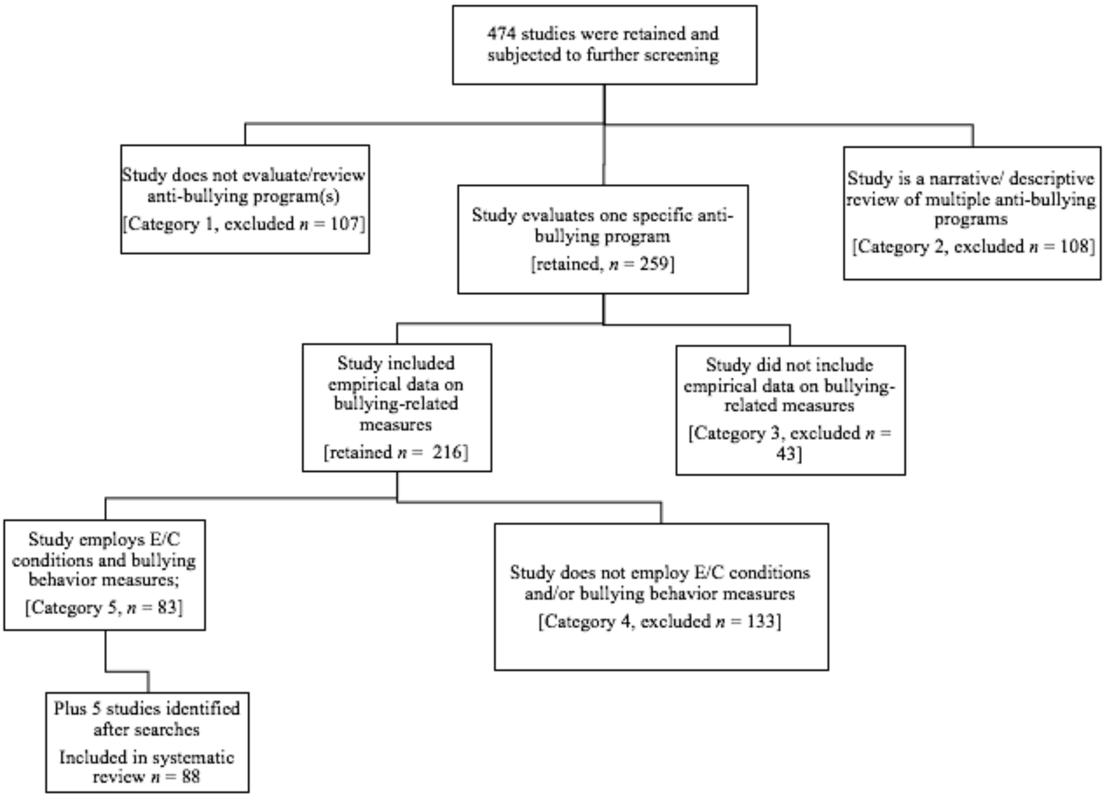
Table 7

Description of relevance scale categories used in first screening wave

Category Name	***Description***
Category 1	Studies were primarily cross-sectional or experimental explorations of factors, constructs or concepts relating to bullying and/or bullying prevention and intervention and implications of findings are discussed in relation to research/development/future anti-bullying programmes.
Category 2	These studies focused more on anti-bullying programmes specifically, either by providing an overview of their effectiveness, theory or implementation or systematically reviewing existing evaluation studies.
Category 3	Studies provided an overview, narrative description of a specific anti-bullying programme or bullying intervention/prevention strategy, however, no evaluation of the effect of implementing the programme is presented.
Category 4	These studies were more relevant to the present review, however, were excluded because they either had methodological issues, the outcomes were not related to a change in actual bullying behaviours (e.g. outcomes related to attitudes towards bullying), or measures related to a construct other than school bullying (i.e. cyberbullying, peer victimisation, or peer aggression).
Category 5 (*included*)	These were evaluation studies of anti-bullying programmes that met all the inclusion criteria for the current review

Figure 1

Screening wave 1: Search results to studies included in the systematic review



4.3 Screening

Our searches of the literature produced 19,877 reports that were screened for eligibility. Based on the title and abstract, a total of 474 primary studies identified as relevant were obtained and subjected to further screening. Studies were allocated to six categories based on their relevance to the current meta-analysis. A description of each category is provided in Table 7.

The initial wave of screening excluded 258 of these primary studies. At this stage, studies were excluded because they: (1) did not evaluate a specific anti-bullying programme (Category 1; n = 107); (2) reviewed several different anti-bullying programmes (Category 2; n = 108); or (3) did not report empirical quantitative data from an evaluation of a specific anti-bullying programme (Category 3; n = 43).

A second wave of screening excluded a further 133 studies (Category 4; see Appendix 1). Primary studies were excluded at this stage because they: (1) reported irrelevant outcomes; (2) did not have an adequate control group; (3) did not meet specified methodological criteria; or (4) did not report independent outcomes (see also section 4.4.2). The screening process is described in detail in Figure 1.

In total, 83 studies published since 2009 were included in our updated systematic review (Category 5). Additionally, five studies were identified during searches conducted for a meta-analytical review of cyberbullying prevention programmes (Gaffney et al., 2019a). These studies were missed during systematic searches for the current review (i.e., Kaljee et al., 2017; Ortega-Ruiz et al., 2012; Ostrov et al., 2015; Silva et al., 2016; Solomontos-Kountouri et al., 2016). One of these studies (i.e., Kaljee et al., 2017) has a publication date outside of the range of our searches. However, it was included because it was available online in 2016.

To provide the most up-to-date analysis of school-based bullying prevention and intervention programmes, therefore, a total of 88 ‘newly identified studies’ are included in the present systematic review. Combining these with the 53 evaluations included by Farrington and Ttofi (2009; i.e., studies published before 2009 or ‘old studies’), a total of 141 studies were included in the systematic review and were eligible for inclusion in the school-bullying meta-analysis.

4.4 Excluded studies

However, a number of these 141 had to be excluded from the meta-analysis and the following sections describe studies that were excluded. Studies were excluded at this stage of the research for one of the following three reasons: (1) inadequate statistical information; (2) non-independent samples; and (3) inadequate evaluation methodology.

4.4.1 Missing information

A certain amount of statistical information is needed in order to produce meaningful effect sizes for a meta-analysis. The effectiveness of anti-bullying programme was estimated as the difference between the experimental and control groups on bullying outcomes, either measured as the percentage of bullies/non-bullies or victims/non-victims or based on mean scores on measurement instruments before and after implementation of the intervention. Therefore, in order to compute an effect size to represent the pre-post intervention effect (see section 2.9) a certain amount of data is needed from the primary study.

However, 21 studies identified by our systematic review did not present sufficient effect size information, and so the primary authors of these publications were contacted. We were able to obtain relevant information for the majority of these studies, but three authors were unable to provide required statistics and seven did not respond to our email communication.

Thus, 10 studies had to be excluded from our meta-analysis because of a lack of information regarding quantitative outcomes. These studies were: Gradinger et al. (2015); Harpin (2011); Kyriakides et al. (2014); Lewis et al. (2013); Lishak (2011); Low and Van Ryzin (2014); van der Ploeg et al. (2016); Şahin (2012); Schroeder et al. (2012); and Wurf (2012). In the previous review by Farrington and Ttofi (2009), 44 out of 53 evaluations provided sufficient information on quantitative outcomes. Thus, nine studies from this previous review were excluded.

4.4.2 Overlapping samples

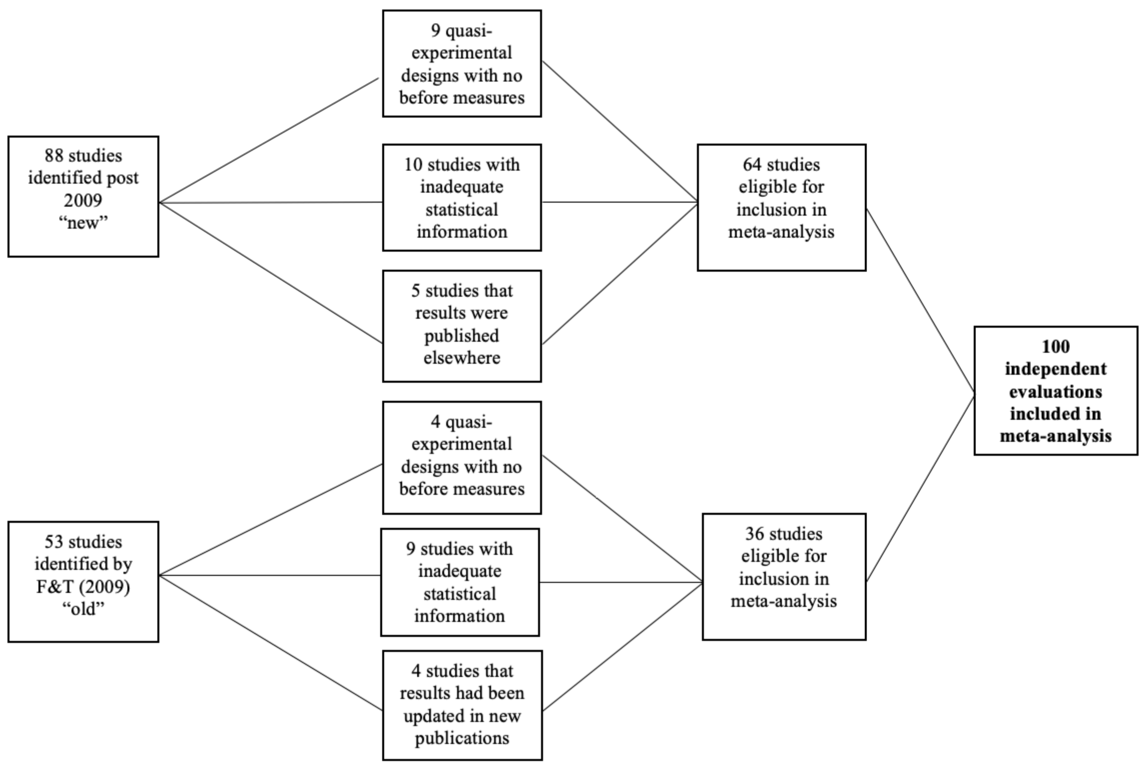
One further stipulation of a meta-analysis is that the final samples must be independent of one another (Ellis, 2010; Borenstein et al., 2009). Overlapping samples are statistically dependent, and thus the variance of the summary effect size produced by the meta-analysis would be under-estimated (Wilson, 2010). Therefore, before conducting our meta-analysis we ensured that all samples were independent of one another.

This issue of non-independent samples was particularly relevant for the multiple evaluations of the KiVa anti-bullying programme. Our thorough systematic searches identified 16 potentially includable studies presenting evaluation data from implementation of the KiVa programme (i.e., Ahtola et al., 2012; Ahtola et al., 2013; Garandeau, Lee, Salmivalli, 2014; Garandeau, Poskiparta, & Salmivalli, 2014; Haataja et al., 2014; Hutchings & Clarkson, 2015; Kärnä et al., 2011a; Kärnä et al., 2011b; Kärnä et al., 2013; Nocentini & Menesini, 2016; Noland, 2011; Sainio et al., 2012; Salmivalli et al., 2012; Williford et al., 2012; Williford et al., 2013; Yang & Salmivalli, 2015).

However, following further screening, only four of the aforementioned studies were subsequently included in the systematic and meta-analytic review (i.e., Kärnä et al., 2013; Kärnä, et al., 2011a; Kärnä, et al., 2011b; Nocentini & Menesini, 2016). These four studies presented independent results of the KiVa programme from the initial nationwide evaluation in Finland and Italy and are included in the meta-analysis. The remaining 12 publications relating to the KiVa programme utilised data from the randomized controlled trial evaluation in Finland (i.e., Kärnä et al., 2013 or Kärnä et al., 2011b) but explored different facets of the programme's effectiveness. These studies are included in 'Category 4' (*n* =12; *N* = 133) and described in Appendix 4.

Figure 2

Screening wave 2: Exclusion of 41 studies



Four studies identified in our systematic searches replaced evaluations included in the earlier review. For example: (1) Menard & Grotpeter (2014) was a continuation of the Menard et al. (2008) evaluation; (2) Cross et al. (2011) was a republication of the Cross et al. (2004) evaluation included in the previous review; (3) Jenson et al. (2013) and Jenson et al. (2010) presented data from additional follow-up points to the Jenson et al. (2007) evaluation; and (4) Frey et al. (2009) used an age cohort design to evaluate follow-up effects from the earlier Frey et al. (2005) study. In cases such as these, the most recent publication, or the publication with the most statistical information, was included in the meta-analysis.

Ten studies (published both before and since 2009) were identified as reporting the effectiveness of an anti-bullying programme from the same sample, or were repeat publications of earlier studies (e.g., DeRosier, 2004 and DeRosier & Marcus, 2005; Domino, 2011 and Domino, 2013; Espelage et al., 2013 and Espelage et al., 2015; Jenson et al., 2013 and Jenson et al., 2010; and Menesini et al., 2012; Study 2 and Palladino et al., 2012). In these instances, the most recent publications were selected, and as a result, five studies were excluded from the meta-analysis.

4.4.3 Inadequate evaluation methodology

In comparison to the previous meta-analysis (i.e., Farrington & Ttofi, 2009), the present review excluded evaluations that were conducted using ‘other experimental-control’ designs. These designs primarily included evaluations whereby quasi-experimental methods were used but bullying outcomes were only measured after the implementation of the intervention. Thus, the effect of the intervention on outcomes of interest cannot be adequately estimated as levels of bullying were not measured before the intervention took place. Overall, nine studies that used this design were omitted from the present meta-analysis.

4.5 Included studies

Therefore, in total 41 studies were excluded from the meta-analysis of the effectiveness of school-based bullying intervention programmes**.** Figure 2 outlines this second wave of screening. Thus, 100 evaluations of anti-bullying programme were included in the meta-analysis, 64 of those were identified by systematic searches conducted for the present research (i.e., post 2009) and 36 were identified by systematic searches by Farrington and Ttofi in 2009.

Table 8 briefly outlines the intervention programmes implemented and evaluated in studies published post 2009 and identified in systematic searches described in the present chapter. For a review of the 36 studies also included in the meta-analysis but identified by searches conducted by Dr. Ttofi please see publications of this earlier meta-analysis (i.e., Farrington & Ttofi, 2009; or Ttofi & Farrington, 2011).

Table 8

Overview of 'new' evaluations included in the school-bullying meta-analysis

Randomized Controlled Trials (*n* = 33 evaluations)			
Project	***Anti-bullying Programme; Key Features***	***Participants***	***Research Design***
Berry & Hunt (2009) Australia	*The Confident Kids Programme*; CBT for anxiety management; Target factors such as: self-esteem, coping strategies, social skills, emotional regulation and internalizing behaviours. 8 weekly sessions led by clinical psychologists.	46 adolescent males (mean age = 13.04) who scored at least 1 SD higher than mean on a pre-test anxiety measure and reported being bullied in the past month.	Participants were assigned to groups based on their grade, and then these groups were randomly assigned to either intervention or waitlist control condition. Child- and parent-report measures completed before, after, and at 3-month follow up.
Bonell et al. (2015) UK	*INCLUSIVE*; Whole-school restorative anti-bullying programme; Action group of staff and students; Needs assessment at baseline informed schools' intervention implementation. Core components: Staff training in restorative practices and student social-emotional skills curriculum.	1,017 Year 8 students aged 12 to 13 years old in English secondary schools.	Matched pairs of schools were randomly assigned to either the intervention (4 schools) or the control (4 schools) condition. Pre- and post-measures of bullying were administered to all participants. Bullying perpetration measured by the self-report AAYP violence scale and bullying victimisation measured by the self-report Gatehouse Bullying Scale.
Brown et al. (2011) US	*Steps to Respect*; Whole-school programme to reduce bullying by increasing staff efficacy, creating positive school climate, and increasing students' social and emotional skills. Classroom curriculum of 10 lessons implemented by trained teachers; Individual bullies and victims received targeted intervention	4,735 staff (*n* = 1,307) and students (*n* = 2,940) from public elementary schools. 128 staff members were teachers. 49% of students were male and 52% identified as white. The mean age of students was 8.9 years.	34 matched school pairs where one of each pair was randomly assigned to the intervention condition, and the other to a waitlist control condition. Teacher-report and self-report measures completed before and after intervention.
Chaux et al. (2016) Germany	*Media Heroes;* Cyberbullying prevention programme; targets empathy, awareness and knowledge about	1,075 students aged 11 – 17 (mean = 13.36) from five schools in Germany.	Schools randomly assigned classrooms to one of three conditions: control; long-version; or short-version. Self-report measures of bullying

	bullying and cyberbullying; provides bystanders with effective intervention and prevention strategies		perpetration and bullying victimisation were administered before and after the intervention.
Cissner & Ayoub (2014) US	*Fourth R: Strategies for Healthy Youth Relationships*; Dating violence prevention programme; Trained teachers implement 21-lesson curriculum targeting: personal safety, healthy growth and sexuality, and substance use/abuse.	517 7th grade students from 10 middle schools.	Students from the 10 schools were randomly assigned to either the experimental or control condition, and all completed self-report bullying measures (secondary outcome) at baseline, post intervention and one-year follow up.
Connolly et al. (2015) Canada	*Youth led programme;* High school students are trained to implement this school violence prevention programme with middle school children; Youth leaders were trained by mental health professionals; Targeted students' knowledge & attitudes of peer aggression and victimisation.	509 7th and 8th grade students from Canadian middle schools, mean age was 12.37 years and 51.4% were female.	Four schools were randomly assigned to either intervention or usual practice control condition. All participants completed self-report bullying measures (from the Safe School Survey) pre- and post-intervention.
Cross et al. (2011); Cross et al. (2004) Australia	*Friendly Schools Project*; Educational techniques based on Social Cognitive Theory; Anti-bullying work implemented at whole-school and community level, and also with students and their families; Trained teachers implemented 9 structured lessons.	1,968 4th grade students from schools in Perth. 51.1% of the intervention condition were female and had a mean age of 8.57 years. 48.3% of students in the control condition were female, and they had a mean age of 8.55 years.	29 schools were randomly assigned to either intervention or standard curriculum control condition. Self-report measures (OBVQ) of bullying perpetration and victimisation was collected at 4 time-points from all participants over the course of the 3-year trial.
Domino (2011); **Domino (2013)** US	*Take the LEAD*; Based on Social-emotional learning and Positive Youth Development theories. 16 weekly lessons covered issues such as: self- and social awareness; self-management; relationship skills; decision making; problem solving and leadership.	323 7th grade suburban middle school students, with a mean age of 12.2 years and 93% were Caucasian.	32 classrooms were randomly assigned to intervention or waitlist control group, and all participants completed self-report bullying measures pre- and post-test.
Espelage et al. (2013); **Espelage et al. (2015)** US	*Second Step: Student Success Through Prevention*; Social-emotional learning middle school programme; Trained teachers implement curriculum in 15 weekly classes, covering issues such as: empathy;	3,658 students from 36 schools in Illinois and Kansas. Mean age was 11 years at the first time-point, 1,961 students received the	36 schools grouped into matched pairs, and schools then randomly assigned to either the intervention condition or a waitlist control condition using a random number table. All

	communication; bullying; emotion regulation; problem solving; and substance abuse prevention	intervention (52.1% male), and 1,697 acted as controls (52.35% male).	participants completed bullying measures at three time points: Wave 1 (pre-test); Wave 2 (post-test; Espelage et al., 2013); and Wave 3 (after 2 years of intervention). Bullying perpetration and victimisation were measured using the self-report Illinois Bully & Victim Scales.
Fekkes et al. (2016) Netherlands	*Dutch Skills for Life;* Universal school-based prevention programme for adolescents; Delivered by trained teachers; 25-lesson curriculum over 2 years; Target: awareness and coping with emotions and feelings; problem-solving; emotional regulation; bullying; friendship; sexuality; and substance abuse; Activities included DVDs, role plays and group discussions	1,394 students in grades 7 – 9 from 26 schools; Aged 13 to 16 years old.	Schools were randomized to the experimental condition (13 schools) or the control group (13 schools). Self-reports of bullying perpetration and victimisation were collected before the intervention (T0), after 1 year of implementation (T1), and at the end of the second year of implementation (T2).
Garaigordobil & Martinez-Valderrey (2015) Spain	*Cyberprogram 2.0*; Cyberbullying intervention programme, traditional bullying also included; 19 lessons aim to raise awareness, outline the consequences of, and develop coping strategies relating to bullying and cyberbullying. Participants are also taught to develop positive social and emotional skills.	176 secondary school students, aged 13 to 15 years old and 56.3% female. 93 students were in the intervention condition, and 83 were in the control condition.	Classrooms from 3 different schools were randomly assigned to either the control or intervention condition and participants from both conditions completed self-report bullying measures pre- and post-implementation.
Holen et al. (2013) Norway	*Zippy's Friends*; Whole-school programme designed to increase coping strategies in order to reduce psychological problems. 24 weekly lessons given by trained teachers; Curriculum based around concept of a character 'Zippy' and his friends as they encounter several relationship problems.	1,483 2nd grade primary school children from 35 schools. 49.3% were female, and the mean age was 7.3 years	Schools were placed in matched pairs and randomly assigned to either the intervention or 'business as usual' control condition. Teacher-reported bullying measured by the Class Climate Survey at pre- and post-intervention.
Jenson et al. (2013); Jenson et al. (2010); US	*Youth Matters*; School violence programme to increase school and peer norms against anti-social behaviours, such as, bullying; 10 modules that aimed to raise awareness, empathy about bullying and social skills	876 6th grade students from public elementary schools. Mean age was 9.82 years old, and 52% were female.	Matched school pairs randomly assigned to intervention and control condition. Self-report measures (OBVQ) administered at 2 time-points: pre-test (baseline) and post-test (12-month-follow up).

Ju et al. (2009) China	*Chinese anti-bullying intervention programme*; Action research framework; Teachers designed and implemented a 5-week intervention for the whole-class, and also specifically for bullies and victims.	354 3rd and 5th grade Chinese primary school children from one school. Two classrooms of each grade participated in evaluation.	Two classrooms were randomly assigned to the intervention condition (one 3rd grade & one 5th grade) and the other two classrooms acted as controls (1 3rd grade & 1 5th grade). Chinese version of the self-report OBVQ employed pre- and post-implementation.
Kaljee et al. (2017) Zambia	*Teachers Diploma Programme;* Situated supported distance learning programme for educators; monthly community of practice meetings to review programme content; target the interaction between psychological and social aspects of participants' lives; focus on self-care, support skills, safe school environment, and positive inter-school relationships.	325 teachers and 1,378 students from 20 experimental and 20 control schools. Mean age of students in 3rd and 4th grade was 10.9 years old and 55.8% were female.	Waitlist randomized controlled design; Students in classes in experimental schools randomly selected; Students in classes in control schools randomly selected; Both teacher-report and self-report measures administered before and after implementation
Kärnä et al. (2011b) Grades 4 - 6 Finland	*KiVa*; Whole-school programme that also targeted individual cases of bullying within a school; Structured curriculum involving class and parent-involved activities; Anti-bullying computer programme for students; Training for teachers on classroom and bullying hotspot supervision/management.	8,237 students from grades 4 – 6 from 275 schools, 429 classrooms, aged 9 to 11 years old.	78 schools were randomly assigned to intervention or control condition. All participants completed self- (OBVQ) and peer-report (Participant Role Questionnaire) measures of bullying perpetration and victimisation at baseline, mid intervention, post-intervention.
Kärnä et al. (2013) Grades 1 - 3 Finland	*KiVa*; See Kärnä et al. (2011b) .	6,927 students from grades 1 – 3 in 74 schools and 397 classrooms.	74 schools were randomly assigned to intervention or control condition. All participants completed self- (OBVQ) and peer-report (Participant Role Questionnaire) measures of bullying perpetration and victimisation at baseline, mid intervention, post-intervention
Kärnä et al. (2013b) Grades 7 - 9 Finland	*KiVa*; See Kärnä et al. (2011b)	16,503 students from grades 7 – 9 in 73 schools and 1,000 classrooms.	73 schools were randomly assigned to intervention or control condition. All participants completed self- (OBVQ) and peer-report (Participant Role Questionnaire) measures of bullying perpetration and victimisation at baseline, mid intervention, post-intervention.

Knowler & Frederikson (2013); UK	*Emotional Literacy intervention*; 12-week programme led by trained professional; Targeted students' emotional literacy skills; Main concepts included: self-awareness; self-regulation; empathy; and social skills.	50 primary school children, aged 8 – 9 identified as being involved in bullying behaviours using a peer nomination measure (Guess Who measure)	Children assigned to intervention (n = 22; 18 male & 4 female) or waitlist control condition (n = 23; 21 male & 2 female). Guess-Who peer nomination measure of bullying perpetration employed to all participants pre- and post-intervention.
Krueger (2010) US	*School Bus anti-bullying intervention*; Intervention materials adopted from "Take a Stand, Lend a Hand, Stop Bullying Now!" online tools; DVD clips about bullying were shown to experimental students each day at the end of school	47 elementary school students that were assigned to one of two possible school buses.	Randomly assigned students to either Bus A, who received the intervention, or Bus B, who were the control group. Data collected from all students prior to the intervention, and 5 days after.
Lewis et al. (2013); **Li et al. (2011)** US	*The Positive Action programme*; School well-being programme; Targets distal (school climate and teacher classroom management) and proximal (students' thoughts & feelings) factors to improve a range of health and behavioural outcomes.	624 grade-3 students were followed over 6-year period.	Matched school pairs randomly assigned to intervention or control group, in a longitudinal design with 8 waves of data collection. Self-reported bullying-related aggression measures employed at each time-point.
McLaughlin (2009) US	*CBT & CBT+media*; Standardized cognitive behavioural therapy (CBT) and an anti-bullying DVD. CBT was delivered in classrooms by a trained professional, and targeted bullying and aggression issues over 4 weekly lessons following a strict outline.	68 6th grade students from 6 classrooms in 3 different schools. Mean age was 11.35 years old and 58.5% were female.	Classrooms were randomly assigned to one of three conditions: (1) CBT only (n = 28); (2) CBT plus media, i.e., the bullying DVD (n = 25); and (3) control group (n = 15). All participants completed self-report measures of bullying perpetration and victimisation (OBVQ) pre- and post-test.

Nocentini & Menesini (2016) Italy	*KiVa*; Whole-school programme that also targeted individual cases of bullying within a school; Structured curriculum involving class and parent-involved activities; Anti-bullying computer programme for students; Training for teachers on classroom and bullying hotspot supervision/management.	2,042 students from 13 Italian schools participated. 1,039 students from 51 classes in 7 schools participated in the intervention, and 1,003 students from 46 classes in 6 schools participated as controls.	7 schools were randomly allocated to intervention condition, and 6 schools were randomly allocated to control condition. The Florence Bullying-Victimisation Scales self-report measure of bullying perpetration and victimisation were employed pre- and post-intervention.
Ostrov et al. (2015) US	*Early Childhood Friendship Project;* classroom-based early childhood intervention; aims to reduce physical and relational aggression; target social-psychological adjustment problems during development; include components on social modelling, problem-solving and conflict resolution, modifying reinforcement contingencies, and social and emotional skills training.	141 participants from six schools accredited for 'Education of Young Children'. 47.5% were female (*n* = 67) and the mean age was 45.53 months old (approximately 3.79 years).	Six classrooms were randomly allocated to the intervention condition (n = 80) and six classrooms were randomly allocated to the control condition (n = 61). Bullying was measured using teacher- and observer-report scale, the PBSM (Preschool Bullying Subscales Measure; Ostrov & Kamper, 2015).
Polanin (2015) US	*Second Step*; Social-emotional learning middle school programme; Trained teachers implement curriculum in 15 weekly classes, covering issues including bullying	55 students in the 5th grade at one middle school. Participants were aged 10 to 11, and 58% identified as Caucasian	Two classrooms were halved, and one half of each classroom were assigned to the intervention and the other hald were assigned to the control condition. Self-reported bullying perpetration and victimisation were measured at 5 time-points.
Stallard et al. (2013) UK	*The Resourceful Adolescent Programme*; Classroom-based CBT programme for depression; 9 lessons outlined in a curriculum manual; Core components include: psychoeducation; helpful thinking; personal strengths; problem solving; and support networks.	1,064 Year 8 – 11 students in UK secondary schools identified at baseline as being 'high risk' for depression. Participants were aged 12 to 16 years old.	Year groups were randomly allocated to one of three possible experimental groups: (1) CBT intervention group; (2) Attention control group 1; and (3) control group 2. OBVQ administered at 3 time-points (baseline, 6 and 12-month follow ups) to assess change in bullying behaviours.

Topper (2011); Study 1 US	*Preventure*; Personality-targeted CBT for high risk students in each of the four domains: hopelessness; anxiety-sensitivity; sensation seeking; and impulsivity. Workshops were implemented by a trained professional.	292 secondary school students from 9 different schools. Mean age was 14 years old, and 67% were female.	Participants were randomly assigned to either intervention (n = 167) or control (n – 125) groups. Self-report bullying measures (OBVQ) were administered at 4 time-points: baseline and 6-, 12- and 18-month follow ups.
Topper (2011); Study 2 US	*Adventure*: extension of Preventure; Intervention followed a similar procedure to the Preventure study, but CBT lessons were implemented by trained teachers.	1,089 secondary school students in years 9 - 11, from 18 different schools. 55.1% of participants were male, and the mean age was 13.71 years.	Schools were randomly assigned to intervention (n = 625) or control (n = 464) condition, and all participants completed self-report bullying (OBVQ) measurement instruments at baseline (pre-intervention) and 6-, 12-, and 18-month follow up time-points.
Trip et al. (2015) Romania	*REBE and ViSC*; Dual components of Rational Emotive Behavioural Education and the ViSC social competence programme; Targets social-emotional factors related to bullying and aggression.	970 6th grade Romanian students from 11 different schools. Mean age was 11.82 years old, and 53% of participants identified as being male.	Schools were randomly assigned to one of three potential conditions according to the order in which they were exposed to the intervention programmes: (1) REBE then ViSC group (n = 385); (2) ViSC then REBE group (n = 270); and (3) control group (n = 315) who were not exposed to either programme. Self-reports of ever being bullied/ever bullied collected pre, during and post intervention.
Tsiantis et al. (2013) Greece	*Greek anti-bullying programme (2);* School-based programme implemented by trained teachers and accompanying programme manual; Ongoing support from mental health professionals; 11 weekly workshops (90 minutes each); Classroom activities included discussion groups, and formation of class anti-bullying rules. Parent information sessions were also held.	666 4th to 6th grade students from 20 elementary schools.	Schools were matched based on prevalence levels of bullying and victimisation. All participants completed the Greek translation of the OBVQ (self-report) pre- and post-implementation.
Waasdorp et al. (2012)	*School-wide Positive Behavioural Interventions and Supports*; Universal behavioural intervention programme	12,334 elementary school students from 37 U.S. public schools. 52.9% of participants	Schools randomly assigned to intervention or waitlist control condition, and teacher-report

US	targeting school-level factors; Focuses on schools' discipline and behavioural management strategies to reduce bullying; Bullying 'hot spots' targeted for increased teacher supervision, and anti-bullying materials spread around the school	were male and 46.1% identified as Caucasian.	(Teacher Observation of Classroom Adaptation-Checklist) of bullying perpetration employed at pre- and post-intervention.
Wölfer & Scheithauer (2014) Germany	*fairplayer.manual*; 15-week curriculum classroom-based anti-bullying programme delivered by either trained teachers or professionals. Aim to reduce bullying by increasing students' social and moral competencies. Lessons target: raising awareness, changing attitudes and encouraging bystander intervention.	328 students in 7th to 9th grades from 2 German secondary schools. 51% were female and the mean age was 13.7 years old.	3 class groups from each school were randomly selected and assigned to the intervention group. The remaining participants acted as waitlist control group. Pre- and post-self-report measures of bullying perpetration and victimisation (OBVQ) were implemented 4 months apart.
Yanagida et al. (2016) Austria	*ViSC;* Training programme led by professionals to increase students' sense of responsibility and competency in conflict; 13 structured lessons; Covered topics such as: impulsivity; reflecting on behaviour; and acting in a socially responsible manner.	2,042 secondary school students from 103 5th to 7th grade classrooms in 26 schools in Vienna. 1,377 were in the intervention group and 665 were in the control group. 47.6% were female and the mean age was 11.7 years old.	13 schools were randomly assigned to the intervention group and 13 schools were randomly assigned to the control group. All participants completed outcome measures for bullying perpetration and victimisation pre- and post-implementation.
Before-After, Experimental-Control designs (*n* = 25 evaluations)			
Project	***Anti-bullying Programme; Key Features***	***Participants***	***Research Design***
Battey (2009) US	*The Bully Prevention Challenge Course Curriculum*; Activity-based anti-bullying programme implemented by Physical Education/Health teachers; Intervention includes warm-up activities, group discussions and raising awareness about bullying.	249 7th grade students from two public middle schools.	Intervention (n = 120) and control (n = 129) students all completed bullying measures pre- and post-implementation.

Bull et al. (2009) Germany	*fairplayer.manual*; Weekly curriculum classroom-based anti-bullying programme delivered by either trained teachers or professionals. Aim to reduce bullying by increasing students' social and moral competencies. Lessons target: raising awareness, changing attitudes and encouraging bystander intervention.	119 7th to 9th grade students from one German secondary school. 64 were female and the mean age was 15.13 years old	Three experimental groups were employed according to the duration of intervention they received: (1) Received 10 weeks of the intervention over the course of 15 to 17 weeks; (2) Received 10 weeks of intervention over 12 months; and (3) Control group that were not exposed to intervention. All participants completed bullying measures, pre, post (+4 months) intervention and at a 12 month follow up.
Elledge et al. (2010) US	*Lunch Buddy mentoring programme*; Victims of bullying are paired with a trained college mentor; Mentors and mentees meet twice a week, over the course of 5/6 months; Mentors sit with mentees during lunchtimes and provide social and emotional support.	36 students from 4 primary schools, grades 4 and 5, whom teacher and peer report indices identified as being victims of bullying. Mean age was 10.36 years old.	Employed 3 experimental groups: (1) Intervention group (n = 12); (2) 'Same' control group who were from the same school as the experimental group (n = 12); and (3) 'Different' control group who were from a different school (n = 12). All participants completed bullying measurement instruments pre- and post-implementation.
Finn (2009) US	*Olweus Bullying Prevention programme*; Whole-school approach; Individual-, peer-, classroom-, teacher-, and school-level factors included.	801 3rd to 5th grade students from 4 elementary schools.	Assigned 2 schools to intervention condition (n = 437) and 2 schools to control condition (n = 383). All participants completed the OBVQ pre- and post-implementation.
Herrick (2012) UK	*Defeat Bullying*; Curriculum-based anti-bullying programme developed by the NSPCC; Targets several key bullying-related issues, such as, attitudes and feelings about bullying, diversity, safety and encouraging bystanders to prevent, or intervene in, bullying.	69 Year 5 students from 3 primary schools.	Utilised a pre/post non-equivalent quasi experimental design. School 1 received the intervention; School 2 received the intervention plus parental involvement; and School 3 acted as a waitlist control school.

Joronen et al. (2011) Finland	*Drama programme;* Based on drama and social cognitive theories; Trained teachers implemented one drama session per month; Themes included: bullying, friendship, loss of a friend, supporting a victim of bullying, tolerance and child abuse.	190 Grade 4 and 5 students from 2 Finnish primary schools.	Schools were purposively allocated to the intervention or control condition, and bullying was measured pre- and post-implementation of the intervention programme.
Losey (2009) US	*OBPP*; Whole-school programme, also included individual-, class-, and community-level factors; School conference held at beginning of programme; Detailed teacher handbook; Parent/Teacher meetings; Class anti-bullying rules.	699 high school students from 2 U.S. schools, 416 were female.	Schools were allocated to intervention (n = 251 students) or control (n = 448 students) by the region's superintendent based on prevalence of bullying. All participants completed the Revised OBVQ pre- and post-test.
Kimber et al. (2008) Sweden	*SET;* Socio-emotional training programme implemented by teachers in classrooms during normal class hours; Components included: teacher manual, student workbook, role-play and take-home exercises; Topics included social problem solution, strong emotions, similarities and differences; values; conflict management; and resisting peer pressure.	1,417 students in grades 1 – 9 in Swedish schools were included at baseline.	Design was quasi-experimental with participants allocated to SET (n = 1,028) condition or no treatment control (n = 389). All participants completed several measures, including one question on bullying, before and after implementation of intervention (t0, t1). Follow-up was included (t2).
Menard & Grotpeter (2014); Menard et al. (2008) US	*Bully-Proofing Your School*; Whole-school programme; Individual support also provided for bullies and victims; Restorative non-punitive disciplinary policies; Classroom curriculum implemented by teachers; Parent information	3,497 3rd to 5th grade students from 6 elementary schools, 52.1% were female.	Assigned schools to either intervention or control conditions in a non-equivalent groups design. All participants completed bullying measures pre- and post-test over 5-year period.

Menesini et al. (2012; Study 1) Italy	*Noncadiamointrappola (Let's Not Fall Into a Trap); NoTrap!;* Web-based peer-led anti-bullying intervention; Selected group of adolescents monitor an online anti-bullying forum; In-class anti-bullying activities	386 secondary school students at 8 Tuscan schools, 20.3% were male, and the mean age was 16.29 years old. 9th to 13th grade students for intervention running from December 2009 – June 2010.	Students were assigned to one of three potential groups: (1) Control group; (2) Intervention group; and (3) Peer educators. Bullying measures were administered pre- and post-test (6 months apart).
Ortega-Ruiz et al. (2012) Spain	*ConRed;* Cyberbullying prevention programme; developed using evidence on effective anti-bullying intervention components; Involves several strategies: (1) proactive policies, procedures and practices; (2) school community key understandings and competencies; (3) protective school environment; (4) school-family-community partnerships	893 high school students, 595 were in the intervention group (45% female) and 298 in the control group (47.6% female). Students were aged 11 – 19, with a mean age of 13.8 years old.	Researchers and teachers allocated classes of students to experimental or control groups; All participants completed the European Bullying Intervention Project Questionnaire (ECIPQ; Brighi et al., 2012) before and after implementation.
Palladino et al. (2012); Menesini et al. (2012; Study 2) Italy	*NoTrap!*; Web-based peer-led anti-bullying intervention; Selected group of adolescents monitor an online anti-bullying forum; In-class anti-bullying activities	375 9th to 13th grade students at 4 Tuscan high schools for year December 2010 – June 2011.	Students were assigned to one of three potential groups: (1) Control group; (2) Intervention group; and (3) Peer educators. Bullying measures were administered pre- and post-test (6 months apart).
Palladino et al. (2016; Trial 1) Italy	*NoTrap!*; Web-based peer-led anti-bullying intervention; Selected group of adolescents monitor an online anti-bullying forum; In-class anti-bullying activities	622 9th grade students from 8 high schools in Tuscany during the school year 2011/2012. 22 classes in 5 high schools were allocated to the intervention condition (*n* = 451; mean age = 14.79; 57% male) and students from 9 classes in 3 high schools participated as controls (*n* = 171; mean age = 15.28; 69% male).	All participants completed the Florence Bullying-Victimisation scales at pre- and post-test. Scale measures the frequency of bullying perpetration and victimisation experienced by respondents during the past 2 months.

Palladino et al. (2016; Trial 2) Italy	*NoTrap!*; Web-based peer-led anti-bullying intervention; Selected group of adolescents monitor an online anti-bullying forum; In-class anti-bullying activities	461 9th grade students from 7 high schools in province of Lucca during the school year 2012/2013). 10 classes from 4 schools were assigned to the intervention condition (n = 234; mean age = 15.6; 28.6% male). Students from 10 classes in 3 schools acted as controls (n = 227; mean age = 15.57; 76.2% male).	All participants completed the Florence Bullying-Victimisation scales at pre- and post-test. Scale measures the frequency of bullying perpetration and victimisation experienced by respondents during the past 2 months.
Pryce & Frederickson (2013) U.K.	*Anti-bullying Pledge Scheme*; Local anti-bullying initiatives implemented in UK schools; Each school assigned an intervention facilitator; Whole-school intervention is tailored to each schools' specific needs	338 students from Years 4,5, and 6 classrooms in 4 UK primary schools. 160 were female and participants were aged 8 to 11 years old.	Two schools were assigned to the intervention condition and two schools acted as a treatment as usual control group. Pre- and post-data collection was conducted with all participants.
Rawana et al. (2011); Canada	*Strengths in Motion;* Strength-based whole-school anti-bullying intervention; Enhancing individuals' strengths; Designated intervention classroom within experimental school; Room used as: (1) Good Start Centre; (2) Cool Down and Prevention; (3) Good Choices Room; and the site of an ambassador's club.	103 4th – 8th grade students from 2 elementary schools; 50 were allocated to experimental condition (mean age = 11.04; 58% female) and 53 were placed in control condition (mean age = 11.53; 45.5% female)	All participants completed the self-report Safe School Survey, which includes a measure of students' experiences of bullying perpetration and victimisation, at baseline, post-implementation (3 months later), and 8-month follow-up. Schools were allocated to experimental or control.
Sapouna et al. (2010) U.K. & Germany	*FearNot!;* Immersive learning intervention; Virtual-learning; 30-minute sessions for 3 weeks; Bullying scenarios acted out by virtual reality characters; Participants required to select appropriate reactions or responses of character.	942 primary school students from the UK (n = 520) and Germany (n = 422). The mean age of UK participants was 9.36 years and in German schools the mean age was 8.34 years.	Schools with up-to-date computer facilities required to administer the intervention were assigned to the intervention condition, whilst the other schools acted as a control group. Pre- and post-intervention measures were employed with all participants.

Silva et al. (2016) Brazil	*Skill-based intervention;* Behavioural cognitive intervention based on social skills; 8 weekly classes for 50 mins led by clinical psychologists; Groups were mixed by gender and bullying-involvement status; Targeted: civility, making friends, empathy, self-control, emotional expressiveness, assertiveness, interpersonal problem-solving; Activities included role-play, dramatization, positive reinforcement, modelling, feedback, videos and homework assignments.	188 6th grade students from six schools. Mean age in intervention group was 11.28 years and the mean age in the control group was 11.21 years.	18 classrooms were randomly assigned to intervention (n = 9 classes) and comparison (n = 9 classes) groups. All participants completed a self-report measure of aggression and peer victimisation before and after intervention.
Sismani et al. (2014); Cyprus	*Daphne III;* International anti-bullying initiative; Educate 5th and 6th grade primary school children about bullying and its many forms; 11 workshops following a structured curriculum manual.	188 5th and 6th grade students from Cypriote primary schools.	All students completed the OBVQ pre- and post-intervention. Students were allocated to either the intervention group or control group.
Solomontos-Kountouri et al. (2016) Cyprus	*ViSC;* Training programme led by professionals to increase students' sense of responsibility and competency in conflict; 13 structured lessons; Covered topics such as: impulsivity; reflecting on behaviour; and acting in a socially responsible manner.	1,652 students from 82 classes in 6 schools. Mean age was 12.6 years old and 48.9% of the sample were female.	30 classes (n = 602 students) of 7th grade and 8th grade students were allocated to the intervention condition, and 52 classes (n = 1,050 students) were allocated to the control condition. Self-report measures of bullying perpetration and bullying victimisation were collected at three time-points, before and after implementation, and follow-up.
Sutherland (2010) Canada	*Beyond the Hurt;* Peer-led anti-bullying programme; High-school programme involving four key components: (1) training of peer facilitators, (2) in-class presentations, (3) teacher workshops, (4) and online training materials for teachers & parents	621 high school students in Canada. 47% were male and 93% reported being Caucasian.	Schools were allocated to the intervention or waitlist control condition and bullying measures were conducted pre- and post-implementation in both groups.

Toner (2010) US	*Bully-Proofing Your School;* Whole-school programme; Individual support also provided for bullies and victims; Restorative non-punitive disciplinary policies; Classroom curriculum implemented by teachers; Parent information	149 6th grade students from 2 suburban public elementary schools. School S – implemented BPYS (*n* = 58) and School U – control (*n* = 91). 63.8% of participants were female and 62.4% were White.	Participants in experimental and control schools completed a self-report measure of direct and indirect bullying perpetration and victimisation, pre- and post-implementation.
Williams et al. (2015) US	*Start Strong;* School-based teen dating-violence prevention programme; Bullying included as secondary violence outcome.	1,517 students from 8 middle schools. Sample was ethnically diverse with 23% identifying as White; 28% African-American; and 33% Latino.	Matched school pairs were created with one school from each pair being allocated to the intervention condition. The remaining schools formed the control group. Data collected pre- and post-intervention.
Wong et al. (2011) Hong Kong	*Restorative Whole-School Approach;* Whole-school anti-bullying programme based on restorative justice principles; Whole-school non-punitive anti-bullying policy and ethos implemented; Curriculum lessons target: empathy, assertiveness, coping, problem-solving and conflict resolution.	1,480 high school students from 4 middle band (based on academic ratings) schools in Hong Kong. Students were aged 12 to 14 years old.	Three experimental groups were utilised: (1) Intervention group; (2) Partial intervention group; and (3) Control group. All participants completed pre- and post-measures of bullying.
Yaakub et al. (2010) Malaysia	*OBPP*; Whole-school programme, also included individual-, class-, and community-level factors; School conference held at beginning of programme; Detailed teacher handbook; Parent/Teacher meetings; Class anti-bullying rules.	3,816 students from 6 secondary schools in Malaysia.	Three schools were assigned to the intervention condition, and the remaining three acted as a control group. Participants from both groups completed bullying measures pre- and post-intervention.

Age Cohort Designs (n = 6 evaluations)			
Project	***Anti-bullying Programme; Key Features***	***Participants***	***Research Design***
Busch et al. (2013) Netherlands	*Utrecht Healthy Schools;* Whole-school health programme; Implement a healthy-school policy; Ensure healthy food options, smoke- and alcohol-free sites and appropriate sports facilities; Parent workshops and take-home tasks; Involve public health services.	336 4th grade students aged 15 to 16 years old.	Fourth grade students before the 3-year intervention were compared with fourth grade students after the implementation.
Kärnä et al. (2011a) Finland	*KiVa*; Whole-school programme that also targeted individual cases of bullying within a school; Structured curriculum involving class and parent-involved activities; Anti-bullying computer programme for students; Training for teachers on classroom and bullying hotspot supervision/management.	Approximately 200,000 students in 888 Finnish schools. 156,634 and 156,629 students comprised the control groups for victimisation and perpetration respectively. 141,103 and 141,099 students comprised the intervention groups for victimisation and perpetration respectively.	Cohort-longitudinal design with adjacent cohorts. All participants completed the Revised Olweus Bully/Victim Questionnaire.
Limber et al. (2017) US	*Olweus Bullying Prevention Programme;* 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	70,998 students from 210 schools in grades 3 to 11.	Extended age cohort design. All students completed the self-report OBVQ measure of bullying perpetration and victimisation.
Olweus; New National Cohorts 1 to 6 Norway	*OBPP;* 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	Six cohorts from a national implementation of the OBPP.	Extended selection cohorts design; Testing began in October 2001, and subsequent measurements at half-year intervals.
Purugulla (2011) US	*OBPP;* School level (e.g., Staff discussion groups; Bullying Prevention Coordinating Committee); Classroom level (e.g., classroom rules); individual level	785 7th grade (n = 399) and 8th grade (n = 386) students in year one of evaluation and 847 7th grade (n = 417) and 8th grade (n = 410) students from one middle school.	Age cohort design, with year one students acting as control for experimental year two students. All participants completed OBVQ measure of

	(e.g., supervision of students); and community level components		bullying and bullying-related discipline records were also obtained.
Roland et al. (2010) Norway	*Zero Programme;* Preventive programme; Emphasis on school staff to ensure a zero tolerance to bullying; Discussion groups about bullying occur in classes; Restorative conflict resolution meetings take place between victims, teachers, parents and then, perpetrators.	20,446 students in Years 2 to 7 from 146 Norwegian schools.	Age equivalent design; Surveys were administered in Spring 2001 and 2004.

5. Data Extraction: School-bullying

5.1 Overview

After identifying studies eligible for inclusion in the present systematic and meta-analytical review detailed information about the anti-bullying programmes, sample involved, and evaluation design were extracted from primary studies. The following chapter outlines the coding framework applied in greater detail.

Table 9 outlines each piece of information extracted. Information was extracted from primary studies under four main headings: (1) Descriptive; (2) Design; (3) Programme; and (4) Outcomes. Under the 'Programme' heading, information relating to the specific intervention components included in primary evaluations is outlined. The results of this data extraction process are included in Table 8 (see Chapter 4, section 4.5) for data on the descriptive, design, and programme level. Raw outcome data extracted from each of the 100 evaluations could not be included due to restrictions on the length of this dissertation.

Additionally, the following section outlines information extracted from primary studies in order to create a risk of bias index. The items utilised to assess risk of bias for each of the methodological designs included in the present report are also outlined.

The data extraction procedure was carried out in consultation with doctoral supervisors, Dr. Ttofi and Professor Farrington. There were a number of studies from the previous Campbell Collaboration report (i.e., Farrington & Ttofi, 2009) for which full texts were unavailable and thus, were excluded from several of the moderator analyses for school-bullying outcomes.

Table 9

Data extraction codebook for the school-bullying meta-analysis

Type	Information extracted	Example
Descriptive	• Sample size	• Total N; *n* experimental; *n* control
	• Age of sample in years	• Mean age/range
	• Grade(s) of sample or range	
	• Sex: % female and % male	
	• Location or country	
	• Publication Year	• 2009 versus 2016
	• Publication Type	• Journal article, book chapter, dissertation, report
Design	• Evaluation method	• RCT; BA/EC; or Age cohort design
	• Measures	• Name of instrument
		• Timeframe
		• Perpetration/ victimisation/ both
		• Type of report
	• Data collection timepoints	• Baseline/Post-intervention/Follow-up
	• Unit of allocation/randomization	
	• N clusters	
	• Matched-groups	
Programme	• Name of programme	• e.g., OBPP or KiVa
	• Intervention length	
	• Core components	• Peer, parent, and teacher involvement
		• Involvement of external stakeholders
		• Intervention activities
		• Curriculum/structure/non-structured
	• Intervention aim and/or target	
	• N workshops	• High, low, possible risk
	• Conflict of Interest	• High, low, medium specificity
	• Specificity	
Outcomes	• Bullying at baseline for exp and control	• Mean, SD, N
		• N and % bullies and/or victims
	• Bullying post-intervention for exp and control	
	• Independent samples	
	• Type of outcome	

	• Multiple measures

Note. N = total sample; *n* = number of participants in groups; RCT = randomised controlled trial; BA/EC = quasi-experiments with before and after measures of bullying (non-randomised); OBPP = Olweus Bullying Prevention Programme; SD = standard deviation; exp = experimental

5.2 Descriptive information

Various pieces of descriptive information were extracted from each of the 100 evaluations included in the present research. Information specific to the evaluation, such as the location or the start/end date, were recorded along with detailed information concerning the sample.

5.2.1 Sample size

The total sample size and also the *n* of the relevant experimental and control groups were recorded. Where reported, the % of females and males included in the evaluation was extracted. Too few primary evaluations reported information regarding the ethnicity, sexuality, or gender identity of participants to be coded in the present review. In order to examine the relationship between sample size and effect size, two variables for sample size were used. Sample size was recorded as both a continuous variable (i.e., total number of participants) and a categorical variable. To create the latter the interquartile range was used.

5.2.2 Age of participants

Age was extracted in three ways as there was inconsistency in how age was reported by included primary evaluations. Studies reported the age of participants either as: (1) a continuous variable representing the mean age of all participants (e.g., mean age = 8.57; Cross et al., 2011); (2) an ordinal variable based on school-grade (e.g., Grades 4 to 6; Karna et al., 2011a); and (3) an ordinal variable based on the range in years of participant ages (e.g., ages 13 to 15 years; Garaigordobil & Martinez-Valderrey, 2015).

For the purpose of analyses therefore, the age of participant variable was transformed into one comparable continuous variable. Therefore, steps were taken to compute an estimate mean age of participants for all primary evaluations. Where age was represented as school grades, online resources[10] from each country were used to establish the ages of students in said grade. Then a mean age for that range was computed. Similarly, where an age range was provided, the minimum and maximum ages were used to estimate a mean value with the assumption that age was evenly distributed in the sample. Two categorical age variables were then created to replicate, but extend, Farrington and Ttofi's (2009) research.

Firstly, a categorical was also created to compare groups of participants based on age. Following previous analyses (i.e., Farrington & Ttofi, 2009), a dichotomous age variable was created. Studies were grouped based on the mean age of participants, and whether this value indicated participants were younger than 11 years old, or older than 11 years old. Thus, the younger category included participants up to age 10.99 years and the older category included participants aged 11.01 years and over. For example, the mean age of participants included in Martin et al. (2005) evaluation was 10.98 years and thus categorised as 'younger'. Similarly, Rawana et al. (2011) included participants with a mean age of 11.04 years old and was included in the 'older' category. This dichotomy reflects the typical age that students leave primary school and enter middle/secondary school (dependent on the location).

Secondly, a more detailed categorical variable was created to reflect age ranges in more detail. The first category included studies that involved participants aged between 4 years and 7 years old (i.e., 'primary one'). The second category included studies that involved participants aged between 8 years and 10 years old (i.e., 'primary two'). The 'middle' category included evaluations conducted with participants aged between 11 and 13 years old and the 'secondary' category included evaluations conducted with participants aged between 14 and 18 years old.

---

[10]For example, in England and Wales: https://www.gov.uk/schools-admissions/school-starting-age; in Canada: https://www.uopeople.edu/blog/understanding-the-canadian-education-system/; or for the USA: https://www.acs-schools.com/egham/admissions/grade-placement#2

5.2.3 Publication type and year

Descriptive information about the publication of the evaluation was also extracted. Specifically, the type of publication and the publication year was recorded. The former represents a categorical moderator to reflect whether or not the evaluation was published via the following channels (in order of hypothesized negative correlation with bias): (1) peer-reviewed journal article; (2) chapter in an edited book/ book; (3) governmental report or similar; (4) correspondence; and (5) unpublished masters or doctoral theses.

Correspondence was included to reflect data obtained from multiple evaluations of the Olweus Bullying Prevention programme sent to Farrington and Ttofi in preparation of their earlier meta-analysis. Where evaluation data had been published in multiple formats, we favoured the category associated with the least potential bias. For example, Domino (2011) reported the results of an evaluation of Take the LEAD programme in a doctoral dissertation, but later published these results in a peer-reviewed journal (i.e., Domino, 2013). In this scenario, the included study was coded as “article”.

5.3 Design level

Included studies were further categorized according to several aspects of the research design used. Information was coded regarding both the measures (i.e., instruments to measure bullying behaviours) and research design.

In relation to measurements of bullying, the following information was recorded: the timeframe (i.e., past 3 months or “ever”) in which participants were asked to report on experiences of bullying; the type of report used (i.e., self-, peer-, or teacher-report); and data collection points (i.e., baseline, post-intervention, 3-month follow-up etc). It was also noted if the measure was a continuous scale or a global item and whether bullying perpetration, victimisation, or both, outcomes were measured.

As for the evaluation design, information regarding the unit of allocation (or unit of randomization for RCTs; see below), the number of 'clusters' included, whether groups were matched at baseline, and the number of experimental or control groups was recorded. For example, Elledge et al. (2010) included multiple control groups: matched controls and non-matched controls.

Information about the evaluation methodology was also extracted from primary reports. The types of evaluation methodologies included in the present report are now described in further detail.

5.3.1 Evaluation methodology

In order to optimize the comparability of effect sizes, primary studies included in a meta-analysis should use the same, or at least conceptually similar, research designs (Wilson, 2010). Following Farrington and Ttofi's (2009) criteria, systematic searches used in the present meta-analysis aimed to identify evaluations using any of the following three[11] research designs:

(1) Randomised controlled trials (RCTs);
(2) Before-After/Quasi-Experimental-Control designs (BA/EC);
(3) Age cohort designs.

These methodologies varied on three key elements: (1) the randomization of participants (or clusters of participants); (2) the use of experimental and control groups; and (3) the administration of quantitative bullying measures before and after intervention. All studies coded as a RCT had to include random assignment to experimental conditions (i.e., intervention and control groups) but did not have to use before and after measures of bullying

[11] Four research designs were included by Farrington and Ttofi (2009), the fourth design being 'other experimental-control' designs. However, as described in Chapter 4 (section 4.4.3) these designs were excluded from the updated meta-analysis due to inadequate methodological rigor.

outcomes and BA/EC studies had to include before and after measures of bullying, but not random assignment.

Randomised controlled trials (RCTs) are considered to be the ‘gold standard’ of experimental evaluations (Weisburd et al., 2001). Random assignment of a large number of units is used as a way in which evaluators can also randomise possible confounding variables between groups. As a result, one can infer that any observed differences result from the experimental manipulation (Farrington, 1983). The assumption is that randomization ensures that both observed and unobserved variables that may impact the results of an evaluation are also randomly distributed between groups. However, problems may arise if the unit-of-allocation/ the unit-of-randomization, and the unit-of-analysis do not align.

Before-After/Quasi-experimental-control (BA/EC) designs, are conceptually similar to RCTs, but they do not involve random assignment to experimental conditions. Instead, participants or clusters of participants may be assigned to the intervention or control group on a self-selected basis (e.g., Menesini et al., 2012), for convenience (e.g., Sapouna et al., 2010), or based on a greater need for intervention (e.g., Losey, 2009). Thus, BA/EC designs may be subject to selection biases (Farrington & Petrosino, 2001) that may reduce the validity of the results. These can be controlled if outcomes are measured before and after the intervention. Studies coded as BA/EC in the present report all used experimental and control groups but did not randomly assign participants to conditions. They also had to measure bullying outcomes before and after implementation of the intervention.

In an age cohort design, students of a particular age X are initially assessed in the first year and serve as the control group for the evaluation of an intervention. Then, all students receive the intervention, and different students of the same age X (in the same school, in the second year) serve as the experimental group (see Kärnä et al., 2013). This design, which is largely used in evaluations of the Olweus Bullying Prevention Programme, deals with some selection effects, since it ensures that experimental and control children are matched on age and school, and it deals with some threats to internal validity (e.g. ageing and maturation). However, this design may be influenced by period and testing effects, and the experimental and control groups may differ on other uncontrolled variables.

5.4 Programme level

Using a socio-ecological systems theory framework (Bronfenbrenner, 1977) and the previous meta-analysis (i.e., Farrington & Ttofi, 2009) as guidelines, information about the specific intervention programme was recorded. General details about the intervention, such as the name of the programme (where relevant) and the aim of the intervention were noted along with more detailed information about the anti-bullying programmes.

5.4.1 Intervention components

Intervention components at multiple levels of the socio-ecological model, specifically: (1) school; (2) classroom; (3) teacher; (4) parent; (5) peer; and (6) the individual student. Intervention components that did not fit with this categorisation were grouped under the label: “intervention-specific”, i.e., they related to the specific intervention materials implemented. For the purpose of the present analysis components at all levels were coded dichotomously, as either being *absent* (0) or *present* (1) in the specified intervention programme. The exceptions to this were the variables relating to the type of programme and the approach to anti-bullying. Further details of our codebook used for intervention component analysis is provided in Table 10.

5.4.1.1 School-level. At the school-level, the presence or absence of a whole-school approach (or universal approach) to anti-bullying and supervision in ‘hot spots’ for bullying was coded. A whole-school approach actively involves all actors within the school environment in anti-bullying activities, and the supervision involved identifying specific areas of the school environment where bullying was more likely to occur and increasing the presence of teachers in these areas. The implementation or use of an anti-bullying policy in intervention programmes was also coded. A typical anti-bullying policy includes clear definitions and examples of what constitutes bullying behaviours and specifies that these behaviours are not accepted, along with evident strategies for dealing with bullying.

5.4.1.2 Classroom-level. At the classroom-level the presence or absence of classroom rules throughout the implantation period of intervention programmes was coded. Similar to the anti-bullying policy intervention component, the classroom rules component refers to interventions where a clearly defined set of rules against bullying were implemented and enforced at the classroom-level. In some studies, these rules were created in conjunction with the participating students. Finally, the inclusion of classroom management techniques in intervention activities was coded. This component describes interventions where a particular focus was placed on teachers identifying and dealing with bullying behaviours in their respective classrooms.

5.4.1.3 Teacher-level. Generally speaking at this level, components refer to the participation of teachers in the anti-bullying programme. However, the degree of teacher involvement varied and this is reflected in the coding of this component. Thus, the TInfo component describes interventions that provided information about the intervention to teachers in participating schools. Information about the intervention could have been provided in the form of intervention packs or short information sessions/ meetings with teachers. Furthermore, the TTrain component refers to whether teachers were trained to specifically facilitate the anti-bullying programme in their respective classrooms or within their respective schools.

5.4.1.4 Parent/Guardian-level. Following the socio-ecological framework of bullying prevention and intervention, parents/guardians are also frequently involved in anti-bullying activities. This may involve take-home letters (e.g., Brown et al., 2011), ‘homework’ lessons on anti-bullying materials to be completed under parental supervision and/or with parental participation, or evening meetings to inform parents about bullying-related issues. As there is some obvious discrepancy in the level of *active* involvement on the part of parents, the present research divided the ‘information for parents’ variable evaluated by Farrington and Ttofi (2009) into the two independent levels.

Firstly, the PInfo level of the parental-involvement component refers to studies that provided parents with information about bullying-related issues or the intervention being evaluated through take-home letters or leaflets. Secondly, the PInvolve component refers to active parent involvement. This dimension of the parental-involvement component refers to programmes where parents were invited to, or attended, meetings held by school staff, or intervention facilitators. During these meetings, bullying and related issues, or the specific intervention programme, was discussed. For example, parents may have been informed about the prevalence of bullying, the associated risk and/or protective factors, or the specific intervention that was being implemented in the respective school. Parents may also have been informed about approaches they may take to prevent, and/or reduce, bullying perpetration or victimisation amongst their own children.

5.4.1.5 Peer-level. In the same way as the parent-level components, the current report added additional levels to peer-related intervention activities in order to explore the effect of peer involvement in more detail. The informal peer involvement component, called ‘Peer1’, refers to the general use of in-class, or group-based, discussion during intervention activities. Discussion is often led by teachers or trained intervention facilitators and occurs between peers. Secondly, a common facet of peer-related components observed in primary studies was the emphasis on engaging bystanders and encouraging of non-involved peers to intervene when they observe bullying situations. Thus the component ‘Peer2’ relates to the absence or presence of encouraging bystanders to prevent bullying, or intervene in bullying situations, throughout intervention activities. Finally, formal peer involvement in intervention activities was coded. Examples of formal peer involvement could include peer-mentoring schemes, peer-led anti-bullying activities, or the training non-involved students to provide active support to participants experiencing bullying (e.g., Palladino et al., 2012; Menesini et al., 2012).

5.4.1.6 Individual-level. This level in the socio-ecological framework refers to factors relating to the individual within the specified population. Intervention components refer to programme elements that relate directly to the students experiencing bullying, either through perpetration or victimisation. The ‘Bull’ component relates to intervention components that involve activities conducted with individual students identified as bullies, and the ‘Vic’ component relates to intervention components that involve activities conducted with individual students identified as victims of bullying. Additionally, the ‘Coop’ element describes the involvement of external professionals in intervention activities. However, this does not include interventions where external partners provided training to teachers, for example. This component only refers to studies in which these external partners worked directly with victims and/or bullies in experimental schools.

5.4.1.7 Intervention-specific. In addition to intervention components at the school, classroom, parent, teacher, peer, and individual levels, there were a number of components coded that are related specifically to the intervention programmes. Based on the previous review and the wider literature, the presence or absence of curriculum materials (‘Curriculum’) and the inclusion of socio-emotional skills (‘SESkills’) or mental health issues (‘CBT/MH’) in intervention programmes was coded. The socio-emotional skills component referred to intervention activities centred around specific social, emotional and psychological concepts, such as empathy, conflict resolution, problem-solving, self-control, decision-making, prosocial or coping skills (e.g., Holen et al., 2013; Trip et al., 2015; Silva et al., 2016). The component 'CBT/MH' refers to the absence or presence of intervention activities that incorporated cognitive-behavioural techniques or strategies and/or mental health issues, such as anxiety or depression (e.g., DeRosier & Marcus, 2005; McLaughlin, 2009; Stallard et al., 2013). In addition, we coded the use of disciplinary measures. This level involved either the presence/absence of punitive disciplinary measures (e.g., formal punitive sanctions for bullying behaviours) or the presence/absence of non-punitive disciplinary measures (e.g., restorative justice or 'No Blame' methods).

Table 10

Codebook for intervention components coded at each level of the socio-ecological framework.

Level	*Component*	*Variable Name*	*Code*	*Description*
School level	**Anti-bullying policy**	ABP	0	Intervention *did not involve* implementation of an anti-bullying policy
			1	Intervention *did involve* implementation of an anti-bullying policy
	Supervision	Sup	0	Intervention *did not incorporate* improving teacher supervision of students
			1	Intervention *did include* incorporate improving teacher supervision of students
	Whole school approach	WSA	0	Intervention *did not employ* a whole school approach to anti-bullying
			1	Intervention *did employ* a whole school approach to anti-bullying
Classroom level	**Classroom management**	CManage	0	Classroom management strategies were *absent*, or study did not refer to teacher's specifically being taught how to identify, manage, or prevent bullying in classroom environment
			1	Classroom management strategies were *present* in intervention, involving instruction on how to identify, manage and prevent bullying in classroom environment
	Class Rules	CRule	0	Intervention *did not involve* development or implementation of a set of rules for behaviour in classrooms
			1	Intervention *did involve* development and implementation a set of rules for behaviour in classrooms
Teacher Level	**Information**	TInfo	0	Teachers *were not involved* in the intervention, or were not provided with information via manuals or intervention packages
			1	Teachers *were involved* in the intervention and were provided with information about the intervention and/or bullying via manuals, letters, or intervention packages
	Training	TTrain	0	Teachers *were not involved* in the intervention, or did not receive formal training

			1	Teachers *were involved* with the intervention received formal training through workshops, seminars, or training days
Parent/ Guardian level	**Information**	PInfo	0	Information about the intervention, beyond parental consent, *was not sent* to parents
			1	Information about the intervention and/or bullying *was* to parents using take-home leaflets or letters
	Active parental involvement	PInvolve	0	Intervention *did not involve* parents, beyond parental consent
			1	Intervention *did involve* parent discussion groups, parent-teacher meetings, or intervention homework to be completed with parent/guardian
Peer level	**Informal**	Peer1	0	Intervention *did not involve* peers
			1	Intervention *did involve* some degree of peer involvement (e.g. through class/group discussions) or vague information is provided
	Bystanders	Peer2	0	Intervention *did not involve* peers
			1	Intervention *did involve* working with peers of bullies and victims in order to encourage bystanders to intervene
	Formal	Peer3	0	Intervention *did not involve* peers
			1	Intervention *did involve* working with peers, e.g. peer-led, peer-support, or peer-mentoring programmes
Individual level	**Work with Bullies**	Bull	0	Intervention *did not include* elements targeted at bullies
			1	Intervention *did include* elements specifically targeted at identified bullies
	Work with Victims	Vic	0	Intervention *did not include* elements targeted at victims of bullying
			1	Intervention *did include* elements specifically targeted at identified victims of bullying
	Cooperative Group Work	Coop	0	Intervention *did not involve* cooperative group work
			1	Intervention *did involve* cooperative group work between school staff, external professionals and individual bullies and/or victims
Intervention-specific	**Curriculum Materials**	Curriculum	0	The use of curriculum materials in intervention was *absent*
			1	The use of curriculum materials in intervention was *present*
		SESkills	0	The intervention *did not include* socio-emotional skills

	Socio-emotional skills		1	The intervention *did include* socio-emotional skills, such as, empathy or problem-solving
	Mental health	CBT/MH	0	Intervention activities *did not include* any issues relating to mental health or cognitive-behavioural techniques
			1	Intervention activities *did include* any issues relating to mental health or cognitive-behavioural techniques
	Disciplinary methods	Punitive	0	The intervention *did not involve* the use of punitive disciplinary methods
			1	The intervention *did involve* the use of punitive disciplinary methods
		Non-punitive	0	The intervention *did not involve* the use of non-punitive disciplinary methods
			1	The intervention *did involve* the use of non-punitive disciplinary methods

5.4.2 Conflict of interest

In addition to specific programme elements included in interventions, data was also extracted in relation to possible sources of bias in evaluations and intervention development. Conflict of interest has previously been reported to impact evaluation results of many interventions and is a growing area of interest (COI; Eisner & Humphreys, 2012) with studies identified as having higher conflict of interest associated with larger overall effect sizes. Eisner and Humphreys outline many other possible sources of COI, such as financial gain to the evaluator, but this information was difficult to obtain for anti-bullying programmes. Thus, a simple indication of potential COI was utilised.

In the context of the current research, COI focused on the overlap between individuals included as author/co-author on the evaluation study, is also included on previous evaluations of the same programme (e.g., NoTrap!; Menesini et al. (2012); Palladino et al. (2012); Palladino et al., 2016) or is in fact referenced as the developer of that particular programme (e.g., Tsiantis et al., 2013). If no reference to a publication relating to the specific programme was included, the conclusion was drawn that the author had developed the programme, and thus, the evaluation was deemed high risk for conflict of interest.

5.4.3 Programme specificity and richness

Programme specificity refers to whether the intervention programme was specifically targeting bullying outcomes, or if other outcomes were also included. Highly specific programmes (i.e., those that only included bullying outcomes and very few others) were coded as ‘high’. Thus, programmes that were less specific and included many other outcomes in addition to bullying measures were labelled ‘low’. A third category was created (i.e., ‘medium’) to include studies that did multiple other outcomes in addition to bullying outcomes, but these additional variables were bullying-related. The number of included components in an intervention programme was also recorded and used to create a 'programme richness' variable. The minimum score possible was zero and the maximum was 15. The total number of intervention components reflected the 'richness' of the intervention, thus, a higher score suggested a richer programme.

5.5 Outcome level

Information also extracted several pieces of statistical information from primary studies that was required for the estimation of effect sizes. Statistics for bullying behaviours, e.g., means and standard deviations or sample sizes and percentage of bullies and/or victims, were extracted for experimental and control groups at baseline and immediately post-intervention timepoints.

Bullying data for additional follow-up timepoints where this information was reported by primary studies was also recorded. Data was extracted and recorded separately for independent samples (i.e., female and male, Palladino et al., 2016; older and younger, Baldry & Farrington, 2004) and different measures. For example, data for both self- and peer-report measures were extracted from Berry and Hunt (2009) and for different forms of bullying (e.g., Frey et al., 2005).

5.6 Risk of bias

As per the Campbell Collaboration reporting guidelines, a risk of bias index was created for the purpose of the present report. The EPOC tool was utilised to assess the risk category of each study on several items relating to the methodological quality of evaluations. Following earlier Campbell reviews (e.g., Valdebenito et al., 2018) this tool was also used for non-randomised studies as other risk of bias measurement instruments were considered inappropriate for non-scientific or medical trials.

Each primary evaluation was measured on the following items: (1) Allocation sequence [AS]; (2) Allocation concealment [AC]; (3) Baseline equivalence on outcomes [BE]; (4) Baseline equivalence on participant characteristics [BC]; (5) Incomplete outcome data [ID]; (6) Contamination protection [CP]; and (7) Selective outcome reporting [SOR]. The applicability of these categories for each of the methodological designs included in the present report is outlined in Table 11. Each study was categorized as being high, low, or unclear (if insufficient information was available) risk on each of these EPOC items. Table 11 provides examples of the application of this tool to included school-based anti-bullying intervention programmes.

Table 11

EPOC risk of bias tool and examples

EPOC item	Design	Risk Category	Criteria
Allocation sequence [AS]	RCTs	Low	Random component in sequence generation process is described (e.g., used a random number table).
		High	A non-random method is used (e.g. date of agreement to participate).
	BA/EC	Low	Matched-pairs design used; Units could not be randomised due to lack of specific intervention-related resources (e.g. computer access) beyond evaluator control
		High	Unmatched design used or unit allocated as a result of specific request due to increased levels, or perceived high levels, of bullying. Units could not be randomised due to failure of schools to agree to participation if in control group/would be randomly assigned to condition.
	AC	Low	No age cohorts were categorized as low-risk, due to the nature of allocation to experimental and control conditions.
		High	All age cohorts were categorized as high risk on this item, due to the non-random nature of allocation.
Allocation concealment [AC]	RCTs	Low	Random allocation was conducted by external body; research team; or prior to screening, or after consenting to participate; Allocation was communicated using sealed envelopes
		High	Random assignment was managed by schools themselves; Randomization occurred after participant screening; Allocation was randomised

			prior to consent to participate, and was communicated to schools in information sheet
	BA/EC	Low	Schools were asked to agree to participation before being allocated to experimental or control condition
		High	Schools were asked to agree to participate *after* being told the experimental condition they were assigned to; Schools specified they would participate on the basis of being allocated to a specific condition.
	AC	Low	No age cohorts were categorized as low-risk, due to the nature of allocation to experimental and control conditions.
		High	All age cohorts were categorized as high risk on this item, due to the non-random nature of allocation.
Baseline equivalence [BE]	ALL	Low	Baseline levels of bullying in experimental and control groups is reported and no significant differences are found; Means and distribution of bullying is similar between experimental and control groups at baseline
		High	Baseline levels of bullying in experimental and control groups is reported and significant differences are found; Means and distribution of bullying are different between experimental and control groups at baseline
Baseline characteristics [BC]	ALL	Low	Balance in participant demographics between experimental and control groups at baseline; Matched pairs of units of allocation
		High	Imbalance in participant demographic between experimental and control groups at baseline; No information of baseline characteristics of participants is reported

Incomplete data [ID]	ALL	Low	Zero attrition is reported; Attrition represents a low percentage of cases; Missingness was equivalent across experimental and control groups; Attrition was reported and an adequate strategy to deal with attrition was applied
		High	High percentage of attrition reported and no strategy to deal with attrition mentioned; List-wise deletion was used to respond to attrition; Attrition impacted the experimental and control groups unequally
Blind outcome assessment [BOA]	ALL	Low	Individuals who were independent of intervention implementation collected outcome data; Individuals collecting data were unaware of experimental condition
		High	Individuals who implemented intervention administered outcome measurement instruments; If individuals collecting data were aware of experimental condition or if observers in observational data were aware of experimental condition
Contamination protection [CP]	ALL	Low	Schools are unit of allocation to intervention or control group; Measures taken to avoid cross-over effects
		High	Classes, or individuals within schools are the unit of allocation to experimental or control group; No measures put in place to avoid cross-over
Selective outcome reporting [SOR]	ALL	Low	Outcomes proposed are outcomes that are reported
		High	Outcomes proposed are not the outcomes that are reported

Note: RCT = randomised controlled trial; BA/EC = Quasi-experimental design with before and after measures of bullying; AC = age cohort designs

6. Results: School-bullying

6.1 Systematic review

Table 8 (see Chapter 4, section 4.5) outlines brief details about each of the evaluations included in the present research that were published after 2009. Information about the intervention programme, sample, and evaluation design is provided. The following sections of this dissertation outline further results of the systematic review of school-based anti-bullying programmes to reduce offline bullying.

A number of moderators and mediators were selected *a priori* for further analysis, under the descriptive label (i.e., location of intervention; publication type; publication year), design label (i.e., evaluation method and unit of allocation/randomization), and the programme heading (i.e., name of intervention, conflict of interest; and programme specificity). Additionally, specific intervention components were coded from primary evaluations.

6.1.1 Descriptive level

As per Chapter 5, several pieces of information regarding descriptive aspects of primary evaluations were coded. The following sections provide the results of the systematic review in relation to these descriptive moderators, namely, the location of the intervention, the age of participants, and both the publication type and year.

6.1.1.1 Location of intervention. Evaluations included in the present analysis were conducted in many different countries around the world. However, there were only a few countries in which multiple evaluations of anti-bullying programmes had been published.

Specifically, in the following countries only one evaluation was included in the present report: Austria (i.e., Yanagida, Strohmeir & Spiel, 2016); Brazil (i.e., Silva et al., 2016); China (i.e., Ju, Shuqiiong, & Wenxin, 2009); Czechoslovakia (modern day Czech Republic and Slovakia; i.e., Rican, Ondrova, & Svatos, 1996); Hong Kong (i.e., Wong et al., 2011); Ireland (i.e., O'Moore & Minton, 2004); Malaysia (i.e., Yaakub, Haron, & Leong, 2010); Romania (i.e., Trip et al., 2015); Sweden (i.e., Kimber Sandell, & Bremberg, 2008); Switzerland (i.e., Alsaker & Valkanover, 2001); South Africa (i.e., Meyer & Lesch, 2000); and Zambia (i.e., Kaljee et al., 2017). If these evaluations were to be included in further moderator analysis, we would be examining the differences based on only one sample and effect size. Therefore, moderator analysis was conducted only between locations in which multiple evaluations of anti-bullying programmes had been conducted.

Of the 100 evaluations included in our meta-analysis of school-based anti-bullying programmes, the majority (79 for perpetration, 82 for victimisation) were conducted in one of 12 different countries. With respect to bullying perpetration outcomes, these countries were as follows: Australia (n = 2); Canada (n = 6); Cyprus (n = 3); Finland (n = 6); Germany (n = 5); Greece (n = 2); Italy (n = 11); Netherlands (n = 3); Norway (n = 8); Spain (n = 3); United Kingdom (n = 4); and US (n = 26). With respect to bullying victimisation outcomes, these countries were as follows: Australia (n = 3); Canada (n = 7); Cyprus (n = 3); Finland (n = 6); Germany (n = 4); Greece (n = 2); Italy (n = 10); Netherlands (n = 3); Norway (n = 7); Spain (n = 3); United Kingdom (n = 6); and US (n = 28).

6.1.1.2 Sample size. The present meta-analysis represents data collected from over 400,000 participants (N = 432,874 for perpetration outcomes and N = 428,057 for victimisation outcomes). Sample size was recorded in two ways in the present research. The total sample size was recorded as a continuous variable, and a categorical variable (see Table 12) was created to reflect the range of sample sizes. The interquartile range was used to create a categorical sample size variable. The quartiles for sample size were as follows for studies (N = 90) reporting bullying perpetration outcomes were: Q1 = 245; Q2 = 699; Q3 = 1,459; Q4 = 297,737. For studies reporting bullying victimisation (N = 93) outcomes, the quartiles were: Q1 = 246; Q2 = 666; Q3 = 1,378; Q4 = 297,737. Therefore, the average value of the corresponding quartiles for both bullying perpetration and victimisation were used (i.e., Q1 = 246; Q2 = 683; Q3 = 1,418; Q4 = 297,737).

To reflect this distribution of sample size amongst our 100 studies, categories around these values were established. A handful of studies were deemed to be outliers in relation to sample size[12] (*n* perpetration = 8 studies; *n* victimisation = 7 studies). Therefore, if they were included in the moderator analysis, they may impact the overall results unduly. In other words, moderator analysis would be confounded by the uppermost category where the outlier studies were incorporated. Thus, the sample size categories were established using 82 studies for bullying perpetration and 86 studies for bullying victimisation. The omission of these large studies also helped to create relatively equal numbers of studies in each subgroup. This is also applicable in analysis of intervention components.

[12] These studies were considered outliers because the *N* was exponentially larger than the other evaluations. This meant that these studies contributed too much weight to the model and would unduly impact the results. This is explained in greater detail in Chapter 12.

Table 12

Categorical variable for sample size; school-bullying meta-analysis

Category	Range of sample sizes	*n studies*	*n studies*
		Bullying perpetration	*Bullying victimisation*
1	$0 \leq N \leq 246$	24	26
2	$247 \leq N \leq 683$	20	20
3	$684 \leq N \leq 1{,}418$	24	23
4	N = 1,419+	15	15

Note. N in the current table is used to denote the total number of participants included in any given evaluation and *n* is used to denote the number of evaluations included in each of these categories.

6.1.1.3 Age of participants. In relation to evaluations that examined reductions in bullying perpetration, the age of participants ranged from 3.79 years old to 16.8 years old. The mean age for studies reporting effects on bullying perpetration outcomes was 11.34 years old. Evaluations that evaluated the impact of anti-bullying programmes on bullying victimisation outcomes included participants between 6 years and 16.8 years old. The mean age for these studies was 11.45 years old.

Overall, there were 38 studies in the ‘younger’ category (i.e., participants were aged 10 years old and younger) and 50 studies in the ‘older’ category (i.e., participants were aged 11 years old and older) for bullying perpetration outcomes. There were 40 studies in the ‘younger’ category and 52 studies in the ‘older’ category, for bullying victimisation outcomes. The majority of studies included participants aged between 8-10 years old (i.e., ‘primary 2’; perpetration n = 31; victimisation n = 38) and between 11-13 years old (i.e., ‘middle’ perpetration n = 35; victimisation n = 39). Only a few studies included younger participants aged between 4-7 years old (i.e., ‘primary 1’; perpetration n = 5; victimisation n = 2) or older participants aged between 14-18 years old (i.e., ‘secondary’; perpetration n = 14; victimisation n = 13).

6.1.1.4 Publication type and year. Overall, the majority of evaluations were published in peer-reviewed journal articles, for both bullying perpetration (*n* = 67) and bullying victimisation (*n* = 72) outcomes. Two evaluations were published in chapters of edited books and both reported effects of a programme on both bullying victimisation and perpetration. No included evaluations were published as entire books. Moreover, 12 unpublished dissertations were identified that published evaluation data for bullying perpetration and bullying victimisation outcomes. Data was also retrieved for both outcomes from three governmental reports. Four of the effect sizes included in the present report were estimated from data emailed to Dr. Ttofi and Professor Farrington in the preparation of the previous meta-analysis (i.e., Olweus/ Bergen 1; Olweus/New National; Olweus/Oslo 1; Olweus/Oslo 2).

Evaluations were categorised according to whether they were included in the previous report (i.e., “2009” studies), or only included in the present report (i.e., “2016” studies). In relation to bullying perpetration outcomes, 37 studies were coded as 2009 studies and 53 studies were coded as 2016 studies. Similarly, more studies were coded as 2016 (*n* = 54) in comparison to 2009 (*n* = 39) studies for bullying victimisation outcomes.

6.1.2 Design level

On the design level, mediators and moderators that were coded included specific information about measurement instruments (e.g., the timeframe for reported bullying behaviours and the type of report), when bullying data was collected in relation to the implementation of the intervention, and the specific evaluation methodology used. Furthermore, information about the unit of allocation was extracted.

6.1.2.1 Measurement instruments. The most common measurement instrument used to measure school-bullying perpetration and victimisation was the Olweus Bully Victim Questionnaire (OBVQ; Olweus, 1986). This questionnaire was used to assess changes in bullying as a result of a wide range of anti-bullying programmes, including the Olweus Bullying Prevention Programme (e.g., Limber et al., 2018). This measure is a self-report measure and participants are asked to indicate the mean frequency of bullying perpetration and victimisation experienced in the past three months. Other measurement tools used more than once included the European Cyberbullying Intervention Project Questionnaire (ECIPQ, Brighi et al., 2012; Herrera-López et al., 2017), the Colorado School Climate Survey (Plog et al., 2006; used by e.g., Beran & Shapiro, 2005; Toner, 2010), the Peer Relations Questionnaire (Rigby, 1996; used by e.g., Domino, 2013; Hunt, 2007; Pryce & Frederickson, 2013), and the Safe School Survey (Totten, Quigley, & Morgan, 2004; used in e.g., Connolly et al., 2015; Rawana et al., 2011). However, no one measurement instrument was used frequently enough, besides the OBVQ, to be able to compare the mean effect sizes for bullying outcomes in relation to the specific instrument used.

Overwhelmingly the majority of included studies used self-report measures to assess bullying perpetration and bullying victimisation. Eight-three evaluations (from a possible maximum of 104[13]) utilised self-report measures. By contrast, only 13 used peer-report measures, two used parent-report and 4 used teacher-report measures of bullying behaviours. In addition, two evaluations used observational data to estimate the prevalence of bullying perpetration and victimisation. Therefore, there was a stark discrepancy between the numbers of studies using other-report measurement instruments compared to those that used self-report measures. Subgroups could not be created to compare studies that used self-report measures (i.e., 'self-report') with studies that did not use self-report measures (i.e., 'other-report') as there would be a clear overlap. In other words, many studies used more than one measure of bullying (e.g., self-report and peer-report: Cross et al., 2011; Fonagy et al., 2009;

[13] This number is higher than the total number of evaluations (i.e., 100) to account for some evaluations using multiple types of measurements.

Kärnä et al., 2011a; 2011b; 2013). The present research prioritised self-report measures where multiple measurement instruments were used, but 9 studies included in the meta-analysis reported outcomes using only ‘other-report’ tools (e.g., observations: Frey et al., 2005; Krueger, 2010; peer-report: Fox & Boulton, 2003; Knowler & Frederickson, 2013; Menesini et al., 2003; Salmivalli et al., 2005; and teacher-report: Holen et al., 2013; Ostrov et al., 2015; Waasdorp et al., 2012). Thus, in the case of these studies, data collected with ‘other-report’ measures were used to estimate effect sizes. Given the unequal numbers, subgroup analyses were not conducted, but comparing effect sizes computed using self-report and other-report measures is an important avenue for future research.

The timeframe within which participants were asked to indicate the occurrence or frequency of bullying behaviours and/or experiences was also recorded. Twenty-two studies asked participants to report bullying in the ‘past couple of months’ (e.g., Busch et al., 2013; Kärnä et al., 2011b; Losey, 2009; Nocentini & Menesini, 2016; Wölfer & Scheithauer, 2014). A total of eight studies used the past month as a timeframe and ten studies used 2-3 months. A few studies (n = 6) used measurement instruments that required participants to indicate the frequency or prevalence of bullying experiences in the past/current academic year (e.g., Battey, 2009; Brown et al., 2011; Ertesvag & Vaaland, 2007; Salmivalli et al., 2005). Overall, 36 studies did not provide participants with a specific timeframe or the time period in which they were asked to respond was unclear. The variation in the timeframe used and the unequal numbers in potential subgroups meant that subgroups analysis analogous to a one-way ANOVA was not conducted for the timeframe variable.

The total number of items used to measure bullying perpetration and victimisation was also extracted from primary studies, where this information was recorded. Of studies that reported this information, the minimum number of items was one for both bullying perpetration and victimisation. This reflects studies that used global items (i.e., Have you ever/How often do you experience bullying victimisation; or Have you ever/How often do you bully others). The maximum number of items was 20 for bullying perpetration (Menesini et al., 2012) and 24 for bullying victimisation (Yaakub et al., 2010). The mean number of items used was 7.41 for bullying perpetration (n = 67) and 7.44 for bullying victimisation (n = 62).

6.1.2.2 Evaluation methodology. The primary moderator chosen for further analysis was the evaluation method; specifically, whether the evaluation was conducted using a randomised controlled trial (RCT), a quasi-experimental design with before and after measures (BA/EC) or an age cohort (AC) design.

Overall, in relation to bullying perpetration outcomes, 36 evaluations used RCT designs, 31 used BA/EC designs and 14 used age cohort designs. However, due to some evaluations reporting data for multiple independent samples, a total of 40 effect sizes were estimated for bullying perpetration outcomes from RCT designs. A further 36 were estimated from BA/EC designs and 14 effect sizes came from evaluations using age cohort designs.

For bullying victimisation outcomes, overall, 33 evaluations used RCT designs that gave 37 independent effect sizes for bullying victimisation and 37 evaluations used BA/EC designs and yielded 42 independent effect sizes. Similar to perpetration outcomes, 14 evaluations used age cohort designs to evaluate the effect of anti-bullying programmes on bullying victimisation outcomes and yielded 14 independent effect sizes.

6.1.2.3 Unit of allocation/randomization. Systematic review findings showed that the one consistent issue arising in included intervention programmes was that the unit of allocation of participants, or clusters of participants, was different from the unit of analysis in most evaluations. Age cohort designs were omitted from this moderator analysis, as the unit of allocation was largely unclear due to the logistics of this experimental design.

The majority of RCT and BA/EC evaluations assigned schools to experimental conditions (perpetration n = 44; victimisation n = 47) yet the unit of analysis was individual students. A number of evaluations (perpetration n = 19; victimisation n = 15) assigned classes to experimental conditions yet the unit of analysis was individual students. Less than 10 evaluations (perpetration n = 7; victimisation n = 9) assigned students to experimental and control conditions. One study randomly assigned districts to experimental conditions, and information was not available for five studies in relation to bullying perpetration and four studies in relation to bullying victimisation.

6.1.3 Programme level[14]: Intervention components

In relation to the specific intervention programmes, a detailed codebook was created to extract information about the intervention components utilised in programmes as described in Chapter 5. The following sections provide systematic review results of the data extraction process in relation to intervention components.

6.1.3.1 School-level. At the school-level, intervention components referred to the inclusion of the whole-school approach, increased supervision, and anti-bullying policies in primary intervention programmes. Intervention programmes that aimed to reduce school-bullying perpetration and school-bullying victimisation were coded separately. In relation to school-bullying perpetration, 43 programmes included a whole-school approach and 39 programmes did not. Additionally, the majority of programmes (n = 61 and n = 57 component absent; n = 21 and n = 25 component present) did not include increased supervision techniques or specifically implement an anti-bullying policy, respectively, in interventions to reduce bullying perpetration. In relation to school-bullying victimisation, 42 programmes implemented a whole-school approach and 44 programmes did not. The

[14] The intervention components analysis was conducted using 82 independent effect sizes for school-bullying perpetration and 86 independent effect sizes for school-bullying victimisation. The justification for omitting effect sizes from over-powered evaluations for perpetration and victimisation outcomes respectively is further outlined in Chapter 12, sections 12.2.2.1 – 12.2.2.3.

majority of programmes (*n* = 65 and *n* = 60 component absent; *n* = 21 and *n* = 26 component present) did not include increased supervision or implement a specific anti-bullying policy, respectively, to reduce school-bullying victimisation.

6.1.3.2 Classroom-level. At the classroom-level intervention components referred to the creation and implementation of classroom rules against bullying and also the inclusion of effective and improved classroom management techniques for teachers. In relation to bullying perpetration, the majority of programmes did not incorporate classroom rules (*n* = 51 absent; *n* = 31 present) or classroom management (*n* = 60 absent; *n* = 22 present) components. Similarly, in relation to school-bullying victimisation, the majority of programmes did not include classroom rules (*n* = 56 absent; *n* = 30 present) or classroom management (*n* = 64 absent; *n* = 22 present) components.

6.1.3.3 Teacher-level. Components on the teacher-level of the socio-ecological framework applied to intervention component coding related to the provision of information (e.g., manuals, handouts, lesson guidelines) for teachers in intervention schools and also the inclusion of formal teacher training sessions. The majority of intervention programmes included these components in relation to the reduction of both bullying perpetration and victimisation. Specifically, 63 programmes and 70 programmes included information for teachers in efforts to reduce perpetration and victimisation respectively (*n* = 16 absent for both outcomes). Moreover, 51 programmes included teacher training components to reduce bullying perpetration (*n* = 31 absent) and 55 programmes included teacher training to reduce bullying victimisation (*n* = 31 absent).

6.1.3.4 Parent-level. Similar to the teacher-level, intervention components at the parent-level referred to the provision of information about the programme or bullying to parents and also the formal involvement of parents. Programmes that included information for parents (*n* = 35 and 36 present) were slightly less common than programmes that did not provide information (such as leaflets or letters) for parents for bullying perpetration and victimisation outcomes respectively. More intervention programmes did not provide information for parents for perpetration (*n* = 47 absent) and victimisation (*n* = 50 absent). The differences were greater for parental involvement. The majority of intervention programmes did not include formal involvement of parents to reduce bullying perpetration (*n* = 61 absent; *n* = 21 present) or victimisation (*n* = 61 absent; *n* = 24 present).

6.1.3.5 Peer-level. Three intervention components were coded at the peer-level: informal peer involvement; encouraging bystanders; and formal peer involvement. The majority of studies included informal peer involvement for both school-bullying perpetration (*n* = 57 present; *n* = 25 absent) and victimisation (*n* = 55 present; *n* = 31 absent). However, fewer studies included intervention components encouraging bystanders (*n* = 25 perpetration; *n* = 25 victimisation) and formal peer involvement (*n* = 13 perpetration; *n* = 15 victimisation). The majority of studies for both perpetration and victimisation outcomes did not include either encouraging bystanders (*n* = 57 perpetration; *n* = 62 victimisation) or formal peer involvement (*n* = 69 perpetration; *n* = 71 victimisation) components.

6.1.3.6 Individual-level. At the individual-level, components were coded in relation to intervention activities with students identified as bullies and students identified as victims. This level of the socio-ecological framework also included the co-operative group work element. This intervention component refers to the involvement of professionals (e.g., school counsellors and psychologists) in anti-bullying activities. The majority of interventions did not include work with individual bullies for either bullying perpetration (*n* = 55 absent; *n* = 27 present) or bullying victimisation (*n* = 58 absent; *n* = 28 present) outcomes. Similarly, fewer interventions included work with individual victims for both bullying perpetration (*n* = 51 absent; *n* = 31 present) and victimisation (*n* = 50 absent; *n* = 36 present) outcomes. Overall, 45 programmes that aimed to reduce bullying perpetration included co-operative group work components (*n* = 37 absent) and equal numbers of programmes that aimed to reduce bullying victimisation included, or did not include, co-operative group work components (*n* = 39 present and absent).

6.1.3.7 Intervention-specific. In relation to more specific elements of the intervention programmes, a number of components were coded that did not fit neatly into other levels of the socio-ecological framework. For example, the majority of studies included curriculum materials in efforts to reduce bullying perpetration (*n* = 69 present; *n* = 13 absent) and victimisation (*n* = 71 present; *n* = 15 present). Twenty-seven programmes included components relating to the use of socio-emotional skills to reduce bullying perpetration (*n* = 55 absent) and 30 programmes included this component to reduce bullying victimisation (*n* = 56 absent). Only a handful of programmes included mental health components to reduce bullying perpetration (*n* = 8 present; *n* = 77 absent) and bullying victimisation (*n* = 8 present; *n* = 78 absent). Few programmes included disciplinary methods to reduce bullying perpetration outcomes, either punitive (*n* = 16 present; *n* = 66 absent) or non-punitive (*n* = 11 present; *n* = 71 absent) approaches. Similarly, few programmes used punitive (*n* = 14 present; *n* = 72 absent) or non-punitive (*n* = 11 present; *n* = 75 absent) disciplinary methods to reduce bullying victimisation.

6.1.4 Programme level: Other

Also, at the programme level, information about any potential conflict of interest and the specificity of the intervention programme was coded.

6.1.4.1 Packaged intervention programmes. Very few specific anti-bullying programmes have been implemented and evaluated more than once using independent samples. Sixty-five different school-based bullying intervention and prevention programmes were included in our meta-analysis, but only eight were repeatedly evaluated. Moderator analysis with respect to the specific intervention programme, therefore, focused on programmes that had been repeatedly evaluated.

In relation to reducing bullying perpetration outcomes the intervention programmes included in the moderator analysis were: Bully Proofing Your School (*n* =3; e.g., Menard & Grotpeter, 2014); fairplayer.manual (*n* = 2; e.g., Bull et al., 2009); KiVa (*n* = 6; e.g., Kärnä et al., 2011b); NoTrap! (*n* = 4; e.g., Menesini et al., 2012); Second Step (*n* = 3; e.g., Espelage et al., 2015); Steps to Respect (*n* = 2; e.g., Frey et al., 2005); ViSC (*n* = 5; e.g., Yanagida et al., 2016).

Similarly, these interventions were included in our moderator analysis in relation to bullying victimisation, with the exception of the fairplayer.manual programme. This intervention was evaluated only twice in relation to bullying perpetration.

Additionally, multiple evaluations of the Olweus Bullying Prevention programme (i.e., OBPP) were included in our meta-analysis. Overall, 12 independent evaluations of this intervention were included in our analysis in relation to bullying perpetration and victimisation outcomes. These are included in our moderator analysis as a collective subgroup and also as further subgroups. Evaluations of the OBPP conducted in the US (perpetration *n* = 6; victimisation *n* = 7) and those conducted in Norway (perpetration *n* = 5; victimisation *n* = 5) were included in the moderator analysis separately. There was one evaluation of the OBPP conducted in Malaysia that is included in the overall category for this programme in relation to bullying perpetration (*n* = 12).

6.1.4.2 Conflict of interest. In the present research, 40 studies, for both bullying perpetration and victimisation, were categorised as high conflict of interest because either the programme developer was the evaluator of the programme or was included as an author on the publication of the evaluation results. A large number of studies (perpetration *n* = 36; victimisation *n* = 39) were considered low conflict of interest, and 14 were categorized as possible conflict of interest. Information concerning conflict of interest was unavailable for 4 evaluations in relation to bullying perpetration outcomes.

6.1.4.3 Programme specificity and richness. Overall, a small number (*n* = 11) of studies included in our analysis were coded as 'low' on the programme specificity variable. The vast majority of evaluations were considered highly specific (i.e., were mostly concerned with only bullying behavioural outcomes; *n* = 59). Additionally, 18 studies were categorised as medium in relation to specificity, where extra outcome variables were measured but these variables were related to bullying (e.g., school climate).

A final continuous variable was created to reflect the total number of intervention components included in primary intervention programmes, with a minimum score of zero, and a maximum score of 15. The mean programme richness variable was 8.28 components (range = 1 – 15 components) for perpetration outcomes (*n* = 90) and 8.17 components (range = 1 – 13 components) for victimisation outcomes (*n* = 93).

6.1.5 Risk of bias

Figure 3 presents the results of the risk of bias analysis for each of the items on the EPOC tool and the additional items we included. The following section describes each of these categories in more detail and examples of high- and low-risk studies are included in Table 11 (p. 105). The main limitation in assessing risk of bias was the lack of information reported by primary studies. Thus, while the best possible efforts were made to categorise each primary evaluation as being high or low risk, a large number of studies were recorded as 'unclear' risk.

As seen in Figure 3, the fewest studies were considered unclear risk on contamination protection and selected outcome reporting. Furthermore, a large number of studies were considered low risk on these items.

For the purpose of analysis, the categories high, unclear, and low risk were transformed into scores of 3, 2, and 0 respectively, as per previous research (Valdebenito et al., 2018). A continuous risk of bias variable was then estimated as the sum total of scores on each of the EPOC items as per previous meta-analyses (Valdebenito et al., 2018). It should be noted that using a sum score for risk of bias may influence the results. It is possible that individual risk of bias items may be related to effect size independently. The lowest possible score a study could be given was zero and the maximum score was 24.

Descriptive statistical analysis showed that risk of bias scores ranged from 0 – 17, with a mean score of 9.62. Meta-regression analysis was conducted to assess the relationship between risk of bias and effect sizes. The result of this analysis is included in section 6.2.4 of this report. The following sections provide more detail about each of the risk categories.

6.1.5.1 Allocation sequence. Allocation sequence refers to the way in which participants, or clusters of participants, were assigned to experimental conditions. For example, low-risk studies were those where a random number generator or another randomization software was used. In total, 30 studies were categorised as high risk on the allocation sequence item. Moreover, 29 studies were low risk and 32 were unclear risk.

6.1.5.2 Allocation concealment. Allocation concealment item refers to whether the method of allocation was concealed from participants or not. In total, 36 studies were categorised as high risk on the allocation concealment item. A further 19 studies were considered low risk, and 34 were unclear risk.

6.1.5.3 Baseline equivalence: Outcome. Baseline equivalence refers to the comparability of experimental and control participants before the intervention has taken place. This item specifically refers to equivalence on relevant outcomes, in this case, school bullying perpetration and victimisation. When experimental and control participants are not statistically significantly different at baseline then we can be more certain that any later differences are a result of the intervention. Overall, 14 studies were categorised as high risk on the baseline equivalence on bullying outcomes. A total of 54 studies were low risk and 21 were unclear risk.

6.1.5.4 Baseline equivalence: Characteristics. Similarly, baseline equivalence on participant characteristics increases the chance that any later difference is a result of the intervention, and not a reflection of different participant characteristics at baseline. Overall, 15 studies were categorised as high risk on the baseline equivalence in participant characteristics item, 64 studies were low risk, and 11 were unclear risk.

6.1.5.5 Incomplete outcome data. Included evaluations were required to incorporate pre- and post-intervention measures of bullying (except if randomization was used). However, because of this, it is likely that there will be some attrition in primary studies. The incomplete outcome data item referred to the risk associated with differential attrition between experimental and control groups and/or ways in which attrition and missing cases were dealt with by primary studies. Twelve studies were categorised as high risk on the incomplete outcome data item. Additionally, 48 studies were low risk and 29 were unclear risk.

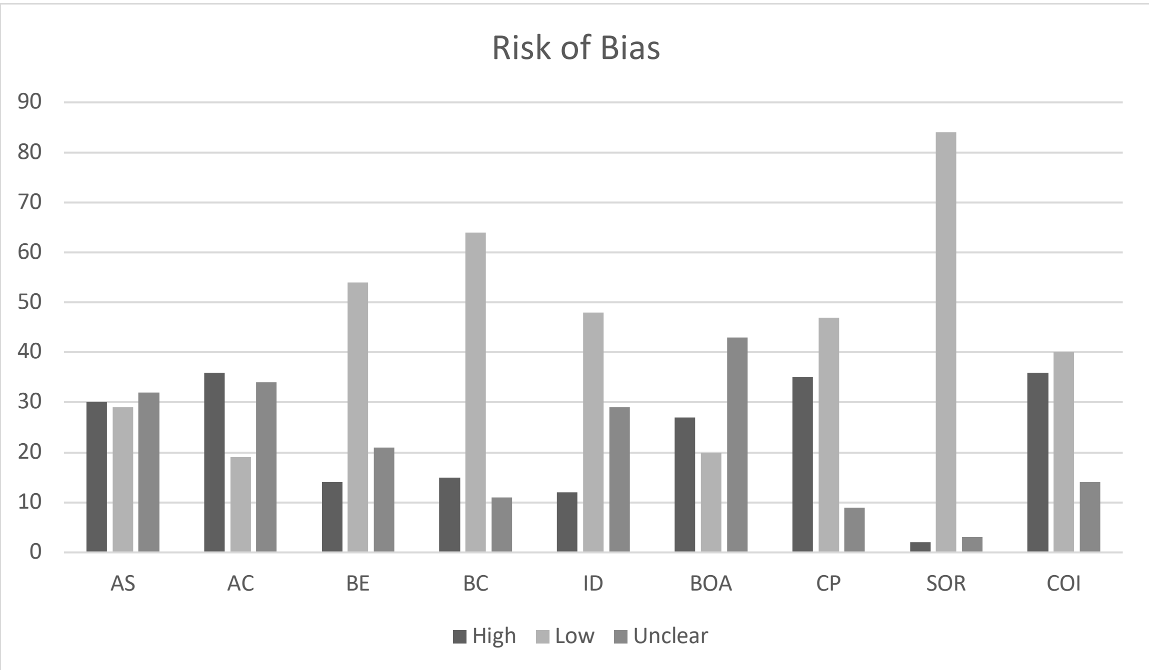
6.1.5.6 Blind outcome assessment. This item assesses the risk associated with any bias which may arise if outcome measurements are not conducted blindly: in other words, whether the individual, or individuals, who administer and collect the measurement instruments are aware of the experimental conditions of participants at the time of measurement. Overall, 27 studies were categorised as high risk on the blind outcome assessment item. 20 studies were low risk and 43 were unclear risk.

6.1.5.7 Contamination protection. Risk of contamination occurs when there is a possibility that experimental and control participants may interact or encounter one another during the course of the evaluation. Thus, the effects of the intervention may 'spill over' to control students and impact the results of the evaluation. In the current analysis, 35 studies were categorised as high risk on the contamination protection item, 47 studies were low risk, and 9 were unclear risk.

6.1.5.8 Selective outcome reporting. Selective outcome reporting occurs when the outcomes reported in an evaluation study differ from the outcomes of interest proposed originally; for example, if a trial protocol proposed different outcomes than those actually reported in the publication of the trial results. Two studies were categorised as high risk on the selective outcome reporting item. 84 studies were low risk, and 3 were unclear risk.

Figure 3

Risk of bias analysis results for school-bullying outcomes



Note: AS = allocation sequence; AC = allocation concealment; BE = baseline equivalence on outcomes; BC = baseline equivalence on participant characteristics; ID = incomplete outcome data; BOA = blind outcome assessment; CP = contamination protection; SOR = selected outcome reporting; COI = conflict of interest.

6.2 Meta-analysis

This section outlines the results of the meta-analysis to evaluate the effectiveness of school-based intervention and prevention programmes to reduce school-bullying perpetration and/or victimisation. In total, 100 primary studies were included in the meta-analysis from which 103 independent effect sizes were estimated. Meta-analysis results are presented using two models of meta-analysis, the random effects model and the MVA model, as outlined in Chapter 2. Analyses were conducted using Comprehensive Meta-Analysis™ software.

6.2.1 School-bullying perpetration

Overall, the results show that anti-bullying programmes significantly reduced bullying perpetration under both computational models of meta-analysis. The effect sizes for each evaluation are presented in Table 13. The mean summary effect sizes were similar under both the multivariance adjustment model (MVA OR = 1.324; 95% CI 1.27 – 1.38; $z = 13.4$; $p < .001$; $I^2 = 81.42$) and the random effects model (RE OR = 1.309; 95% CI: 1.24 – 1.38; $z = 9.88$; $p < 0.001$; $tau^2 = 0.044$).

This result indicates that participants in primary studies who received an anti-bullying intervention were less likely to bully others after completing the programme in comparison to control students who did not partake in the programme.

Analysis of the funnel plot (Figure 4) suggests that publication bias is not likely to be a threat to findings, as studies are symmetrically distributed around the mean effect size. In addition, point estimates did not vary using Duval and Tweedie's trim and fill procedure under a random effects model (in both cases: OR = 1.308; 95% CI 1.240 – 1.380). Based on these results, it was reasonable to assume that publication bias was not likely.

Table 13

Meta-analysis results: School-bullying perpetration outcomes

Study	*OR*	*CI*	*z*	*p*
Randomised Controlled Trials (*n* = 36 evaluations; 40 effect sizes)				
Baldry & Farrington (2004); Older	2.237	0.940 – 5.327	1.820	0.069
Baldry & Farrington (2004); Younger	0.495	0.203 – 1.207	-1.546	0.122
Beran & Shapiro (2005)	1.234	0.571 - 2.669	0.535	0.593
Boulton & Flemington (1996)	0.871	0.443 - 1.712	-0.400	0.689
Brown et al. (2011)	1.192	1.034 – 1.375	2.425	0.015
Chaux et al. (2016)	1.620	1.123 – 2.336	2.583	0.010
Cissner & Ayoub (2014)	0.793	0.459 – 1.370	-0.832	0.406
Cross et al. (2011)	0.803	0.552 – 1.168	-1.147	0.252
DeRosier & Marcus (2005)	1.208	0.769 – 1.897	0.819	0.413
Domino (2013)	3.417	2.167 – 5.390	5.286	< .001
Espelage et al. (2015); Illinois	1.108	0.823 – 1.493	0.678	0.498
Espelage et al. (2015); Kansas	1.052	1.093 – 1.274	4.245	0.000
Fekkes et al. (2006)	1.105	0.620 – 1.970	0.339	0.735
Fekkes et al. (2016)	2.514	1.264 – 5.003	2.627	0.009
Fonagy et al. (2009)	1.248	0.946 – 1.646	1.564	0.118
Frey et al. (2005)	1.058	0.813 – 1.376	0.419	0.675
Garaigordobil & Martinez-Valderrey (2015)	4.828	2.440 – 9.554	4.521	< .001
Holen et al. (2013)	2.127	1.688 – 2.679	6.400	< .001
Hunt (2007)	1.431	0.876 – 2.337	1.431	0.152
Jenson et al. (2013)	1.099	0.551 – 2.190	0.267	0.789
Kaljee et al. (2017)	0.592	0.496 – 0.707	-5.780	< .001
Kärnä et al. (2011b); Grades 4 – 6	1.101	1.000 – 1.212	1.963	0.050
Kärnä et al. (2013); Grades 2 – 3	1.165	1.021 – 1.328	2.270	0.023
Kärnä et al. (2013); Grades 8 – 9	1.075	0.987 – 1.171	1.667	0.096
Krueger (2010)	2.423	0.621 – 9.456	1.274	0.203
Li et al. (2011)	2.221	1.350 – 3.654	3.142	0.002
McLaughlin (2009)	0.845	0.262 – 2.721	-0.283	0.777

Meyer & Lesch (2000)	0.880	0.432 – 1.793	-0.351	0.726
Nocentini & Menesini (2016); Middle	1.562	1.184 – 2.062	3.154	0.002
Nocentini & Menesini (2016); Primary	1.332	1.009 – 1.757	2.026	0.043
Ostrov et al. (2015)	2.049	1.030 – 4.077	2.044	0.041
Polanin (2015)	1.543	0.448 – 5.316	0.687	0.492
Rosenbluth et al. (2004)	1.001	0.652 – 1.538	0.005	0.996
Sprober et al. (2006)	0.654	0.285 – 1.499	-1.004	0.315
Stallard et al. (2013)	1.057	0.774 – 1.443	0.346	0.729
Trip et al. (2015)	1.243	0.868 – 1.780	1.188	0.235
Tsiantis et al. (2013)	1.914	0.570 – 6.425	1.050	0.294
Waasdorp et al. (2012)	1.282	1.173 – 1.401	5.480	< .001
Wölfer & Scheithauer (2014)	0.790	0.479 – 1.304	-0.922	0.357
Yanagida et al. (2016)	1.399	0.699 – 2.798	0.949	0.343
Random Effects: RCTs	*1.240*	*1.118 – 1.375*	*4.069*	*< .001*
MVA model: RCTs	*1.171*	*1.082 – 1.268*	*3.913*	*< .001*
Before-After/Experimental-Control designs (*n* = 31 evaluations; 36 effect sizes)				
Alsaker & Valkanover (2001)	1.134	0.579 – 2.222	0.367	0.713
Andreou et al. (2007)	1.956	1.305 – 2.934	3.246	0.001
Bergen 2/Olweus	1.770	0.974 – 3.218	1.872	0.061
Bull et al. (2009)	2.455	0.343 – 17.563	0.894	0.371
Ciucci & Smorti (1998)	1.198	0.581 – 2.470	0.491	0.624
Evers et al. (2007); High	1.745	1.136 – 2.681	2.543	0.011
Evers et al. (2007); Middle	1.547	0.909 – 2.630	1.609	0.108
Finn (2009)	1.162	0.853 – 1.584	0.954	0.340
Gini et al. (2003)	0.762	0.151 – 3.846	-0.329	0.742
Gollwitzer et al. (2006)	0.968	0.451 – 2.079	-0.084	0.933
Joronen et al. (2011)	1.210	0418 – 3.509	0.352	0.725
Losey (2009)	0.903	0.618 – 1.322	-0.523	0.601
Martin et al. (2005)	2.560	0.333 – 19.656	0.904	0.366
Melton et al. (1998)	1.519	1.248 – 1.849	4.172	< .001
Menard & Grotpeter (2014)	1.085	0.855 – 1.377	0.672	0.502
Menesini et al. (2003)	1.594	0.952 – 2.669	1.772	0.076
Menesini et al. (2012; Study 1)	0.549	0.336 – 0.896	-2.399	0.016

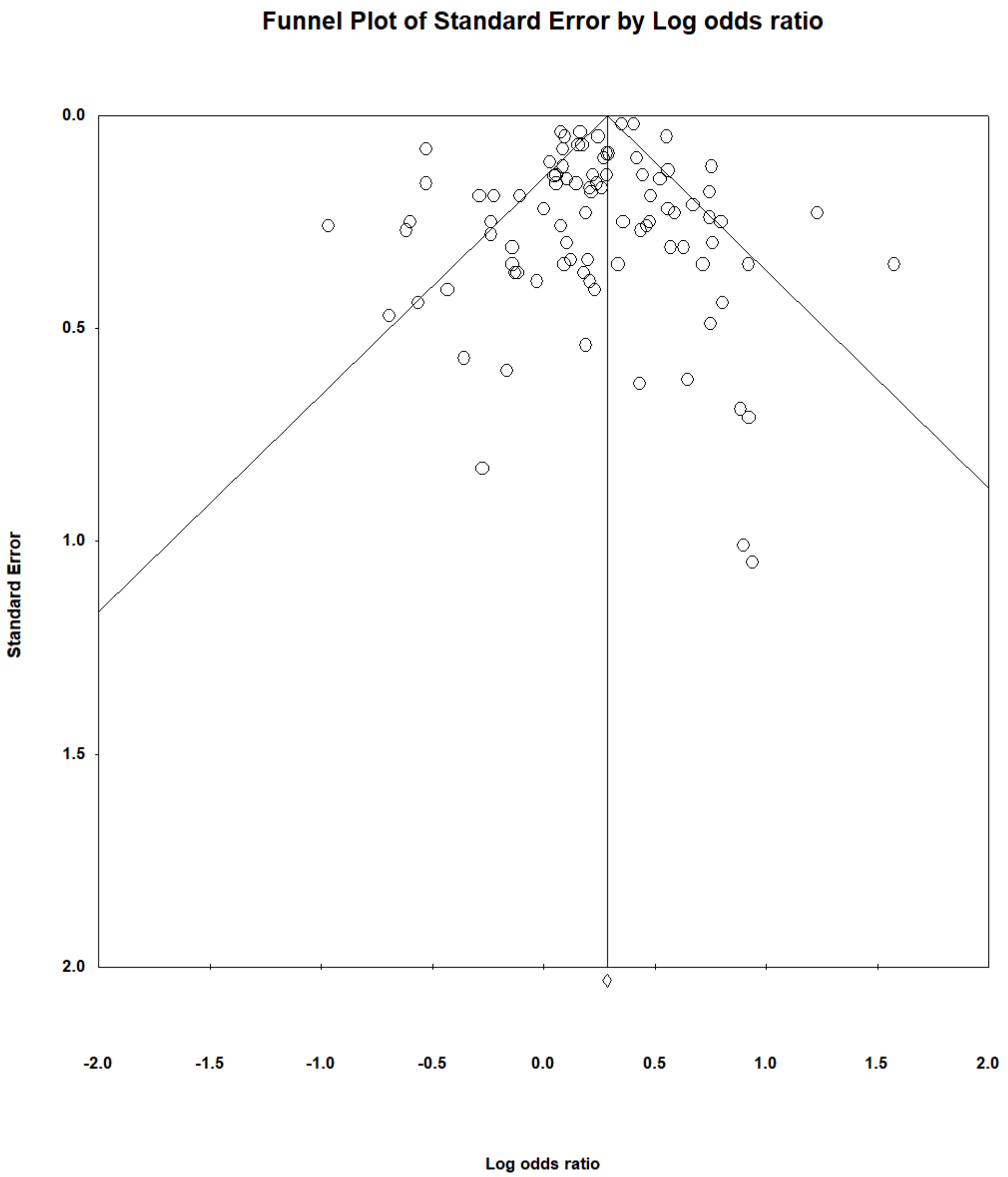
Ortega-Ruiz et al. (2012)	1.230	0.893 – 1.693	1.268	0.205
Palladino et al. (2012)	1.611	0.987 – 2.632	1.906	0.057
Palladino et al. (2016; Trial 1)	1.803	1.148 – 2.832	2.559	0.010
Palladino et al. (2016; Trail 2)	2.107	1.305 – 3.401	3.048	0.002
Pepler et al. (2004)	1.883	1.030 – 3.444	2.055	0.040
Pryce & Frederickson (2013)	0.543	0.324 – 0.909	-2.324	0.020
Rahey & Craig (2002); Senior	1.223	0.629 – 2.378	0.594	0.553
Rahey & Craig (2002); Junior	1.075	0.654 – 1.769	0.286	0.775
Rawana et al. (2011)	0.565	0.240 – 1.330	-1.307	0.191
Rican et al. (1996)	2.522	0.638 – 9.964	1.320	0.187
Sapouna et al. (2010)	0.867	0.465 - 1.617	-0.450	0.653
Silva et al. (2016)	1.259	0.562 – 2.822	0.559	0.576
Sismani et al. (2014)	0.699	0.231 – 2.116	-0.634	0.526
Solomontos-Kountouri et al. (2016); 7th grade	1.029	0.832 – 1.274	0.267	0.790
Solomontos-Kountouri et al. (2016); 8th grade	0.593	0.431 – 0.817	-3.200	0.001
Sutherland (2010)	0.754	0.519 – 1.095	-1.482	0.138
Toner (2010)	0.890	0.427 – 1.859	-0.309	0.757
Wong et al. (2011)	2.111	1.480 – 3.013	4.120	< .001
Yaakub et al. (2010)	1.085	0.935 – 1.260	1.071	0.284
Random Effects: BA/EC	*1.183*	*1.040 – 1.345*	*2.564*	*0.010*
MVA model: BA/EC	*1.171*	*1.049 – 1.307*	*2.812*	*0.005*
Age Cohort Designs (*n* = 14 evaluations; 14 effect sizes)				
Busch et al. (2013)	0.380	0.226 – 0.639	-3.653	< .001
Ertesvåg & Vaaland (2004)	1.340	1.133 – 1.587	3.407	0.001
Kärnä et al. (2011a); Nationwide	1.180	1.093 – 1.274	4.245	< .001
Limber et al. (2017); OBPP Pennsylvania	1.503	1.427 – 1.582	15.474	< .001
Olweus/Bergen 1	1.690	1.252 – 2.282	3.431	< .001
Olweus/New National	1.744	1.575 – 1.931	10.717	< .001
Olweus/Oslo 1	2.140	1.182 – 3.876	2.512	0.012
Olweus/Oslo 2	1.751	1.354 – 2.263	4.275	< .001
O'Moore & Minton (2004)	2.119	0.809 – 5.547	1.530	0.126

Pagliocca et al. (2007)	1.300	0.926 – 1.824	1.514	0.130
Purugulla (2011)	1.274	0.923 – 1.758	1.473	0.141
Roland et al. (2010)	1.417	1.368 – 1.468	19.430	< .001
Salmivalli et al. (2005)	1.310	1.068 – 1.606	2.596	0.009
Whitney et al. (1994)	1.330	1.113 – 1.589	3.132	0.002
Random Effects: Age Cohorts	*1.474*	*1.392 – 1.560*	*13.416*	*< .001*
MVA model: Age Cohorts	*1.422*	*1.359 – 1.487*	*15.563*	*< .001*
Overall: Random Effect model	**1.308**	**1.239 – 1.380**	**9.792**	**< .001**
Overall: MVA model	**1.324**	**1.271 – 1.379**	**13.403**	**< .001**

Note: Sig = statistically significant; MVA = multiplicative variance adjustment; OR = odds ratio; CI = 95% confidence intervals; RCTs = randomised controlled trials; BA/EC = before-after/experimental control designs

Figure 4

Publication bias analysis: School-bullying perpetration



6.2.2 School-bullying victimisation

Overall, the results show that anti-bullying programmes significantly reduced bullying victimisation under both computational models of meta-analysis. The effect sizes for each evaluation are presented in Table 14. The mean summary effect sizes were very similar under both the multiplicative variance adjustment model (MVA OR = 1.248; 95% CI 1.21 – 1.29; $z = 12.06$; $p < .001$; $I^2 = 78.327$) and the random effects model (RE OR = 1.244; 95% CI: 1.19 – 1.31; $z = 8.92$; $p < 0.001$; $tau^2 = 0.032$).

This result suggests that students who participated in an anti-bullying programme were significantly less likely to report being bullied by others after receiving the intervention, in comparison to students who did not receive the intervention.

The funnel plot in Figure 5 indicates that publication bias is not likely to threaten findings in relation to bullying victimisation effect sizes, as the studies fall symmetrically around the mean effect size. Duval and Tweedie's trim and fill procedure highlighted some minor differences between observed effect sizes (OR = 1.245; 95% CI 1.186 – 1.306; $Q = 460.97$) and adjusted effect sizes (OR = 1.241; 95% CI 1.182 – 1.303; $Q = 473.43$). However, this difference is negligible. Based on these results, it was reasonable to assume that publication bias was not likely.

6.2.3 Analysis of heterogeneity

In a meta-analysis, heterogeneity (Q) is the between-study spurious variance that occurs partly because of true variation in effect sizes, but also as a result of random error (Borenstein et al., 2009). Heterogeneity is estimated as the excess variation that exists when the total amount of between-study variance and within-study random error is compared.

In the present meta-analysis, there was significant heterogeneity between studies for both bullying perpetration ($Q = 323.392$; $df = 85$; $p < 0.001$; $I^2 = 73.716$) and bullying victimisation (Q = 387.255; df = 87; p < 0.001; I^2 = 77.534) outcomes. Multiple subgroup analyses were conducted to explore possible explanations for this heterogeneity.

6.2.4 Risk of bias analysis

Scores on each of the risk of bias items were summed to estimate a total risk of bias score. This continuous variable was then used to examine the relationship between effectiveness and risk of bias in meta-regression models.

For perpetration outcomes, risk of bias was not associated with effect size under a random effects model of meta-regression (b = 0.003; SE = 0.006; z = 0.50; p = .621) or under the MVA model (b = 0.014; SE = 0.014; z = 1.01; p = .156). Similarly, risk of bias scores did not significantly predict bullying victimisation effect sizes under a random effects meta-regression (b = 0.007; SE = 0.005; z = 1.30; p = .195) or the MVA model (b = 0.012; SE = 0.012; z = 1.006; p = .157).

Table 14

Meta-analysis results: School-bullying victimisation outcomes

Study	*OR*	*CI*	*z*	*p*
Randomised Controlled Trials (*n* = 33 evaluations; 37 effect sizes)				
Baldry & Farrington (2004); Older	2.874	1.207 – 6.842	2.385	0.017
Baldry & Farrington (2004); Younger	1.011	0.425 – 2.407	0.025	0.980
Berry & Hunt (2009)	9.865	3.129 – 31.102	3.907	< .001
Bonell et al. (2015)	1.000	0.761 – 1.315	0.000	1.000
Brown et al. (2011)	1.212	1.051 – 1.397	2.650	0.008
Chaux et al. (2016)	1.236	0.857 – 1.783	1.136	0.256
Cissner & Ayoub (2014)	0.632	0.342 – 1.167	-1.466	0.143
Connolly et al. (2015)	0.917	0.638 – 1.317	-0.471	0.638
Cross et al. (2011)	1.202	0.884 – 1.635	1.172	0.241
DeRosier & Marcus (2005)	0.878	0.559 – 1.378	-0.567	0.571
Domino (2013)	5.305	3.342 – 8.422	7.077	< .001
Espelage et al. (2015); Illinois	0.733	0.542 – 0.991	-2.091	0.043
Espelage et al. (2015); Kansas	0.934	0.607 – 1.438	-0.309	0.757
Fekkes et al. (2006)	1.006	0.672 – 1.506	0.029	0.977
Fekkes et al. (2016)	2.430	1.188 – 4.970	2.433	0.015
Fonagy et al. (2009)	1.182	0.895- 1.559	1.179	0.238
Frey et al. (2005)	1.117	0.859 – 1.453	0.824	0.410
Garaigordobil & Martinez-Valderrey (2015)	2.213	1.171 – 4.182	2.447	0.014
Hunt (2007)	1.259	0.771 – 2.056	0.920	0.357
Jenson et al. (2013)	1.309	0.785 – 2.183	1.031	0.303
Ju et al. (2009)	1.669	0.752 – 3.700	1.260	0.208
Kaljee et al. (2017)	0.878	0.735 – 1.048	-1.440	0.150
Kärnä et al. (2011b); Grades 4 – 6	1.273	1.156 – 1.401	4.926	< .001
Kärnä et al. (2013); Grades 2 – 3	1.148	1.028 – 1.282	2.452	0.014
Kärnä et al. (2013); Grades 8 – 9	0.937	0.860 – 1.020	-1.500	0.134
Knowler & Frederickson (2013)	0.573	0.196 – 1.669	-1.022	0.307
McLaughlin (2009)	1.458	0.453 – 4.697	0.632	0.527
Nocentini & Menesini (2016); Middle	1.668	1.264 – 2.201	3.615	< .001

Nocentini & Menesini (2016); Primary	1.600	1.212 – 2.111	3.321	0.001
Polanin (2015)	1.214	0.352 – 4.184	0.307	0.758
Rosenbluth et al. (2004)	0.699	0.515 – 0.949	-2.295	0.022
Sprober et al. (2006)	1.031	0.450 – 2.361	0.073	0.942
Topper (2011); Adventure	1.230	0.949 – 1.594	1.562	0.118
Topper (2011); Preventure	0.762	0.480 – 1.209	-1.154	0.249
Trip et al. (2015)	1.028	0.718 – 1.471	0.149	0.882
Tsiantis et al. (2013)	1.857	0.749 – 4.602	1.337	0.181
Yanagida et al. (2016)	3.725	1.656 – 8.377	3.180	0.001
Random Effects: RCTs	*1.200*	*1.078 – 1.336*	*3.331*	*0.001*
MVA model: RCTs	*1.117*	*1.027 – 1.215*	*2.571*	*0.010*
Before-After/Experimental-Control designs (*n* = 37 evaluations; 42 effect sizes)				
Alsaker & Valkanover (2001)	3.114	1.609 – 6.029	3.371	0.001
Andreou et al. (2007)	1.376	0.918 – 2.064	1.544	0.123
Battey (2009)	0.773	0.352 – 1.696	-0.643	0.521
Bauer et al. (2007)	1.013	0.793 – 1.294	0.100	0.92
Beran et al. (2004)	1.101	0.657 – 1.843	0.366	0.715
Bergen 2/Olweus	1.438	0.956 – 2.161	1.745	0.081
Bull et al. (2009)	2.366	0.357 – 15.680	0.892	0.372
Ciucci & Smorti (1998)	1.234	0.595 – 2.558	0.565	0.572
Elledge et al. (2010)	0.492	0.138 – 1.751	-1.095	0.273
Evers et al. (2007); High	0.915	0.565 – 1.482	-0.362	0.718
Evers et al. (2007); Middle	2.257	1.288 – 3.953	2.846	0.004
Finn (2009)	1.031	0.757 – 1.405	0.195	0.845
Fox & Boulton (2003)	0.739	0.174 – 3.139	-0.410	0.682
Gini et al. (2003)	0.405	0.116 – 1.414	-1.417	0.157
Gollwitzer et al. (2006)	0.968	0.451 – 2.079	-0.084	0.933
Herrick (2012)	0.661	0.205 – 2.137	-0.691	0.490
Joronen et al. (2011)	2.482	0.894 – 6.890	1.745	0.081
Kimber (2008)	1.833	1.122 – 2.993	2.420	0.016
Losey (2009)	0.831	0.568 – 1.216	-0.953	0.340
Martin et al. (2005)	1.970	0.231 – 16.781	0.620	0.535
Melton et al. (1998)	1.058	0.869 – 1.287	0.559	0.576

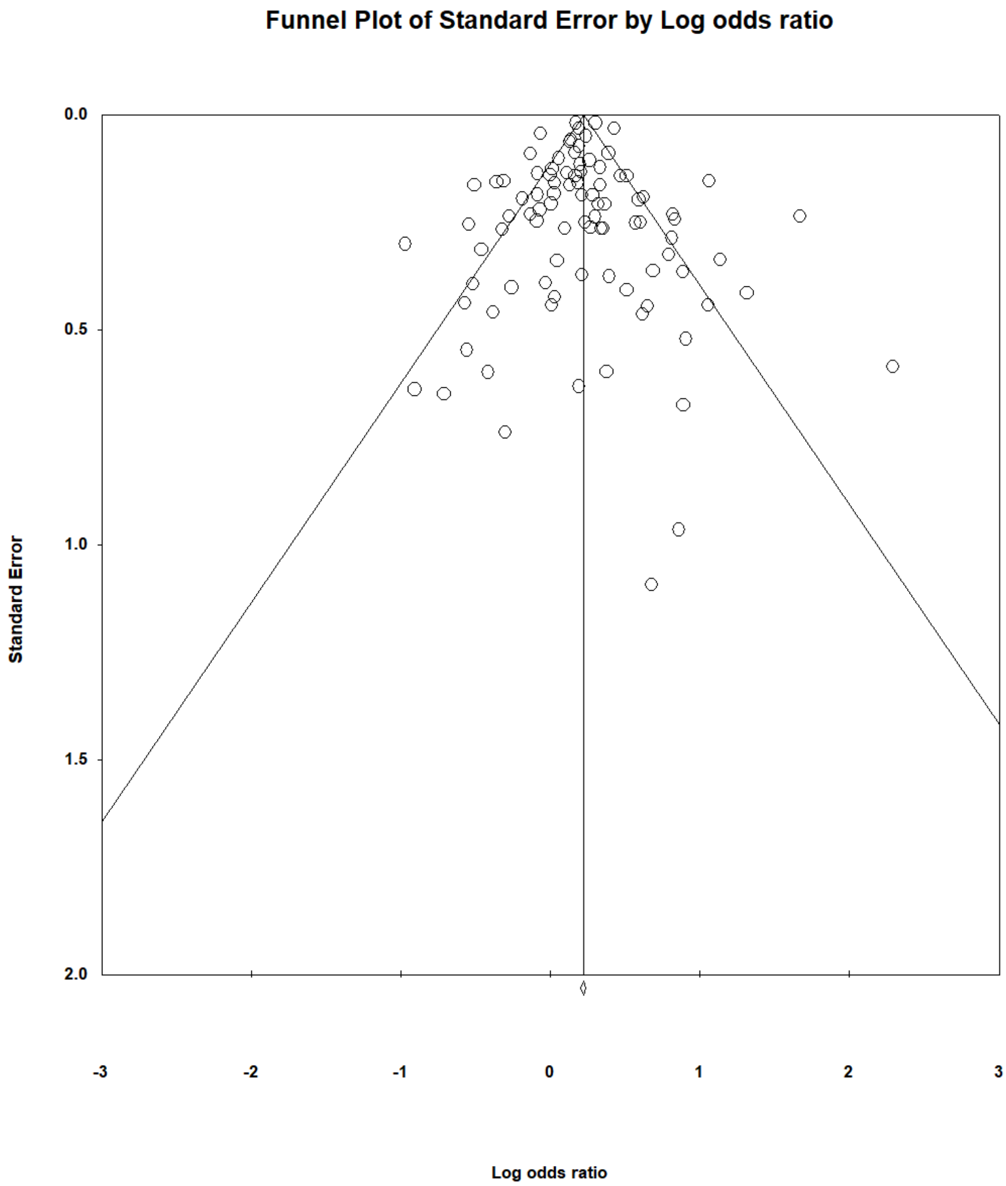
Menard & Grotpeter (2014)	1.395	1.099 – 1.770	2.739	0.006
Menesini et al. (2003)	1.422	0.849 – 2.381	1.338	0.181
Menesini et al. (2012; Study 1)	0.596	0.276 – 1.290	-1.313	0.189
Ortega-Ruiz et al. (2012)	1.394	1.012 – 1.918	2.036	0.042
Palladino et al. (2012)	1.771	1.084 – 2.892	2.283	0.022
Palladino et al. (2016; Trial 1)	2.270	1.445 – 3.566	3.559	< .001
Palladino et al. (2016; Trial 2)	2.306	1.432 – 3.712	3.437	0.001
Pepler et al. (2004)	0.724	0.430 – 1.219	-1.214	0.225
Pryce & Frederickson (2013)	1.406	0.840 – 2.355	1.297	0.195
Rahey & Craig (2002); Junior	1.048	0.539 – 2.038	0.139	0.889
Rahey & Craig (2002); Senior	0.582	0.354 – 0.958	-2.129	0.033
Rawana et al. (2011)	0.565	0.240 – 1.330	-1.307	0.191
Rican et al. (1996)	2.438	0.650 – 9.134	1.322	0.186
Sapouna et al. (2010)	1.351	0.849 – 2.150	1.270	0.204
Silva et al. (2016)	0.683	0.278 – 1.680	-0.830	0.407
Sismani et al. (2014)	1.917	0.802 – 4.587	1.463	0.143
Solomontos-Kountouri et al. (2016); 7th grade	1.142	0.829 – 1.572	0.811	0.417
Solomontos-Kountouri et al. (2016); 8th grade	0.603	0.438 – 0.830	-3.100	0.002
Sutherland (2010)	1.868	1.286 – 2.714	3.279	0.001
Toner (2010)	1.482	0.710 – 3.094	1.048	0.294
Williams et al. (2015)	1.326	0.921 – 1.909	1.516	0.129
Random Effects: BA/EC	*1.226*	*1.085 – 1.385*	*3.278*	*0.001*
MVA model: BA/EC	*1.188*	*1.066 – 1.325*	*3.104*	*0.002*
Age Cohort Designs (*n* = 14 evaluations; 14 effect sizes)				
Busch et al. (2013)	0.380	0.211 – 0.684	-3.227	0.001
Ertesvåg & Vaaland (2004)	1.181	0.995 – 1.400	1.908	0.056
Kärnä et al. (2011a); Nationwide	1.210	1.137 – 1.287	6.045	< .001
Limber et al. (2017); OBPP Pennsylvania	1.189	1.148 – 1.232	9.655	< .001
Olweus/Bergen 1	2.889	2.141- 3.900	6.935	< .001
Olweus/New National	1.533	1.441 – 1.632	13.497	< .001
Olweus/Oslo 1	1.809	1.230 – 2.662	3.010	0.003

Olweus/Oslo 2	1.480	1.243 – 1.762	4.404	< .001
O'Moore & Minton (2004)	1.990	0.977 – 4.053	1.895	0.058
Pagliocca et al. (2007)	0.920	0.705 – 1.201	-0.610	0.542
Purugulla (2011)	1.221	0.975 – 1.529	1.737	0.082
Roland et al. (2010)	1.355	1.308 – 1.404	16.925	< .001
Salmivalli et al. (2005)	1.300	1.058 – 1.596	2.495	0.013
Whitney et al. (1994)	1.140	1.004 – 1.295	2.015	0.044
Random Effects: Age Cohorts	*1.302*	*1.230 – 1.378*	*9.092*	*< .001*
MVA model: Age Cohorts	*1.289*	*1.288 – 1.353*	*10.218*	*< .001*
Overall: Random Effects model	**1.242**	**1.183 – 1.304**	**8.767**	**< .001**
Overall: MVA model	**1.248**	**1.204 – 1.294**	**12.06**	**< .001**

Note: Sig = statistically significant; MVA = multiplicative variance adjustment; OR = odds ratio; CI = 95% confidence intervals; RCTs = randomised controlled trials; BA/EC = before-after/experimental control designs

Figure 5

Publication bias analysis: School-bullying victimisation



6.3 Moderator and mediator analyses

Additional analyses were conducted to examine potential reasons for the heterogeneity observed between effect sizes for school-bullying perpetration and school-bullying victimisation outcomes. As explained in Chapter 2, section 2.7, subgroup analyses analogous to a one-way ANOVA were computed to examine differences between subgroups of primary studies.

Effect sizes are presented as weighted odds ratios for dichotomous categorical variables for each intervention component, i.e., evaluations of programmes in which the component was included (present) compared to evaluations of programmes in which the component was not included (absent). The 95% confidence intervals are also reported, along with the Q_B heterogeneity test and relevant p value. This test indicates the statistical significance of the differences observed between two weighted mean odds ratios. Results for subgroup analyses are presented following the structure for data extraction outlined in Chapter 5.

6.3.1 Descriptive level

On the descriptive level, studies were compared according to the location of the intervention evaluation, the age of participants, and both the publication type and year.

6.3.1.1 Location of the intervention. Mean effects for bullying perpetration and bullying victimisation according to the location of the intervention are presented graphically in Figures 6 and 7 respectively. Table 15 shows the mean effects for the 12 countries for both bullying perpetration and victimisation outcomes under both the MVA model and the random effects model.

Evaluations conducted in Greece were associated with the largest effect sizes for bullying perpetration outcomes, followed by Norway, Italy, US, and Finland under the MVA model of meta-analysis. Evaluations conducted in Italy were associated with the largest mean effect sizes in relation to bullying victimisation, followed by Spain, Norway, US, and Finland under the MVA model of meta-analysis. Additionally, evaluations conducted in Germany and the United Kingdom had significant mean effects when computed using the MVA model. Under the random effects model, Greek evaluations were similarly associated with the largest effect sizes for bullying perpetration, followed by Spanish and Norwegian evaluations. Evaluations conducted in Italy and the US were also associated with significant mean effects for reductions in bullying perpetration. In relation to bullying victimisation, evaluations conducted in Spain and Italy were associated with very similar mean effect sizes under the RE model and were the largest of the 12 effect sizes, followed by evaluations conducted in Norway. Evaluations conducted in Australia were also associated with significant mean effects in reducing bullying victimisation ($p < .05$) and evaluations conducted in Finland and the US were nearly statistically significant ($p = .05$ and .06 respectively) under the random effects model.

Figure 6

Forest plot of effect sizes by location: School-bullying perpetration

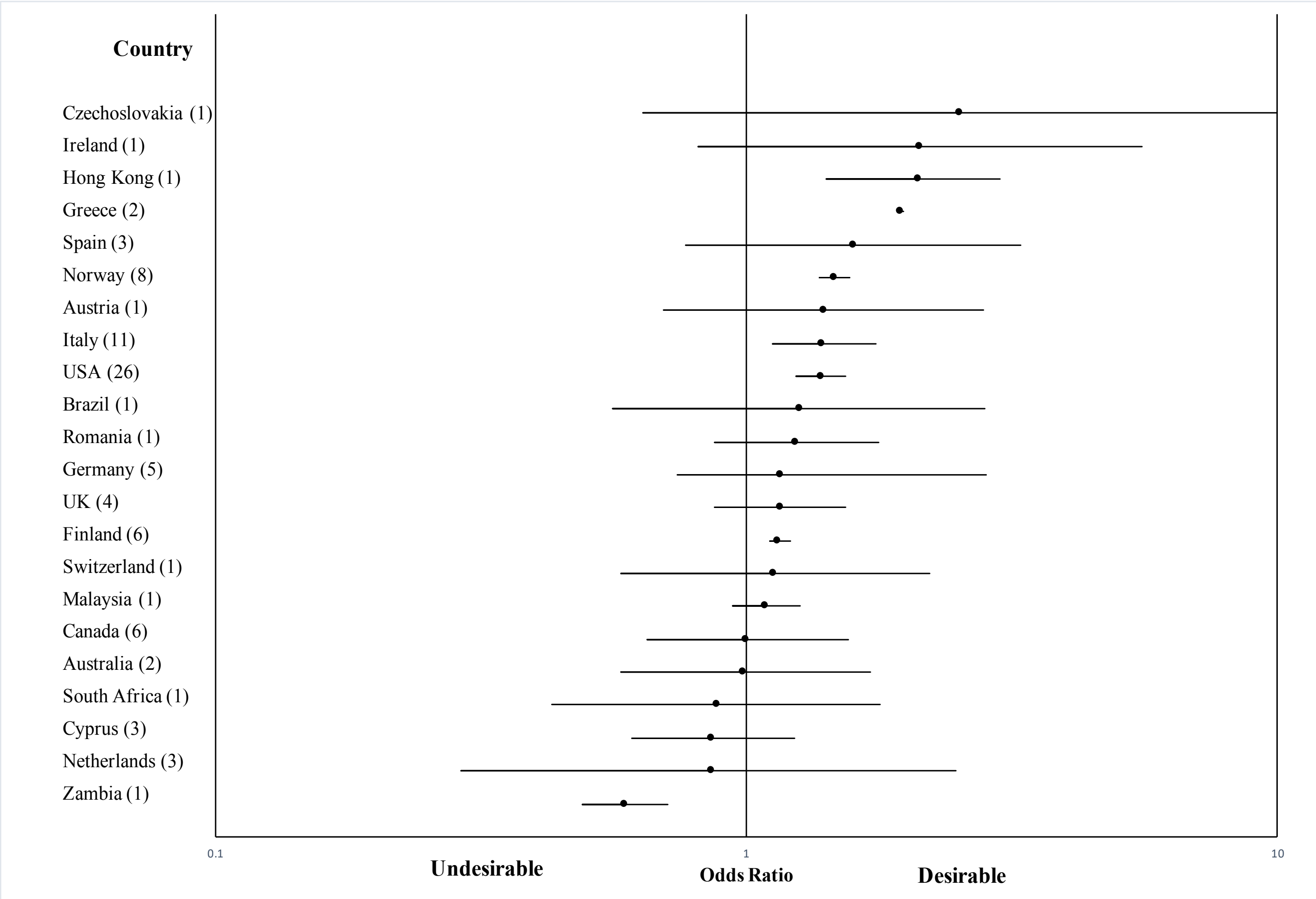


Figure 7

Forest plot of effect sizes by location: School-bullying victimisation

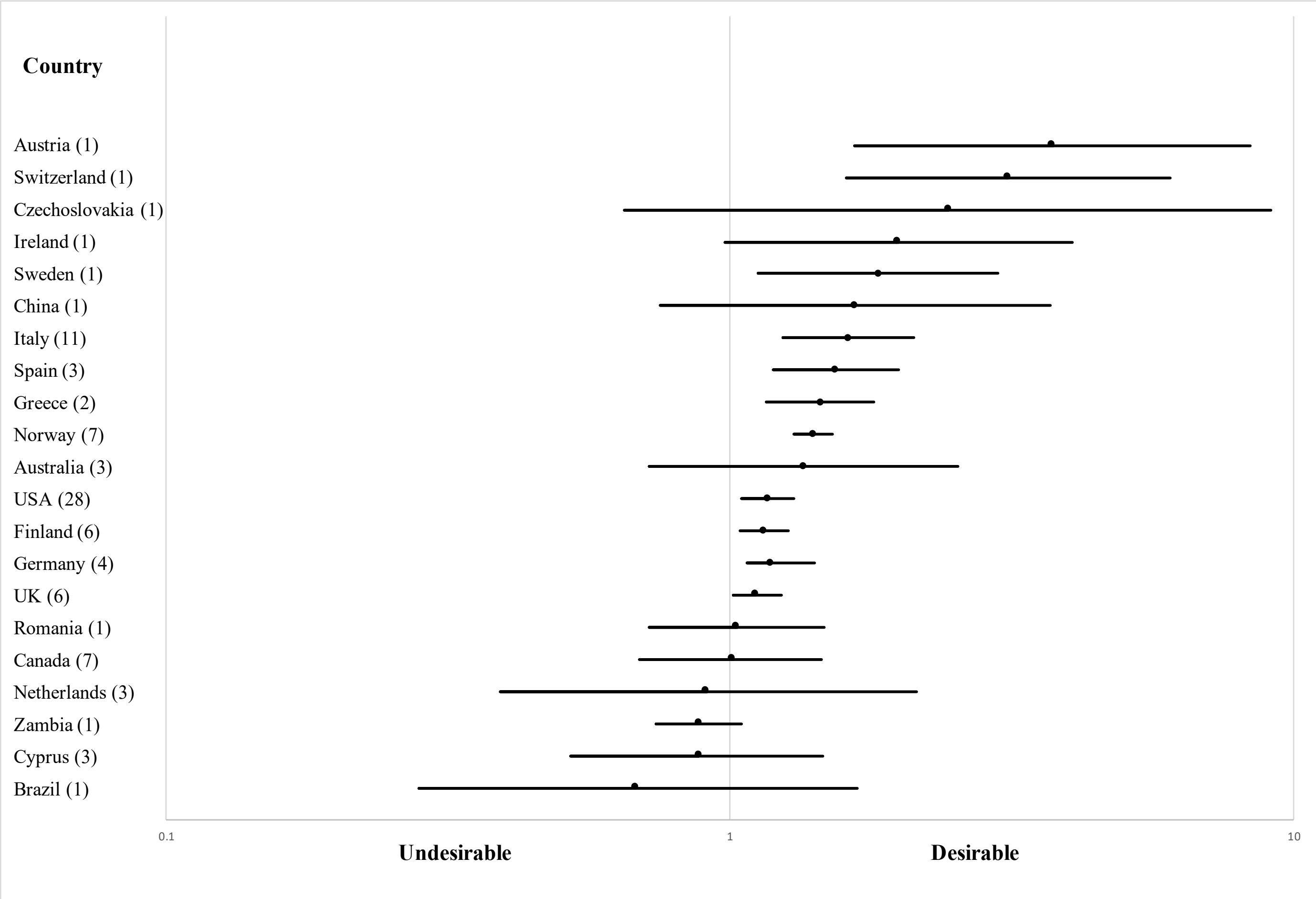


Table 15

School-bullying moderator analysis results: Location of evaluation

	MVA Model					Random effects model			
Location (*n*)	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
School bullying perpetration (*n* = 79 effect sizes)									
Australia (2)	0.994	0.58 – 1.71	.980	3.364 (.067)	70.273	1.020	0.699 – 1.489	.916	.059
Canada (6)	1.00	0.65 – 1.56	.99	3.950 (.413)	26.582	0.919	0.683 – 1.235	.574	.021
Cyprus (3)	0.86	0.61 – 1.23	.42	8.660 (.013)	76.905	0.854	0.648 – 1.127	.266	.035
Finland (6)	1.15	1.11 – 1.21	< .001	4.982 (.418)	0.361	1.158	0.994 – 1.348	.059	.003
Germany (5)	1.16	0.74 – 2.83	.52	8.779 (.118)	54.437	1.062	0.796 – 1.416	.685	.021
Greece (2)	1.95	1.93 – 1.98	< .001	0.001 (.973)	NA	1.949	1.209 – 3.145	.006	.212
Italy (11)	1.39	1.12 – 1.75	.004	26.349 (.003)	62.048	1.370	1.141 – 1.643	.001	.056
Netherlands (3)	0.86	0.29 – 2.48	0.78	19.548 (< .001)	89.769	0.892	0.606 – 1.313	.563	.593
Norway (8)	1.47	1.37 – 1.57	< .001	30.430 (< .001)	76.996	1.659	1.436 – 1.918	< .001	.002
Spain (3)	1.59	0.77 – 3.29	.21	12.859 (.002)	84.447	1.791	1.222 – 2.624	.003	.490
UK (4)	1.16	0.87 – 1.54	.32	11.618 (.009)	74.178	1.029	0.807 – 1.313	.816	.036
US (26)	1.38	1.24 – 1.54	< .001	65.804 (< .001)	62.008	1.293	1.171 – 1.428	< .001	.004
School bullying victimisation (*n* = 82 effect sizes)									
Australia (3)	1.349	0.721 – 2.529	.351	12.15 (.002)	83.539	1.463	1.029 – 2.078	.034	0.316
Canada (7)	1.052	0.691 – 1.452	.982	17.121 (.004)	64.955	1.016	0.792 – 1.304	.902	0.069
Cyprus (3)	0.875	0.520 – 1.462	.614	10.982 (.004)	81.788	0.912	0.666 – 1.249	.564	0.095
Finland (6)	1.149	1.044 – 1.273	.008	32.574 (<.001)	84.650	1.180	1.004 – 1.388	.045	0.001
Germany (4)	1.229	1.068 – 1.414	.01	1.169 (.883)	156.629	1.220	0.886 – 1.678	.223	0.076
Greece (2)	1.446	1.161 – 1.803	< .001	0.349 (.555)	186.533	1.475	0.924 – 2.355	.104	0.092

Italy (10)	1.632	1.237 – 2.122	< .001	19.198 (.038)	53.120	1.592	1.314 – 1.928	< .001	0.035
Netherlands (3)	0.911	0.389 – 2.136	0.833	15.947 (< .001)	87.458	0.914	0.631 – 1.326	.636	0.415
Norway (7)	1.404	1.302 – 1.515	< .001	39.737 (< .001)	84.901	1.548	1.326 – 1.809	< .001	0.014
Spain (3)	1.537	1.19 0– 1.987	< .001	1.670 (.434)	19.760	1.610	1.091 – 2.377	.016	0.053
UK (6)	1.110	1.011 – 1.229	.041	4.056 (.541)	23.274	1.060	0.831 – 1.352	.639	0.017
US (28)	1.168	1.050 – 1.303	.005	90.373 (< .001)	70.124	1.105	0.996 – 1.227	.059	0.019

6.3.1.2 Sample size. A categorical variable was created to compare groups of studies based on sample size. The results of subgroup analyses based on sample size are presented in Table 16. Evaluations in 'Category 2' (i.e., sample size was between 247 and 683 students) had the largest mean effect sizes for both bullying perpetration and victimisation outcomes. Moreover, this was observed when mean effect sizes were computed under both the MVA model and the random effects model. Given the marginal differences between mean effect sizes for subgroups based on sample size, further comparisons (i.e., for statistical difference) was not conducted. However, the pattern of results suggests a potential curvilinear relationship between sample size and effect size for both perpetration and victimisation outcomes.

Sample size was also represented as a continuous variable in the present research and meta-regression analyses was performed. Meta-regression analyses found no significant relationship between sample size and bullying perpetration outcomes under both the random effects model (p = .964) and the MVA model (p = .862). Similarly, meta-regression showed no significant relationship between evaluation sample size and bullying victimisation effect sizes under the MVA model (p = .121) and the random effects model (p = .923).

6.3.1.3 Age of participants. Similar to sample size, the age of participants included in primary evaluations was coded as both a continuous variable and a categorical variable. Meta-regression analyses used the continuous variable to examine the relationship between age and effect sizes for bullying perpetration and victimisation outcomes. Under the MVA model, age of participants did not significantly predict bullying perpetration outcomes (B = -.0116, p = .3750 or bullying victimisation outcomes (B = -.023, p = .123). Similarly, under a random effects meta-regression model, age of participants did not significantly predict either bullying perpetration (B = -.005, p = .719) or bullying victimisation (B = -.003, p = .834) outcomes.

The mean summary effect sizes for subgroups of studies based on age are outlined in Table 17 (dichotomous variable) and Table 18 (categorical variable). In relation to school-bullying perpetration, studies that evaluated interventions implemented with 'older' adolescents gave a larger mean effect size under both the MVA and RE models, but the difference was only statistically significant under the MVA model (Q_B = 19.514, $p < .001$). There were only minor differences between subgroups on school-bullying victimisation using this dichotomous variable to represent age. No clear result was identified, as 'younger' studies gave a *slightly* larger mean effect size under the MVA model, but 'older' studies gave a larger mean effect size under the RE model. Furthermore, the comparisons were statistically significant at $p < .05$ level under the MVA model (Q_B = 6.269, $p = .012$) but was not statistically significant under the RE model (Q_B = 0.933, $p = .334$).

Age was also represented as a four-level categorical variable in the present analyses. Under the MVA model of meta-analysis, the mean effect sizes of subgroups can be ordered from largest to smallest in correspondence to youngest to oldest for both school-bullying perpetration and victimisation outcomes. In other words, the subgroup of studies labelled 'primary 1' were associated with the largest effect sizes, followed by 'primary 2' and 'middle', with studies labelled 'secondary' associated with the smallest effect sizes. Under the MVA model the differences between groups were statistically significant for school-bullying perpetration (Q_B = 31.324, $p < .001$) and victimisation (Q_B = 27.862, $p < .001$). In addition, under the MVA model, studies that included the youngest participants (i.e., 'primary 1', ages 4 – 7) and the oldest (i.e., 'secondary', ages 14 – 18) were collectively not statistically significant in reducing bullying victimisation.

Results under the RE model were not as clear. Most notably, studies labelled 'primary 1' were associated with the largest effect sizes under the RE model for both perpetration and victimisation outcomes. Interestingly, studies labelled 'primary 2' were associated with the smallest mean effect sizes for both perpetration and victimisation outcomes under the RE model. However, differences between subgroups were not statistically significant under the RE model for either bullying perpetration ($Q_B = 7.012$, $p = .072$) or bullying victimisation outcomes ($Q_B = 3.349$, $p = .329$).

Table 16

School-bullying subgroup analysis results: Sample size

	MVA Model					*Random effects model*			
Category	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
School bullying perpetration (83 effect sizes)									
$0 \leq N \leq 246$ (24)	1.206	0.980 – 1.484	.077	36.94 (*p* = .033)	37.733	1.206	0.970 – 1.500	.093	.103
$247 \leq N \leq 683$ (20)	1.356	1.072 – 1.715	.011	91.13 (*p* < .001)	79.151	1.308	1.027 – 1.666	.029	.231
$684 \leq N \leq 1{,}418$ (24)	1.101	0.952 – 1.273	.194	137.38 (*p* < .001)	83.301	1.192	1.022 – 1.391	.025	.113
$N \geq 1{,}419$ (15)	1.221	1.106 – 1.348	< .001	47.58 (*p* < .001)	70.573	1.272	1.133 – 1.428	< .001	.030
School bullying victimisation (87 effect sizes)									
$0 \leq N \leq 246$ (26)	1.251	1.012 – 1.546	< .001	37.549 (*p* = .028)	38.747	1.245	0.993 – 1.561	.057	.111
$247 \leq N \leq 683$ (20)	1.443	1.135 – 1.834	< .001	136. 422 (*p* < .001)	83.141	1.334	1.040 – 1.711	.024	.304
$684 \leq N \leq 1{,}418$ (23)	1.190	1.069 – 1.324	< .001	72.793 (*p* < .001)	69.777	1.204	1.073 – 1.350	.002	.049
$N \geq 1{,}419$ (15)	1.156	1.059 – 1.261	< .001	42.144 (*p* < .001)	64.408	1.112	1.004 – 1.232	.042	.023

Note. N represents the total sample size of primary studies; Categories were created using the interquartile range values of the total sample size.

Table 17

School-bullying subgroup analysis results: Age of participants as a dichotomous variable

	MVA Model					Random effects model			
Category (*n*)	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
School bullying perpetration (86 effect sizes)									
Younger (36)	1.159	1.069 – 1.259	< .001	138.07 (*p* < .001)	74.651	1.181	1.067 – 1.306	.001	.051
Older (50)	1.285	1.215 – 1.359	< .001	410.28 (*p* < .001)	88.057	1.301	1.198 – 1.414	< .001	.044
School bullying victimisation (92 effect sizes)									
Younger (40)	1.283	1.225 – 1.345	< .001	102.99 (*p* < .001)	62.132	1.183	1.078 – 1.297	< .001	.022
Older (52)	1.224	1.157 – 1.295	< .001	312.96 (*p* < .001)	83.704	1.255	1.162 – 1.356	< .001	.046

Note. Younger = participants under 11 years old; Older = participants aged 11 years and older.

Table 18

School-bullying subgroup analysis results: Age of participants as a categorical variable

	MVA Model					*Random effects model*			
Category (*n*)	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
School bullying perpetration (86 effect sizes)									
Primary 1 (5)	1.836	1.400 – 2.408	< .001	8.069 (*p* = .089)	38.034	1.649	1.193 – 2.280	.003	.077
Primary 2 (34)	1.289	1.212 – 1.372	< .001	172.381 (*p* < .001)	80.856	1.149	1.033 – 1.279	.010	.041
Middle (35)	1.249	1.163 – 1.340	< .001	284.366 (*p* < .001)	88.044	1.335	1.204 – 1.480	< .001	.053
Secondary (14)	1.103	0.928 – 1.310	.134	73.639 (*p* < .001)	82.346	1.195	1.002 – 1.424	.047	.163
School bullying victimisation (92 effect sizes)									
Primary 1 (2)	1.817	0.627 – 5.268	.136	5.174 (*p* = .023)	80.673	1.815	1.057 – 3.115	.031	.476
Primary 2 (38)	1.282	1.223 – 1.342	< .001	95.684 (*p* < .001)	61.331	1.168	1.063 – 1.282	.001	.009
Middle (39)	1.244	1.174 – 1.318	< .001	223.639 (*p* < .001)	83.008	1.252	1.148 – 1.365	< .001	.038
Secondary (13)	1.039	0.863 – 1.249	.344	69.859 (*p* < .001)	82.823	1.269	1.066 – 1.510	.007	.199

Note. Primary 1 = participants aged between 4 and 7 years old; Primary 2 = participants aged between 8 and 10 years old; Middle = participants aged between 11 and 13 years old; Secondary 1 = participants aged between 14 and 18 years old.

6.3.1.4 Publication type and year. Table 19 outlines the mean summary effect sizes for each of the publication types for bullying perpetration and victimisation outcomes. Evaluations for which data was received via email correspondence from evaluators had the largest mean effect sizes for both bullying perpetration and bullying victimisation. Differences in the mean effect sizes for evaluations reported via unpublished dissertations, either masters or doctoral theses, were the smallest for both bullying perpetration and victimisation outcomes. Subgroup analysis was not conducted further using these categorizations due to the imbalance in numbers of evaluations in each category (i.e., evaluations were overwhelmingly published in peer-reviewed journal article format).

However, additional analysis was conducted to examine any potential differences between peer reviewed and non-peer reviewed evaluations. Therefore, the above categories were collapsed, and evaluations reported by dissertation, chapter, correspondence and governmental reports (perpetration $n = 23$; victimisation $n = 21$) were compared to evaluations published via peer-reviewed journal articles.

Under the MVA model, non-peer-reviewed evaluations gave a larger (OR = 1.493; 95% CI 1.266 – 1.761; $p < .001$) mean effect size than peer-reviewed evaluations (OR = 1.315, 95% CI 1.251 – 1.383, $p < .001$) for bullying perpetration outcomes. Moreover, subgroup analysis indicated that the difference was statistically significant ($Q_B = 12.861$; $df = 1$; $p < .001$). However, under the random effects model, both groups gave similar effect sizes for bullying perpetration outcomes, and the difference between peer-reviewed (OR = 1.230; 95% CI 1.146 – 1.321, $p < .001$) and non-peer-reviewed (OR = 1.309; 95% CI 1.137 – 1.508; $p < .001$) was not statistically significant ($Q_B = 0.595$; $df = 1$; $p = .441$).

For bullying victimisation outcomes, similar results were obtained. Under the MVA model, non-peer-reviewed evaluations gave statistically significant larger mean effect sizes (OR = 1.403; 95% CI 1.262 1.560; $p < .001$) than peer-reviewed evaluations (OR = 1.223, 95% CI 1.176 – 1.272, $p < .001$; $Q_B = 27.197$; $df = 1$; $p < .001$). Yet, there was a marginal difference under the random effects model between peer-reviewed (OR = 1.209, 95% CI 1.137 – 1.286, $p < .001$) and non-peer-reviewed (OR = 1.231; 95% CI 1.059 – 1.431; $p = .007$) and the difference was not statistically significant ($Q_B = 0.048$; $df = 1$; $p = .827$).

Table 19

School-bullying subgroup analysis results: Publication Type

	MVA Model					Random effects model			
Publication Type (*n*)	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
School bullying perpetration (*n* = 90 effect sizes)									
Article (67)	1.315	1.251 – 1.383	< .001	409.65 (*p* < .001)	83.89	1.230	1.146 – 1.321	< .001	.044
Chapter (2)	1.278	0.909 – 1.796	.158	3.98 (*p* = .264)	24.58	1.321	0.926 – 1.885	.125	.033
Correspondence (4)	1.745	1.692 – 1.799	< .001	0.51 (*p* = .972)	0.00	1.745	1.602 – 1.901	< .001	.000
Dissertation (12)	1.040	0.878 – 1.232	.649	7.74 (*p* = .356)	9.59	1.037	0.870 – 1.237	.686	.006
Gov Report (3)	1.311	0.969 – 1.773	.079	7.241 (*p* = .027)	72.38	1.154	0.805 – 1.654	.435	.070
School bullying victimisation (*n* = 93 effect sizes)									
Article (72)	1.223	1.176 – 1.272	< .001	297.08 (*p* < .001)	76.10	1.209	1.137 – 1.286	< .001	.027
Chapter (2)	1.267	0.316 – 5.083	.738	11.55 (*p* = .001)	91.34	1.480	0.354 – 6.179	.591	.972
Correspondence (4)	1.568	1.367 – 1.799	< .001	17.41 (*p* = .001)	82.77	1.791	1.419 – 2.261	< .001	.042
Dissertation (12)	1.107	0.962 – 1.274	.156	18.04 (*p* = .081)	39.01	1.073	0.934 – 1.280	.267	.026
Gov Report (3)	1.006	0.848 – 1.194	.946	2.46 (*p*< .001)	18.67	0.993	0.826 – 1.193	.939	.006

The mean summary effect size for “2009” studies was OR = 1.487 (95% CI 1.430 – 1.546; $p < .001$) under the MVA model and OR = 1.411 (95% CI 1.315 – 1.513; $p < .001$) under the random effects model for bullying perpetration outcomes. Across both computational models these summary effects were larger than those for studies labelled “2016” on bullying perpetration for the MVA model (OR = 1.243; 95% CI 1.667 – 1.324; $p < .001$) and the RE model (OR = 1.184; 95% CI 1.087 – 1.289; $p < .001$). Subgroup analysis analogous to the ANOVA showed that this difference was statistically significant ($Q_B = 76.412$; $df = 1$; $p < .001$) under the MVA model and the random effects model ($Q_B = 9.676$; $df = 1$; $p = .002$).

In relation to bullying victimisation, the mean summary effect size for studies labelled “2009” was larger (OR = 1.322; 95% CI 1.220 – 1.432; $p < .001$) under the MVA model than the mean summary effect size for studies labelled “2016” (OR = 1.229; 95% CI 1.175 – 1.285; $p < .001$). Subgroup analysis analogous to the ANOVA found that this difference was statistically significant ($Q_B = 10.115$; $df = 1$; $p = .001$) but the difference between odds ratios was marginal. However, under the random effects model the “2009” studies (OR = 1.215; 95% CI 1.094 – 1.350; $p < .001$) were not statistically different from the “2016” studies (OR = 1.223; 95% CI 1.139 – 1.313; $p < .001$; $Q_B = 0.010$; $df = 1$; $p = .920$).

6.3.2 Design level

In relation to design features, data was extracted concerning the measurement instruments used to collect data on bullying behaviours, the evaluation methodology used, the unit of allocation used in the evaluation.

6.3.2.1 Measurement instruments. With respect to the measurement instruments used, not much of the data extracted could be used to conduct appropriate subgroup analyses. For example, there was too much variation in the specific instruments used to measure bullying to assess how effect sizes may vary across measures. Similarly, too few of the studies used measures that were *not* self-report measures, and so subgroup analyses were not conducted.

6.3.2.2 Evaluation methodology. Subgroup analyses were conducted to further investigated the effectiveness of anti-bullying programmes in relation to the methodological designs used by evaluation studies. The breakdown of results by methodological design is shown in Tables 13 and 14 for bullying perpetration and victimisation outcomes respectively.

Primary studies employing age cohort (AC) designs were associated with the largest effect sizes for both bullying perpetration (OR = 1.474; 95% CI 1.39 – 1.56; $p < 0.001$) and bullying victimisation (OR = 1.302; 95% CI 1.230 – 1.378; $p < 0.001$) under a random effects model. Similarly, AC studies were associated with the largest effect sizes under the MVA model (perpetration OR = 1.422; 95% CI 1.36 – 1.46; $p < .001$ and victimisation OR = 1.289; 95% CI = 1.29 – 1.35; $p < .001$).

Under the MVA model of meta-analysis, mean effect sizes were the same for randomised controlled trial (RCT) evaluations (OR = 1.171; 95% CI 1.08 – 1.27; $p < .001$) and before-after quasi-experimental-control (BA/EC) evaluations (OR = 1.170; 95% CI 1.05 – 1.31; $p = .005$) for bullying perpetration outcomes. Moreover, the differences between RCT evaluations (OR = 1.117; 95% CI 1.03 – 1.22; $p = .01$) and BA/EC evaluations (OR = 1.188; 95% CI 1.07 – 1.33; $p = .002$) were marginal for bullying victimisation outcomes under the MVA model.

In relation to bullying victimisation outcomes, BA/EC designs gave the second largest mean effect size (OR = 1.225; 95% CI 1.085 – 1.383; $p = 0.001$), followed by RCTs (OR = 1.210; 95% CI 1.091 – 1.342; $p < 0.001$) under a random effects model. However, the result was the opposite for bullying perpetration outcomes under a random effects model (RCT OR = 1.244; 95% CI 1.123 – 1.379; $p < 0.001$; BA/EC OR = 1.187; 95% CI 1.044 – 1.350; $p = 0.009$).

Due to the marginal differences and lack of clear pattern in regard to which method was associated with the largest effect sizes (between RCT and BA/EC) further subgroup analyses were not conducted.

6.3.2.3 Unit of allocation/randomization. Table 20 outlines the mean effects for subgroups of studies according to how participants were allocated to experimental and control groups. Results are presented for bullying perpetration and victimisation outcomes for all studies that allocated studies in classes, schools, or individual students. The mean effects for RCT and BA/EC according to the allocation unit are also presented.

In relation to bullying perpetration outcomes, under the MVA model, studies that assigned participants in classes were associated with the largest effect sizes. However, the difference between the mean effects for all evaluations that used classes or schools as the unit of allocation were close to statistical significance ($Q_B = 3.705$, $df = 1$, $p = .054$). Under the random effects model, evaluations that assigned students to experimental conditions were associated with the largest effect size for bullying perpetration outcomes when all designs were included, and for RCT evaluations and BA/EC evaluations individually. However, the mean effect size for many of the subgroups were not collectively statistically significant overall under the random effects model.

Similarly, under the MVA model, evaluations conducted using a RCT design, and that assigned classes to conditions, were associated with the largest effect size for bullying perpetration, although the mean effect size for this subgroup was not statistically significant. Moreover, subgroup analysis analogous to the ANOVA found that the mean effect size for RCT designs that assigned classes to experimental or control conditions was not statistically different from RCT designs that assigned schools to experimental or control conditions ($Q_B = 1.140$, $df = 1$, $p = .286$).

In relation to BA/EC designs, evaluations that assigned students to experimental conditions were associated with the largest mean effect size, although the effect was not statistically significant for bullying perpetration outcomes. However, the difference between the mean effect for BA/EC evaluations that assigned classes and those that assigned schools to conditions was statistically significant under the MVA model ($Q_B = 4.551$, $df = 1$, $p = .033$).

For bullying victimisation outcomes, studies where the unit of allocation was classes were associated with the largest effect sizes, followed by schools and individual students under the MVA model. The difference between studies that allocated classes and studies that allocated schools was statistically significant ($Q_B = 12.450$, $df = 1$, $p < .001$). This pattern was observed when all designs were included, and for the subgroup of RCT evaluations and the subgroup of BA/EC evaluations. Thus, when participants were assigned in classes the mean effect size for these RCT evaluations was significantly larger ($Q_B = 13.590$, $df = 1$, $p < .001$) for reductions in bullying victimisation than for RCT evaluations that assigned schools. Yet the difference between the mean effect sizes for BA/EC evaluations that assigned classes was not statistically significant ($Q_B = 3.359$, $df = 1$, $p = .067$) compared with BA/EC evaluations that assigned schools to experimental conditions.

Table 20

School-bullying moderator analysis results: Unit of allocation/randomization

	MVA Model					Random effects model			
Unit of allocation (*n*)	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
School bullying perpetration (*n* = 70 effect sizes)									
All designs									
Classes (19)	1.319	1.087 – 1.601	< .001	44.763 (< .001)	59.788	1.286	1.044 – 1.586	.018	.338
Schools (44)	1.163	1.091 – 1.240	< .001	136.032 (< .001)	68.390	1.188	1.098 – 1.286	< .001	.185
Students (7)	0.725	0.489 – 1.074	.109	47.208 (< .001)	87.290	1.465	0.749 – 2.865	.265	.771
Randomised controlled trials (n = 39 effect sizes)									
Classes (11)	1.295	0.952 – 1.761	.099	36.998 (< .001)	72.972	1.246	0.892 – 1.740	.197	.460
Schools (22)	1.184	1.107 – 1.266	< .001	57.455 (< .001)	63.450	1.242	1.141 – 1.352	< .001	.135
Students (6)	0.720	0.471 – 1.101	.129	45.737 (< .001)	89.068	1.407	0.699 – 2.835	.339	.776
Quasi-experimental designs with before and after measures (n = 31 effect sizes)									
Classes (8)	1.353	1.109 – 1.651	< .001	7.648 (.365)	8.473	1.349	1.099 – 1.655	.004	.008
Schools (22)	1.091	0.942 – 1.263	.244	75.193 (< .001)	72.072	1.108	0.940 – 1.305	.223	.095
Students (1)	2.046	0.340 – 17.807	.373	NA	NA	2.460	0.340 – 17.807	.373	.001
School bullying victimisation (*n* = 71 effect sizes)									
All designs									
Classes (15)	1.529	1.168 – 2.001	< .001	50.377 (< .001)	72.210	1.523	1.138 – 2.038	.005	.462

Schools (47)	1.164	1.063 – 1.275	< .001	132.738 (< .001)	65.345	1.181	1.068 – 1.305	.001	.261
Students (9)	0.940	0.717 – 1.232	.654	27.401 (.001)	70.804	1.157	0.771 – 1.734	.482	.455
Randomised controlled trials (n = 32 effect sizes)									
Classes (7)	1.716	0.967 – 3.046	.065	39.039 (< .001)	84.631	1.637	0.876 – 3.058	.122	.568
Schools (19)	1.156	1.028 – 1.300	< .001	49.942 (< .001)	63.958	1.165	1.025 – 1.324	.019	.046
Students (6)	0.943	0.677 – 1.314	.729	25.486 (< .001)	80.381	1.203	0.777 – 1.863	.407	.220
Quasi-experimental designs with before and after measures (n = 38 effect sizes)									
Classes (8)	1.418	1.144 – 1.757	< .001	9.662 (.209)	27.551	1.422	1.130 – 1.789	.003	.029
Schools (28)	1.175	1.016 – 1.358	< .001	82.710 (< .001)	67.356	1.186	1.013 – 1.389	.034	.107
Students (2)	0.943	0.193 – 3.335	.762	1.825 (.177)	45.205	0.917	0.203 – 4.133	.910	.558

Note. In relation to randomised controlled trials, the above moderator refers to the unit of randomization and where quasi-experimental designs with before and after measures are of interest, the moderator used was the unit of allocation.

6.3.3 Programme level: Intervention components

As previously stated, intervention components from several levels of a socio-ecological framework were coded. The results of coding intervention components for each primary evaluation are provided in Appendix 5. Where effect sizes were coded for separate groups in an evaluation included in the meta-analysis, the table in Appendix 5 just presents the primary evaluation as a whole, unless the intervention components varied for independent groups. Subgroups were created on the basis of the presence of the specific intervention components and subsequent subgroup analyses compared groups of studies where the specific component was present with studies where the component was absent. The results of this analysis are outlined in Table 21 for school-bullying perpetration outcomes and in Table 22 for school-bullying victimisation outcomes. Results for intervention component analysis are presented *only* using the MVA model. The justification for this is discussed further in Chapter 12.

6.3.3.1 School-bullying perpetration. Under the MVA model of meta-analysis, the *presence* of the following components was significantly correlated with larger mean effect sizes for school-bullying perpetration outcomes: whole-school approach; anti-bullying policy, classroom rules; information for parents; informal peer involvement; work with victims; co-operative group work; and mental health approaches. Studies where these components were present produced a larger weighted mean effect size in comparison to studies where these components were absent.

Moreover, the inclusion of the following intervention components: classroom management (p = .039) and punitive disciplinary measures (p = .046) gave larger mean subgroup effect sizes, but the differences between groups were only marginally significant. Interestingly, the *absence* of socio-emotional skills was significantly correlated with larger subgroup summary effect sizes for school-bullying perpetration outcomes.

6.3.3.2 School-bullying victimisation. Under the MVA model of meta-analysis, the presence of only two intervention components, informal peer involvement and information for parents, was significantly correlated with larger subgroup summary effect sizes for school-bullying victimisation outcomes. Additionally, the absence of socio-emotional skills was significantly correlated with larger subgroup summary effect sizes for school-bullying victimisation outcomes. At a less conservative level ($p < .05$) of statistical significance, there were also differences between groups that included or excluded the 'encouraging bystanders' intervention component. Namely, studies that did not include this component were correlated with larger mean effect sizes ($p = .044$).

Table 21

Intervention component analyses for school-bullying perpetration outcomes (N = 82)

	Component Present			Component Absent				
Intervention Component	*n*	OR	95% CI	*n*	OR	95% CI	Q_B	*p*
School-level								
Whole-school approach	43	**1.263**	1.159 – 1.377	39	1.095	0.955 – 1.256	10.291	.001*
Increased supervision	21	1.238	1.117 – 1.371	61	1.194	1.073 – 1.329	.812	.368
Anti-bullying policy	25	**1.288**	1.167 – 1.422	57	1.150	1.013 – 1.282	7.992	.005*
Classroom-level								
Classroom rules	31	**1.289**	1.205 – 1.379	51	1.137	1.290 – 1.002	9.787	.002*
Classroom management	22	**1.265**	1.166 – 1.372	60	1.165	1.038 – 1.307	4.222	.039**
Teacher-level								
Information for teachers	66	1.219	1.124 – 1.321	16	1.155	0.894 – 1.492	.533	.465
Teacher training	51	1.194	1.089 – 1.309	31	1.292	1.118 – 1.492	2.501	.114
Parent-level								
Information for parents	35	**1.280**	1.177 – 1.392	47	1.141	1.078 – 1.209	8.149	.004*
Involvement of parents	21	1.149	0.964 – 1.370	61	1.226	1.125 – 1.335	1.368	.242
Peer-level								
Informal peer involvement	57	**1.294**	1.199 – 1.396	25	1.022	0.948 – 1.102	27.440	.001*
Encouraging bystanders	25	1.170	1.066 – 1.285	57	1.237	1.178 – 1.298	1.729	.188
Formal peer involvement	13	1.324	1.129 – 1.553	69	1.194	1.096 – 1.301	3.544	.059
Individual-level								
Work with Bullies	27	1.147	1.116 – 1.179	55	1.166	1.045 – 1.301	0.163	.686
Works with Victims	31	**1.285**	1.177 – 1.404	51	1.151	1.025 – 1.292	7.593	.006*
Co-operative group work	37	**1.329**	1.207 – 1.464	45	1.148	1.029 – 1.279	12.619	.001*

Intervention-specific								
Curriculum materials	69	**1.263**	1.172 – 1.361	13	0.980	0.762 – 1.260	21.343	.001*
Socio-emotional skills	27	1.027	0.866 – 1.218	55	**1.307**	1.217 – 1.403	30.733	.001*
Mental Health	8	**1.523**	1.157 – 2.004	77	1.163	1.091 – 1.239	11.201	.001*
Punitive disciplinary methods	16	1.279	1.162 – 1.409	66	1.178	1.066 – 1.302	3.966	.046**
Non-punitive disciplinary methods	11	1.284	1.125 – 1.466	71	1.196	1.096 – 1.306	1.994	.158

Note. * = $p < 0.001$; ** = $p < 0.05$, i.e. the difference between mean effect sizes for subgroups is statistically significant at the respective p level. Odds ratios presented in **bold** were the significantly larger subgroup mean summary effect size.

Table 22

Intervention component analyses for school-bullying victimisation outcomes (N = 86)

	Component Present			**Component Absent**				
Intervention Component	*n*	OR	95% CI	*n*	OR	95% CI	Q_B	*p*
School-level								
Whole-school approach	42	1.186	1.096 – 1.307	44	1.226	1.065 – 1.412	0.575	.448
Increased supervision	21	1.215	1.077 – 1.371	65	1.179	1.071 – 1.297	0.607	.436
Anti-bullying policy	26	1.219	1.101 – 1.351	60	1.169	1.051 – 1.300	1.158	.282
Classroom-level								
Classroom rules	30	1.236	1.125 – 1.358	56	1.152	1.033 – 1.285	3.209	.073
Classroom management	22	1.196	1.114 – 1.285	64	1.159	1.038 – 1.294	0.646	.420
Teacher-level								
Information for teachers	70	1.249	1.199 – 1.301	16	1.151	0.904 – 1.465	1.205	.272
Teacher training	55	1.192	1.091 – 1.303	31	1.211	1.065 – 1.377	0.115	.735
Parent-level								
Information for parents	36	**1.246**	1.132 – 1.371	50	1.125	1.007 – 1.257	6.492	.011*
Involvement of parents	24	1.197	0.979 – 1.463	62	1.196	1.111 – 1.289	0.001	.992
Peer-level								
Informal peer involvement	55	**1.246**	1.138 – 1.363	31	1.096	0.975 – 1.232	9.36	.002*
Encouraging bystanders	25	1.199	1.049 – 1.369	62	**1.293**	1.225 – 1.366	4.042	.044**
Formal peer involvement	15	1.263	1.087 – 1.466	71	1.178	1.085 – 1.279	2.151	.143
Individual-level								
Work with Bullies	28	1.203	1.073 – 1.349	58	1.191	1.082 – 1.311	0.071	.791
Works with Victims	36	1.214	1.129 – 1.305	50	1.178	1.072 – 1.295	0.581	.446
Co-operative group work	43	1.213	1.089 – 1.349	43	1.184	1.072 -1.307	0.385	.535

Intervention-specific								
Curriculum materials	71	1.192	1.049 – 1.354	15	1.118	0.976 – 1.281	1.481	.224
Socio-emotional skills	30	1.039	0.884 – 1.221	56	**1.252**	1.161 – 1.349	16.859	.001*
Mental Health	8	1.103	0.811 – 1.501	78	1.201	1.114 – 1.294	0.775	.378
Punitive disciplinary methods	14	1.257	1.092 – 1.447	72	1.169	1.073 – 1.273	3.044	.081
Non-punitive disciplinary methods	11	1.242	1.126 – 1.370	75	1.182	1.084 – 1.289	1.211	.271

Note. * = $p < 0.001$; ** = $p < 0.05$, i.e. the difference between mean effect sizes for subgroups is statistically significant at the respective *p* level. Odds ratios presented in **bold** were the significantly larger subgroup mean summary effect size.

6.3.3.3 Programme specificity and richness. The majority of evaluations included in our meta-analysis were of highly specific intervention programmes (i.e., those that targeted bullying behaviours and no other outcomes). Consistently across computational model and both perpetration and victimisation outcomes, these studies were associated with the largest mean effect sizes. These results are presented in Table 23.

Highly specific programmes were the only subgroup of evaluations that gave a statistically significant mean summary effect under both the MVA model and the random effects model for bullying victimisation outcomes. In relation to bullying perpetration outcomes, the subgroup of evaluations that were coded as ‘medium’ on the programme specificity moderator were associated with a statistically significant mean effect size under the MVA model ($p < .001$) and the random effects model ($p = .036$).

Multiple models of meta-regression were conducted for school bullying perpetration and victimisation outcomes. The continuous variable programme richness, which indicated the total number of intervention components included, did not significantly predict either school-bullying perpetration (B = 0.007; SE = 0.003) or school-bullying victimisation (B = -0.003; SE = 0.003) outcomes.

Moreover, when all intervention components were included in a meta-regression model, no components significantly predicted either school-bullying perpetration and/or victimisation outcomes under the MVA model. Thus, the second planned meta-regression analysis, in which only significant predictors would have been included, was not conducted.

6.3.4 Programme level: Other

Beyond the intervention components, subgroup analyses were also conducted to examine potential differences between packaged intervention programmes (i.e., KiVa, OBPP, and Steps to Respect). Furthermore, the impact of conflict of interest on weighted mean effect sizes and the differences between intervention programmes based on specificity were investigated.

6.3.4.1 Packaged intervention programmes. The mean summary effect sizes for 10 different intervention programmes in relation to reducing bullying perpetration behaviours and 9 different intervention programmes in relation to reducing bullying victimisation behaviours were investigated. Table 24 outlines the effectiveness of specific anti-bullying programmes in reducing both school-bullying perpetration and victimisation. The effectiveness of these programmes varied greatly.

In relation to school-bullying perpetration outcomes, the OBPP was associated with the largest mean effect sizes. In addition, evaluations of the OBPP in Norway were associated with larger summary effect sizes than evaluations of OBPP conducted in the USA. However, the difference was not statistically significant for school-bullying perpetration outcomes when moderator analysis analogous to the ANOVA was conducted ($Q_B = 3.65$; $df = 1$; $p = 0.06$).

Other programmes were significantly effective in reducing school-bullying perpetration behaviours, for example KiVa, Second Step, and Steps to Respect. Positive effect sizes (i.e., OR > 1) were also observed for the BPYS and NoTrap! programmes, but these effects were not statistically significant in relation to reduction in bullying perpetration outcomes. Negative effects were found for two anti-bullying programmes, the fairplayer.manual and ViSC, although these effects were not statistically significant.

In relation to school-bullying victimisation outcomes, NoTrap! was associated with the largest mean effect size, followed by the Bully Proofing Your School Programme, and then the OBPP. The analysis showed that other anti-bullying programmes were also significantly effective in reducing school-bullying victimisation, for example, Steps to Respect and KiVa.

Again, effect sizes for the OBPP varied between evaluations conducted in Norway and evaluations conducted in the USA for bullying victimisation outcomes. Moreover, our analysis found that the difference in the magnitude of these effect sizes was statistically significant ($Q_B = 74.95$; $df = 1$; $p < 0.001$). Our analysis also identified negative effects of the Second Step programme in relation to bullying victimisation outcomes. Evaluations of the ViSC programme also had a negative effect on bullying victimisation, although this effect was not statistically significant.

6.3.4.2 Conflict of interest. Conflict of interest (COI) was a categorical moderator variable with three levels: high-risk (H), low-risk (L); and possible-risk (P). Moderator analysis analogous to the ANOVA was conducted in order to assess the differences between evaluations on each level. Studies categorized as possible-risk on the COI variable were excluded from subgroup comparisons to establish the differences between evaluations that were clearly high-risk and evaluations that were clearly low-risk. Table 25 shows the mean summary effects for each group for both bullying perpetration and bullying victimisation outcomes.

Subgroup analyses found that the difference between high-risk and low-risk studies on the COI variable was statistically significant for bullying perpetration outcomes under both the MVA model ($Q_B = 50.129$; $df = 1$; $p < .001$) and the random effects model ($Q_B = 4.900$; $df = 1$; $p = .027$). This suggests that evaluations considered to have high COI were associated with larger overall effect sizes for bullying perpetration. Similarly, high-risk COI studies were significantly associated with slightly larger effect sizes for bullying victimisation in comparison to low-risk COI studies when compared under both the MVA model ($Q_B = 16.127$; $df = 1$; $p < .001$) and the random effects model ($Q_B = 4.449$; $df = 1$; $p = .03$).

Table 23

School-bullying moderator analysis results: Programme specificity

	MVA Model					Random effects model			
Specificity (*n*)	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
School bullying perpetration (*n* = 85 effect sizes)									
High (66)	1.343	1.285 – 1.403	< .001	279.036 (< .001)	76.706	1.295	1.209 – 1.388	< .001	.004
Medium (14)	1.208	1.038 – 1.404	< .001	108.843 (< .001)	88.056	1.165	1.009 – 1.343	.036	.013
Low (5)	1.014	0.625 – 1.645	.955	24.652 (.001)	83.774	0.996	0.761 – 1.303	.976	.135
School bullying victimisation (*n* = 88 effect sizes)									
High (63)	1.262	1.210 – 1.317	< .001	328.981 (< .001)	81.154	1.292	1.212 – 1.377	< .001	.007
Medium (16)	1.022	0.889 – 1.173	.763	33.055 (.005)	54.621	1.061	0.919 – 1.225	.422	.010
Low (9)	1.059	0.824 – 1.347	.676	25.746 (.001)	68.927	1.008	0.833 – 1.219	.937	.050

Table 24

School-bullying moderator analysis results: Packaged intervention programmes

	MVA Model					*Random effects model*			
Intervention (*n*)	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
School bullying perpetration (*n* = 36 effect sizes)									
BPYS (3)	1.065	0.950 – 1.193	.279	0.252 (.616)	693.651	1.054	0.787 – 1.412	.724	.061
Fairplayer.manual (2)	0.846	0.498 – 1.439	.539	1.198 (.274)	16.528	0.855	0.507 – 1.443	.557	.093
KiVa (6)	1.143	1.075 – 1.215	< .001	9.347 (.096)	46.507	1.180	1.063 – 1.309	.002	.001
NoTrap! (4)	1.378	0.764 – 2.483	.286	18.301 (< .001)	83.607	1.374	1.059 – 1.782	.017	.246
OBPP: Overall (12)	1.532	1.438 – 1.631	< .001	22.292 (.014)	55.141	1.501	1.358 – 1.659	< .001	.002
OBPP: USA (6)	1.473	1.374 – 1.579	< .001	10.604 (.060)	52.848	1.349	1.185 – 1.535	< .001	.002
OBPP: Norway (5)	1.749	1.695 – 1.804	< .001	0.498 (.974)	703.213	1.759	1.503 – 2.059	< .001	.018
Second Step (3)	1.101	1.027 – 1.181	< .001	0.304 (.859)	557.895	1.107	0.879 – 1.395	.387	.029
Steps to Respect (2)	1.160	1.052 – 1.279	< .001	0.609 (.435)	64.204	1.142	0.934 – 1.397	.197	.001
ViSC (5)	0.952	0.730 – 1.241	.714	12.237 (.016)	67.312	0.949	0.785 – 1.149	.596	.045
School bullying victimisation (*n* = 35 effect sizes)									
BPYS (3)	1.349	1.189 – 1.530	< .001	0.734 (.693)	172.48	1.323	0.962 – 1.819	.085	.036
KiVa (6)	1.160	1.033 – 1.302	< .001	41.222 (< .001)	90.296	1.240	1.063 – 1.447	.006	.021
NoTrap! (4)	1.836	1.150 – 2.931	< .001	9.929 (.019)	69.785	1.772	1.296 – 2.424	< .001	.165
OBPP: Overall (12)	1.264	1.158 – 1.379	< .001	102.667 (< .001)	89.286	1.285	1.137 – 1.451	< .001	.039

OBPP: Norway (5)	1.172	1.122 – 1.224	< .001	10.141 (.119)	60.556	1.053	0.899 – 1.233	.522	.017
OBPP: USA (7)	1.566	1.391 – 1.762	< .001	17.579 (.002)	65.868	1.726	1.424 – 2.092	< .001	.016
Second Step (3)	0.807	0.666– 0.977	< .001	1.249 (.535)	60.128	0.832	0.593 – 1.168	.289	.024
Steps to Respect (2)	1.190	1.113 – 1.272	< .001	0.287 (.592)	248.432	1.171	0.884 – 1.551	.273	.008
ViSC (5)	0.952	0.635 – 1.429	.813	20.146 (.001)	80.145	1.004	0.781 – 1.291	.975	.190

Table 25

School-bullying moderator analysis results: Conflict of interest

	MVA Model					Random effects model			
COI-risk (*n*)	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
School bullying perpetration (*n* = 86 effect sizes)									
High (40)	1.375	1.309 – 1.444	< .001	196.882 (< .001)	80.191	1.330	1.232 – 1.435	< .001	.025
Possible (10)	1.390	1.185 – 1.631	<.001	13.468 (.142)	33.175	1.445	1.182 – 1.766	.844	.030
Low (36)	1.146	1.024 – 1.282	.017	214.119 (< .001)	83.654	1.123	0.988 – 1.277	.077	.106
School bullying victimisation (*n* = 89 effect sizes)									
High (40)	1.270	1.213 – 1.329	< .001	218.053 (< .001)	82.114	1.324	1.232 – 1.422	< .001	.022
Possible (10)	1.090	0.957 – 1.241	.192	16.538 (.056)	45.581	1.087	0.908 – 1.301	.365	.030
Low (39)	1.129	1.010 – 1.262	.033	162.359 (< .001)	76.595	1.132	0.997 – 1.285	.056	.101

Note. Four studies and six studies were excluded from the present moderator analysis for perpetration and victimisation outcome respectively as not enough information was available

Part Three: Cyber-bullying

7. Literature Review: Cyberbullying

7.1 Overview

In the past decade, research interest in cyberbullying perpetration and victimisation has grown exponentially (Smith & Berkkun, 2017). With the rapid developments in information communication technologies (ICTs), and their ever-growing presence in our daily lives, research into forms of aggressive behaviours in these environments has been very important. For example, a recent report highlights the prevalence of internet-ready and smart devices amongst individuals of all ages in the United Kingdom (Ofcom, 2017). Of individuals who took part in the survey, 88% reported having access to the internet and 76% reported ownership of a smartphone, in comparison to 77% and only 27% respectively in 2011. Furthermore, 90% of younger users (aged 18 to 24 years old) reported using at least one social media app, and so did 69% of participants over the age of 54 years.

In addition, a recent systematic review of children's rapidly increasing access to ICTs reported that, in the UK, the use of the internet at home increased with age, from 37% of 3-4 year olds, to 58% of 5-7 year olds, 87% of 8-11 year olds, and 95% of 12-15 year olds (Livingstone & Smith, 2014). The ownership of mobile phones, particularly smartphones, and other Internet-ready devices such as tablets, music players and games consoles, is also on the rise, with 62% of 12 to 15-year-olds in 2012 reporting ownership. Thus, as our interpersonal communications move into the online sphere, it is only to be expected that these platforms will increasingly be used for aggressive forms of behaviours (Asam & Samara, 2016). Amongst these behaviours, that occur within the 'global playground' of the online world is cyberbullying (Li, Cross, & Smith, 2012).

The present chapter attempts to provide a brief overview of just some of the recent literature on cyberbullying, with a particular focus on cyberbullying amongst youth in the United Kingdom and Ireland. The literature review presented in this chapter has been published in an edited volume that is concerned with international perspectives on cyberbullying (Baldry, Blaya, & Farrington, 2018).

The prevalence rates of reported cyberbullying behaviours vary greatly between international studies, from 10% to 72% (Marczak & Coyne, 2015). When referring to the United Kingdom, it is worthwhile to note that this term includes England, Scotland, Wales and Northern Ireland. When discussing cyberbullying in Ireland, this is in reference to the Republic of Ireland. The following sections of this paper review cyberbullying research with respect to the definitional criteria, prevalence, risk and protective factors and the associated outcomes, amongst adolescents in the UK and Ireland.

7.2 Definition

Cyberbullying has been defined by UK academics as an aggressive, intentional act carried out by a group or individual, *using electronic forms of contact* repeatedly and over time against a victim who cannot easily defend himself or herself (Smith et al., 2008). However, in contrast to the literature surrounding school-bullying, also often referred to as traditional or offline bullying, there remains a lack of consensus regarding the definition of cyberbullying (Bauman, 2013; Betts, 2016).

There are several aspects of cyberbullying that make it difficult to formulate an accurate and consistent definition. For example the following features are unique to cyberbullying in comparison to traditional school-bullying, such as: the ability of the perpetrator to remain relatively anonymous; the lack of physical and social cues in online communication; the added complexity of the bystanders' roles in cyberbullying; and the fact that there is 'no place to hide' (Marczak & Coyne, 2015; p. 149). Smith, Barrio, and Tokunaga (2013) provide a full and comprehensive overview of the definitional issues, including the permeance of data and the indefinite number of potential bystanders.

For the purposes of the present research, we utilised a definition of cyberbullying that incorporated the three key elements of the definition of school-bullying: (1) intention to harm; (2) repetitive nature; and (3) clear power imbalance (Centers for Disease Control and Prevention, 2014). In addition, definitions of cyberbullying must refer to the occurrence of these aggressive behaviours using ICTs, for example, mobile phones (text messages/phone calls) or the Internet (e.g., email, social networking sites/social media).

However, these three elements may not be as straightforward when applied to cyberbullying in comparison to school-bullying. For example, in comparison to school-bullying (or: offline bullying/ face-to-face bullying/ traditional bullying), cyberbullies may only be perpetrators on a single occasion, but the experience of victimisation could be recurring for the victim. For example, a cyberbully may share or distribute embarrassing images of a victim on one occasion but, as others ‘like’ or share the content further, the victimisation is repeated but the perpetration of the act is not. Therefore, the concept of repetition is difficult to define in cyberbullying, as it may have different implications for cyberbullies and cyber-victims. Moreover, a qualitative study conducted with youth in the UK indicated that children were aware that cyberbullying occurred typically as an extension or continuation of offline bullying (Betts & Spenser, 2017). Participants in this study reported how, in comparison with school bullying that has a clear cut-off point (typically when the victim goes home from school), cyberbullying experiences had the potential to occur at any time of the day or night because of constant access to, and engagement with, technology (Betts & Spenser, 2017).

The complexity of cyberbullying is partially attributable to the significantly large number of different behaviours that it may encompass (Marczak & Coyne, 2015). For example, Willard (2006) identified seven potential categories of cyberbullying behaviours: flaming, online harassment, cyberstalking, denigration, masquerade, outing, and exclusion.

However, these categories, although proposed only 11 years ago, may already be outdated or incomplete because of the rapid rise and development of social media platforms and sharing apps that could facilitate cyberbullying. For example, when this typology was suggested, the vastly popular picture-sharing app Snapchat was not in existence.

More recent research suggests the need to identify a wider array of potential cyberbullying behaviours. For example, Rivers and Noret (2010) identified ten categories of behaviours: threat of physical violence, abusive or hate-related, sexual acts, demands or instructions, threats to damaging existing relationships, threats to family or home, and menacing chain messages. Moreover, Nuccitelli (2012) proposes over thirty-six different behaviours that could be considered cyberbullying. Other studies propose broader categories, such as: sexting, trolling, and ‘griefing’ (Slonje, Smith, & Frisén, 2013); or direct and indirect cyberbullying (Langos, 2012). Direct cyberbullying includes behaviours that occur exclusively between the perpetrator(s) and the victim(s), for example aggressive content sent via text/instant messages and/or phone calls, or exclusion from online groups. Indirect cyberbullying occurs in the public online environment, for example, publicly posting hurtful or embarrassing posts and/or pictures about an individual or the creation of public forums targeting the victim specifically.

7.3 Prevalence

Given that there are still differences in operational definitions and methods of measuring cyberbullying, estimating accurate prevalence rates of cyberbullying perpetration and victimisation is hard. However, recent reviews have suggested that cyberbullying is quite prevalent amongst school-aged populations globally, for example, in Canada (Riddell, Pepler, & Craig, 2018); the UK and Ireland (Gaffney & Farrington, 2018); and the US (Espelage, Hong, & Valido, 2018). The prevalence of cyberbullying victimisation and perpetration can vary across a number of demographic (e.g., gender, age, race/ethnicity, sexuality, disability), and individual (e.g., experiences with traditional bullying, personality, weight status, technology use), peer, family, and school risk and protective factors (Kowalski, Limber, & McCord, 2018).

A recent meta-analysis of 80 studies found that, while prevalence rates of cyberbullying were lower than those for traditional school-bullying, there were significant correlations between these types of aggressive behaviours (Modecki et al., 2014). Moreover, the perceived impact of cyberbullying has been frequently reported to be worse than the impact of face-to-face or traditional school-bullying, but this relationship may vary according to the type of cyberbullying experienced (Smith et al., 2006). In the Republic of Ireland, participants also thought that all forms of cyberbullying behaviours had more impact than traditional school-bullying, with the exception of bullying via email (Cotter & McGilloway, 2011).

An early study of cyberbullying prevalence in the UK was conducted in 2002, with Year 8 students (N = 779; mean age = 12 years old) reporting how often they had received nasty emails or text messages (Oliver & Candappa, 2003). This seminal study reported that 4% of children reported receiving nasty text messages, and 2% reported receiving nasty emails. Subsequent research studies identified higher prevalence rates from data collected between 2002 and 2006. Rivers and Noret (2010) reported the prevalence of ‘receiving nasty text messages or emails’ in a sample of British adolescents aged 11 to 14 years old. The results are presented for this five-year study independently for each year of data collection and for males and females separately. The results show a steady increase in the rate of cyberbullying victimisation experienced by girls from the first point of data collection in 2002 (14.1%) to 2005 (21.3%). The rates declined slightly in 2006 to 20.8% of girls reporting receiving nasty text messages or emails. The figures for boys were less consistent.

In 2005, a study conducted with UK adolescents aged 11 to 16 years old (Smith et al., 2008; Study 1) found that a maximum of 1.1% of children reported cyberbullying others via phone calls, texts, emails and/or instant messages outside school, and 2.3% reported cyberbullying others via email inside school more than once or twice. Cyberbullying victimisation varied from 1.1% (via websites outside school), to 3.3% (via phone calls, texts, emails inside school) and 10.9% reporting bullying victimisation via phone calls outside school more than once or twice in the past couple of months (Smith et al., 2008). This pilot study was subsequently followed up and revealed higher incidence rates of children reporting having 'ever' cyberbullied someone (from 1% via chat rooms to 5.3% via instant messaging) or having ever been a victim of cyberbullying (from 2.5% via chatrooms to 9.5% via phone calls and 9.9% via instant messaging; Smith et al., 2008).

The prevalence of cyberbullying perpetration and victimisation has varied greatly in more recent studies, from 2.5% (mobile bullying perpetration), 3% (internet bullying perpetration), 4.1% (mobile bullying victimisation) and 6.6% (internet bullying victimisation) for 2,227 Year 8 to 12 students in 2008 (Genta et al., 2012) to 13.5% of 1,144 Year 8 students reporting engaging in cyberbullying perpetration less than once a week (Fletcher et al., 2014). Also, in 2008, 5% and 20.5% of primary school children in England, aged 7 to 11 years old, self-identified as cyber-bullies and cyber-victims respectively (Monks et al., 2012). Ackers (2012) discovered that 11% of 325 Year 7 to 9 students from one secondary school in the UK responded that they had been cyberbullied, while 7% of the sample responded that they had cyberbullied someone else. Research conducted with older adolescents (90 students aged 16 to 18 years old) found that 13.5% and 16.2% of children reported cyberbullying perpetration and victimisation respectively (Brewer & Kerslake, 2015).

National data collected via the Longitudinal Study of Young People in England (of Year 10 students in 2014) discovered that 11% of children reported cyberbullying victimisation (Lasher & Baker, 2015). Moreover, the international EU Kids Online survey concluded that 8% of UK children reported cyberbullying victimisation (Livingstone et al., 2011). An exploratory study of cyber-aggression and cyber-victimisation found that 31.5% and 56.2% of 339 Year 7 to 9 students in one UK secondary school reported engaging in and experiencing cyber-aggression and cyber-victimisation respectively (Pornai & Wood, 2010). A more recent study (Bevilacqua et al., 2017) discovered that, among a sample of Year 7 students from 40 English schools, 1.6% and 6.4% of children reported cyberbullying perpetration and victimisation respectively. Among older children, prevalence rates of cyberbullying perpetration (1.9%) and victimisation (7.9%) were slightly higher (West, 2015).

Other studies have categorized children according to their self-reported involvement in cyberbullying. Del Rey and colleagues (2015) found that 0.9% of 737 UK students were categorized as aggressors of cyberbullying, 2.0% were categorized as bully-victims, and 6.4% were victims of cyberbullying. In this large-scale European study, the prevalence of cyberbullying perpetration among UK adolescents was relatively low in comparison to the overall sample that included children from countries such as Italy, Greece, Poland, Spain and Germany. However, the number of children in the UK who reported cyberbullying victimisation was in line with the mean prevalence reported by the total sample (6.4% compared with 6.8%). Wolke and colleagues (2017) found that 1.1% of 2,754 UK adolescents aged 11 to 16 years old were classified as ‘pure cyber-victims’, in other words, being victimized online only.

In Northern Ireland, a government report concluded that 1.1% and 3.5% of primary and secondary school students reported cyberbullying perpetration and victimisation respectively (McClure Watters, 2011). Additionally, 3.7% of 425 Year 9 to 11 students from two secondary schools reported experiencing cyberbullying victimisation. Results from a nationally disseminated survey (Kids Life and Times) showed that 13.8% of Northern Irish adolescents reported cyberbullying victimisation (Devine & Lloyd, 2012). Moreover, among a sample of nearly 3,500 children, aged 11, attending 217 Northern Irish primary schools, 10.3% reported experiencing cyberbullying victimisation (McGuckin et al., 2010).

Seminal research on cyberbullying in the Republic of Ireland in 2011 found that 9% and 17% of secondary school students reported cyberbullying perpetration and victimisation, respectively (Cotter & McGilloway, 2011). In addition, this study found that the majority of cyberbullying perpetration and victimisation reported by participants was experienced outside school. Furthermore, O'Moore (2012) reported that 4.4% of over 3,000 secondary school students were classified as pure cyber-bullies, 4.1% were categorized as bully-victims, and 9.8% were categorized as pure cyber-victims. International studies, that have included Irish children, have found that 4% of adolescents reported cyberbullying victimisation (Livingstone et al., 2011). Similarly, Corcoran and colleagues (2012) reported that 2.6% of post-primary Irish adolescents reported cyberbullying perpetration, and 6.3% reported cyberbullying victimisation. A recent study conducted in the Republic of Ireland concluded that 9.8% of Irish adolescents aged 15 to 18 years old (N = 318) had experienced cyberbullying victimisation (Callaghan et al., 2015).

Previous research has found that cyberbullying is associated with several undesirable psychological, behavioural and health-related outcomes. For example, studies conducted in Europe have discovered that cyber-victims report higher levels of emotional and social problems, psychological difficulties, headaches, abdominal pain, and sleeping difficulties (Sourander et al., 2010). In addition, the cyber-bullies who were identified in this study reported higher frequencies of conduct problems, hyperactivity, smoking and alcohol use. Additionally, cyberbullying victimisation is correlated with several undesirable mental health outcomes, such as depression, anxiety, and suicidal ideation (Betts, 2016). Thus, given the prevalence of cyberbullying amongst youth in the UK and Ireland, cyberbullying is an important area for research.

7.4 Outcomes and impact

Media reports in the UK and Ireland have covered several tragic cases of teenage suicide, attributed to experiences of victimisation online, have heightened public awareness and concern about cyberbullying. For example, Felix Alexander, aged 17 from Worcester, tragically committed suicide in 2016 after years of being bullied. Felix's mother wrote that online bullying had exacerbated the effect that victimisation had on her son, and that in an effort to prevent the online attacks he had removed himself from multiple social media sites. However, in doing so, this increased his feelings of social isolation (The Guardian, October 5, 2016).

Since the emergence of cyberbullying as an important topic for research, many studies have concluded that both cyberbullying perpetration and victimisation may lead to undesirable behavioural and health-related outcomes. Studies conducted in Europe found that adolescents who reported experiences of cyber-victimisation had high levels of emotional and social problems, headaches, abdominal pain, and trouble sleeping (Sourander et al., 2010). Moreover, this study also showed that adolescents categorized as cyber-bullies were more likely to report high levels of conduct problems, hyperactivity, smoking cigarette, alcohol consumption, and headaches. Betts (2016) found that cyber-victimisation too was correlated with high levels of mental health problems such as depression, anxiety, and suicidal ideation. Numerous studies have investigated a plethora of outcomes relating to cyberbullying (e.g., Brewer & Kerslake, 2015; Brown, Demaray, & Secord, 2014; Calvete, Orue, & Gámez-Guadix, 2016), but, in comparison to school-bullying, we do not yet know the longitudinal outcomes of cyberbullying (Baldry, Farrington, & Sorrentino, 2015).

Recent international empirical studies have demonstrated the impact that cyberbullying can have on adolescent mental health when the interactive nature of online and offline bullying was controlled. Baier and colleagues (2019) found that 'psychological cyberbullying' (e.g., rumours or negative gossip spread online) significantly predicted higher depression and anxiety scores, as did 'sexual cyberbullying' (e.g., victim was sent unwanted sexual images online or told to engage in non-consensual online sexual activity) in German adolescents. In relation to the latter finding, there were stark gender differences. Namely, sexual cyberbullying experiences significantly predicted worse mental health outcomes for female participants but not male participants (Baier et al., 2019). Experiences of psychological and sexual cyberbullying also significantly predicted greater incidences of somatic symptoms amongst the participants (Baier et al., 2018).

There is also an important impact of the overlap of online and offline victimisation on the relationship between cyberbullying and mental health outcomes. For example, in the US, studies have found that whilst both forms of victimisation (i.e., online and offline) significantly predicted negative mental health outcomes, the relationship between cyberbullying and outcomes did not remain significant when controlling for offline bullying (Hase et al., 2015). However, the inverse interaction was significant. In other words, when controlling for cyberbullying, offline victimisation did significantly predict negative mental health outcomes.

Yet despite potential inconsistent research in the primary research, recent meta-analytical reviews have suggested a strong impact of cyberbullying experiences and an array of worrying outcomes. For example, meta-analysis has found that cyber-victims were more than two times more likely to also report self-harm (OR = 2.35, 95% CI 1.65 – 3.34), have suicidal thoughts (OR = 2.15, 95% CI 1.70 – 2.71), exhibit suicidal behaviours (OR = 2.10, 95% CI 1.73 – 2.55), and attempt suicide (OR = 2.57, 95% CI 1.69 – 3.90) in comparison to their non-victimized peers (John et al., 2018). Perpetrators of cyberbullying were also more likely to report suicidal behaviours (OR = 1.02 – 1.44) and suicidal ideation (OR = 1.23, 95% CI 1.10 – 1.37) in comparison to non-cyberbullies (John et al., 2018).

There is clearly a significant need for effective intervention and prevention programmes, but there is a significant lack of research in this area. In developing effective intervention and prevention programmes to combat the issue of cyberbullying, an understanding of the risk and protective factors associated with these aggressive behaviours is needed.

7.5 Risk and protective factors

A large-scale review assessed risk factors associated with cyberbullying perpetration and victimisation as measured by 53 studies conducted in various international locations (Baldry, Farrington, & Sorrentino, 2015). This review categorized factors according to a socio-ecological framework (Bronfenbrenner, 1979), with risk factors identified at the individual (e.g., technology use, personality traits, values), peer and family (e.g., pro-social peers, peer rejection, parental support), and school (e.g., lack of teacher support, negative school climate) levels. This theoretical framework is commonly used to explain risk factors associated with cyberbullying (e.g., Cross et al., 2015).

This section will further explore the potential risk factors and predictors that are associated with cyberbullying perpetration and victimisation as measured in UK or Irish samples. The included studies measured mainly individual-level factors, such as gender, ethnicity, demographics, traditional bullying perpetration and victimisation, and various psychological and cognitive constructs. In addition, some school-level variables have been studied. Because of the general lack of longitudinal studies, it is difficult to draw conclusions about prediction or about causal effects.

7.5.1 Gender

Assessing the prevalence rates reported by studies conducted with samples in the United Kingdom and Ireland, it appears that girls report, on average, higher rates of cyber-victimisation than boys, and boys report, on average, higher rates of cyberbullying perpetration. Bevilacqua and colleagues (2017) concluded that 1.13% of males and 0.45% of females reported frequent cyberbullying perpetration, and 1.94% of males and 4.48% of females reported frequent cyberbullying victimisation. In Northern Ireland, female adolescents (15%) reported statistically significant higher rates of cyberbullying victimisation compared to their male peers (11%; χ^2 =18.45, df = 2, $p < 0.001$; Devine & Lloyd, 2012).

Similar results were found by Pornai and Wood (2010), with females (58.8%) reporting higher rates of cyber-victimisation compared to males (53.2%). Of the children who were categorized as 'pure cyber-victims' (i.e., those reporting experiencing bullying victimisation online only) in Wolke and colleagues' study (2017), 58.1% were female. Ackers (2012) concluded that there was a significant main effect for gender in self-reported cyberbullying victimisation, with females being more likely to report being victimized. The frequency of receiving nasty or threatening text messages or emails varied between 10.3% and 12% for boys, but it was higher for girls, varying between 14.1% and 21.3% (Rivers & Noret, 2010). An exploratory study of 1,144 year 8 students in UK secondary schools concluded that males (14.7%) were more likely than females (13.4%) to report engaging in cyberbullying perpetration (OR) = 0.91; 95% CI 0.64 - 1.28), although this difference was not statistically significant (Fletcher et al., 2014).

In the Republic of Ireland, however, one study found that boys reported more cyberbullying victimisation than females, 10.3% and 9.2% respectively (Callaghan et al., 2015). In comparison, employing a sample of Irish adolescents, aged 12 to 16 years old, O'Moore (2012) classified more girls (15.6%) as pure victims of cyberbullying than boys

(6.9%). This study categorized more boys (4.9%) as pure bullies than girls (3.5%), however, more girls (4.5%) were classified as bully-victims than boys (3.9%). Similarly, Pornai and Wood (2010) found that girls were more likely to report cyber-aggression perpetration than boys (OR = 1.66, $p < 0.05$). However, some studies found no significant association between gender and cyberbullying perpetration or victimisation (e.g., Monks et al., 2012).

7.5.2 Ethnicity and demographic variables

A few of the studies conducted in the UK and Ireland considered the impact of several demographic and sociodemographic variables. One study found that males of mixed ethnicity (4.5%) and females identifying as Black or Black British (0.8%) were more likely to report engaging in frequent cyberbullying perpetration (Bevilacqua et al., 2017). Both males and females identifying as White 'Other' (3.4% and 5.2% respectively) were more likely to report frequent cyberbullying victimisation. An analysis of the relationship between ethnicity and cyberbullying perpetration in another study suggested that, in comparison with children identifying as White British (11.6%), those of dual heritage (20%; OR = 1.92; 95% CI 1.09 - 3.40) and other ethnicity (19.1%; OR = 1.76; 95% CI 1.03 - 3.00) were more likely to report cyberbullying perpetration (Fletcher et al., 2014). The differences between children identifying as White British and those identifying as Asian or Asian British (9.9%; OR = 0.83, 95% CI 0.47 - 1.48) or as Black or Black British (17.2%; OR = 1.55, 95% CI 0.97 - 2.48) were not statistically significant. However, due to the lack of research the relationship, if any, between cyberbullying and BAME youth in the UK and Ireland is not fully understood.

Fletcher and colleagues (2014) found no differences in cyberbullying perpetration reported by students according to family structure (i.e., living with two parents, one parent, or other). Adolescents who reported having unemployed parents (21.1%) were more likely to engage in cyberbullying perpetration than students with parents in employment (13.8%; OR = 1.6; 95% CI 0.98 - 2.6), although this effect was not quite statistically significant.

7.5.3 Traditional school-bullying and victimisation

The most common finding by studies conducted with children in the United Kingdom and Ireland is that there is a significant relationship between school bullying perpetration and victimisation offline and cyberbullying perpetration and victimisation. Previous research has found that there is a distinct overlap between offline and online victimisation, with individuals participating in both acts as pure offline and pure online bullies and victims, but also various combinations of online and offline bullies, victims and bully-victims (Schultze-Krumbholz et al., 2015). For example, in the study conducted by Wolke and colleagues (2017), 8.1% and 5.8% of children reported experiencing victimisation as a result of direct and relationship bullying respectively, while only 1.1% of children reported experiencing only cyber-victimisation. In addition, 5.1% of children reported experiencing direct, relational and online bullying victimisation. Typically, reports of offline bullying perpetration and victimisation are higher than those reported for online perpetration and victimisation (e.g., Bevilacqua et al., 2017; Cotter & McGilloway, 2011; Livingstone et al., 2011; Monks et al., 2012; O'Moore, 2012).

Pornai and Wood (2010) concluded that, among a sample of UK adolescents, high levels of traditional aggression correlated with an increased likelihood of an adolescent being a cyber-bully ($B = 0.24$; SE = 0.03; $p < 0.001$). Similarly, high levels of traditional victimisation correlated with an increased likelihood of being a cyber-victim ($B = 0.10$; SE = 0.02; $p < 0.001$), but with a decreased likelihood of being a cyber-bully ($B = -0.09$; SE = 0.03; $p = 0.001$). Fletcher et al. (2014) also investigated the relationships between self-reported aggressive behaviour at school and the frequency of cyberbullying perpetration. Students who reported higher levels of aggressive behaviour in school were significantly more likely to also report cyberbullying perpetration (37.1%; OR = 14.35; 95% CI 7.96 - 25.86), in comparison with those reporting lesser degrees of in-school aggression. In a sample of primary school children, being a traditional victim was a significant predictor of being a cyber-victim, but not a cyber-bully. Furthermore, being a traditional bully was a significant predictor of being a cyber-bully, but not a cyber-victim (Monks et al., 2012). However, when age is taken into consideration, the relationship between traditional bullying and cyber-bullying may change. For example, O'Moore (2012) found that 32% of post-primary cyber-bullies reported traditional bullying victimisation, while 28.9% of cyber-victims reported engaging in traditional bullying perpetration.

7.5.4 Cognitive and psychological factors

Five studies reviewed in this chapter (i.e., Brewer & Kerslake, 2015; Corcoran et al., 2012; Fletcher et al., 2014; Wolke et al., 2017) investigated the relationship between cyberbullying and different cognitive or psychological factors. Because of the infrequency of longitudinal studies, it is unclear whether these are risk factors for, or outcomes of, cyberbullying perpetration and victimisation, but the results are important to guide future research.

In an adjusted multi-level regression model, Wolke and colleagues (2017) found that pure cyber-victimisation was significantly related to lower self-esteem (B = -2.19, p = 0.004) and higher levels of self-reported behavioural difficulties (B = 4.13, p > 0.001). Furthermore, when effect sizes were adjusted for demographic variables, interesting relationships were observed between self-reported cyberbullying perpetration of UK adolescents and their psychological functioning, overall mental wellbeing and several aspects of mental and physical health (Fletcher et al., 2014[15]). Based on a measure of psychological functioning

[15] Only statistically significant relationships are reported here. For a full overview see Fletcher et al., 2014, table 3, p. 1396

(the Strengths and Difficulties Questionnaire; Goodman, 2006), Fletcher et al. (2014) suggested that children with greater overall difficulties (OR = 2.32; 95% CI 1.97 - 3.24) and greater conduct problems (OR = 1.3; 95% CI 1.08 - 1.55) were more likely to report bullying others online in comparison with children reporting fewer overall difficulties or fewer conduct problems. Significant negative relationships were observed between cyberbullying perpetration and the quality of life (OR = -3.51; 95% CI -5.7 - -0.1), psychosocial health (OR = -5.04; 95% CI -7.26 -1.6), emotional functioning (OR = -5.6; 95% CI -9.03 - -0.18) and school functioning (OR = -7.35, 95% CI 9.27 - -4.95; see Fletcher et al., 2014).

In Northern Ireland, Devine and Lloyd (2012) observed that adolescents who experienced cyberbullying victimisation reported significantly poorer overall psychological well-being (t (1, 3382) = 10.77, $p < 0.001$). In the Republic of Ireland, Corcoran and colleagues (2012) discovered interesting relationships between aspects of participants' self-concepts, measured using the Piers-Harris 2 (Piers & Herzberg, 2002) instrument. Cyber-victims scored lower on overall general self-concept and the 'freedom of anxiety' subscale, in comparison to non-involved groups. For UK adolescents, Brewer and Kerslake (2015) found that cyberbullying victimisation was significantly and positively correlated with loneliness (r = 0.8, $p < 0.01$) and negatively correlated with self-esteem (r = -0.42, $p < 0.01$). In addition, cyberbullying perpetration was significantly negatively correlated with loneliness (r = -0.38, $p < 0.01$) and self-esteem (r = -0.22, $p < 0.01$). Based on standard regression models, low self-esteem was significantly related to cyberbullying victimisation. Low levels of empathy and self-esteem were also significantly related to cyberbullying perpetration (Brewer & Kerslake, 2015).

Pornai and Wood (2010) also conducted exploratory analyses of several individual cognitive factors and cyber-aggression perpetration and victimisation among UK adolescents. The results indicated that the moral justification facets of moral disengagement were related to cyber-aggression perpetration ($B = 0.20$, SE = 0.04; $p < 0.001$). Moreover, hostile attribution bias was significantly related to cyber-aggression victimisation ($B = 0.12$, SE = 0.04, $p < 0.05$). Finally, Corcoran and colleagues (2012) investigated the relationship between cyberbullying perpetration and victimisation and personality, as measured by the Junior Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975). Significant differences were observed between groups (i.e., cyber-bullies, cyber-victims, traditional-bullies, traditional-victims, non-involved) on both psychoticism and neuroticism scores. Specifically, the cyber-victim group reported significantly higher scores on the neuroticism scale compared to the non-involved group.

7.5.5 School-level factors

Bevilacqua and colleagues (2017) further investigated the relationship between several school-level variables and the frequency of self-reported cyberbullying perpetration and victimisation. Effect sizes, adjusted for all individual-level variables, such as gender and ethnicity, evaluated the relationship between the proportion of children eligible for free school meals, the Income Deprivation Affecting Children Index score, and the most recent overall Ofsted rating, and the prevalence of cyberbullying perpetration and victimisation. Moreover, school type (e.g., community – funded by local authorities; voluntary-aided – funded by a charity and partially by local authorities; sponsor-led academies and foundation schools), size and sex composition were also investigated in relation to cyberbullying and cyber-victimisation. Significant relationships were found for the impact of the proportion of students eligible for free school meals (adjusted OR = 1.02, 95% CI 1.002 - 1.05), community schools (adjusted OR = 4.25, 95% CI 1.54 - 11.71), foundation schools (adjusted OR = 4.73, 95% CI 1.83 - 12.26), and the 'requires improvement' Ofsted rating (adjusted OR = 4.01, 95% CI 1.05 - 15.24) versus cyberbullying perpetration. These results suggested that cyberbullying perpetration was more likely to occur in schools with lower socioeconomic demographics and poor national ratings. In relation to cyberbullying victimisation, no statistically significant effects were found.

The majority (74.2%) of pure cyber-victims, categorized by Wolke and colleagues (2017), were from schools that were not eligible for the pupil premium (an indicator of deprivation and special assistance within schools). This study also investigated the relationship between cyber-bullying and parental education. The majority of pure cyber-victims reported that their parents had 12 to 13 years of education; 32.3% reported parental education of more than 13 years, and 6.5% reported that their parents had spent less than 11 years in full-time education (Wolke et al., 2017). In Northern Ireland, the prevalence of cyber-victimisation was higher among students attending a school in an urban location (4.3%) compared to those attending a smaller school in a rural location (2.7%; Purdy & York, 2016). These results suggest that adolescents who self-report cyberbullying perpetration are also more likely to report a wide range of psychological and social problems. This is an important observation to better inform cyberbullying intervention and prevention programmes in the UK and Ireland.

7.6 Effective intervention

Walker, Craven, and Tokunaga (2013) pointed out that there is currently a pressing need for meta-analyses that evaluate the effectiveness of intervention and prevention programmes on cyberbullying outcomes. To date, there has only been one published meta-analysis that aimed to estimate the effectiveness of cyberbullying intervention programmes (i.e., Mishna et al., 2011). However, this review was concerned with 'cyber abuse' and not specifically with cyberbullying behaviours. Furthermore, as searches were conducted in 2009, only three eligible studies were included, and the evaluated programmes mainly focused on issues of internet safety (Mishna et al., 2011).

Narrative reviews have pointed to potential strategies for cyberbullying intervention, including strategies such as empathy training (Ang, 2015), educational campaigns (Chisholm, 2014), or programmes developed through collaborative work with adolescent participants (Ashtorab & Vitak, 2016). While meta-analyses of risk factors have suggested techniques and approaches for preventing and/or reducing cyberbullying perpetration and victimisation (e.g., Chen, Ho, & Lwin, 2017), there is a current gap in the literature for an extensive systematic and meta-analytical review of the effectiveness of cyberbullying intervention programmes.

Researchers have indicated the importance of utilising evidence-based research on school-bullying interventions to better inform cyberbullying intervention and prevention (Pearce et al., 2011). Furthermore, reviews of theoretical frameworks have highlighted top-down methods of cyberbullying intervention and prevention (e.g., the Barlett and Gentile Cyberbullying Model; Barlett, 2017), and the effectiveness of implementation of anti-cyberbullying programmes remains unclear. Therefore, the aim of the present research is to address this gap in the literature and evaluate the effectiveness of cyberbullying intervention and prevention programmes.

7.6.1 Legal aspects

In comparison to the United States, there is currently no law in place in the UK or Ireland that criminalizes cyberbullying behaviours (Marczak & Coyne, 2010). Some researchers have described cyberbullying as being in a state of legal limbo (Asam & Samara, 2016). However, current legislation in the United Kingdom specifies that all schools must have a clearly defined anti-bullying policy (Marczak & Coyne, 2010; the *School Standards and Framework Act, 1998*). Furthermore, the Education and Inspections Act (2006) gives teaching professionals powers to regulate students' behaviour in school, including the ability to confiscate personal ICTs (Asam & Samara, 2016). As pointed out in this chapter, there is quite frequently an overlap in experiencing traditional and cyber- bullying. Therefore, it is pertinent for UK schools to incorporate elements targeting cyberbullying into these anti-bullying policies. In addition, teachers are key players in cyberbullying intervention and prevention. By removing ICTs from a student's possession, they are able to physically stop cyberbullying perpetration from taking place in school.

There are ways in which online aggression, that may amount to cyberbullying, can be prosecuted in the UK. For example, online hate crimes have recently received media attention, with sources specifying that the Crown Prosecution Service in the UK will start to seek harsher penalties for abuse perpetrated online via social media sites such as Twitter and/or Facebook. The Director of Public Prosecutions stated recently that the criminal justice system in the United Kingdom must start to handle cases of online hate crimes as seriously as it handles offences that occur face-to-face (Dodd, 2017). Recent news stories have highlighted the extreme levels of hate and abuse that those in the public eye receive online. For example, Olivia Attwood, who appeared on a reality-style dating show aired on ITV2, received abuse that was so bad that she could not disclose it on live television (BBC, 2017). Celebrities are not the only ones who are subject to such abuse. Cyber-bullying and general cyber-aggression are becoming increasingly common in our society, as communications rapidly increase in the online sphere.

7.6.2 School-based intervention and prevention

School-based anti-bullying programmes have been widely researched internationally, with results indicating that they can be effective in reducing traditional bullying perpetration and victimisation (e.g., Farrington & Ttofi, 2009). Thompson and Smith (2012) conducted a large-scale review of anti-bullying policies in UK schools. Their evaluation found that anti-bullying efforts in UK schools occurred at several different levels, including whole-school, classroom, and playground strategies. As the current chapter has shown, in the UK and Ireland, traditional and online bullying commonly overlap, so that it is important that schools in the UK and Ireland integrate cyberbullying into their existing anti-bullying policies. More recently, a content analysis of anti-bullying policies in schools in Northern Ireland revealed that the majority of schools incorporate elements targeting cyberbullying (Purdy & Smith, 2016). Additionally, the ‘Quality Circles’ approach has been employed in schools in order to tackle the problem of cyberbullying (Paul, Smith, & Blumberg, 2010).

Large numbers of parents in the Republic of Ireland report that they are aware of the risk posed by cyberbullying, and they are either worried or unsure about whether their children are exposed (O’Higgins Norman, O’Moore, & McGuire, 2016). Moreover, head teachers of secondary schools in Northern Ireland and the Republic of Ireland report that cyberbullying is prevalent in their schools and that they are frustrated with their attempts to handle this complex problem (Purdy & McGuckin, 2015). Research has investigated the factors that predict teachers’ intention to intervene in bullying, including cyberbullying, scenarios. Boulton and colleagues (2014) concluded that the three significant predictors of willingness to intervene were ratings of empathy, coping, and severity of the behaviours. Therefore, the inclusion of parents and teachers in school-based cyberbullying intervention and prevention efforts is very important.

7.7 Conclusions

The effectiveness of several widely disseminated anti-bullying programmes in reducing cyberbullying perpetration and victimisation have been evaluated internationally, including the KiVa programme in Finland (Williford et al., 2013) and the NoTrap! programme in Italy (Menesini, Nocentini, & Palladino, 2012; Palladino, Nocentini, & Menesini, 2016). However, the overall effectiveness of cyberbullying intervention and prevention programmes is not yet understood. Given the increased likelihood of those involved in cyberbullying also experiencing negative psychological and behavioural outcomes, there is a definite need for more research in this field. There are also several factors associated with cyberbullying that have been outlined in this chapter which could be targeted by intervention and prevention programmes. It is noteworthy, that several of these factors make also be considered outcomes, or as occurring as a result of, online bullying. However, given the lack of existing longitudinal research in this area, this remains unclear.

8. Systematic Review: Cyberbullying

8.1 Inclusion and exclusion criteria

Inclusion criteria applied in the systematic review of school-bullying intervention programmes were adapted slightly to identify evaluations of anti-bullying programmes that explicitly targeted cyberbullying behaviours. Specifically, the following criteria were used. To be included in the systematic review of cyberbullying behaviours, primary studies must:

(1) Describe an evaluation of an anti-bullying programme and/or a programme designed to reduce cyberbullying implemented in schools with school-age participants (depending on the site of evaluation, ages may vary between 4 – 18 years of age);
(2) Utilise an operational definition of cyberbullying that coincides with accepted definitions (e.g., Smith et al., 2008);
(3) Measure cyber-bullying perpetration and/or victimisation using quantitative measures, such as, self-, peer-, or teacher-report questionnaires; and
(4) Use an experimental or quasi-experimental design, with one group receiving the intervention and another (control group) not receiving the intervention; and
(5) Have been published from 2000 onwards.

Previous research has suggested that Finkelhor, Mitchell, and Wolak (2000) were potentially the first to discuss ‘online harassment’, and that most cyberbullying research has been conducted since the turn of the millennium (Smith & Berkkun, 2017; Völlink et al., 2016). Existing systematic reviews of cyberbullying issues have pointed out that studies relating to cyberbullying intervention and prevention began to emerge from 2011 onwards (Zych, Ortega-Ruiz, & Del Rey, 2015). Therefore, the searches for evaluations of cyberbullying intervention programmes were limited to peer-reviewed and unpublished studies during and after the year 2000.

Moreover, as previously discussed, there is still much debate in the literature regarding an agreed-upon definition of cyberbullying, and about what behaviours should be encompassed under this definition. For the purpose of the present meta-analysis, cyberbullying was defined as an: aggressive, intentional act, carried out by a group or individual, using electronic forms of contact (usually mobile phones or the Internet) repeatedly and over time against a victim who cannot easily defend themselves (Smith et al., 2008). However, variations on this definition were also considered for inclusion. For example, definitions of cyberbullying could also be tailored to include characteristics that are unique to online environments, for example, anonymity and publicity (e.g., Menesini et al., 2012b). The present systematic review excluded studies if outcomes related to: Internet harassment (e.g., Ybarra & Mitchell, 2004); online harassment (e.g., Finkelhor et al., 2000); or electronic aggression (e.g., Pyżalski, 2012) as these are difficult to distinguish from cyberbullying behaviours. Outcomes of traditional school-bullying perpetration and/or victimisation were also excluded, as these were included separately in the school-bullying meta-analysis.

The present systematic review is concerned with evaluating anti-cyberbullying programmes implemented with school-aged children and adolescents. Therefore, to be included, studies must discuss the implementation of an intervention programme with school-aged participants (i.e., typically aged between 4 and 18 years). As a result, studies using University students, juvenile delinquents, or clinical samples were excluded. In addition, studies that targeted cyberbullying occurring amongst adults in the workplace, or elsewhere, were excluded.

8.2 Searches

Multiple searches were conducted using combinations of the following keywords: *cyber; bully*; victim*; "cyberbullying"; "cyber-victimisation"; "cyber aggression"; "electronic bullying"; "online bullying"; intervention; prevention; programme*; evaluation; effective*.* Searches were conducted for the time period 2000 to end December 2017 on various online databases (i.e., Web of Science, Scopus, PsychINFO, PsychARTICLES, Google Scholar, DARE, ERIC, and ProQuest).

In addition to searching online databases, past issues of specific journals were also hand-searched for relevant studies, as there has been a rapid increase in the number of academic journals dedicated to research conducted on online environments (Livingstone & Smith, 2014; Smith & Berkkun, 2017). For example, past issues of the following journals were hand-searched for potentially includable studies: *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*; *Cyberpsychology, Behaviour, and Social Networking*; *Journal of Children and Media*; *Computers in Human Behaviour*; and *Computers and Education*.

Finally, the searches of the literature identified a number of existing systematic reviews of issues relating to cyberbullying (e.g., Zych et al., 2015). Therefore, studies included (and excluded) by these reviews were also screened for potentially includable primary evaluations of cyberbullying intervention and prevention programmes.

8.3 Screening

Searches for studies evaluating the effectiveness of intervention programmes implemented in schools to reduce cyberbullying behaviours returned a total of 3,994 results. The title and abstract of each result were screened for potential eligibility, and 192 studies were retained for further screening. Included in this number were eight studies identified by searches conducted for, but not included in, the systematic meta-analytical review of school-bullying intervention and prevention programmes (see Chapter 4). Additionally, five studies identified in searches for, and included in, this systematic review of school-bullying intervention programmes were also eligible for inclusion in the present meta-analysis. This screening process is represented using a flowchart (Figure 8).

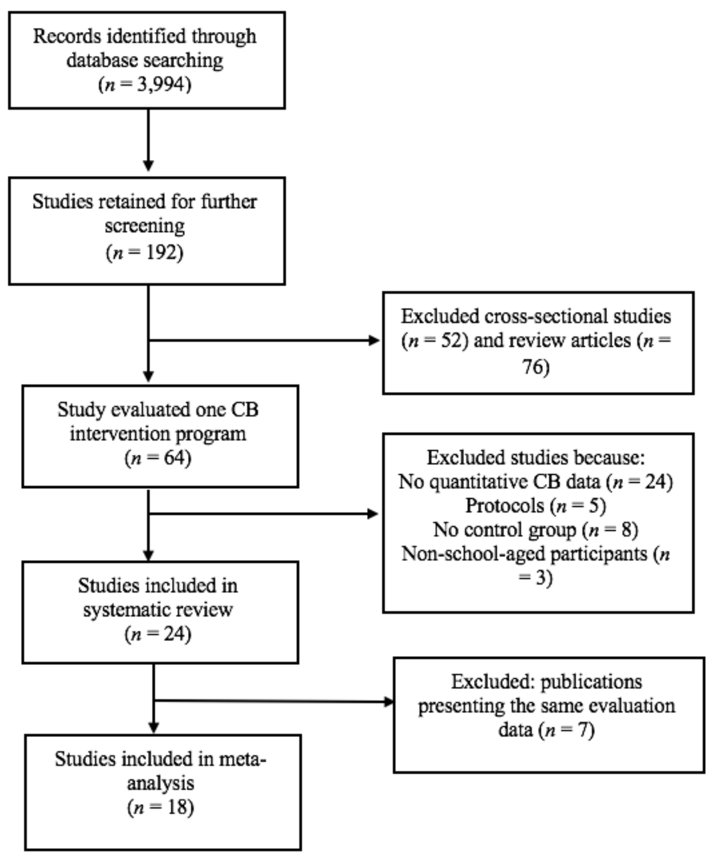
Further screening of the 192 retained studies indicated that 128 of them had to be excluded from the systematic review. These excluded studies comprised 26 narrative and 6 systematic reviews of various cyberbullying-related issues (e.g., Zych et al., 2015). Four meta-analyses were identified (i.e., Chen et al., 2017; Gardella, Fisher, & Teurbe-Tolon, 2017; Guo, 2016; Modecki et al., 2014) by the searches, but none of these evaluated the effectiveness of cyberbullying programmes. Nineteen qualitative or theoretically based studies were excluded (e.g., Barlett, 2017; Cross et al., 2015; Gradinger, Strohmeier, & Spiel, 2017), in addition to 13 studies excluded for ‘other’ reasons. Moreover, eight identified studies provided detailed descriptions of existing cyberbullying interventions but did not provide details regarding the programme’s implementation or evaluation (e.g., Chan & Wong, 2017; Sapouna et al., 2015; Singh et al., 2017).

The majority of excluded studies (*n* = 52) were classified as empirical, cross-sectional studies that assessed various aspects of cyberbullying behaviour and made reference to the implications of results for future intervention and prevention programmes. Twenty of these cross-sectional studies explored various correlates of cyberbullying behaviours. A wide variety of factors (e.g., moral influences: Allison & Bussey, 2017), were investigated by these studies in relation to cyberbullying, but in the present systematic review they are all referred to as correlates or associated factors, as often no clear causal model was employed (Barlett, 2017). Related factors varied from those that could easily be applied to cyberbullying intervention programmes (e.g., resilience: Hinduja & Patchin, 2017; empathy: Brewer & Kerslake, 2015), to potential social and emotional predictors (e.g., Sourander et al.,

2010) and concurrent negative or risky behaviours, both online and offline (e.g., Gámez-Guadix et al., 2013).

Figure 8

Flowchart of the screening process for the cyberbullying systematic review



The remainder of the cross-sectional studies were excluded for a variety of reasons: 11 studies assessed the overlap between offline and online bullying (e.g., Baldry, Farrington, & Sorrentino, 2017); six were concerned with the patterns and/or prevalence of cyberbullying behaviours (e.g., Barlett & Chamberlain, 2017); and five investigated the potential outcomes associated with cyberbullying behaviours (e.g., depression and/or anxiety: Rose & Tynes, 2015). Finally, ten excluded studies targeted various aspects of bystander behaviours, intentions, and characteristics in cyberbullying experiences (e.g., cultural issues: Ferreira et al., 2016; impulsivity and helping behaviours: Erreygers et al., 2016).

8.4 Excluded studies

A total of 64 studies were therefore subjected to a second wave of more detailed screening. These studies were originally thought to be includable in the present systematic review but, upon further screening, 40 studies were excluded. A brief outline of these studies is provided in Table 26. The majority of these (*n* = 25) were excluded either because they did not include cyberbullying-related outcomes (e.g., cyber dating-abuse: Foshee et al., 2015; offline bullying behaviours: Ostrov et al., 2015), or because their cyberbullying-related outcomes were not measuring cyberbullying behaviours. For example, Barkoukis and colleagues (2016) and Lee et al. (2013) utilised measures of 'behavioural intentions' relating to cyberbullying. Nine studies were excluded because aspects of their evaluation methodology did not meet the inclusion criteria, for example, studies that used University/College student samples (e.g., Doane, Kelley, & Pearson, 2016; Leung, Fung, & Farver, 2017), did not utilise a control group (e.g., Timmons-Mitchell et al., 2016), or non-randomised studies that did not measure cyberbullying before and after implementation of an intervention (e.g., Tangen & Campbell, 2010).

Table 26

Descriptions of studies excluded from the cyberbullying meta-analysis

Study	***Description*** ***➛Reason for exclusion***
Appelqvist-Schmidlechner et al. (2017)	Evaluation of the feasibility of the 'Together at School' intervention programme and also the programme's effectiveness at reducing several socio-emotional problems. Self-reported school climate and school satisfaction questionnaires included items on bullying experiences, but only referred to school-bullying and not cyberbullying. ➛*No cyberbullying outcomes*
Avșar & Alkaya (2017)	Empirical evaluation of an assertiveness training programme on school-bullying and level of assertiveness. ➛*No cyberbullying outcomes*
Barkoukis et al. (2016)	Evaluation of a cyberbullying intervention programme that targets the psychosocial risk factors for cyberbullying in adolescence, however, the outcome measure refers to intent to cyberbully others and not actual cyberbullying behaviours committed. ➛*No cyberbullying outcomes*
Camelford & Ebrahim (2016)	Description of the rationale and setting for a pilot evaluation of a psychoeducational intervention for high school girls to prevent cyberbullying. No actual evaluation data is presented. ➛*No evaluation data – protocol only*
Chaux et al. (2017)	Evaluation of the impact of the Classrooms in Peace programme on several violence-related outcomes, including victimisation and aggression analogous to bullying, amongst a sample of Colombian youth. ➛*No cyberbullying outcomes*
Clarkson et al. (2016)	Report on the study protocol for the planned implementation of the KiVa bullying prevention programme in Welsh schools. ➛*No evaluation data – protocol only*
Cleemput et al. (2016)	Description of the development of a 'serious game' intervention to reduce cyberbullying, however, only report results of focus group interviews regarding the feasibility and applicability of the programme. ➛ *Methodology*
Dillon & Bushman (2015)	Outline of an experimental assessment of the Bystander Intervention Model in relation to cyberbullying.

	➛*Methodology*
Doane et al. (2016)	Evaluation of cyberbullying prevention video, developed using the theory of reasoned action, with University students in the USA. ➛*Sample*
Farmer et al. (2017)	Evaluation of the impact of altering a school's play environment to include more challenging and interactive play in order to reduce reports of traditional school-bullying. ➛*No cyberbullying outcomes*
Foshee et al. (2015)	Exploration of the impact of the programme: Moms and Teens for Safe Dates on several victimisation outcomes, including cyber dating-abuse, however, this was not analogous to cyberbullying behaviours. ➛*No cyberbullying outcomes*
Garaigordobil & Martínez-Valderrey (2014)	Evaluation of the effectiveness of the Cyberprogram 2.0 intervention on reducing school-bullying victimisation. ➛*No cyberbullying outcomes*
Garaigordobil & Martínez-Valderrey (2015b)	Evaluation of the effectiveness of the Cyberprogram 2.0 intervention on forms of conflict resolution and participants' self-esteem. ➛*No cyberbullying outcomes*
Garaigordobil et al. (2017)	Evaluation of the effectiveness of the Cyberprogram 2.0 intervention programme in a single-case study employing one 14-year-old male aggressor. ➛*Methodology*
Guo et al. (2015)	Evaluation of the effectiveness of the Positive Action programme on several behavioural outcomes. ➛*No cyberbullying outcomes*
Hicks et al. (2016)	Description of the development and pilot implementation of a solution-focused dramatic empathy training programme to reduce cyberbullying. ➛*No evaluation data – protocol only*
Jacobs et al. (2016)	Detailed overview of the development and theoretical foundations of the Online Pestkoppenstoppen programme and plans for implementation and evaluation. ➛*No evaluation data – protocol only*
Juvonen et al. (2016)	Finnish national evaluation of the KiVa anti-bullying programme to evaluate its' effectiveness of outcomes such as depression and self-esteem. ➛*No cyberbullying outcomes*

Kaljee et al. (2017)	Evaluation of the impact of a teacher-training programme on several outcomes for Zambian students, including reports of traditional bullying. ➛*No cyberbullying outcomes*
Lee et al. (2013)	Evaluation of a cyberbullying intervention programme in Taiwan but outcomes refer to the intention to cyberbullying others. ➛*No cyberbullying outcomes*
Leff et al. (2016)	Evaluation of the impact of the Friend to Friend programme on teacher-student relationships, prosocial behaviours, and aggressive behaviour. ➛*No cyberbullying outcomes*
Leung et al. (2017)	Evaluation of a cyberbullying prevention programme with University students in Hong Kong. ➛*Sample*
McCuddy & Esbensen (2017)	Report on a longitudinal evaluation of the GREAT programme (Gang Resistance Education and Training) in middle schools on predictors of school-bullying victimisation and cyber-victimisation. ➛*No evaluation data for cyberbullying*
McElearney et al. (2008)	Examination of the impact of 'befriending' peer support programmes on bullying and cyberbullying using a case-study approach. ➛*Methodology*
Midgett et al. (2017)	Description of the experiences of students involved in an anti-bullying programme as trained student-advocates. ➛*Methodology*
Ostrov et al. (2015)	Evaluation of the Early Childhood Friendship Project intervention programme on outcomes of traditional physical and relational peer victimisation. ➛*No cyberbullying outcomes*
Ploeg et al. (2016)	Exploration of the impact of the support group approach as part of the Dutch implementation of the KiVa anti-bullying programme on participant reports of change in victimisation. Only refer to traditional victimisation, not cyberbullying victimisation. ➛*No cyberbullying outcomes*
Roberto et al. (2017)	Evaluation of the impact of a parent-training programme on several aspects of parental involvement with their children's potential cyberbullying behaviours, such as, their perceived susceptibility and behavioural intentions in response to the experience.

	➛*No cyberbullying outcomes*
Savage et al. (2017)	Evaluation of the impact of an anti-cyberbullying victimisation intervention message amongst a sample of US University students using a post-test quasi-experimental design. ➛*Sample*
Silva et al. (2016)	Evaluation of the effects of a cognitive-behavioural therapy based social and emotional skill development programme on reports of school-bullying perpetration and victimisation in a sample of Brazilian adolescents. ➛*No cyberbullying outcomes*
Sullivan et al. (2017)	Examination of the moderating effects of disability status and gender on the effectiveness of two anti-bullying programmes (i.e., the OBPP and Second Step) on various violence-related outcomes, including traditional bullying. ➛*No cyberbullying outcomes*
Tangen & Campbell (2010)	Cross-sectional comparison of rates of cyberbullying in two primary schools, one that implements a Philosophy for Children (P4C) and another control that does not. However, no pre- and post-test measures are employed. ➛*Methodology*
Tanrikulu et al. (2015)	Evaluation of the Sensibility Development Programme against Cyberbullying, however, cyberbullying-related outcomes refer to behavioural intention and not actual cyberbullying behaviours committed or experienced. ➛*No cyberbullying outcomes*
Timmons-Mitchell et al. (2016)	Pilot evaluation of StandUp, an online anti-bullying programme, but no control group was employed. ➛*Methodology*
Toshack & Colmar (2012)	Description of a small-scale evaluation of a cyberbullying intervention with female students, however, no control group was utilised. ➛*Methodology*
Usó et al. (2016)	Present the results from an evaluation of a peer mediation programme on traditional bullying. No cyberbullying outcomes are included. ➛*No cyberbullying outcomes*
Wang & Goldberg (2017)	Evaluation of the impact of the Bullying Literature Project-Moral Disengagement Version programme to reduce traditional bullying perpetration and victimisation. ➛*No cyberbullying outcomes*

Wexler et al. (2017)	Examination of the impact of the Youth Leaders Programme on several outcomes in a rural Alaskan community, however, no outcomes relating to either traditional bullying or cyberbullying are employed. ➛*No cyberbullying outcomes*
Williford et al. (2012)	Examination of the impact of the KiVa anti-bullying programme on adolescents' reports of depression, anxiety, and perception of peers. No outcomes relating to cyberbullying are incorporated. ➛*No cyberbullying outcomes*
Ybarra et al. (2016)	Description of the development of an intervention programme, BullyDown, to combat text messaging bullying. ➛*No evaluation data – protocol only*

8.5 Included studies

Twenty-four publications were included in the meta-analysis to evaluate the effectiveness of cyberbullying intervention and prevention programmes for school-aged children and adolescents. The majority of these publications described evaluations using randomised controlled trials (n = 15). Furthermore, the majority of studies used before and after intervention measures of cyberbullying, with the exception of an evaluation conducted by Roberto et al. (2014). This RCT used a post-test control-group design. The remaining nine publications described evaluations that used quasi-experimental designs with before and after intervention measures.

At this point a distinction is made between publications and 'studies', where publications refer to the articles published, and studies refers to the evaluations of an anti-cyberbullying programmes in independent samples. For example, of these 24 included publications, two presented results from multiple evaluations (i.e., Menesini, Nocentini, & Palladino (2012a) – *study 1 and study 2*; Palladino, Nocentini, & Menesini (2016) – *trial 1 and trial 2*). Therefore, while 24 publications are included in the systematic review, 26 evaluations are included. A brief overview of these evaluations is provided in Table 27.

Table 27

Overview of studies included in the systematic review of cyberbullying intervention programmes

Project(s)	*Intervention*	*Participants*	*Research Design*
Randomised Controlled Trials (n = 15 publications; n = 11 studies)			
Athanasiades et al. (2015) Greece	Tabby project – “a pilot short-term intervention against cyber-risks and cyberbullying was designed to be implemented by teachers in the classroom”	314 Greek secondary school students aged 13 to 14 years old.	Students randomly assigned to experimental (*n* = 123) or control (*n* = 140) group. All participants completed the ‘Tabby Checklist’ measure pre- and post-test
Chaux et al. (2016); Wölfer et al. (2014) Germany	Media Heroes – “theoretically (theory of planned behaviour and participant-roles approach to bullying) based preventive intervention programme developed in Germany for the school context”	1,075 German students. Mean age was 13.36 years and 51.8% were female.	35 classes from 5 schools were randomly assigned to one of three experimental conditions: (1) long-version (*n* = 12 classes); (2) short-version (*n* = 7 classes); and (3) control group (*n* = 16 classes). All participants completed the self-report ECIPQ measure of cyberbullying perpetration and victimisation pre- and post-intervention.
Cross et al. (2016); Shaw et al. (2015) Australia	Cyber Friendly Schools – “whole-school programme to enhance the capacity of school staff, students, and families to respond effectively to reduce cyberbullying behaviour”	3,382 Grade 8 and 9 students from 35 schools in Perth, Australia, aged 13 to 15 years old.	Schools were randomly assigned to experimental (*n* = 19 schools) or control (*n* = 16 schools) conditions. All participants completed two 11-item scales measuring cyberbullying perpetration and victimisation.
DeSmet et al. (2018) Belgium	Friendly Attac – “a serious game intervention was designed to promote positive bystander behaviour and reduce negative bystander behaviour”	227 students from two schools. 58.5% of intervention group were female and the mean age was 13.52 years old. 65.3% of control group were female and the mean age was 13.47 years.	One school was randomly allocated to the intervention condition and another randomly allocated to the waitlist control condition. All 8th classes participated and completed a self-report measure of cyberbullying behaviours in past 6 months at baseline, following intervention and 4-weeks later.
Espelage et al. (2015) US	Second Step – “a universal, curricular classroom intervention… through skill building and skill practice, this	3,651 participants from 36 schools. 52% were male, and the mean age at baseline was 11 years.	18 schools were randomly assigned to the experimental (*n* = 2,341 students) and 18 were randomly assigned to the

	comprehensive programme targets risk and protective factors lined to aggression, violence, and substance-use"		control (*n* = 2,074) condition. All participants completed a 4-item self-report measure of cyberbullying perpetration.
Fekkes et al. (2016) Netherlands	Skills for Life – "universal school-based prevention programme aimed at reducing behavioural and health problems in adolescents"	1,394 students from 27 schools. 51% of the control group were male and the mean age was 14.4 years 53% of the experimental group were male and the mean age was 14 years old.	13 schools were randomly assigned to the experimental (*n* = 1,107 students) condition, and 13 schools (*n* = 481 students) to the control condition. All participants were asked how often they had been bullied via the Internet or SMS in the past three months, before and after intervention.
Garaigordobil & Martínez-Valderrey (2015a); Garaigordobil & Martínez-Valderrey (2016) Spain	Cyberprogram 2.0 - "an intervention programme to prevent and reduce cyberbullying"	176 Spanish adolescents aged 13 to 15 years old. 56.3% were female.	93 students were randomly assigned to the experimental condition, and 83 were randomly assigned to the control group. All completed a self-report measure of cyberbullying perpetration and victimisation pre- and post-intervention.
Gradinger et al. (2015); Gradinger et al. (2016) Austria	ViSC - "a primary preventive programme including secondary preventive elements to (1) reduce aggression and bullying, and (2) promote social and intercultural competencies in schools"	2,042 Austrian students in 5th to 7th grade. 47.6% were female and the mean age was 11.7 years.	13 schools were randomly assigned to the experimental condition (*n* = 1,377 students) and 13 schools were randomly assigned to the control group (*n* = 665). All completed Smith et al. (2008) measure of cyberbullying perpetration and victimisation pre- and post-intervention.
Roberto et al. (2014) US	Social Networking Safety Promotion and Cyberbullying Prevention Promotion - "the Arizona Attorney General's Social Networking Safety Promotion and Cyberbullying Prevention presentation" "45-minute presentation that was an example of fear appeal and contained both threat and efficacy components"	425 students from a US middle school in the 6th, 7th and 8th grade. 53% were female and the mean age was 12.58 years old.	21 classes from one school were randomly assigned to intervention (*n* = 11 classes) or control (*n* = 10 classes) condition. Study used a post-test only control-group randomised design. Cyberbullying perpetration and victimisation behaviours were measured with two dichotomous yes/no questions about experiences in the current school year.

Schultze-Krumbholz et al. (2016) Germany;	Empathy training - "a universal, modularized, and theoretically, based preventive intervention for the school context"	897 German secondary school students from 35 classrooms in 5 schools. Mean age was 13.36 years old and 46.3% of the sample were male.	Classes were randomly assigned to either the control group (*n* = 350 students), the short intervention group (*n* = 136) or the long intervention group (*n* = 228). All participants completed the ECIPQ measure of cyberbullying perpetration pre- and post-intervention.
Williford et al. (2013) Finland	KiVa - "…focuses on enhancing the empathy, self-efficacy, and anti-bullying attitudes of bystanders, who are neither bullies nor victims"	18,412 students enrolled in a large national evaluation of the KiVa in Finland. Mean age of Grades 4 – 6 was 11.25 years and 49% were male. Mean age of Grades 8 & 9 was 13.98 and 48% were male.	78 schools were randomly assigned either to the intervention group (*n* = 9,914 students; *n* = 39 schools) or the control group (*n* = 8,498; *n* = 25 schools). All participants completed a modified version of the OBVQ to measure cyberbullying perpetration and victimisation pre- and post-intervention.
Before-After/Experimental-Control designs (n = effect sizes, n = publications)			
Menesini et al. (2012a); Study 1 Italy	NoTrap! - "development of a website to promote peer-to-peer content against bullying and cyberbullying"	386 secondary school students at 8 Tuscan schools, 20.3% were male, and the mean age was 16.29 years old. 9th to 13th grade students for intervention running from December 2009 – June 2010.	Students were assigned to one of three potential groups: (1) Control group; (2) Intervention group; and (3) Peer educators. Bullying measures were administered pre- and post-test (6 months apart).
Ortega-Ruiz et al. (2012); Del Rey et al. (2012); Del Rey et al. (2016) Spain	ConRed - "an evidence-based intervention programme"	893 Spanish students aged 11 – 19 years old. 45.9% were female.	595 participants were in the experimental group and 298 were in the control group. All participants completed the ECIPQ measure of cyberbullying perpetration and victimisation pre- and post-intervention.
Palladino et al. (2012); Menesini et al. (2012a; Study 2) Italy	NoTrap! - "enriched the first edition by adding additional online and offline components"	375 9th to 13th grade students at 4 Tuscan high schools for year December 2010 – June 2011.	Students were assigned to one of three potential groups: (1) Control group; (2) Intervention group; and (3) Peer educators. Bullying measures were administered pre- and post-test (6 months apart).
Palladino et al. (2016); Trail 1 – 2011/2012	NoTrap! - "aimed to standardize the face-to-face activities led by peer educators"	622 9th grade Italian students from 8 schools and 31 classrooms. 60.29% were male and the mean age in the	451 participants were in the experimental group and 171 students were in the control group.

Italy		experimental group was 14.79 years and the mean age in the control group was 15.28 years.	
Palladino et al. (2016); Trail 2 – 2012/2013 Italy	See Palladino et al. (2016; Trial 1)	461 9th grade Italian students from 7 schools and 20 classrooms. 52.06% were male and the mean age in the experimental group was 15.6 years and the mean age in the control group was 15.38 years.	234 participants were in the experimental group and 227 students were in the control group. All participants completed the Florence Cyberbullying/Cybervictimisation scales pre- and post-intervention.
Pieschl et al. (2017); Pieschl & Urbasik (2013) Australia	Surf-Fair - "is a less comprehensive and shorter [programme] and is based on student-centred and constructivist anchored instruction"	150 students from two Australian schools and 9 classrooms. The mean age was 11.31 years old and 52.67% of the sample were male.	74 students from 5 classrooms were allocated to the experimental group and 76 students from 4 classrooms were allocated to the control group. All participants completed the Revised Olweus Bullying Questionnaire (Olweus, 2012) before and after the intervention
Solomontos-Kountouri et al. (2016) Cyprus	ViSC - "both on the school and the class level, the ViSC programme aims to create an environment in which it is less likely that aggressive behaviour, bullying and other victimisation will occur"	1,652 7th and 8th grade students in Cyprus. Mean age was 12.6 years and 48.9% of the total sample were female.	6 schools (82 classes) were allocated to either the intervention (n = 602 students; n = 30 classes) or control group (n = 1,050 students; n = 52 classes). All participants completed self-report measure of cyberbullying perpetration and victimisation (Smith et al., 2008) pre- and post-implementation.

Note: (1) Some studies may have included additional follow-ups (e.g., DeSmet et al. (2018) analysed cyberbullying outcomes at baseline, post-intervention, and 4-weeks post-intervention). However, the present meta-analysis was only interested in measurements taken immediately post-intervention; (2) Where two studies are provided this represents examples of overlapping samples. In these cases, studies written in **bold** were used for meta-analysis; (3) * - study authors were contacted but did not supply additional statistical information required for meta-analysis and therefore, is included here in the systematic review, but was excluded from the meta-analysis; (4) n = number of participants in groups

9. Data Extraction: Cyberbullying

9.1 Overview

In relation to the systematic review and meta-analysis of cyberbullying intervention and prevention programmes the data extraction process echoed that applied for the school-bullying meta-analysis. Information was extracted from primary reports on four different levels: (1) Descriptive; (2) Design; (3) Programme; and (4) Outcome. Table 28 outlines each of these levels and provides examples of the type of information recorded at each level in relation to cyberbullying.

The results of this data extraction process are outlined in Table 27 (see Chapter 8, section 8.5). The same risk of bias index (see Chapter 5, Table 11) was utilised to assess potential risk of bias in cyberbullying evaluations, and thus, an explanation of the process is not repeated in the present chapter.

It was expected that less data would be extracted from cyberbullying evaluations for many reasons. Firstly, there were fewer evaluations included, and thus it was anticipated that there would not be sufficient information for extensive subgroup analyses. Moreover, many of the intervention programmes that were evaluated were the same interventions as those included in the school-bullying meta-analysis and as such the information and coding would be the same.

9.2 Descriptive

On the descriptive level, information regarding the location of the evaluation and the sample was coded. Specifically, in relation to the sample, the total number of participants that were involved in the evaluation, i.e., the number of individuals included in the experimental and the control conditions, was coded as a continuous variable. Where the breakdown of gender in the sample was reported this was also recorded. Information on the ethnic, sexual or gender identity was not reported by enough primary studies to be included in the present review, despite the need for more research in this area. Furthermore, the age of participants was coded. Similar to the school-bullying meta-analysis, age was coded as it was reported by the primary evaluations and then transformed to a continuous variable for the purpose of analysis. The year of publication of each evaluation and the type of publication was also recorded.

Table 28

Data extraction codebook for the cyberbullying meta-analysis

Type	Information extracted	Example
Descriptive	• Sample size • Age of sample in years • Grade(s) of sample or range • Sex: % female and % male • Location or country • Publication Year • Publication Type	• Total N; *n* experimental; *n* control • Mean age/range • 2012, 2014, 2016 etc • Journal article, book chapter, dissertation, report
Design	• Evaluation method • Measures • Data collection timepoints • Unit of allocation/randomisation • N clusters	• RCT or quasi-BA/EC • Name of instrument • Timeframe • Perpetration/ victimisation/ both • Type of report • Baseline/Post-intervention/Follow-up
Programme	• Name of programme • Intervention aim and/or target • Conflict of Interest • Specificity	• e.g., KiVa, ConRed, Cyberprogram 2.0 • High, low, possible risk • Targeted just cyberbullying, or targeted offline and online bullying
Outcomes	• Cyberbullying at baseline for exp and control • Cyberbullying post-intervention for exp and control	• Mean, SD, N • N and % bullies and/or victims

Note. N = total sample; *n* = number of participants in groups; RCT = randomised controlled trial; BA/EC = quasi-experiments with before and after measures of bullying (non-randomised); SD = standard deviation; exp = experimental group

9.3 Design

On the design level, information regarding the evaluation methodology was the primary piece of information recorded. Included evaluations used two types of evaluation methodology, namely, randomised controlled trials and before and after quasi-experimental-control designs (BA/EC). A detailed description of these methodologies is provided in Chapter 5 (see section 5.3.1).

Information about the measurement instruments used to measure cyberbullying perpetration and/or victimisation was also extracted. The name of the measurement instrument was recorded, along with the type of report. Potential report types included, self-report, peer-report, or teacher-report. If participants were asked to report cyberbullying incidences within a specific timeframe (e.g., past 3 months, current school year) this was recorded.

9.4 Programme

The name of each intervention programme that was evaluated in studies included in the cyberbullying meta-analysis was noted. As there were a significant number of the evaluations were also included in the school-bullying meta-analysis, the same coding framework for intervention components was used in the cyber-bullying meta-analysis, where appropriate. Detailed information about specific components is provided in Chapter 5 (see section 5.4.1) and outlined in Table 10.

In contrast to the school-bullying meta-analysis, additional details about intervention components were recorded. Instead of recording just the presence or absence of specific components, the specific ways in which the intervention components were recorded. This will enable a more detailed systematic review of the composition of cyberbullying intervention programmes. As there is significantly less literature on cyberbullying intervention programmes, this will constitute a greater contribution to the field. In addition to specific intervention components, a variable reflecting the offline/online bullying overlap was recorded. Specifically, whether or not the intervention programme targeted only cyberbullying, only school-bullying, or both, was recorded. Similar methods of assessing conflict of interest were also employed (see Chapter 5, section 5.4.2).

9.5 Outcome

Finally, information required to compute effect sizes for cyberbullying perpetration and victimisation was extracted. Primary evaluations of cyberbullying intervention programmes reported perpetration and/or victimisation outcomes at baseline, post-intervention and possibly additional follow-up points. Data was extracted as either a percentage of participants reporting cyberbullying perpetration/victimisation of the total sample or as a continuous measure. Where continuous measures were used, the mean cyberbullying perpetration and/or victimisation score was extracted as well as the standard deviation and sample size.

9.6 Risk of Bias

The Campbell Collaboration recommended EPOC tool was used to assess risk of bias in the meta-analysis of cyberbullying intervention programmes. Primary evaluations were assessed on several factors that may impact bias, such as: (1) Allocation sequence [AS]; (2) Allocation concealment [AC]; (3) Baseline equivalence on outcomes [BE]; (4) Baseline equivalence on participant characteristics [BC]; (5) Incomplete outcome data [ID]; (6) Contamination protection [CP]; and (7) Selective outcome reporting [SOR]. Table 11 (see Chapter 5, section 5.6) outlines the coding process for each of the risk of bias items as well the criteria for an evaluation being classified as low or high risk for all design methodologies included in the meta-analysis.

10. Results: Cyberbullying

10.1 Systematic review

This section presents the results of the systematic review of cyberbullying intervention and prevention programmes. Table 27 (see Chapter 8, section 8.2) presents the studies included in the cyberbullying systematic review and the information extracted from each evaluation on the following levels: Descriptive; Design; Programme; and Outcome. In comparison to the school-bullying (or traditional/offline bullying) meta-analysis, far fewer different interventions were included in the cyberbullying systematic review and meta-analysis. Thus, there was less information to extract.

10.1.1 Descriptive

As described in Chapter 9, a lot of detail and information was extracted from primary evaluations. The following section describes the systematic review of this information. On the descriptive level, information concerning the location of the intervention, the total sample size, the age of participants and the type of publication was recorded.

10.1.1.1 Location of intervention. All of the included evaluations of cyberbullying intervention programmes were conducted in high-income countries. In fact, the majority of evaluations were conducted in European countries; specifically: Austria (e.g., Gradinger et al., 2015), Belgium (e.g., DeSmet et al., 2018), Cyprus (e.g., Solomontos-Kountouri et al., 2016), Finland (e.g., Williford et al., 2013), Greece (e.g., Athanasiades et al., 2015), Germany (e.g., Chaux et al., 2016; Schultze et al., 2016), Italy (e.g., Palladino et al., 2012), the Netherlands (e.g., Fekkes et al., 2016), and Spain (e.g., Ortega-Ruiz et al., 2010). Only one included evaluation was conducted in Australia (i.e., Pieschl et al., 2017) and two were conducted in the US (i.e., Espelage et al., 2015; Roberto et al., 2014).

Germany (n = 2), Italy and the US were the only locations where multiple evaluations have been conducted. In Italy, multiple evaluations of the NoTrap! programmes were included in the cyberbullying meta-analysis (i.e., Menesini et al., 2012; Palladino et al., 2012; Palladino et al., 2016). Therefore, given that there were insufficient numbers of studies and different countries where interventions were evaluated, the location of the intervention was not included in further meta-analytical subgroup analyses.

10.1.1.2 Sample Size. Approximately 35,000 participants are represented in the cyberbullying meta-analysis in total. The 18 effect sizes for cyberbullying intervention programmes that assessed the impact on perpetration behaviours were computed using data from approximately 34,826 participants. Similarly, 19 effect sizes for cyberbullying victimisation were computed using data from approximately 35,637 participants. The sample sizes of included evaluations ranged from 150 participants (i.e., Pieschl et al., 2017) to 18,412 participants (Williford et al., 2013), with a mean sample size of $N = 2,030$ (median = 757.5).

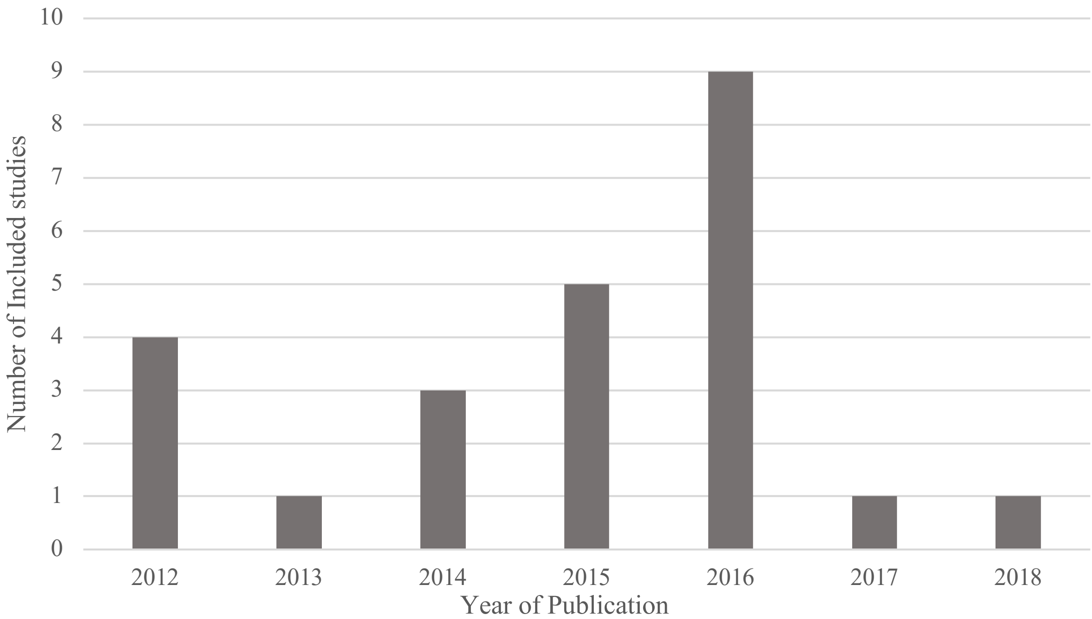
A dichotomous variable to represent variation in sample size was not included in subsequent moderator analyses due to the uneven numbers of studies in subgroups. For example, in relation to cyberbullying perpetration, 12 studies included sample sizes that were below the mean value and 3 studies included sample sizes over the mean value. Regarding cyberbullying victimisation outcomes, 13 studies included samples under this mean value and 3 studies included samples over this mean value. The relationship between sample size and evaluation outcome was thus only examined using meta-regression.

10.1.1.3 Age of participants. Age was coded as either the range of ages, the mean age, or the school grade of participants in primary studies. Where the range of participant ages was reported by a primary evaluation the mean age was estimated using the lower and upper boundaries. For example, Cross et al. (2016) reported that participants were aged between 13 and 15 years old. This group includes children from 13.00 years (i.e., the day of their 13th birthday) to 15.99 years old (i.e., the day before their 16th birthday). Thus, the mean age is estimated as 14.45 years old.

The overall mean age was 13.75 years old, with the youngest participants being 11 years old (i.e., Espelage et al., 2015) and the oldest participants being 16.8 years old (i.e., Palladino et al., 2012). In addition to the continuous age variable, a dichotomous variable was created to assess the differences in effectiveness estimates between younger and older adolescents. Younger adolescents were categorized as participants between 11 and 13 years old and older adolescents were categorized as participants aged 14 years and older. If the mean age was 13.5 years old (e.g., DeSmet et al., 2018), the study was categorized as being ‘younger’. Similarly, if the mean age was 14.4 years old (e.g., Fekkes et al., 2016), the study was categorized as being ‘older’. The subgroups of studies based on this categorical age variable was relatively even in numbers. In relation to cyberbullying perpetration outcomes, 9 studies included younger adolescents and 7 studies included older participants. In relation to cyberbullying victimisation outcomes, 8 studies were categorized as including younger adolescents and 8 studies were categorized as including older adolescents.

Figure 9

Number of studies included in the meta-analysis by year of publication ($N = 24$).



10.1.1.4 Publication year and type. The publication type (e.g., peer-reviewed journal articles, unpublished dissertations, book chapters) of each primary evaluation was also recorded, as well as the year in which the study was published.

All of the evaluations included in the present analyses were published in peer-reviewed journal articles for both cyberbullying perpetration ($n = 18$) and cyberbullying victimisation ($n = 19$). Thus, this factor was not included in the subgroup analyses. The evaluations were published between 2012 and 2018, with the largest number of evaluations published in 2016. Figure 9 represents the years of publication graphically. Given the short time period and small number of studies, the year of publication was not included as a moderator in the cyberbullying meta-analysis.

10.1.2 Design

At the design level, information regarding the method used to evaluate each cyberbullying intervention programme, and the measurement instrument used to assess cyberbullying behaviours were recorded.

Two evaluation methods were eligible for inclusion in the cyberbullying meta-analysis, namely, randomised controlled trials (RCTs) and before-after/quasi-experimental-control designs (BA/EC). An equal number of studies examined the impact of interventions on cyberbullying perpetration using an RCT design ($n = 9$) or a BA/EC design ($n = 9$). Regarding cyberbullying victimisation, 10 studies used an RCT design and 9 studies used a BA/EC design. The number of studies in groups based on evaluation methodology was balanced, and this variable was included in further subgroup analyses.

In total, seven different measurement instruments were used to assess cyberbullying perpetration and/or victimisation, namely: a modified OBVQ (DeSmet et al., 2018; Williford et al., 2013); the European Cyberbullying Intervention Project Questionnaire (Chaux et al., 2016; Ortega-Ruiz et al., 2012); the Illinois Bully Scale (Espelage et al., 2015); the Tabby Checklist (Athanasiades et al., 2015); and Cyberbullying: Screening of Peer Harassment (Garaigordobil & Martínez-Valderrey, 2015). Evaluations of the NoTrap! intervention programme used a scale developed by the authors (Menesini et al., 2012; Palladino et al., 2012; Palladino et al., 2016). Three evaluations used a cyberbullying scale developed by Smith et al. (2008) which was designed using the OBVQ as a model (Cross et al., 2016; Gradinger et al., 2015; Solomontos-Kountouri et al., 2016).

In addition, two evaluations used global items to assess cyberbullying behaviours, i.e., one question asked participants if they had ever experienced cyberbullying victimisation and another question asked if they had perpetrated cyberbullying. All of the measures were self-report and the majority of instruments asked participants to report cyberbullying experiences in the past couple of months. This timeframe varied from the past one month to the past 6 months. Two evaluations asked participants to report cyberbullying behaviours in the past year (Garaigordobil & Martínez-Valderrey, 2015; Roberto et al., 2014). Given the lack of variability and the lack of sufficient numbers of studies in potential subgroups based on aspects of the measurement instrument, this was also not included in subgroup analyses.

10.1.3 Programme

As described in Chapter 9, information about the specific interventions that were evaluated in studies included in the meta-analysis was also extracted. Hence, on the programme level, specifics such as the name of the intervention programme, the intervention components, and any potential conflict of interest were coded.

In total, 14 different intervention programmes were included in the cyberbullying meta-analysis. Only one programme, NoTrap!, was evaluated multiple times (i.e., Menesini et al., 2012; Palladino et al., 2012; Palladino et al., 2016). Five effect sizes for cyberbullying perpetration and five effect sizes for cyberbullying victimisation were included in the meta-analysis from evaluations of this programme. Additionally, the ViSC programme was evaluated twice in different locations. Gradinger et al. (2015) evaluated the effectiveness of the ViSC programme in Austria, and Solomontos-Kountouri et al. (2016) examined the effect on cyberbullying in Cyprus. In light of the small number of studies, and the lack of multiple evaluations of an appropriate number of interventions, moderator analyses were not conducted for specific intervention programmes.

10.1.3.1 Intervention components. In light of the smaller number of primary evaluations included in the cyberbullying meta-analysis, compared to the school-bullying review, there was not enough studies to create comparable subgroups in relation to specific intervention components. For example, only one study (Williford et al., 2013) included ‘classroom rules’ in the intervention activities, one study included components that directly targeted bullies (Williford et al., 2013) and two studies included components that directly targeted victims (Solomontos-Kountouri et al., 2016; Williford et al., 2013). Furthermore, given the lack of previous research on the content of cyberbullying intervention programmes, it was decided that the best contribution to the literature would be to conduct a narrative review of existing programmes and their content. Components were not coded as present or absent, as in the school-bullying meta-analysis. The intervention components for included cyberbullying intervention components are outlined in further detail in Tables 29 and 30. Roberto et al. (2014) is excluded from these tables as only one intervention component was included; namely, this intervention involved a once-off presentation by an external organisation to outline the problems and risks associated with Internet use and cyberbullying.

One aspect of cyberbullying intervention programmes that was recorded as present or absent was whether or not the content of the intervention included both online and offline bullying. This variable was dichotomous, and primary studies were grouped based on whether they did (i.e., “yes”) include both online and offline content in programme activities and materials, or whether they only included content on cyberbullying (i.e., “no”). In relation to cyberbullying perpetration outcomes, 13 effect sizes were estimated from studies that did include both online and offline content and five studies were concerned only with cyberbullying. The majority of effect sizes (i.e., $n = 12$) for cyberbullying victimisation outcomes were estimated from evaluations of interventions that did include both online and offline content. In total, seven studies (for cyberbullying victimisation outcomes) included only online bullying content. Subgroup analyses were conducted for this variable.

Similarly, whether or not the evaluation included both online and offline bullying outcomes (i.e., “yes”) or just online bullying outcomes (i.e., “no”) was coded. In relation to cyberbullying perpetration outcomes, 13 effect sizes were estimated from evaluations that did include both online and offline bullying outcomes. Five effect sizes for cyberbullying perpetration were estimated from evaluations that only included online bullying outcomes and five effect sizes for cyberbullying victimisation were estimated from evaluations that only included online outcomes. Fourteen effect sizes for cyberbullying victimisation were estimated from evaluations that assessed the impact of an intervention on both online and offline bullying behaviours. The differences between studies included in each subgroup for this variable were not that different from studies included in the moderator variable relating to the inclusion of school-bullying and cyberbullying content in interventions. Thus, moderator analyses were conducted but the differences between the mean effect sizes for these two variables was not expected to differ greatly.

Conflict of interest (COI) was also measured for cyberbullying intervention programmes, as described in Chapter 9. The majority of cyberbullying evaluations included in the present research were deemed to have ‘high’ COI (n = 14 studies). In most cases, this was because the programme developer was also the evaluator or a named author on the evaluation publication (e.g., DeSmet et al., 2018; Fekkes et al., 2016; Gradinger et al., 2015; Schultze et al., 2016). Two studies were labelled as ‘possible’ COI (i.e., Anthanasiades et al., 2015; Ortega-Ruiz et al., 2012) and two studies were labelled as ‘low’ COI (i.e., Roberto et al., 2014; Solomontos-Kountouri et al., 2016). Given the discrepancy in the numbers of studies in subgroups based on the COI variable, subgroup analyses were not conducted.

Table 29

Systematic review of the intervention components that are included in cyberbullying intervention programmes: Part one

Study & Intervention	***Programme focus***	***Whole-school approach & Anti-bullying policies***	***Classroom rules & Classroom management***	***Peer Involvement***	***Parent Involvement***	***Teacher Involvement***
Athanasiades et al. (2015) Tabby project	Cyber & Internet use	Age-appropriate guidelines about safe use of the Internet, but also the emotional and legal effects of cyberbullying	Pilot programme was designed to be implemented in the classroom by teachers, but no clear mention of specific classroom-based rules or management strategy	Intervention involved four videos, followed by group discussions	There was no reference to parental involvement in the intervention, beyond parents and guardians providing consent.	Teachers attended a 9-hour training seminar that covered topics relevant to cyberbullying and also the role of the school in prevention efforts.
Chaux et al. (2016) Media Heroes	Internet risks and safety	No specific reference to using a whole-school approach or anti-bullying policies is made.	Programme was implemented in the classroom, in both a long and short version, but no specific mention of classroom rules or management techniques are mentioned.	Intervention activities included peers in informal (e.g., in-class role playing, debates, and cooperative learning) and formal ways. The intervention used a participant-role approach and promoted assertive ways for bystanders to intervene in cyberbullying incidences.	Intervention activities included student-parent presentations.	It is unclear whether or not teachers were involved in intervention activities, beyond assisting with data collection and gathering consent forms.
Cross et al. (2016) Cyber Friendly Schools	Positive use of technology; cyber safety	Whole-school programme designed using an ecological systems approach, so many school-level factors were targeted. The intervention also included a review of school policies on bullying.	The intervention activities were implemented in the classroom, but specific rules or behavioural management strategies are not outlined.	Peers were involved formally, through the use of trained student-leaders to promote positive use of technology, and informally through classroom-based activities that included highlighting students' rights and responsibilities online, particularly as a positive bystander.	Parent cyberbullying prevention training activities were included in the intervention.	Teachers implemented the intervention in their respective classrooms and were invited to take part in a 3-hour training programme in the first 2 years of the intervention.

DeSmet et al. (2018) Friendly Attac	Promote positive bystander behaviour	Intended to be implemented in a whole school programme but evaluated in this study as a separate component.	No reference to any specific classroom activities in the evaluation study.	The intervention specifically targeted adolescent bystanders in cyberbullying and positive bystander behaviour options were reinforced.	Parents were not involved.	Teachers were not involved in intervention activities beyond facilitating the implementation.
Espelage et al. (2015) Second Step	Skill building and practice; socio-emotional skills	The programme is described as a universal curricular classroom intervention.	Programme was implemented in classrooms through structured lessons, but no reference to specific classroom rules or management is made.	Lessons include interactive components where students would engage in small-group discussion, dyadic exercises, whole-class instruction and individual projects. Bullying lessons also included information on 'responding'.	Parents were not included in the intervention activities specifically, beyond providing consent for student participation.	Teachers participated in a 4-hour training session and completed online implementation logs after implementing each lesson.
Fekkes et al. (2016) Skills for life	Enhancing social, emotion, and moral skills	The programme is described as a universal school-based programme.	Lessons are delivered in the classroom, but no specific classroom-rules or management strategies were implemented.	Students learn from each other in the classroom and intervention activities include role-playing, discussion, and feedback.	No direct parent involvement is mentioned in the evaluation study.	Lesson plans were provided to teachers and they took part in 2 periods of 3-day training and 2 follow-up sessions over the course of the evaluation.
Garaigordobil et al. (2015a) Cyberprogram 2.0	Bullying, cyberbullying, coping strategies and social skills	The intervention is delivered in schools, to groups of adolescents, but not specifically a ‘whole-school’ approach. No reference is made to any specific anti-bullying policies.	The programme was implemented in classrooms, and the authors note that keeping the group, location and time of implementation consistent was important, but classroom rules and management was not specifically a component.	Intervention activities included role playing, brainstorming, and guided group discussion and so peers were informally involved. The analyses looked at the consequences for all roles, including observers, but targeting bystanders was not a specific aim.	Parents were not involved in the intervention beyond providing parental consent.	Teachers lead the intervention and are provided with detailed implementation manuals and lesson structures.

Gradinger et al. (2015) ViSC	A general antibullying programme can also prevent cyberbullying	The intervention programme was designed using a socio-ecological model and aims to target risk and protective factors on multiple levels, including the school but specific school-level intervention elements are not described.	Teachers are trained to identify bullying in their classrooms and also how to intervene and prevent bullying on both the school and the class levels. Students engage in a class project, which aims to empower students to be responsible for incidences in the classroom.	Student-centred instruction in the class project encourages participants to work together to finds ways to prevent aggression. A second, smaller project allows students to work together to achieve a common positive goal, not necessarily related directly to bullying prevention.	Parents were not involved in the intervention beyond the provision of parental consent for participation.	Teachers are trained by ViSC researchers and then proceed to train students on anti-bullying intervention and prevention.
Menesini et al. (2012) NoTrap!	Web-based discussion forum for bullying involved students	Before implementation of the intervention, a school-wide launch event was held to raise awareness and discuss cyberbullying and bullying.	An awareness-raising meeting with classes are held but authors do not refer to a specific set of classroom rules.	Peer-led model where peer educators are trained to moderate an online forum where classmates could post questions and engage in discussions about cyberbullying. Peers also and also led in-class awareness raising presentations	Parents were not involved in the intervention beyond providing consent for their children to participate.	Teachers were not necessarily involved in the implementation of intervention activities.
Ortega-Ruiz et al. (2012) ConRed	Internet and social media safety and Cyberbullying	Whole-school approach and proactive bullying policies are implemented as part of the intervention. A holistic approach was used to include all members of the school, but intervention activities primarily involved students. Leaflets, posters and other media were used to raise awareness on the school-level.	Neither classroom management nor classroom rules are specified as intervention activities.	In-class debates in response to relevant videos or news items were led by the researcher and so there was a level of information peer involvement.	Information for parents regarding safe Internet use was provided as part of the intervention and advice given to parents on how best to protect their children. The intervention also included the creation of safe space for student-parent-teacher cooperation' ran concurrent sessions with families	Information for teachers regarding safe Internet use was provided and advice on cyberbullying behaviours was given to teachers.

Palladino et al. (2012) NoTrap!	Web-based discussion forum for bullying involved students	The intervention did not adopt a whole-school approach, but a launch event to present the intervention programme and raise awareness about cyberbullying was held prior to implementation.	Classroom rules and classroom management were not included in the intervention.	The intervention programme was peer-led and self-nominated peer educators were trained to moderate an online forum dedicated to bullying issues. Peer educators worked collaboratively with teachers to produce a final class project, for example, a short movie, a peer counselling service, a new set of ICT guidelines or a poster advertising the project. The second edition of the programme also incorporated more content on involving bystanders.	Parents were not involved in the intervention.	In-class activities administered in conjunction with teachers.
Palladino et al. (2016) NoTrap!	Web-based discussion forum for bullying involved students	The intervention did not adopt a whole-school approach.	No specific classroom rules or classroom management techniques were included.	Peers were involved on many levels, and the intervention was formally peer-led as trained peer educators moderated an online forum and also led in-class activities. Intervention activities in the 3rd edition focused on co-operative group work with classmates and increased involvement of bystanders.	Parents were not involved in the intervention.	In-class activities administered in conjunction with teacher.
Pieschl et al. (2017) Surf-Fair	Cyberbullying intervention	Authors describe the intervention as using a student-centred and constructivist anchored instruction.	No classroom rules or management strategies are referred to.	Emphasis on group work in intervention activities, where students are asked to approach a cyberbullying scenario from the perspective of each cyberbullying role. The curriculum also included a dedicated bystander unit.	Parents were not involved in the intervention activities beyond providing consent for their children to participate.	Detailed intervention manual was provided for teachers and early research focused on the feasibility of using teachers to implement the intervention.

Schultze et al. (2016) Media Heroes	Changes in attitudes about cyberbullying and promoting empathy for victim; Improving empathy, social and online skills, media literacy, specific action alternatives	Universal, modularised and theoretically informed programme implanted in schools. School-level activities included raising awareness about cyberbullying.	The intervention was classroom-based but no specific classroom rules or management was involved.	Informal peer involvement occurred through classroom-based role plays and discussions during intervention activities.	Parents were not involved in the intervention.	Teachers received 8 hrs of training over 2 days and a programme manual was supplied.
Solomontos-Kountouri et al. (2016) ViSC	A general antibullying programme can also prevent cyberbullying	Whole-school programme and included a school-level process to increase the shared responsibility between teachers to prevent bullying.	Teachers are trained to deal with bullying in their respective classrooms.	On the class-level the intervention aimed to increase students' sense of responsibility for negative behaviours and also encourage positive bystander responses.	The intervention involved the delivery of a 2-hour presentation for parents.	5 units of teacher training were implemented, and teachers then delivered the programme materials in classrooms.
Williford et al. (2013) KiVa	Offline and online bullying and role of bystanders	Implemented a whole-school anti-bullying policy.	KiVa elements involved the creation and enforcement of anti-bullying rules in classrooms and classroom teachers were trained to identify incidences of bullying	In-class activities include role play, discussion, group work and peers were involved informally. The intervention also included activities to encourage peer support for victims. Classroom teachers were also trained to encourage prosocial high-status peers to support victims of bullying	Parents receive a guide containing information about bullying	KiVa manual provided and teachers participate in 2 days of face-to-face training. They also receive ongoing support throughout implementation of the programme.

Table 30

Systematic review of the intervention components that are included in cyberbullying intervention programmes: Part two

Study Intervention	***Curriculum materials***	***Work with bullies & Work with victims***	***Co-operative group work***	***Socio-emotional skills & Mental Health materials***	***Online/Offline overlap***
Athanasiades et al. (2015) Tabby project	Teachers were provided with a TABBY toolkit that included the videos to be shown to participants, an implementation checklist, a teachers' booklet and information about the project's website.	No specific activities were implemented with students involved in cyberbullying.	Authors specify that intervention was designed to be easily adopted by teachers *without* help from people like the school psychologist	There was no reference to any specific socio-emotional skills or mental health content.	Project content and measurement instruments dealt solely with cyberbullying.
Chaux et al. (2016) Media Heroes	The programme curriculum focused on providing information about the definitions of cyberbullying, Internet risks and safety issues and the legal consequences of cyberbullying.	Programme focused on participant roles, but no specific work with those involved in cyberbullying was included.	There was no reference to cooperative group work.	Programme aims to prevent cyberbullying by promoting empathy, amongst other skills.	The authors specifically address the possible overlap of offline and online bullying and hypothesize that there may be a spill-over effect of this cyberbullying prevention program to offline bullying. Both online and offline outcomes are included in the evaluation.
Cross et al. (2016) Cyber Friendly Schools	Classroom teaching and learning programme led by classroom teachers aimed to reduce harm by targeting contexts, contacts, confidentiality, conduct, and content.	No reference to specific activities with participants identified as being involved in cyberbullying.	Pastoral care staff were involved in the intervention to design school policies and 'audit tool' which was used to evaluate the schools' ability to implement preventative measures.	The programme aimed to enhance participants' online social skills and emphasized positive communication, resilience, self-management, conflict resolution and social responsibility.	The authors specify that the cyberbullying intervention programme was directly informed by earlier research on offline bullying prevention.
DeSmet et al. (2018) Friendly Attac	The game presents a 'ugly people page' to resemble a hate page on a common social media network and the player	Intervention did not include specific activities with identified cyber-	The storyline included in the game was written by professional	No reference to any specific socio-emotional skills or mental health approaches in evaluation study.	Programme focused on cyberbullying only.

	is virtually sent into the future to address the issue by talking to students and respond to the cyberbullying behaviours using positive bystander intervention.	bullies or cyber-victims	storywriters and in collaboration with end users.		
Espelage et al. (2015) Second Step	Teachers implemented structured lesson plans in the classroom. Lesson content was accompanied by DVDs that included videos of student interviews and demonstrations of specific skills.	Authors do not describe specific intervention or prevention activities with children identified as bullies or victims.	Not specified.	Intervention lessons involved topics on empathy and communication including group-work, disagreeing respectfully and assertiveness. Coping with stress, emotion regulation and problem solving were topics that were also included.	Second Step is a well-established anti-bullying programme that has also been evaluated in relation to offline bullying outcomes. The authors do not specify how, or if, the content dealt with both online and offline bullying.
Fekkes et al. (2016) Skills for life	Intervention curriculum was comprised of structured lesson plans that focus on specific skills and issues, such as giving and seeking help, dealing with bullying, norms, sexuality, and conflict with teachers and peers. Intervention activities included DVD and active enactment.	Intervention did not include specific activities with identified cyber-bullies or cyber-victims	No specific mention of co-operative group work.	Based on social-learning theory and included rational emotive behaviour therapeutic elements. Skills targeted included pro-social behaviour, self-awareness, social awareness, self-control, interpersonal skills and ethical decision making.	No specific reference to content relating to offline bullying is made, but both online and offline bullying outcomes are included in the evaluation.
Garaigordobil et al. (2015a) Cyberprogram 2.0	Structured lesson plans were included in the intervention and in total the curriculum was delivered in 19 one-hour sessions and is structured into three main modules.	The intervention programme did not include any specific activities to work with individual bullies or victims.	The programme is designed to be implemented by either teachers, psychologists or school pedagogue.	Along with bullying outcomes, the programme also targeted many socio-emotional skills such as, empathy, active listening, social skills, control anger-impulsivity, coping strategies and conflict resolution.	The programme content focuses on online bullying and other risky Internet behaviours, but the intervention is evaluated for its impact on offline bullying also.
Gradinger et al. (2015) ViSC	The Viennese Social Competence program is a structured intervention programme that includes detailed lessons outlines.	Intervention activities did not include specifically target individual	Implemented a train-the-trainer model but did not strictly involve cooperative group work.	Project aims to allow students to improve social skills and to practice these skills in collaborative group projects.	Authors specify that the purpose of the evaluation is to assess the effectiveness of a general anti-bullying programme on cyberbullying behaviors, but do

		cyberbullies or cyber-victims.			not outline how content about online and offline bullying was included. Both online and offline outcomes were included in the evaluation.
Menesini et al. (2012) NoTrap!	No curriculum is provided as the main intervention component involves a peer educator moderating an online support forum. Each educator worked for a period of two weeks and were assigned on a rotation schedule. Peer educators moderated the forum and also posted content for new threads of discussion.	Individuals who had experienced bullying could post on the online forum and get support from their peers and also trained peer educators.	Peer educators also participated in discussions with local administrators and police to make their city safer and prepared a TV programme about bullying and cyberbullying for a local network.	Peer educators were trained on communication, problem-solving and social skills both in face-to-face and online environments.	No restrictions were set on the content of the online forum, and as such, both online and offline bullying could be discussed. The first edition of the intervention focused exclusively on cyberbullying but both online and offline outcomes were included in the evaluation.
Ortega-Ruiz et al. (2012) ConRed	The curriculum was made up of three units; targeting issues such as privacy and control online, healthy engagement with social networks online and the problems associated with using the Internet or social networks in a malicious way.	No specific intervention activities with individual bullies or victims was included.	External experts were directly involved in the intervention to train students and worked in collaboration with the school climate planning team.	Social skills or mental health approaches were not directly targeted by the intervention.	The content of the intervention focused primarily on cyberbullying and safe Internet activity, but the evaluation included both online and offline bullying outcomes.
Palladino et al. (2012) NoTrap!	In-class curriculum and classroom activities were led by trained peer educators. In the second edition of the programme a Facebook group to compliment the webpage forum was created.	Victims and/or bullies could seek support or advice from trained peers using the online forum.	Student received training from external experts.	The second edition of the intervention programme included information about coping strategies that individuals could use if bullied online or offline. Peer educators were also trained on communication skills and empathy.	Both online and offline bullying issues could be discussed in the online forum, and the authors do not specify if the offline activities also incorporated offline bullying behaviours. Both school- and cyber-bullying outcomes were included in the evaluation.
Palladino et al. (2016) NoTrap!	The same curriculum as previous editions of the programme (i.e., Menesini et al., 2012; Palladino et al.,	Anyone who was experiencing bullying or had concerns about bullying could post in	First phase was led by psychologists who then train the peer educators.	Offline intervention activities focused on socio-emotional skills such as empathy and problem solving.	Both online and offline outcomes were included in the evaluation study.

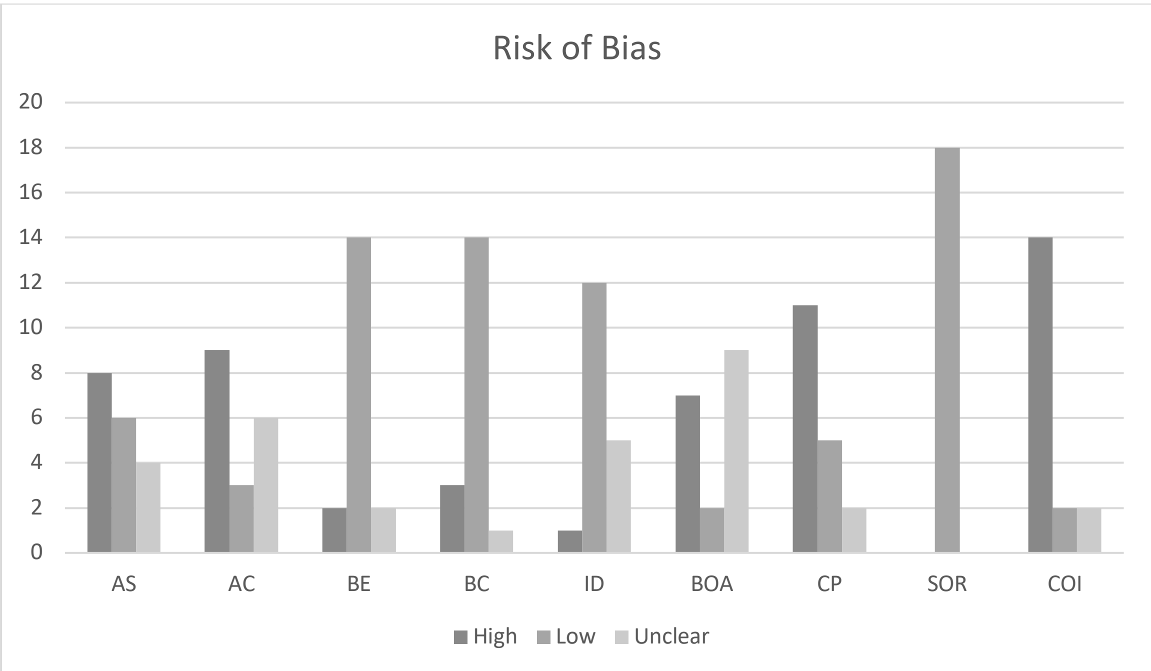
	2012) was used but offline peer-led activities were standardized in the 3rd edition.	the online forum to get support and advice from trained peer educators.			
Pieschl et al. (2017) Surf-Fair	The intervention did include specific curriculum materials and is described as being less comprehensive and shorter than other intervention programmes.	Work with individual victims or bullies was not specifically an intervention component.	Trained research assistants implemented the intervention in the first wave of implementation and teachers observed and asked questions.	Not specified.	Intervention content was specific to cyberbullying and only cyberbullying outcomes was included in the evaluation.
Schultze et al. (2016) Media Heroes	The intervention curriculum is implemented in classrooms and can be provided in a long (ten weeks of 90-minute sessions per week) or a short (total of 4 sessions of 90-minutes) version.	Participant roles in cyberbullying were targeted, but the intervention activities did not specifically include individual bullies or victims.	No co-operative group work was involved in the intervention.	Cognitive and affective empathy were targeted in intervention activities and cognitive-behavioural methods such as positive reinforcement and moral reasoning were used.	Cyberbullying was the primary focus of intervention activities and only online bullying outcomes were included in the evaluation.
Solomontos-Kountouri et al. (2016) ViSC	Teachers were provided with a detailed programme manual with structured lesson plans.	The intervention also tried to empower victims of bullying and build resilience	Trained counsellors, psychologists, and teacher-researchers were also included in the evaluation.	Social and cultural competencies were targeted during intervention activities.	The content overlap was not specified but both online and offline bullying outcomes were included in the evaluation.
Williford et al. (2013) KiVa	Classroom-based lessons were outlined in a detailed intervention manual.	Staff members were trained to conduct discussions with bullies and victims	Cooperative group work was not involved in the intervention.	Bystander empathy & self-efficacy were targeted during social skills training.	It is unclear how the content overlapped to include both traditional bullying and cyberbullying. This study reported only cyberbullying outcomes but is part of a larger evaluation of the KiVa programme on bullying.

10.1.4 Risk of bias

The method of assessing risk of bias was identical to that used in the school-bullying meta-analysis. Figure 10 presents the results of the risk of bias analysis for each item on the EPOC tool as applied to included evaluations of cyberbullying intervention programmes. As the graph suggests, the distribution of studies considered to be high risk on each of the items varied. The mean risk of bias score was 9.72, with a range from a minimum score of 2 to a maximum score of 16. Meta-regression was used to test the relationship between risk of bias and cyberbullying outcomes.

Figure 10

Risk of bias analysis for cyberbullying outcomes



Note: AS = allocation sequence; AC = allocation concealment; BE = baseline equivalence on outcomes; BC = baseline equivalence on participant characteristics; ID = incomplete outcome data; BOA = blind outcome assessment; CP = contamination protection; SOR = selected outcome reporting; COI = conflict of interest.

10.2 Meta-analysis

The following section of this chapter presents the results of the meta-analysis of cyberbullying intervention and prevention programmes. The weighted mean effect sizes for cyberbullying perpetration and victimisation are presented, as well as the results from analysis of heterogeneity and publication bias tests. Subgroup analyses for cyberbullying outcomes are also presented.

10.2.1 Cyberbullying perpetration

The effect sizes for cyberbullying perpetration outcomes are presented in Table 31, and graphically in Figure 11. Overall, the results indicate that cyberbullying intervention programmes were effective in reducing cyberbullying perpetration under a random effects model (OR = 1.233; 95% CI 1.04 – 1.46; $z = 2.41$; $p = 0.02$). Under the MVA model, the mean effect size was OR = 1.144 (95% CI 0.99 – 1.33; $z = 1.79$; $p = 0.07$). This result suggests that participants who received an anti-cyberbullying programme were less likely to report engaging in cyberbullying perpetration in comparison to control participants who did not receive the programme.

Analysis of the publication bias funnel plots (Figure 12) of effect sizes for cyberbullying perpetration suggests that there were potentially missing studies to the left of the mean effect size (i.e., negative or undesirable intervention effects). This could suggest publication bias, as research has found that studies that report non-statistically significant or negative effects are less likely to be published (Easterbrook et al., 1991). Therefore, Duval and Tweedie's (2000a) trim and fill procedure was applied to the data in the current meta-analysis. This approach 'trims' asymmetric effects and then re-estimates the 'true centre' of the funnel plot, or in other words, re-estimates the mean effect size (Duval & Tweedie, 2000b).

In the cyberbullying meta-analysis, Duval and Tweedie's trim and fill computation trimmed three effect sizes for cyberbullying perpetration outcomes, and the adjusted mean effect size was OR = 1.092 (95% CI 1.02 – 1.17) under the MVA model and OR = 1.086 (95% CI 0.90 – 1.31) under a random effects model. These findings suggest that there was a potential publication bias in the present analysis.

Table 31

Meta-analysis results: Cyberbullying perpetration outcomes

Study(s)	*OR*	*95% CI*	*z*	*p*
Randomised Controlled Trials (*n* = 9)				
Chaux et al. (2016); Wölfer et al. (2014)	1.71	1.20 – 2.44	2.93	0.003
Cross et al. (2016); Shaw et al. (2015)	1.10	0.94 – 1.29	1.16	0.25
DeSmet et al. (2018)	2.05	0.15 – 28.07	0.54	0.59
Espelage et al. (2015)	0.96	0.75 – 1.21	-0.37	0.71
Garagordobil & Martínez-Valderrey (2015a; 2016)	4.05	2.08 – 7.87	4.12	0.001
Gradinger et al. (2015); Gradinger et al. (2016)	1.35	1.11 – 1.66	2.91	0.004
Roberto et al. (2014)	0.69	0.33 – 1.46	-0.97	0.33
Schultze-Krumbholz et al. (2016)	1.69	1.17 – 2.42	2.82	0.005
Williford et al. (2013)	1.02	0.86 – 1.21	0.23	0.82
Random Effects: RCTs	1.34	1.09 – 1.64	2.76	0.006
MVA model: RCTs	1.18	0.99 – 1.40	1.90	0.06
Quasi-experiments with before/after measures (*n* = 9)				
Menesini et al. (2012a; Study 1)	0.85	0.40 – 1.80	-0.43	0.69
Menesini et al. (2012a; Study 2); Palladino et al. (2012)	1.27	0.80 – 2.02	1.02	0.31
Ortega-Ruiz et al. (2012); Del Rey et al. (2012; 2015; 2016)	1.08	0.79 – 1.49	0.48	0.63
Palladino et al. (2016; Trial 1)	2.50	1.56 – 3.99	3.83	0.001
Palladino et al. (2016; Trial 2); Males	2.39	1.25 – 4.56	2.64	0.008
Palladino et al. (2016; Trial 2); Females	1.44	0.72 – 2.91	1.03	0.31
Pieschl et al. (2017)	0.83	0.43 – 1.62	-0.54	0.59
Solomontos-Kountouri et al. (2016); 7th grade	1.00	0.72 – 1.39	0.000	1.00
Solomontos-Kountouri et al. (2016); 8th grade	0.58	0.42 – 0.79	-3.32	0.001
Random Effects: BA/EC designs	1.17	0.85 – 1.62	0.95	0.34
MVA model: BA/EC designs	1.07	0.79 – 1.45	0.41	0.69
Overall: Random Effects model	**1.233**	**1.04 – 1.46**	**2.41**	**0.02**
Overall: MVA model	**1.144**	**0.99 – 1.33**	**1.79**	**0.07**

Note. n = number of independent effect sizes; RCT = randomised controlled trials; BA/EC = before-after/experimental control designs; MVA = multiplicative variance adjustment

Figure 11

Forest plot of odds ratio effect sizes for cyberbullying perpetration outcomes

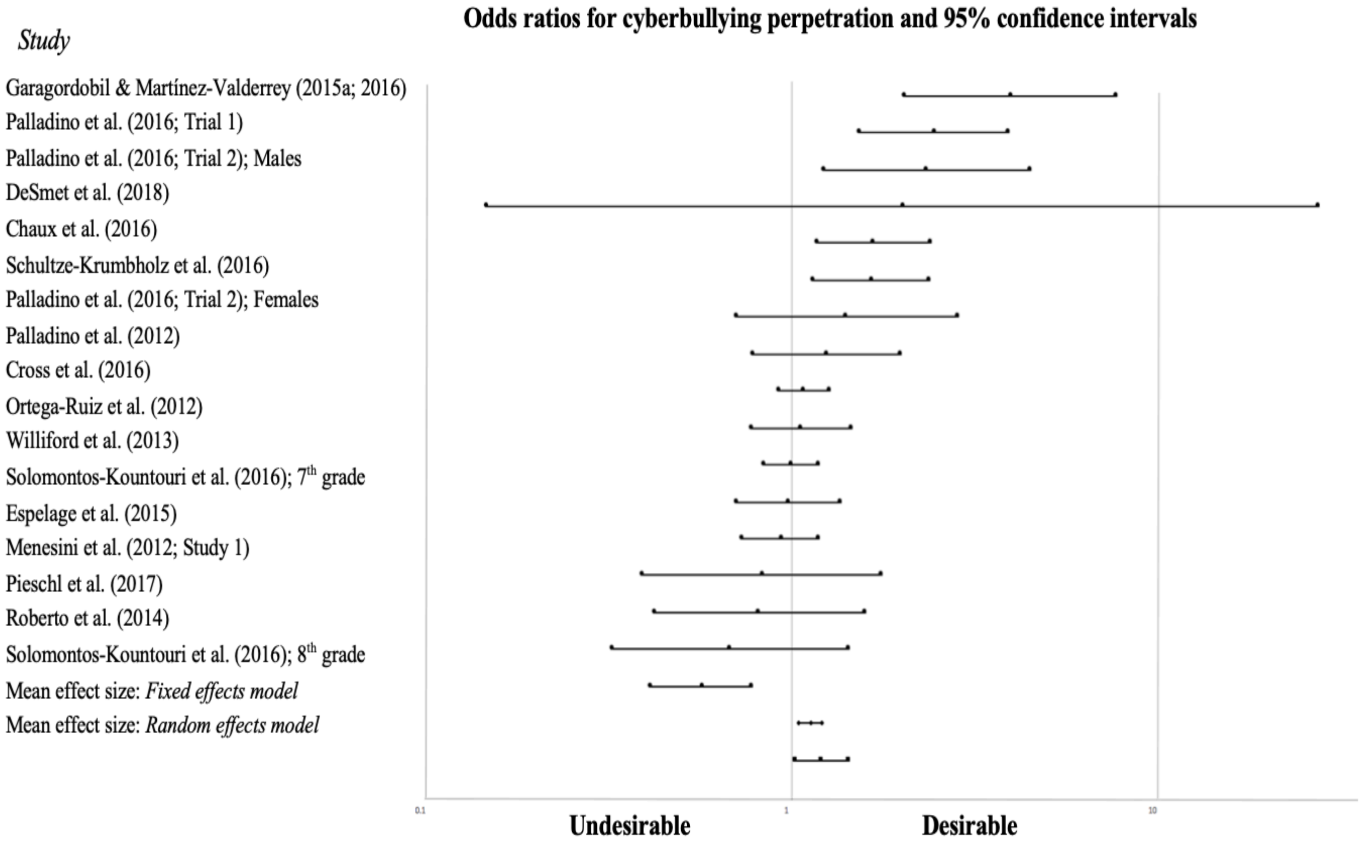
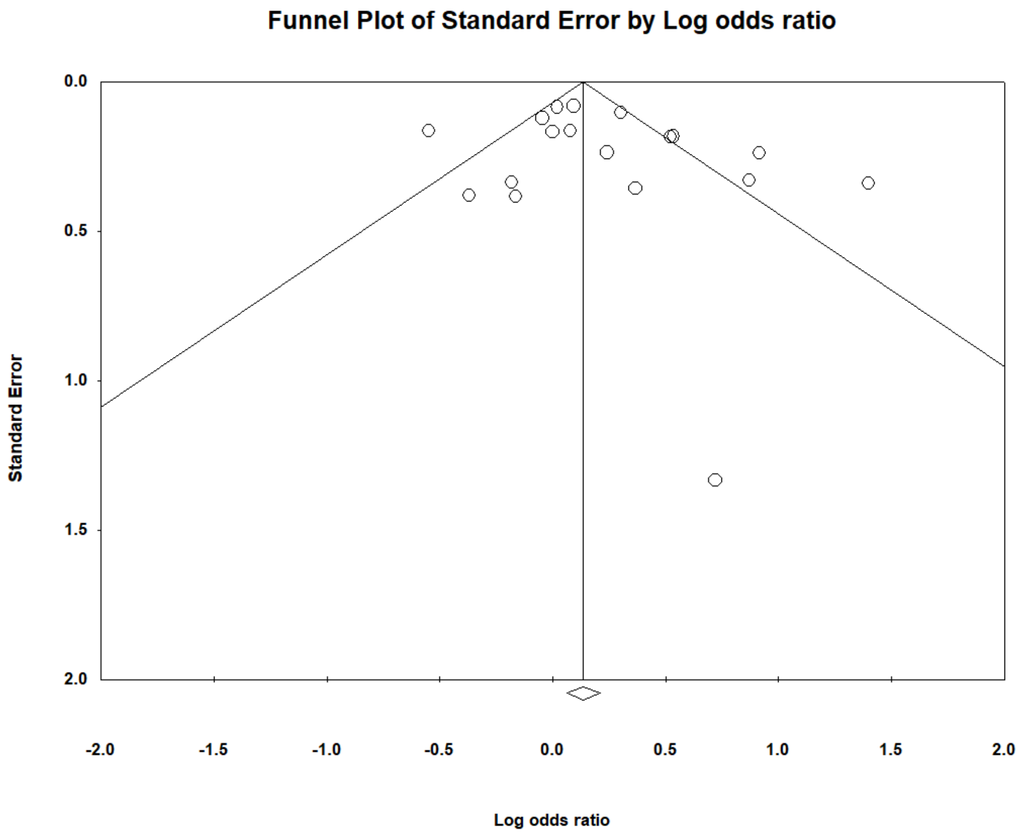


Figure 12

Funnel plot for cyberbullying perpetration outcomes



10.2.2 Cyberbullying victimisation

The effect sizes for cyberbullying victimisation outcomes are presented in Table 32, and graphically in Figure 13. Our meta-analysis also found that cyberbullying interventions were effective in reducing cyberbullying victimisation. Under a random effects model, the mean effect size (OR = 1.227, 95% CI 1.05 – 1.44, z = 2.53; p = 0.01) suggests that participants who participated in a cyberbullying intervention programme were less likely to report experiencing cyberbullying victimisation in comparison to control participants who did not engage with the programme.

Under the MVA model the summary effect size also suggests that intervention programmes are effective at reducing cyberbullying victimisation (OR = 1.227 (95% CI 1.08 – 1.40; z = 3.15; p = 0.002).

Similar to cyberbullying perpetration outcomes, analysis of the publication bias funnel plots (Figure 14) of effect sizes for cyberbullying victimisation suggests that there were potentially missing studies to the left of the overall mean effect size and therefore publication bias may possibly be present. Duval and Tweedie's trim and fill computation trimmed two effect sizes for cyberbullying victimisation outcomes. The mean effect size for cyberbullying victimisation outcomes changed when two studies were trimmed using Duval and Tweedie's trim-and-fill procedure. Under the MVA, the adjusted mean effect was OR = 1.213 (95% CI 1.13 – 1.30) and under a random-effects model the adjusted mean effect was OR = 1.171 (95% CI 1.01 – 1.37). These effect sizes were quite different to the summary mean effect size when all primary evaluations were included.

Table 32

Meta-analysis results: Cyberbullying victimisation outcomes

Study(s)	*OR*	*95% CI*	*z*	*p*
Randomised Controlled Trials (*n* = 10)				
Athanasiades et al. (2015)	1.26	0.74 – 2.13	0.85	0.39
Chaux et al. (2016); Wölfer et al. (2014)	1.23	0.87 – 1.75	1.15	0.25
Cross et al. (2016); Shaw et al. (2015)	1.20	1.03 – 1.41	2.27	0.02
DeSmet et al. (2018)	1.83	0.25 – 13.24	0.59	0.55
Espelage et al. (2015)	0.83	0.66 – 1.06	-1.52	0.13
Fekkes et al. (2016)	0.70	0.25 – 1.98	-0.67	0.50
Garagordobil & Martínez-Valderrey (2015a; 2016)	2.53	1.31 – 4.86	2.78	0.005
Gradinger et al. (2015); Gradinger et al. (2016)	1.31	1.07 – 1.60	2.58	0.01
Roberto et al. (2014)	1.49	0.86 – 2.57	1.42	0.16
Williford et al. (2013)	1.40	1.24 – 1.58	5.59	0.001
Random Effects: RCTs	1.228	1.05 – 1.44	2.51	< 0.001
MVA model: RCTs	1.262	1.12 – 1.42	3.84	< 0.001
Quasi-experiments with pre/post measures (*n* = 9)				
Menesini et al. (2012a; Study 1)	0.77	0.36 – 1.63	-0.69	0.49
Menesini et al. (2012a; Study 2); Palladino et al. (2012)	1.53	0.95 – 2.47	1.76	0.08
Ortega-Ruiz et al. (2012); Del Rey et al. (2012; 2015; 2016)	1.29	0.94 – 1.79	1.59	0.11
Palladino et al. (2016; Trial 1)	2.03	1.28 – 3.23	2.99	0.003
Palladino et al. (2016; Trial 2); Males	2.25	1.18 – 4.31	2.46	0.01
Palladino et al. (2016; Trial 2); Females	0.85	0.42 – 1.71	-0.46	0.64
Pieschl et al. (2017)	2.05	1.06 – 3.99	2.12	0.03
Solomontos-Kountouri et al. (2016); 7th grade	1.10	0.79 – 1.52	0.55	0.58
Solomontos-Kountouri et al. (2016); 8th grade	0.50	0.36 – 0.69	-4.17	0.001
Random Effects: BA/EC designs	1.220	0.86 – 1.74	1.10	0.27
MVA model: BA/EC designs	1.109	0.79 – 1.55	0.61	0.55
Overall: Random Effects model	**1.227**	**1.05 – 1.44**	**2.53**	**0.011**
Overall: MVA model	**1.231**	**1.08 – 1.40**	**3.15**	**0.002**

Note. *n* = number of independent effect sizes; RCT = randomised controlled trials; BA/EC = before-after/experimental control designs; MVA = multiplicative variance adjustment

Figure 13

Forest plot of odds ratio effect sizes for cyberbullying victimisation outcomes.

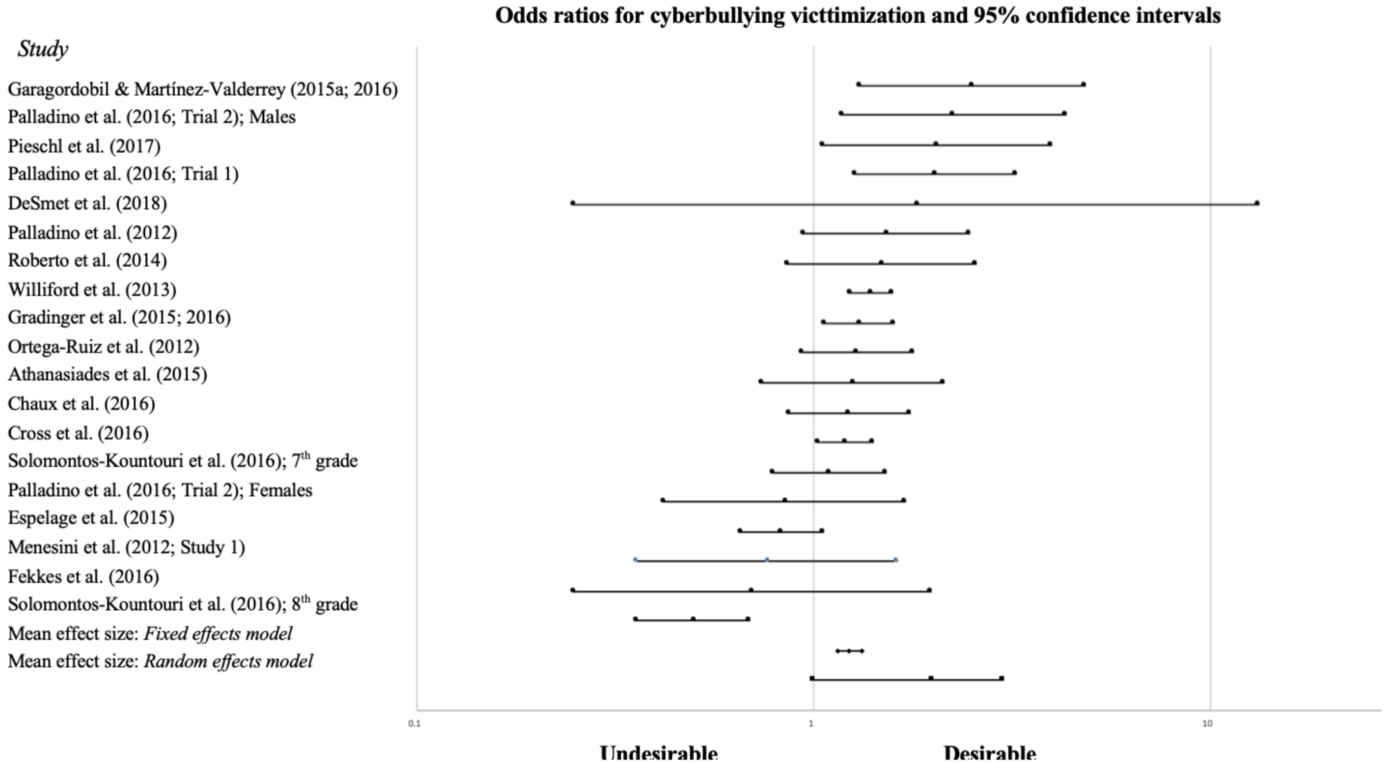
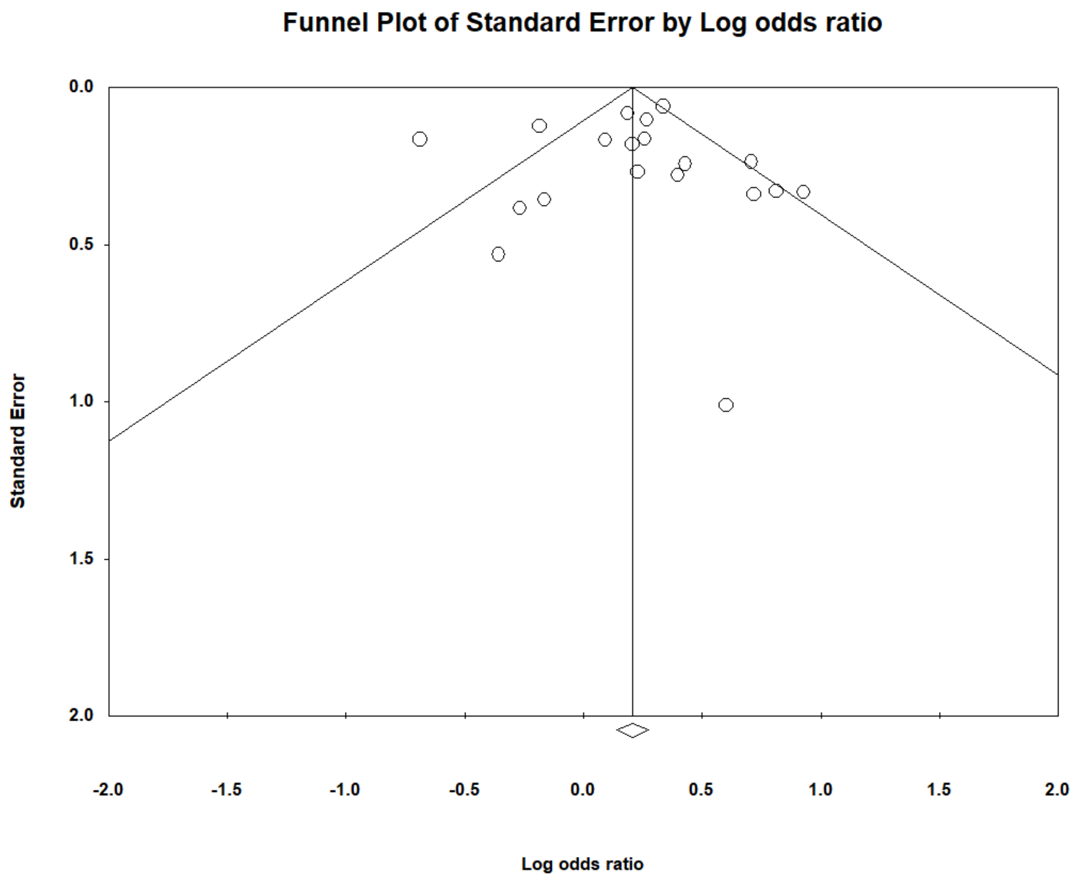


Figure 14

Funnel plot for cyberbullying victimisation outcomes



10.2.3 Analysis of heterogeneity

Heterogeneity is estimated as the excess variation that might exists when the total amount of between-study variance and within-study random error is compared. In the cyberbullying meta-analysis, there was significant heterogeneity between observed effect sizes for both cyberbullying perpetration (Q = 67.49; df = 17; p < 0.001; I^2 = 74.81) and cyberbullying victimisation (Q = 65.31; df = 18; p < 0.001; I^2 = 72.44) outcomes.

A number of variables were explored as potential factors that might explain the significant heterogeneity between primary evaluations. However, in comparison with the school-bullying meta-analysis, fewer variables were included given that far fewer primary evaluations were included in the cyberbullying meta-analysis. This limitation is discussed further in Chapter 11. Variables coded and compared in the cyberbullying meta-analysis included: evaluation design, location of the intervention, age of participants, sample size, measurement instrument, and whether or not the intervention targeted both offline and online bullying simultaneously.

10.3 Moderator and mediator analyses

Subgroup analyses analogous to a one-way ANOVA were computed to examine differences between subgroups of primary studies. This method is explained in detail in Chapter 2 (see section 2.7). The following sections describe the subgroup analyses that were conducted in relation to cyberbullying perpetration and victimisation outcomes.

10.3.1 Sample size

The possible relationship between the number of participants included in primary evaluations of cyberbullying intervention programmes and the effect size was examined using meta-regression. Analysis[16] suggested that there was no relationship between the total sample size and cyberbullying perpetration outcomes under either the MVA model (p = .127) or the random effects model (p = .928) of meta-regression. Similarly, the sample size did not significantly predict variations in cyberbullying victimisation outcomes under the MVA model (p = .227) or the random effects model (p = .462).

10.3.2 Age of participants

The relationship between age of participants and cyberbullying outcomes was investigated using meta-regression and subgruop analyses analogous to the ANOVA with one dichotomous variable (i.e., agecat) and one continuous variable (i.e., the mean age). The results of subgroup analyses with the dichotomous age variable are presented in Table 33.

---

[16] In all cases, the regression coefficient and the standard error were < .0001 and were thus not reported here.

In relation to cyberbullying perpetration outcomes, interventions that were implemented with older participants (aged 14 years old and over) were associated with a statistically significant larger mean effect size under the MVA model (Q_B = 7.309, *df* = 1, *p* = .008) and a random effects model (Q_B = 5.898, *df* = 1, *p* = .015). Moreover, the mean effect size for the subgroup of studies that evaluated interventions implemented with participants aged 11 to 13 years old was not statistically significant under both the MVA model and the random effects model. With respect to cyberbullying victimisation outcomes, the difference between groups based on age was not statistically significant under either the MVA model (Q_B = 1.317, *df* = 1, *p* = .251) or the random effects model (Q_B = 1.466, *df* = 1, *p* = .226). Similar to cyberbullying perpetration outcomes, the mean effect size for evaluations implemented with participants aged 11 to 13 years old was not significant overall under the MVA model, or the random effects model.

The relationship between the mean age of participants in each primary study and cyberbullying outcomes was also explored using meta-regression. Under the MVA model the mean age did not significantly predict cyberbullying perpetration (*B* = 0.057, SE = .052, *p* = .273). Similar results were found under the random effects model (*B* = 0.099, SE = .054, *p* = .063). The mean age of participants also did not significantly predict cyberbullying victimisation outcomes under either the MVA model (*B* = 0.049, SE = .050, *p* = .309) or the random effects model (*B* = 0.057, SE = .051, *p* = .265).

Table 33

Cyberbullying moderator analysis results: Age of participants as a dichotomous variable

	MVA Model					Random effects model			
Category (*n*)	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
Cyberbullying perpetration (*n* = 18 effect sizes)									
Younger (9)	1.056	0.878 – 1.270	.281	29.282 (*p* < .001)	72.679	1.007	0.795 – 1.277	.952	.052
Older (9)	1.293	1.032 – 1.621	.013	30.899 (*p* < .001)	74.109	1.542	1.203 – 1.977	.001	.116
Cyberbullying victimisation (*n* = 19 effect sizes)									
Younger (9)	1.191	0.972 – 1.460	.092	47.699 (*p* < .001)	83.228	1.117	0.888 – 1.405	.346	.099
Older (10)	1.295	1.108 – 1.514	< .001	16.416 (*p* = .059)	45.175	1.374	1.076 – 1.755	.012	.040

Note. Younger = participants aged between 11 and 13 years old; Older = participants aged 14 years and older.

10.3.3 Evaluation methodology

The summary effect sizes for cyberbullying perpetration and victimisation were compared across the evaluation design used. Thus, evaluations using randomised controlled trials (RCTs) and before-after/quasi-experimental-control (BA/EC; see Tables 32 and 33) were compared to investigate potential explanation for the heterogeneity observed between effect sizes.

In relation to cyberbullying perpetration outcomes, RCT designs yielded larger effect sizes (OR = 1.34; 95% CI 1.09 – 1.64) than quasi-experimental designs (OR = 1.17; 95% CI 0.85– 1.62) under a random effects model. The same was observed for cyberbullying victimisation outcomes, with RCTs producing a larger summary effect size (OR = 1.24; 95% CI 1.07 – 1.44) than quasi-experimental designs (OR = 1.22; 95% CI 0.86 – 1.74). Moreover, summary effect sizes under a fixed effects model, and the MVA model, followed the same pattern (see Tables 32 and 33).

There were also significant differences between methodological designs. For cyberbullying perpetration outcomes, there was significant differences between RCT studies (Q = 34.69; df = 8; p < 0.0001; I^2 = 76.94) and BA/EC studies (Q = 31.59; df = 8; p < 0.001; I^2 = 74.68). For cyberbullying victimisation outcomes, there was significant heterogeneity between RCT studies (Q = 22.28; df = 9; p = 0.011; I^2 = 57.71) and BA/EC studies (Q = 41.57; df = 8; p < 0.001; I^2 = 80.76). The total within-groups heterogeneity ($Q_{W;}$ Lipsey & Wilson, 2001) for cyberbullying perpetration (Q_W = 66.28; df = 16; p < 0.001) and cyberbullying victimisation was also significant (Q_W = 62.85; df = 17; p < 0.001). Moreover, for both cyberbullying perpetration and cyberbullying victimisation outcomes the heterogeneity between groups (Q_B; Lipsey & Wilson, 2001) was not statistically significant (Q_B = 1.21; df = 1; p = 0.27 and Q_B = 2.46; df = 1; p = 0.12 respectively).

10.3.4 The offline-online overlap

Two variables relating to the overlap of online and offline bullying were included in the subgroup analyses: the inclusion of both online and offline bullying content and the inclusion of online and offline bullying outcomes. In relation to the content variable the results are outlined in Table 34. Under the MVA model, both subgroups of studies, i.e. both those that included online and offline bullying content and those that didn't, gave mean effect sizes that indicated an undesirable change in cyberbullying perpetration behaviours. However, these mean effect sizes were not statistically significant. The difference between the mean effect sizes was marginal and also not statistically significant (Q_B = 0.367; *df* = 1; *p* = .545). The results were similar under the random effects model but the mean effect size for studies that did include both online and offline bullying content indicated a desirable change in cyberbullying perpetration outcomes. Yet, the differences between the mean effect sizes for subgroups was not statistically significant (Q_B = 0.426; *df* = 1; *p* = .514).

Under the MVA model, the mean effect size for studies that did not include both online and offline bullying content in intervention activities were associated with a desirable effect on reports of cyberbullying victimisation, whereas studies that did include both online and offline content had an undesirable effect. Additionally, the difference between groups was statistically significant (Q_B = 7.949; *df* = 1; *p* = .005). The same result was observed under the random effects model, also but using this computational model the differences were not statistically significant (Q_B = 0.958; *df* = 1; *p* = .328).

Comparisons were also conducted using subgroup analyses for subgroups of studies that did, or did not, include both online and offline bullying outcomes. The results of the moderator analyses for this specific variable is presented in Table 35. Under the MVA model of meta-analysis, the mean effect sizes for both subgroups of studies indicated undesirable effect on cyberbullying perpetration outcomes, but these were not statistically significant.

Moreover, the marginal differences between the mean effect sizes were not statistically significant either (Q_B = 0.771; *df* = 1; *p* = .380). Comparatively, under the random effects model, studies that included both online and offline bullying outcomes in the evaluation had a mean effect size that suggested a marginally desirable change in cyberbullying perpetration outcomes. However, mean effect sizes for both subgroups were not statistically significant, and neither was the difference (Q_B = 0.099; *df* = 1; *p* = .753).

In relation to cyberbullying victimisation, studies that did not include both online and offline bullying outcomes in the evaluations had an overall positive mean effect size that was statistically significant under the MVA model. However, those that did include both online and offline bullying outcomes had a statistically significant mean effect size that suggested an undesirable intervention impact. Moreover, these differences were statistically significant (Q_B = 9.959; *df* = 1; *p* = .002). The differences between mean effect sizes were marginal and was not statistically significant under the random effects model (Q_B = 2.299; *df* = 1; *p* = .129).

10.3.5 Risk of bias analysis

The relationship between the overall risk of bias score and cyberbullying outcomes was assessed using meta-regression. Under the MVA model, the total risk of bias score did not predict variations in cyberbullying perpetration outcomes (*B* = -0.013, SE = 0.011, *p* = .218). Similarly, under the random effects model of meta-regression the relationship was not statistically significant either (*B* = 0.004, SE = 0.022, *p* = .871). Moreover, under the MVA model, the total risk of bias score did not predict variations in cyberbullying victimisation outcomes (*B* = .023, SE = 0.014, *p* = .097). This relationship was not statistically significant under a random effects model of meta-regression either (*B* = 0.019, SE = 0.027, *p* = .475).

Table 34

Cyberbullying moderator analysis results: Inclusion of both online and offline bullying content

	MVA Model					Random effects model			
Category (*n*)	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
Cyberbullying perpetration (*n* = 18 effect sizes)									
Yes (13)	0.964	0.797 – 1.166	.708	35.506 (*p* = .001)	88.734	0.929	0.685 – 1.262	.639	.001
No (5)	0.907	0.763 – 1.078	.268	4.943 (*p* = .545)	142.768	1.052	0.852 – 1.298	.638	.025
Cyberbullying victimisation (*n* = 19 effect sizes)									
Yes (12)	0.841	0.778 – 0.909	< .001	29.369 (*p* = .002)	62.546	0.987	0.866 – 1.639	.283	.072
No (7)	1.186	0.995 – 1.413	.057	3.349 (*p* = .764)	79.158	1.191	0.809 – 1.204	.898	.012

Note. Yes = studies that *did* include both online and offline bullying content in the intervention; No = studies that *did not* include both online and offline bullying content in the intervention.

Table 35

Cyberbullying moderator analysis results: Inclusion of both online and offline bullying outcomes

	MVA Model					Random effects model			
Category (*n*)	*OR*	*95% CI*	*p*	*Q (p)*	I^2	*OR*	*95% CI*	*p*	Tau^2
Cyberbullying perpetration (*n* = 18 effect sizes)									
Yes (13)	0.900	0.759 – 1.066	.229	33.759 (*p* = .001)	64.454	1.031	0.833 – 1.276	.782	.002
No (5)	0.983	0.796 – 1.213	.871	6.287 (*p* = .179)	36.377	0.973	0.726 – 1.302	.852	.012
Cyberbullying victimisation (*n* = 19 effect sizes)									
Yes (14)	0.840	0.781 – 0.903	< .001	29.993 (*p* = .005)	56.657	0.965	0.802 – 1.161	.702	.011
No (5)	1.266	1.121 – 1.431	< .001	0.715 (*p* = .949)	459.441	1.301	0.926 – 1.826	.129	.093

Note. Yes = studies that *did* include both online and offline bullying outcome in the intervention evaluation; No = studies that *did not* include both online and offline bullying outcome in the intervention evaluation.

Part Four: Discussion

11. Discussion: School- and cyber-bullying

11.1 Overview

The current chapter of this dissertation will present a detailed discussion of the findings from the two meta-analytical reviews undertaken to address the question of 'what works' in school- and cyber-bullying intervention and prevention programmes. A summary and synopsis of the findings from each of the meta-analyses will be presented, along with the results of the post-hoc moderator and mediator subgroup analyses. The implications of these findings are then discussed in detail, and the limitations and avenues for future research are outlined. The interpretation of results from both systematic and meta-analytical reviews are presented alongside one another instead of separately, as per the previous chapters of the dissertation. It is hoped that this will highlight some of the similarities and differences between school-based anti-bullying programmes and school-based cyberbullying intervention and prevention programmes. This chapter will thus conclude with a discussion on the applicability of separating the two phenomena in future research.

11.2 Summary of findings: School-bullying

Overall, the school-bullying meta-analysis found that school-based anti-bullying programmes are effective in reducing both school-bullying perpetration and victimisation. For school-bullying perpetration the weighted mean OR = 1.324 under the Multiplicative variance adjustment model of meta-analysis (MVA) and OR = 1.309 under a random-effects model (RE). Applying the transformation described in Appendix 3 these odds ratios correspond to approximately a 19% to 20% decrease in bullying perpetration. Similarly, the weighted mean OR under a random effects model corresponds to a reduction in bullying perpetration of roughly 18 – 19%.

In comparison, the weighted mean ORs for bullying victimisation outcomes were 1.248 and 1.242 under the MVA and the random effects models respectively. These mean effect sizes correspond to an approximate reduction in bullying victimisation of 15 – 16%. These results suggest that the included interventions might have been slightly more effective at reducing school-bullying perpetration than school-bullying victimisation.

The results of this meta-analysis are consistent with findings from most of the previous reviews that indicate that anti-bullying programmes are effective, as outlined in Chapter 1. There are however some variations in the overall results, and these are largely attributable to the methodological differences in inclusion and exclusion criteria of previous meta-analyses (Ttofi et al., 2014). The mean effect sizes in this dissertation are also consistent with Farrington and Ttofi's (2009; Ttofi & Farrington, 2011) review, although the differences further indicate that moderator variables, such as methodological design or intervention components, may be responsible for the variability. For example, the weighted mean effect sizes for both bullying perpetration and bullying victimisation outcomes estimated in the earlier meta-analysis were slightly larger than those estimated in the present report. Publication year was included as a categorical moderator variable in the present review. The results showed that more recent studies (i.e., those that were *not* included by Farrington & Ttofi, 2009) were significantly different from studies that were included in the earlier review. Namely, recent studies were associated with significantly smaller effect sizes for both bullying perpetration and victimisation outcomes (see Chapter 6, section 6.3.1.4).

The fact that studies considered to have utilised less scientifically rigorous methodological designs were excluded from this updated meta-analysis may explain the differences in the weighted mean effect sizes. Specifically, evaluations conducted using "other experimental-control designs", described in the earlier review as evaluations in which participants were assigned to experimental and control conditions but bullying outcomes were only measured after implementation of the intervention, were excluded from the current updated analyses. This is because attributing any change in behaviours to the intervention is potentially risky in these studies because there may be other reasons why a positive effect of the intervention was observed. For example, the experimental and control groups may not have been comparable at baseline, but this remains unknown as no measure of bullying was obtained prior to implementation. Farrington (2003) emphasized the importance of methodological quality in evaluation research, and in the context of the present meta-analysis it could be argued that not all of the identified evaluations necessarily meet these standards. Thus, the inclusion of these less methodologically rigorous evaluations may explain why the weighted mean effect sizes reported in the earlier review were larger than those reported in the current report. This was evident in the present research also in relation to evaluations conducted using an age cohort design, a less rigorous design, that were associated with statistically significantly larger mean effect sizes for perpetration and victimisation (Flay et al., 2005). Furthermore, this result was seen under both computational models.

Overall however, the results of both meta-analyses suggest that the evaluation methodology is not the only reasonable explanation for differences between effect sizes. Investigation of further moderators is key to an in-depth understanding of variability between effect sizes (Lipsey, 2003). Therefore, a number of different moderators were explored. The following sections of this chapter will aim to discuss the findings obtained by subgroup analyses and also the strengths and limitations of the current analyses and potential avenues for future research.

School-based anti-bullying programmes were also coded for the presence and absence of a number of potential moderator/mediator variables. These variables were classified on four levels: (1) Descriptive; (2) Design; (3) Programme and (4) Outcome. The first three of these are included in the present discussion. Where appropriate, subgroup analyses analogous to the ANOVA were conducted to examine the possible relationship between the presence (or absence) of these variables and the overall effectiveness estimates for school-bullying perpetration and victimisation. In relation to school-bullying outcomes, the results of these analyses are presented in greater detail in Chapter 6 (see section 6.3). The following sections of the current dissertation critically discuss the meaning of the results.

11.2.1 Descriptive level

On the descriptive level, included variables were the location of the evaluation, the sample size of primary evaluations, the age of participants in primary evaluations, to examine the impact of developmental factors on effectiveness, and finally, the publication type and year of primary evaluations.

11.2.1.1 Location of intervention. Overall, the results of the school-bullying meta-analysis are consistent with previous findings and show that school-based anti-bullying programmes have a significant effect in reducing bullying behaviours. However, the present meta-analysis included evaluations of anti-bullying programmes from a wide range of countries and many specific intervention programmes, far more than previous meta-analyses (e.g., Cantone et al., 2015; Chalamandaris & Piette, 2015; Evans et al., 2014; Jiménez-Barbero et al., 2012; Jiménez-Barbero et al., 2016). As a result, the findings of this meta-analysis are robust and have implications for bullying research globally.

The present analysis shows that anti-bullying programmes worldwide are effective in reducing school-bullying perpetration and victimisation by significant amounts, but that evaluations in different countries appear to vary in effectiveness. In Greece, where evaluations included in the meta-analysis were associated with the largest effect sizes, school-bullying perpetration behaviours were reduced by approximately 40%. Evaluations conducted in the Norway, Italy and the US were also effective in reducing bullying perpetration by approximately 21 – 25%.

Anti-bullying programmes implemented and evaluated in Italy were associated with the largest reduction in school-bullying victimisation in the present meta-analysis, with the odds ratio effect size corresponding to an approximate reduction of 31%. Moreover, evaluations conducted in Spain and Norway reduced school-bullying victimisation by approximately 28% and 23%, respectively. Evaluations conducted in Finland, Germany and the UK were also significantly effective, although less so, reducing school-bullying victimisation by approximately 8 – 12%.

There are many potential explanations for the differences in effectiveness observed between countries. For example, definitions of school-bullying, and behaviours that constitute bullying, differ between countries. Previous research conducted by Peter Smith and colleagues (2000; 2016) showed that school-bullying is perceived differently across different countries and cultures and this may explain the variability in effect sizes. For example, Smith, Kwak, and Toda (2016) showed that school bullying in Eastern cultures manifests more often as exclusion or isolation of an individual victim. In comparison, school bullying in Western cultures comprises a wider range of physical, verbal and relational forms of aggression.

The school-bullying meta-analysis included some examples of cases where the same intervention programme was evaluated in different countries (e.g., KiVa programme in Finland (Kärnä et al., 2011a; 2011b; 2013) and in Italy (Nocentini & Menesini, 2016)). Whilst societal practices, educational systems, and individual lifestyles may differ greatly, some argue that there may be differential cross-national applicability of specific intervention programmes. However, there is a current lack of existing research comparing the effectiveness of specific anti-bullying programmes in specific countries. There is increasing evidence in the transportability of interventions from one culture or country to another and in the factors influencing their successful implementation in new contexts (e.g., Gardener et al., 2016; Webster-Stratton et al., 2012).

The research on transportability in the area of school-bullying intervention is still in its infancy and the current research shows that implementing a programme designed in one country may not be as successful when implemented in another. This is particularly evident when observing the variations in effect sizes for the Olweus Bullying Prevention Programme (OBPP; e.g., Olweus, 1993a; Olweus, 1993b) and the KiVa anti-bullying programme. These programmes may be the most well-known anti-bullying programmes that are commercially available, and as such are the only examples in the present review of interventions evaluated in completely different locations.

The OBPP programme was originally designed and implemented in Norway, and it is therefore not surprising that the OBPP programme appears to be effective in reducing both school-bullying perpetration and victimisation when evaluated in Norway, compared to evaluations in the US. Whilst the programme was still significantly effective in the US, the percentage decrease in school-bullying perpetration was roughly 25% and in school-bullying victimisation was roughly 11%. These figures are less than the decreases in bullying behaviours seen in Norwegian evaluations (approximately 35% perpetration; 29% victimisation). These differences could be attributed to different evaluation methodologies (see Gaffney et al., 2019), but they most likely reflect cultural and societal differences between youth in Norway and youth in the US.

When the OBPP was evaluated in six Malaysian secondary schools, with a sample size of approximately 3,816 students, the programme was not significantly effective in reducing school-bullying victimisation (Yaakub et al., 2010; OR = 1.09, p = 0.28). This could be a result of the different manifestations of school-bullying victimisation in Eastern societies. As previously stated, researchers (e.g., Smith et al., 2016) have shown that bullying manifests differently in Eastern and Western cultures. This may explain why, in Malaysia, the OBPP was seemingly ineffective in reducing bullying victimisation. It may be that the programme itself was not tailored to the specific experiences and/or behaviours demonstrated by Malaysian students.

Interestingly, the opposite is observed with the KiVa programme. When KiVa was evaluated in Finnish samples, the programme was effective in reducing school-bullying perpetration by approximately 4-5% and school-bullying victimisation by approximately 6% (Kärnä et al., 2011; 2013a; 2013b). However, when evaluated in Italian primary and secondary schools, the effect sizes were much larger. Nocentini and Menesini (2016) found that KiVa was effective in reducing school-bullying perpetration by approximately 15 – 20% and school-bullying victimisation by approximately 25%.

In the case of KiVa, each of the evaluations used the same methodology (i.e., RCT), but varied greatly in the sample size which can greatly influence the summary effect size and the conclusions drawn from primary evaluation research (Sullivan & Feinn, 2012). Thus, further research is needed to explain why some interventions (e.g., OBPP or KiVa) appear to be more effective in some samples compared to others. The programmes are still effective, but the variations in effect size could be attributable to a number of different methodological and implementation factors that warrant further exploration.

11.2.1.2 Sample size. In primary evaluation research it is generally accepted that the number of participants included in the study can have a major impact on the significant of the result. For example, if a sample is too small, an effect size may not reach statistical significance yet in a sample that is too large, a very small effect size may be statistically significant. In a practical capacity too, sample size may impact the overall effectiveness estimate. It may be that an evaluation study faces more implementation problems if the sample size is too large. Meta-analysis is one way in which researchers can overcome these issues of primary research, but sample size may also have an impact on the outcomes of a meta-analysis (Cumming, 2014; Olejuik & Aljua, 2000).

As suggested in Chapter 2 (see section 2.6.3) and discussed further in Chapter 12 (see section 12.3) the variations in sample size of primary evaluations largely influences the outcome of a meta-analysis through the way in which weights are assigned. However, sample size was also included as a variable in subgroup analyses of the school-bullying meta-analysis. Meta-regression analysis using the total number of participants as a continuous variable found that the sample size of primary evaluations did not significantly predict effect sizes for either school-bullying perpetration or school-bullying victimisation outcomes. Moreover, when the mean sample size was represented as a categorical variable, the differences between subgroups of studies were marginal. Generally, for both perpetration and victimisation outcomes, studies that included an average of approximately 250 to 680 participants were associated with the largest mean effect sizes. The mean odds ratio for this group of studies equated to an approximate 15% reduction in bullying others and an approximate reduction of 18% in being bullied.

This finding has important implications for future research and evaluation studies of anti-bullying programmes. There are many difficulties involved in conducting primary research within education systems, let alone implementing evaluation research across different education systems (Phillips & Ertl, 2003). For example, often participant recruitment can be difficult. Schools around the world are often restricted in the time and resources that they can dedicate to evaluation research. Furthermore, participant attrition can have a detrimental impact on the outcomes of primary evaluations. Therefore, future evaluation studies of anti-bullying should consider the results of the present research. Whilst a study needs to have a certain minimum number of participants in order to have sufficient power, time and resources could be saved by recruiting participants in large numbers.

11.2.1.3 Age of participants. The age of participants and its relationship to effect size was a controversial finding of Farrington and Ttofi's (2009) earlier review. This previous meta-analysis found that programmes that were implemented and evaluated with participants aged 11 years old and over were significantly associated with greater effectiveness for school-bullying perpetration and victimisation outcomes. As previously mentioned in Chapter 1 (see section 1.5.1), prominent researchers in the field of bullying research criticized this simplistic dichotomy and suggested instead that age should be examined using within-programme effects.

Therefore, to address these issues, age was included in the present research as two different categorical variables and also one continuous variable. Firstly, when the mean age of included participants was included as a continuous variable for meta-regression analyses, no significant relationship was identified. This suggests that the mean age of participants did not influence the effectiveness of the intervention overall. However, using the mean age in this way may be too crude to fully understand the relationship between age and effect size, which may be non-linear.

Using categorical variables to represent age for subgroup analyses analogous to the ANOVA, the results of the present research were slightly more informative. Specifically, for school-bullying perpetration outcomes, evaluations that were conducted with participants aged 11 and older (i.e., similar to the approach used by Farrington & Ttofi, 2009) were significantly associated with greater effectiveness. Yet no such relationship was identified for school-bullying victimisation outcomes, and the results were also influenced by the computational model chosen to compute the weighted mean effect size.

Potentially the most interesting finding of the present analysis of the relationship between age and effect size was when age was represented as a categorical variable with additional levels (see section 6.3.1.3). Adopting this approach, the results were consistent with more recent meta-analyses. For example, Yeager and colleagues (2015) used a hierarchical meta-analysis of within-study effects by participant age and found that bullying was effectively prevented when a programme was implemented with participants aged 13 years old and younger[17]. When programmes were implemented with older participants this study found that the effectiveness estimate was reduced greatly (Yeager et al., 2015).

Under the MVA computational model, the results of the present meta-analysis are consistent with this earlier review. It was observed that the mean effectiveness of anti-bullying programmes decreased as the categorical age of participants also decreased. Moreover, these differences were statistically significant. Yet, future research is still needed to better understand the relationship between age and programme effectiveness. Bullying is said to be influenced greatly by child and adolescent developmental stages, and therefore a clear understanding of which programmes, and which specific intervention activities, are the most effective with particular age groups is fundamental. Zych and Farrington (*in press*)[18] have recently assessed bullying, and cyberbullying, from a developmental perspective and concluded that whilst prevalence rates of bullying change with age, additional large-scale evaluations of intervention programmes are needed. Fundamentally, to truly understand the relationship between effectiveness and age, evaluations of specific anti-bullying programmes need to include participants from all age groups, as the within-program effects are evidently more informative.

11.2.1.4 Publication type and year. With respect to the results on analysis of the relationship between publication type and effect size, this is covered in section 11.2.4.2, alongside a discussion of the impact of conflict of interest. Comparisons of effect size by publication year are discussed in section 11.2.

11.2.2 Design level

On the design level, potential moderating variables were concerned with features of the evaluation methodology. Specifically, the type of experimental methodology used by the

---

[17] This is a US study and the authors report that programmes were effective when implemented with participants in the 7th grade or below. Online sources indicate that students in the 7th grade in US schools are can be aged between 11 and 13 years old. https://www.k12academics.com/school-grades/7th-grade

[18] Shared via email communication with Dr. Izabela Zych, 9th February 2020.

primary evaluation (i.e., randomised controlled trials, quasi-experimental designs with pre and post measures, and age cohort designs) was investigated. Also, subgroup analyses were conducted to examine the possible relationship between the bullying measurement instrument used in the primary evaluation and the unit of allocation/randomisation.

11.2.2.1 Evaluation methodology. Under both the MVA and random effects models of meta-analysis, evaluations conducted using age cohort designs were found to be, collectively, the most effective, or at least associated with the largest mean effect sizes. This is consistent with Farrington and Ttofi's (2009) review. This methodological design was first introduced as an evaluation design for the Olweus Bullying Prevention Programme (Olweus, 1991). This approach has been criticized for the potential threats to internal validity, history and testing effects (Farrington & Ttofi, 2009, p. 15). However, it has been suggested that this design avoids the threats of aging and maturation effects, as individuals within the same school act as a control group for same-aged experimental participants (Olweus, 2005a). Yet, this design is vulnerable to cross-contamination between experimental and control participants which would impact the overall effectiveness.

Notably, intervention researchers have tested the Olweus Bullying Prevention Programme (OBPP) with other methodological designs (e.g., Bauer et al., 2007) which resulted in smaller effects. These could be explained by an interaction effect between the methodology used and the location of the evaluation, as age cohort designs were primarily used in Norwegian evaluations and experimental designs were primarily used elsewhere. Furthermore, as previously demonstrated, the location of the intervention may influence the overall effectiveness, so it is also possible that there is an interactive influence of methodology and location on outcomes. Future research is needed to understand this interaction in more detail.

Interestingly, the pattern of differences between RCTs and BA/EC designs was less clear. In relation to bullying victimisation outcomes, evaluations using BA/EC designs appear to yield more effective results than evaluations using RCT designs. However, for bullying perpetration outcomes, evaluations using RCT designs appear to yield more effective results than evaluations that utilised BA/EC designs. However, the nature of these analyses is correlational and the differences between effect sizes are marginal. Thus, no concrete conclusion can be drawn in relation to the association between randomised versus non-randomised quasi-experimental designs and effect size in the present context. Further research should aim to clarify this difference. as there can be implications for example, willingness of schools to participate in evaluation research.

11.2.2.2 Measurement instrument. Experts in the area of school-bullying research have pointed out how there still remain issues of comparability in the assessment of school-bullying perpetration and victimisation (Volk, Veenstra, & Espelage, 2017). For example, a recent systematic review shows clearly how the time frame measurement and other methodological issues (e.g., whether data are continuous or dichotomous) greatly impact the prevalence rates of bullying and cyberbullying (Vivolo-Kantor et al., 2019).

Studies included in the present meta-analysis used a wide variety of quantitative measures of school-bullying behaviours, including self-report measures (e.g., the Revised Olweus Bully/Victim Questionnaire - Olweus, 1986, 1996), or peer-report measures (e.g., the Participant Role Questionnaire – Salmivalli et al., 1996). One issue that arises is that the time frame within which participants are required to indicate the frequency of bullying can vary greatly. One scale may ask about bullying experiences within the last three months, whilst another may ask about ever having experienced, or participated in, school-bullying. Moreover, included studies utilised a mixture of continuous and dichotomous measures of school-bullying, and the cut-off points used to categorize someone as either a bully, victim, or not-involved also varied. Therefore, further research is needed to examine how such aspects of measurement instruments (e.g., timeframe or type of variable) could impact estimates of effectiveness using findings from recent research as a guideline (Vivolo-Kantor et al., 2019). It is essential to have appropriate measurements of bullying behaviours to be able to correctly estimate how effective programmes are in reducing said behaviours.

When conducting systematic searches for the school-bullying review, restrictions were not set on the measurement instruments used, other than including quantitative measures of school-bullying behaviours. However, types of reports for example, could influence the overall effectiveness effect size. As highlighted in Chapter 6 (see section 6.1.2.1) there was considerable consistency in measurement instruments used by primary evaluations to evaluate the effectiveness of an anti-bullying programme. Moreover, even though a large number of specific measurement tools were used in primary evaluations, it was often observed that measures were greatly influenced or based on existing measurement instruments, especially the Olweus Bully/Victim Questionnaire.

With regard to consistency in the use of self-report measures, this may possibly explain why the school-bullying meta-analysis found that programmes are more effective in reducing bullying perpetration outcomes in comparison to bullying victimisation outcomes. For example, if programmes are concerned with raising awareness about bullying and the associated negative impact on victims, participants who reported bullying perpetration before the intervention may be less likely to self-report bullying behaviours after completing the programme. As a result, the intervention may be perceived as being effective, but the change in reports of bullying may have been a result of increased social desirability responding (He et al., 2015; Rigby & Johnson, 2006). Conversely, raising awareness about the negative impact of school bullying may lead to increased reporting of victimisation due to sensitization effects (Stevens, de Bourdeaudhuji, van Oost, 2000). Sensitization effects due to raised awareness may affect not only self-report data but also peer nomination data and teacher reports (Smith, Ananiadou, Cowie, 2003) but arguably less so. The present meta-analysis identified only 13 evaluations that had utilised peer-report measures for example, and these measures were predominantly used alongside self-report measures. Given the possible cross-over of social desirability effects and the unequal numbers in established subgroups of primary evaluations, subgroup analysis was not appropriate. Further research is needed in this area, as using other-report types, such as peer-report measures of bullying perpetration, could give a more accurate indication of effectiveness.

11.2.2.3 Unit of allocation/randomisation. In theory, RCTs are the best method of evaluation of interventions because random allocation ensures that any observed differences between experimental and control groups occur as a result of chance variation, thus giving the best possible internal validity (Farrington, 1983; Farrington, 2003). However, the unit of random allocation can have an impact on internal validity. For example, we assume that individuals are randomly assigned to experimental and control conditions, so that RCT designs adequately account for the random variation that occurs in real-world research (Weisburd, 2003). However, in practice, evaluations of anti-bullying programmes may be more likely to assign groups of individuals, for example in terms of classrooms or schools, to experimental conditions rather than individual students.

This is true for both randomised (e.g., classrooms, Chaux et al., 2016; or schools, Espelage et al., 2015) and non-randomised (e.g., classrooms, Ortega-Ruiz et al., 2012; or schools, Rawana et al., 2011) methodologies. When this is the case, large numbers of units assigned to ensure adequate statistical conclusion validity and avoid issues of selection effects and differential attrition (Farrington & Ttofi, 2009; Ttofi & Farrington, 2011) are needed. There was a lot of variation in the unit of allocation in included primary studies, which may explain why the results did not find that one methodological design was more effective than another.

Moreover, the majority of included evaluations did not use the same unit for allocation and analysis, thus, posing a threat to the validity of the results. The results must therefore with be interpreted with caution, favouring more conservative estimates. Furthermore, the relationship between the unit of randomisation/allocation moderator variable and the effect sizes for bullying perpetration and victimisation outcomes was unclear. Whether or not the differences between subgroups of evaluations that assigned classes or schools to experimental conditions were statistically significant or not depended on the computational model used and the bullying outcome in question. For bullying perpetration, the differences between studies based on unit of allocation were not statistically significant for randomised and non-randomised studies. For bullying victimisation outcomes, studies where classes were the unit of allocation were associated with the largest effect sizes when all designs where included and for randomised evaluations, but not for non-randomised evaluations.

Risk of bias analysis also found that a large number of RCT studies were categorized as being high risk for allocation-related items on the EPOC tool. Therefore, the differences observed between primary evaluations in the current meta-analysis may be due to the observation that often the unit of allocation and the unit of analysis were not the same in primary studies. However, further analysis and investigation is needed to better understand these results.

11.2.3 Programme level: Intervention components

The results of the subgroup analyses suggest that many components of existing anti-bullying programmes are effective in reducing both school-bullying perpetration and victimisation. Under the MVA model of meta-analysis, the presence or absence, of numerous specific components was associated with larger summary effect sizes. Overall, the results presented in the current report provide good evidence for a socio-ecological system-based approach to anti-bullying programmes. It should also be noted that neither the presence nor the absence of any intervention component was significantly associated with undesirable intervention results, namely, an increase in bullying outcomes.

Earlier research highlighted how varying levels of implementation of each intervention component may explain the variability in intervention outcomes (Bloom, Hill, & Riccio, 2003). Interestingly, a narrative review by Smith and colleagues (2004) reported that, although 14 whole-school antibullying programmes obtained modest effects overall, those that monitored implementation obtained twice the mean effects on self-reported rates of bullying and victimisation than those that did not monitor implementation. Additionally, the current findings are largely consistent with earlier subgroup analyses (i.e., Farrington & Ttofi, 2009; Ttofi & Farrington, 2011). In the previous review, Farrington and Ttofi conducted detailed coding of interventions and evaluations and analysed how effect sizes varied according to components and features of primary studies. For example, parent training, playground supervision, and more intense and longer programmes were significantly correlated with larger reductions in bullying perpetration (Ttofi & Farrington, 2011). Moreover, several intervention components were associated with larger reductions in bullying victimisation (e.g., videos, disciplinary methods, co-operative group work and more intense and longer programmes).

More detail has been extracted in the present analysis and different computational models were used so exact comparisons are difficult, yet the overall findings are relatively consistent. Specifically, in Farrington and Ttofi’s intervention component analyses, the presence of classroom rules, the whole-school approach and use of co-operative group work (i.e., the inclusion of external professionals) were associated with greater reductions in bullying perpetration. These components were also significantly associated with greater reductions in bullying perpetration in the present, or ‘updated’, analyses. There were no components that were associated with increases in bullying perpetration in both the present and the previous analyses. However, the earlier analyses suggested that there was an undesirable relationship between specific intervention components and bullying victimisation. This is discussed in further detail later in this section.

The findings indicate that various components and anti-bullying activities can be implemented to reduce bullying in schools. Moreover, meta-regression analyses suggest that programme richness does not significantly predict more desirable outcomes. In other words, interventions that included many, or all, of the intervention components did not result in significantly greater effectiveness when compared to interventions that included fewer components. This finding will be useful to schools around the world that wish to implement measures to prevent or reduce bullying quickly and efficiently, but also, in the development of future intervention programmes. Many of the intervention components can be costly, both in monetary terms and in the time commitment for school staff. Thus, the current findings highlight multiple intervention components that can be implemented to have a desirable impact on bullying behaviours.

When interpreting these results, the reader should be aware that the analysis is correlational and also could be influenced by unequal numbers of studies in subgroups (e.g., curriculum materials: present in 66 evaluations and absent in 16 evaluations for school-bullying perpetration outcomes). Furthermore, the findings suggest that there are more components associated with effectiveness for reducing bullying perpetration in comparison to bullying victimisation. This may be influenced by social desirability, as previously mentioned.

In terms of the consistency between outcomes, there were some components that were significantly related to larger summary effect sizes for both perpetration and victimisation outcomes. For example, the presence of both informal peer involvement (e.g., class/group discussions or role-playing activities) and information for parents (e.g., letters/leaflets about bullying or intervention sent home to parents and guardians) were significantly associated with greater effectiveness in reducing both school-bullying victimisation and perpetration. Notably, the absence of socio-emotional skills training was statistically correlated with larger reductions in both school-bullying perpetration and victimisation. In other words, programmes that did not specify that the intervention programme incorporated elements relating to social-emotional skills (e.g., empathy, conflict resolution, or resilience), whether through specific intervention activities or dedicated intervention modules, were associated with greater effectiveness in our analyses.

Generally, the findings of the present analysis show that components of anti-bullying programmes that involve instituting and encouraging informal social control between all members of school communities are associated with greater effectiveness. This is in line with the vast evidence base on how collective efficacy and informal social control are key factors in reducing antisocial behaviours (e.g., Sampson, 1986; Silver & Miller, 2004; Williams & Guerra, 2011). For example, the presence of informal peer involvement was significantly associated with greater overall effectiveness. Studies that included informal peer involvement reduced bullying perpetration (by approximately 12.5%) and bullying victimisation (by approximately 9%) significantly more than studies that did not incorporate informal peer involvement (1% and 4.5% respectively).

This is contrary to previous findings by Farrington and Ttofi (2009), that work with peers was associated with increases in bullying victimisation. However, this conflicting result is most likely explained by the more detailed coding system applied in the present report.

Peer involvement was coded on three non-mutually exclusive levels in the present analyses, in comparison to one single component in the previous meta-analysis. Moreover, this result is consistent with the large body of bullying research which highlights the important yet complex role of peers in bullying amongst children and adolescents (e.g., Salmivalli, 1996; 2010).

The presence of informal peer involvement refers to studies in which intervention activities incorporated the peer group of bullies and victims through implicit means, such as whole-class or small group discussion. In this way, individual bullies or victims were not singled out, yet bullying experiences, attitudes, and behaviours were discussed within the peer group. Other informal peer involvement activities included: teaching students' assertiveness and encouraging them to intervene as bystanders when they witness bullying occurring (e.g., Menard & Grotpeter, 2014), or online forums monitored and peer-led by groups of trained students (e.g., Palladino et al., 2012; Menesini et al., 2012).

Interestingly, the *exclusion* of encouraging bystanders was significantly associated with larger effect sizes for victimisation outcomes. Moreover, the *inclusion* of formal peer involvement was very nearly significantly associated with a greater overall reduction in bullying perpetration, despite the differences in the numbers of studies that incorporated this component. Dissecting the involvement of peers in this way, our results are consistent with previous research that peer involvement may be beneficial in reducing bullying, but the key issue is the nature of peer involvement.

The results of the school-bullying meta-analysis can provide better understanding of the mechanisms of change involved in anti-bullying programmes. Many previous studies have emphasized the importance of understanding mechanisms of change in the development of problem behaviours in general (e.g., van Lier, Vuijk, & Crijnen, 2005). For example, it may be that by increasing awareness of bullying behaviours and involving all individuals in the classroom environment creates a social space that is less conducive to bullying. Additionally, having systems in place to hold bullies accountable for their behaviour, such as classroom rules, may lead to larger reductions in bullying perpetration. Giving teachers the skills to manage child behaviour in classrooms can also contribute to greater overall reductions in reports of bullying perpetration behaviours. This is consistent with previous studies that have emphasized the importance of utilising student- and classroom-level mechanisms of change to further the development of bullying prevention research (Saarento, Boulton, & Salmivalli, 2015).

Most anti-bullying programmes will incorporate peers and teachers in some form, especially if they are school-based and implemented during school hours. Therefore, we must strive to better understand how intervention components at all levels of a socio-ecological framework contribute to the overall effectiveness of an anti-bullying programme. The findings of the present research demonstrate some of the ways in which anti-bullying programmes can utilise a whole-school approach to prevent and reduce bullying. For example, the presence of classroom rules and the whole school approach were significantly associated with larger summary effect sizes, with studies that included these components reducing bullying perpetration by approximately 11% each. Comparatively, studies that did not include these intervention components reduced bullying perpetration by approximately 5 – 6% each.

In addition, information for parents was significantly associated with greater reductions in both school-bullying perpetration and victimisation. This may indicate that communicating information about bullying and the intervention with parents and guardians via letters/leaflets may be a more appropriate method in which future anti-bullying programmes can involve parents in comparison to the more costly method of providing workshops or information evenings. It is plausible to assume that, if anti-bullying information is provided to parents through letters or leaflets via their children, it may be less likely to be passed on to parents, as these letters or leaflets may well stay in a child's schoolbag. However, providing information in this way is more cost-effective and the result of this meta-analysis show that parent information in this format can help to reduce bullying perpetration and victimisation.

This further supports the proposal that the most effective components of anti-bullying programmes are those in which informal social control is established, particularly in relation to bullying perpetration. Furthermore, the establishment of accountability as a component of social control, whereby others are made aware of bullying behaviours (i.e., parents through information leaflets sent home and teachers and/or peers enforcing classroom rules against bullying) is an important aspect of bullying perpetration prevention.

However, the involvement of parents in anti-bullying programmes more officially, for example, by conducting information evenings for parents to attend, is not significantly associated with increases in the effectiveness of the intervention. It may be that, when parents are involved in anti-bullying programmes in this way, the 'right' parents do not engage. That is to say that possibly the parents of children involved in bullying do not voluntarily participate in the anti-bullying programme. However, possibly additional intervention components relating to the ways in which parents can be involved in anti-bullying programmes could be included in future research to better understand this relationship.

Components that targeted the actual students were also significantly associated with greater effectiveness in reducing bullying perpetration outcomes, such as working with victims and including cognitive-behavioural and mental health techniques in the intervention. Previous research has suggested that being bullied is independently related to child and adolescent mental health, and also that experiencing internalizing and externalizing problems can increase the risk of being bullied (Arsenault, Bowes, & Shakoor, 2010).

11.2.3.1 Programme specificity and richness. Another moderator that was coded was the specificity of the intervention programme. In other words, each intervention programme was evaluated on how specifically it related to bullying behaviours. Unsurprisingly, the findings suggest that programmes that were specifically dedicated to bullying prevention and intervention were associated with the largest overall effect sizes, although the significance of the differences between subgroups was not computed due to the large discrepancies between the numbers of evaluations included in each subgroup. Evaluation researchers have stressed the importance of program richness and it has been suggested that there is a dose-response relationship with bullying reductions even in the presence of heterogenous evaluations (e.g., Prochaska et al., 2007). However, programme richness (i.e., the number of individual intervention components included in the programme) did not significantly predict effectiveness in reducing either school-bullying perpetration or victimisation. This has important implications for future research, as often implementing very rich programmes with lots of different intervention activities is time-intensive and requires a lot of resources, as previously discussed.

However, the inclusion criteria for the current report were strictly concerned with school-bullying intervention programmes and behavioural outcomes of bullying. Therefore, effective programmes that only included non-behavioural outcomes of bullying (e.g., attitudes towards bullying, awareness of bullying) or other problem behaviours (e.g., peer aggression or victimisation, mental health issues, juvenile delinquency etc) that occur amongst young people in schools may have been overlooked. Changes in these behaviours may also impact bullying, either directly or indirectly, but more research is needed to understand this potential effect. Most obvious in the present report is how programmes that target specifically school-bullying may impact cyber-bullying, and vice versa, given the significant overlap in the prevalence of these behaviours (Baldry, Farrington, & Sorrentino, 2017).

11.2.4 Programme level: Other

Beyond the specific intervention components included in the programmes compared in the present analysis, a number of additional programme-related variables were coded and included in subgroup analyses. This section will discuss the results of subgroup analyses relating to the differences between specific packaged intervention programmes and the presence of possible conflict of interest.

11.2.4.1 Packaged intervention programmes. The relationship between effectiveness estimates and specific anti-bullying programmes was evaluated. The four most widely disseminated anti-bullying programmes included in the school-bullying review were the KiVA programme, NoTrap!, the Olweus Bullying Prevention Programme (OBPP), and ViSC. When comparing the effectiveness of these interventions, the OBPP was apparently the most effective in reducing school bullying perpetration. Across 11 evaluations, the OBPP reduced bullying perpetration by approximately 26%, which was larger than any other widely disseminated programme. In relation to school-bullying victimisation outcomes, the NoTrap! programme was the most effective, reducing victimisation by around 37%. NoTrap! also reduced bullying perpetration by a considerable amount, approximately 22%, but this effect was not statistically significant. The KiVa programme significantly reduced school bullying perpetration by approximately 9% and school bullying victimisation by approximately 11%. The ViSC programme was the only programme to increase bullying perpetration (by roughly 4%) and bullying victimisation (by roughly 4%), although these effects were not statistically significant.

11.2.4.2 Conflict of interest and publication type. Possibly the most conclusive results from the school-bullying subgroup analyses were observed in relation to conflict of interest and publication type. Firstly, across both computational models and outcomes, studies that were categorised as being high-risk for COI were associated with significantly larger reductions in bullying perpetration and victimisation. Secondly, under the MVA model of meta-analysis, non-peer-reviewed evaluations were associated with significantly larger reductions in both bullying perpetration and victimisation outcomes. However, the same results were not observed under the random effects model.

In the present research, conflict of interest was examined in terms of the involvement of the programme developer in the evaluation. The results from the school-bullying meta-analysis may indicate possible sources of bias. For example, it may be that when the individual, or team, that are credited with developing an anti-bullying programme are also involved in the evaluation of the said intervention, biases such as confirmation bias may impact the results. However, it may not be a perceptibly 'negative' source of bias. Perhaps, when the programme developer is involved in the implementation of the programme, the intervention is simply delivered better and more effectively. There are a number of other factors that could also be affected and could in turn impact the effect size, such as teacher and staff efficacy and motivation to participate in the programme. Therefore, the results may reflect differences in the quality of programme implementation rather than troublesome biases.

However, there are more sophisticated measures of COI (e.g., Eisner & Humphreys, 2012) that include elements such as whether or not the evaluator could potentially benefit financially from the intervention programme. Further indicators of conflict of interest are thus needed to better understand its impact on evaluation results. Additional research is needed.

11.3 Summary of findings: Cyberbullying

Overall, the cyberbullying meta-analysis found that school-based cyberbullying intervention programmes are effective in reducing both cyberbullying perpetration and victimisation. For cyberbullying perpetration, the weighted mean OR = 1.144 under the Multiplicative variance adjustment model of meta-analysis (MVA) and OR = 1.233 under a random-effects model (RE). Applying the transformation described in Chapter 2 (see section 2.9) these odds ratios correspond to approximately a 9% to 15% decrease in cyberbullying perpetration.

In comparison, the weighted mean ORs for cyberbullying victimisation outcomes were 1.231 and 1.227 under the MVA and the random effects models respectively. These mean effect sizes correspond to an approximate reduction in cyberbullying victimisation of 14 – 15%. These results suggest that the included interventions might have been slightly more effective at reducing cyberbullying victimisation than cyberbullying perpetration.

The results of this meta-analysis are consistent with findings from the only other meta-analytical review of cyberbullying intervention programmes. Specifically, Cleemput and colleagues (2014) found that the mean effect size for cyberbullying perpetration was positive and significant (OR = 1.275[19], *n* evaluations = 6). Moreover, this previous meta-analysis found that the mean effect for cyberbullying victimisation was also positive and significant (OR = 1.277, *n* evaluations = 6).

These results have important implications for both future research and public policy. Given the significant impact of cyberbullying experiences on adolescent health and mental wellbeing, anti-cyberbullying programmes should be considered for significant funding resources and national-level implementation. Additional analyses were undertaken to increase the understanding of what works in cyberbullying intervention and prevention programmes. However, in comparison to the school-bullying meta-analysis, the emphasis is on systematically reviewing the specific intervention activities that were included in

[19] The online calculator was used to transform Hedges' *g* mean effect sizes reported by Cleemput et al. (2014) to odds ratios. https://www.psychometrica.de/effect_size.html

cyberbullying intervention and prevention programmes. Some additional moderator variables were included, and the following sections of the present chapter will discuss the results of these analyses in more detail.

11.3.1 Sample size and age of participants

Given that fewer primary studies were included in the cyberbullying meta-analysis, in comparison to the school-bullying meta-analysis, fewer moderators were included in the post-hoc subgroup analyses. Sample size was included in the cyberbullying meta-analysis as a continuous variable only, and meta-regression analyses suggested that there was no statistically significant relationship between the total number of participants included and the effectiveness of cyberbullying intervention programmes. Additional primary evaluations are needed to be able to categorize and compare cyberbullying interventions on the basis of sample size.

The age of participants was also included in subgroup analysis for cyberbullying outcomes. Meta-regression analysis using the mean age of participants suggested that age did not significantly predict greater, or lesser, reductions in either cyberbullying perpetration or cyberbullying victimisation outcomes. Again, given that fewer evaluations were included for cyberbullying outcomes, age could only be categorized using a dichotomous variable, i.e., ‘younger’ versus ‘older’. Furthermore, the youngest participants were only aged 11 years old in included evaluations of cyberbullying intervention and prevention programmes. Subgroup analyses suggested that evaluations of programmes implemented with participants aged 14 years or older were associated with a significantly greater reduction in cyberbullying perpetration outcomes. No relationship was seen for cyberbullying victimisation outcomes. However, as is the case for most of the comparisons conducted in the cyberbullying meta-analysis, more primary studies are required to better understand the relationship between age and effectiveness.

11.3.2 Evaluation methodology

Mean effect sizes for the effectiveness of intervention programmes for both cyberbullying perpetration and victimisation outcomes were compared according to the methodological design of the evaluation. In the cyberbullying meta-analysis, included evaluations used randomised controlled trials (RCTs) and before-after/quasi-experimental-control designs (BA/EC). Under both MVA and RE models of meta-analysis, RCTs yielded larger effects in comparison to BA/EC designs in relation to cyberbullying perpetration and cyberbullying victimisation outcomes. This is contrary to previous meta-analyses in criminological research, as randomised controlled trials often produce smaller effect sizes (Chalamandris & Piette, 2015). Further research is needed to expand on this observation particularly given current debate about the 'gold standard' label previously assigned to randomised controlled trials (Farrington et al., 2020; Nagin & Sampson, 2019). Given the small number of studies included in the cyberbullying meta-analysis and the issue surrounding unit-of-randomisation and unit-of-analysis, it is hypothesized that the larger effect size for RCTs in the cyberbullying meta-analysis could be attributed to another moderator.

Randomised controlled trials are considered the best method of evaluating intervention effectiveness, because random assignment ensures that any observed differences between groups occur as a result of chance variations (Farrington, 1983) and can be argued to be because of experimental manipulation. Interventions that were evaluated using RCTs reduced, on average, cyberbullying perpetration by roughly 11-12% (MVA), or 19 - 20% (RE), and cyberbullying victimisation by roughly 16% (MVA), or 14 - 15% (RE). In comparison, interventions that were evaluated using a BA/EC reduced, on average, cyberbullying perpetration by roughly 4 - 5% (MVA), or 10 - 11% (RE), and cyberbullying victimisation by roughly 7 - 8% (MVA), or 8 - 9% (RE).

This analysis (published in Gaffney et al., 2019a) was the first published[20] meta-analytical review of cyberbullying interventions, there is limited previous research with which to compare the findings. For the purpose of comparison, the meta-analysis of school-bullying intervention programmes found less consistent results when the relationship between methodology and effect size was explored, as previously mentioned (see section 11.2.2.1). For school-bullying perpetration outcomes, evaluations conducted using quasi-experimental designs with before and after measures were less effective overall than evaluations conducted using RCT designs. The school-bullying meta-analysis found that the opposite was true for school-bullying victimisation outcomes, as previously outlined in the current chapter.

11.3.3 The offline-online overlap

Two important moderators that were included in the cyberbullying analysis related to the inclusion of online and offline content and outcomes, in primary evaluations of intervention programmes. Given the increasing attention that the overlap of offline and online bullying is receiving, it is imperative that we understand how best to combat these harmful behaviours. In primary research, children and adolescents are often categorised using the following labels: pure-bully, pure-victim, and bully-victim roles. Furthermore, young people involved in bullying can be classified as: (1) offline-only pure-bullies, (2) offline-only pure-victims; (3) offline-only bully-victims; (4) online-only pure-bullies; (5) online-only pure-victims; and (6) online-only bully-victims. Individuals can also be classified as not involved in bullying. However, for the purpose of intervention and prevention, what may be more important for the development of effective intervention programmes is the significant overlap commonly identified in the occurrence of offline and online bullying. In other words, there is a need for research on the following roles: (1) online *and* offline pure-bullies; (2) online *and*

[20] Cleemput et al. (2014) presented the findings from a systematic and meta-analytical review of cyberbullying intervention programmes at a conference, “Etmaal van de Communicatiewetenschap”, in Wageningen, Netherlands. The results have not been published in a peer-reviewed journal at the time of preparing this dissertation [checked in February 2020].

offline pure-victims; (3) online-victims *and* offline-bullies; and (4) online-bullies *and* offline-victims. To add further complexity, bully-victims could also be categorized in this manner. However, a full exploration of these typologies is beyond the scope of the present research.

For example, Baldry, Farrington, & Sorrentino (2017) found that, particularly for male participants, those involved as offline pure-bullies were also significantly more likely to be involved as online pure-bullies, and individuals involved as offline pure-victims were significantly more likely to be classified as online pure-victims. Whilst these specific role overlaps were not identified for female participants in this study, both male and female offline bully-victims were at a greater risk of also being online bully-victims. Future research is needed to investigate these typologies, as participants could also be offline pure-bullies but online pure-victims.

This discussion is beyond the scope of the present chapter, but this observation has important implications for intervention and prevention programmes. If an intervention programme targets offline bullying only, but the perpetrators or victims of offline bullying are also being targeted online, then this poses a difficult task for intervention programmes. Using the data from the cyberbullying meta-analysis two moderator variables were created to explore the effectiveness of existing interventions to reduce online and offline bullying concurrently.

Firstly, whether or not an intervention programme included *both* online and offline content within intervention activities and curriculum materials was examined for a potential impact on online bullying behaviours. Surprisingly, subgroups based on this variable were found to be associated with a collective negative impact on cyberbullying perpetration (i.e., an increase in cyberbullying others). The difference between studies that did include both online and offline content and studies that only included cyberbullying content was not statistically significant. This poses interesting implications for future research. These results must be interpreted with caution, as it is quite possible that interventions may have included content on both online and offline bullying but that this information was not published when reporting the results of the evaluation. Further discussion of the separation of online and offline bullying behaviours is presented in Chapter 12 (see section 12.2).

11.4 Limitations and future research

The following sections of the current chapter outline the limitations of the school-bullying meta-analysis and the cyberbullying meta-analysis. As such, avenue for future research are also examined.

A limitation shared by both the school-bullying and cyberbullying meta-analyses was the absence of important demographic variables, such as ethnicity, sexuality and gender identity, as possible mediators. The results therefore demonstrate the lack of knowledge on 'what works' for particularly vulnerable groups, such as LGBTQ+ youth, children and adolescents of minority ethnic groups, and those with physical or intellectual disabilities. There are few primary evaluations that investigate the effectiveness of specific anti-bullying programs with such groups, even though the impact of bullying on vulnerable individuals can have long-lasting undesirable effects.

Similarly, the risk of bias analyses in both systematic reviews may not adequately reflect the true relationship between bias and effect size. As a sum score was used to conduct risk of bias analyses, the results of the present dissertation did not examine the possible relationship between individual risk of bias items (e.g., allocation sequence, selective outcome reporting etc) and effect sizes for bullying and cyberbullying perpetration and victimisation. While the sum score suggested there was no significant relationship with effect size, using individual items may provide a better picture of the influence of specific aspects of bias and effect size. This is an important consideration for future research.

Finally, both meta-analyses relied primarily on data from self-report measures. This may influence the accuracy of results, as self-report measures are known for the prevalence of biases, such as social desirability (Rigby & Johnson, 2006). Many researchers in the field have called for better measures of bullying behaviours (Furlong et al., 2010; Volk et al., 2017). Most commonly multiple types of reports are used to evaluate anti-bullying programmes (e.g., self- and peer-reports: Kärnä et al., 2011a; 2011b; 2013; Elledge et al., 2010). Future meta-analyses could aim to compare effectiveness estimates across different types of measurement instruments. Yet caution is needed, as whilst research shows that the two generally correlate, albeit with small effects, the relationship between self-reported and peer-reported incidences of bullying is complex (see Volk et al., 2017 for a full review). Further research into the reliability and validity of using self-reported, peer-reported, or both, measurements of bullying in intervention research is needed.

11.4.1 School-bullying meta-analysis

Like most meta-analyses, the current research is largely limited by the lack of understanding as to what is the 'true effect'. When comparing mean effect sizes between moderators for example, it is difficult to determine the validity of the result. Throughout the discussion of results, it is highlighted that one subgroup of studies was associated with larger or smaller effect sizes than another, and the statistical significance of these differences. Thus, saying studies in subgroup A (e.g., evaluations conducted in Greece) are more effective than studies in subgroup B (e.g., evaluations conducted in Italy) is avoided. Due to the correlational nature of the moderator analyses causal inferences cannot be made.

This limitation is similar to that of most primary anti-bullying research; correlation is not causation. Conducting subgroup analysis using meta-analytical techniques is limited by the correlational nature of the comparison and by the nature of the comparison groups. In order to better understand any potential causal link, evaluators of anti-bullying programmes should vary the implementation of components systematically between experimental groups in future research. Some evaluations included in the present report did vary aspects of the implementation of different intervention components. For example, Trip et al. (2015) incorporated two experimental intervention groups, with one group receiving the REBE intervention activities first, followed by the ViSC anti-bullying activities. The order was then reversed for the second intervention group. Stallard et al. (2013) compared the effectiveness of a classroom-based CBT programme to two forms of control group; the first control group completed the schools' usual PSHE curriculum delivered by teachers, and the second control group also completed the PSHE curriculum, but lessons were delivered by a teacher who was assisted by two trained facilitators.

However, very few included studies compared experimental groups based on the implementation of specific intervention components. Polanin (2015) evaluated the impact of the Second Step programme, with additional cultural-awareness lessons, but did not compare the effectiveness of the Second Step programme with the effectiveness of the Second Step programme plus additional components. If future evaluations of anti-bullying interventions were to systematically vary the implementation of specific intervention components, it would become clearer as to what *actually* works in anti-bullying programmes and where there are differences in outcome according to the specific intervention component implemented. However, this adds yet another layer of complexity to evaluation methodology, and the implementation fidelity of components would need to be held constant across groups. Moreover, large studies are needed because, for example, if varying 4 intervention components, a minimum of 16 experimental conditions would be needed. This would be time consuming and costly, but an important factor for evaluators to consider in future research.

Additionally, the implementation fidelity and sustainability of intervention results need to be explored in greater detail. The present results are estimated using data before intervention and immediately post-intervention. Few studies include additional follow-up timepoints or quantitative measurement of implementation fidelity within the evaluation. The importance of including long-term follow-ups in experimental studies and implications for evaluation research is only recently beginning to be addressed (e.g., Farrington & MacKenzie, 2013).

Therefore, the long-term effectiveness of anti-bullying programmes is unclear. Another limitation is that the present research relied on information published about included intervention programmes, so that there may well be interventions that included a particular component but did not explicitly report this in reports of evaluation studies. Where possible, however, additional publications of included interventions were consulted. Thus, the present analyses may not adequately represent every component included in anti-bullying programmes and it is also not known how well and consistently the components coded were implemented. Future primary research on the effectiveness of anti-bullying interventions should aim to include and specify all relevant components of an intervention programme, although, this may be difficult because space is often limited in peer-reviewed publications.

This research would benefit too from deductive qualitative data that asks school staff and teachers to comment on the reality of implementing specific intervention components. It is equally, or more, important that these reductions in bullying are sustainable and maintained beyond the evaluation of the intervention programme. Therefore, a component may be statistically associated with greater reductions in bullying behaviours, but, if such a component is not feasible for schools to implement after the official evaluation has stopped, then this needs to be addressed.

Finally, any meta-analysis is impacted by the computational model chosen to assign weights to primary studies and limited by existing meta-analytical tests. The overall meta-analyses of school-based anti-bullying programmes present findings using two computational models of meta-analyses: the random effects model and the multiplicative variance adjustment (MVA) model. Whilst the random effects model is often suggested as the preferred model for meta-analyses in social sciences, for reasons already discussed (see Chapter 2), this approach is also limited. However, even though many meta-analyses in medical sciences (e.g., Ayieko et al., 2014; Dorjee et al., 2018; Woolf-King et al., 2013) have used the MVA model as an alternative method of accounting for between-study heterogeneity in weighted mean effect sizes, this model is yet to be widely accepted in behavioural sciences. A number of recent publications (e.g., Portnoy & Farrington, 2015; Zych, Viejo, Vila, & Farrington, 2019) have begun to use the MVA model. Additionally, further support for the MVA model is presented in Chapter 12 using data from the school-bullying meta-analysis.

It is evident in the current research that the results are influenced by the computational model used. The overall mean effect sizes for bullying perpetration and victimisation were not that different under both models but the results of subgroup analyses were greatly influenced by how the between-study heterogeneity was accounted for. Further research is needed in order to examine the reasons for this and also to evaluate how best to choose an appropriate computational model when conducting a meta-analysis.

Moreover, in the subgroup analyses both the MVA and the random effects model were deemed inadequate as previously discussed. The random effects model assigned too little weight to larger evaluations and the MVA model assigned too much weight to larger evaluations. Therefore, the decision to omit over-powered studies from the subgroup analysis means that results are presented under an appropriately weighted computational model (MVA) and better reflect the distribution of intervention components between multiple programmes. However, this analysis did not take account of all studies so future research should explore alternative approaches.

The analysis demonstrates that existing programmes are definitely effective anti-bullying initiatives, but packaged interventions are often quite expensive to purchase or require high levels of training and staff commitment (e.g., KiVa). Therefore, while packaged anti-bullying programmes are a viable and reliable option to reduce bullying, the present analysis provides interested stakeholders with a detailed breakdown of specific intervention activities that are shown to be associated with greater effectiveness. These results should be helpful in developing new programmes. Analyses also show that programmes that are more intensive and include a larger number of intervention components were not necessarily more effective. This suggests that there are options other than extensive multi-component and packaged interventions for schools that want to tackle bullying.

The current subgroup analysis is limited as intervention components were treated as being mutually exclusive. The strength of the socio-ecological theoretical approach is that it generally allows for exploration of the dynamic interactions between factors on all levels. However, interaction effects could not be explored in the present analyses. Recent meta-analyses have used advanced statistical tests (e.g., the 'three-level' meta-analysis by Yeager et al., 2015) to examine moderator effects. Future research should aim to utilise such advanced statistical tests to better our understanding of 'what works' in anti-bullying programmes, specifically in relation to the potential combinations of intervention components.

Two key limitations of the present review are the omission of cyberbullying behaviours and bully/victims. There are many complex participant roles in school-bullying, of which the bully/victim role is particularly complex (Salmivalli, 2010). Bully/victims include individuals who bully others but are also victims of bullying themselves. Moreover, prominent researchers in the area have argued that cyberbullying behaviours do not warrant a completely separate line of study, because of the significant overlap between offline and online bullying (Olweus & Limber, 2017). The second meta-analysis that comprises this dissertation concluded that cyberbullying intervention and prevention programmes can be effective, and Chapter 12 presents a detailed discussion of these findings. As illustrated in the cyberbullying review, there is a need for future research to assess the effectiveness of intervention programmes that target both online and offline bullying concurrently. As a result of the significant overlap of not only cyberbullying and school-bullying (e.g., Waasdorp & Bradshaw, 2015), but also bullying perpetration and victimisation offline (e.g., Baldry et al., 2017), it is important for policy makers, researchers, and programme developers to know what works in combating these forms of youth aggression.

11.4.2 Cyberbullying meta-analysis

The previous limitations discussed in relation to the school-bullying meta-analysis are similarly applicable to the cyberbullying meta-analysis. Briefly, the results are correlational and largely influenced by the quality and rigor of primary evaluation studies, but also by the content included in intervention programmes and reported by publications. Moreover, the results are impacted by the computational model chosen to estimate a mean effect size and assign weights to primary studies[21]. Nevertheless, there are a few limitations specific to the cyberbullying meta-analysis that should be highlighted.

The main limitation of the present systematic and meta-analytical review of cyberbullying intervention programmes is the number of primary studies included in the analysis. In comparison to previous attempts to synthesize the effectiveness of interventions (i.e., Mishna et al., 2011), significantly more evaluations of cyberbullying-specific programmes are included. However, in comparison to the school-bullying meta-analysis that included 100 evaluations of intervention programmes, there are relatively few evaluations (*n*

[21] The weighted mean effect sizes for cyberbullying outcomes were quite similar under both the MVA and the random-effects models of meta-analyses.

= 18 for cyberbullying perpetration and *n* = 19 for cyberbullying victimisation) included in the present report. Yet, this is not necessarily just a limitation of the present meta-analyses, but also of the cyberbullying literature. As has been previously discussed, research on cyberbullying is still relatively new, but it does seem to be growing exponentially in recent years. Moreover, there were a number of intervention protocols identified in the searches conducted for this dissertation and therefore, this meta-analysis should be updated in the near future. Tackling cyberbullying is an ongoing project, and although we are yet to understand the long-term effects, preventing this form of aggressive behaviour is highly important.

Another limitation of the present review is the exclusion of non-school-aged participants. Previous studies have found that cyberbullying is prevalent in samples other than children and adolescents. For example, cyberbullying perpetration and victimisation is prevalent amongst University student samples (Cowie et al., 2013) and there have been cyberbullying intervention programmes implemented and evaluated with University student samples (e.g., Doane et al., 2016; Leung et al., 2017; Savage et al., 2017). In addition, researchers in the field of cyberpsychology have suggested that the ‘digital age divide’, frequently discussed in relation to online activity (i.e., the idea that younger people are more active online than older individuals), is in fact narrowing (Attrill, 2015). Therefore, the number of individuals who are potentially at risk of exposure to online aggressive behaviour is no longer restricted to children and adolescents.

Nonetheless, there is a significant overlap between offline and online victimisation, specifically amongst adolescent populations (Olweus & Limber, 2017). In a US study of over 28,000 participants, with a mean age of 15.93 years old, 50.3% reported both online and offline bullying victimisation, in multiple forms, namely cyber, relational, physical, and verbal bullying (Waasdorp & Bradshaw, 2015). Thus, evaluating the effectiveness of anti-cyberbullying programmes with school-aged populations is an important avenue for research.

Given the number of school-bullying intervention programmes and the apparent comorbidity of offline and online bullying amongst adolescents, future research should aim to investigate whether these types of behaviour should be targeted simultaneously. In other words, should school-based programmes target cyberbullying and school-bullying concurrently or separately? Potential future studies could explore the differences in effectiveness of programmes that incorporate offline and online bullying, and offline-specific and online-specific intervention components. This analysis was attempted in the subgroup analyses of the cyberbullying meta-analysis but given the relatively small numbers of studies the results are not reliable. The results are indicative but should be treated with caution because of the small number of primary evaluations. Related analyses could be conducted using the data from the school-bullying meta-analysis and, given the larger number of primary evaluations, the findings of such research would be very important for future research, programme development, and educational policy.

Finally, in the school-bullying meta-analysis the intervention components analysis found that some specific intervention activities were more effective than others and were associated with a larger decrease in school-bullying perpetration and victimisation. In relation to cyberbullying aspects of the present dissertation, a systematic review of the contents of included programmes is provided. However, if similar moderator analysis could be conducted in future research, this would have important implications for the development of future programmes. If we were able to identify which components of intervention programmes are most effective in reducing cyberbullying perpetration and victimisation, then we would be better equipped to effectively prevent the negative outcomes associated with these behaviours. To achieve this aim more primary evaluations are needed.

11.5 Post-hoc considerations

The current dissertation has addressed the question of 'what works' in anti-bullying programmes to reduce both offline and online bullying but future research is needed. When, or if, the current research is updated, additional exploration of how effective programmes that target the overlap of bullying roles is warranted and perhaps using advanced analytical tests one could improve our understanding of the effectiveness of anti-bullying programmes even further.

Additionally, our understanding of how bullying behaviours are expressed and experienced is constantly evolving, and even though the current research examined online and offline behaviours separately, future research may, or should, not. Primarily, a common observation was that the separation of school- and cyber-bullying may not be appropriate anymore. Upon setting out to undertake this research, it was decided that two *separate* meta-analyses would be produced, one to evaluate the effectiveness on school-bullying outcomes, and another on cyberbullying outcomes. This decision was made for two reasons. Firstly, if cyberbullying behaviours were examined separately, then the school-bullying meta-analysis would be a more direct and comparable update of the earlier meta-analysis conducted by Farrington and Ttofi (2009). This updated review was intended to be undertaken as a Campbell Collaboration review and thus the protocol had already been submitted. As cyberbullying was not included in the first meta-analysis it was decided that online bullying would also not be included in the updated review. Secondly, the decision to keep these reviews separate was justified in that there had been no previous meta-analysis of cyberbullying intervention programmes and so the current research addresses a pressing gap in the literature. At the time of writing, only one previous study was available online (i.e., Cleemput et al., 2014) that conducted a meta-analysis of cyberbullying intervention programmes, but this was only a paper presented at a Dutch conference and not published in a high-impact peer-reviewed journal.

However, it became clear over the course of disseminating the findings of this research that increasingly children and adolescents do not distinguish between their offline and online worlds in the same way that academic researchers separate the two phenomena. Qualitative and student-led research is needed to outline exactly how the individuals impacted by bullying distinguish between the online and offline and qualitative evaluations are needed to examine whether a 'catch-all' type of programme can effectively impact school- and cyberbullying concurrently. Furthermore, given the significant overlap of these behaviours, future research is needed to investigate whether targeting offline bullying behaviours can also impact cyberbullying behaviours.

Yet the findings of this dissertation demonstrate that existing anti-bullying programmes are effective in reducing both online and offline forms of these pervasive aggressive behaviours. This is an important finding and a significant contribution to the literature which should greatly benefit future research and educational practice and policy.

11.6 Conclusions

In conclusion, the current research satisfactorily addresses the research questions proposed in Chapter 1; namely, 'what works' in existing school-based intervention and prevention programmes to reduce online and offline bullying amongst children and adolescents. Using methods of systematic review and meta-analysis, this dissertation has shown that current intervention and prevention programmes *are effective* and can work to improve the lives of young people by reducing rates of school- and cyber-bullying. Therefore, we have the tools to effectively reduce these forms of aggressive behaviours that are associated with such negative and harmful outcomes in the short- and long-term (Farrington et al., 2012). Moreover, the current research has thoroughly examined what works when, with whom, and in what context but the findings are limited. Several intervention activities were found to be associated with greater reductions in school-bullying and, this should have a great impact on the development of future programmes. This is a significant contribution to the literature, specifically in relation to the cyberbullying meta-analysis. Many facets of anti-bullying research will be impacted by the application of the current findings, in particular the development of future effective anti-bullying programmes and future evaluations of implemented interventions.

Of course, future research is still needed. The present research is limited in the ability to inform policy and future programmes to effectively reduce bullying with particularly vulnerable groups, such as LGBTQ+ communities and ethnic minorities. These groups may be more susceptible to bullying, either in school or online, and yet primary evaluations do not disaggregate their data for these specific groups. Similarly, not enough primary evaluations report gender-disaggregate data to be able to ascertain what works best for female and male students (Criado Perez, 2019). We know what works, now we need to know what works, for whom, when, and under what circumstances.

12. References

12.1 Studies included in school-bullying meta-analysis

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). Guilford.

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. https://doi.org/10.1080/01443410601159993

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. https://doi.org/10.1002/ab.20000

Battey, G. J. L. (2009). *Can bullies become buddies? Evaluation of and theoretical support for an experiential education bully prevention curriculum with seventh grade students.* (Unpublished doctoral dissertation). Oregon State University.

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. https://doi.org/10.1016/j.jadohealth.2006.10.005

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. https://doi.org/10.2307/4126451

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. https://doi.org/10.1177/082957350401900105

Berry, K., & Hunt, C. J. (2009). Evaluation of an intervention program for anxious adolescent boys who are bullied at school. *Journal of Adolescent Health, 45*, 376 – 382. https://doi.org/10.1016/j.jadohealth.2009.04.023

Bonell, C., Fletcher, A., Fitzgerald-Yau, N., Hale, D., Allen, E., Elbourne, D., Jones, R., Bond, L., Wiggins, M., Miners, A., Legood, R., Scott, S., Christie, D., & Viner, R. (2015). Initiating change locally in bullying and aggression through the school environment (INCLUSIVE): A pilot randomised controlled trial. *Health Technology Assessment, 19*(53). https://doi.org/10.3310/hta19530

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. https://doi.org/10.1177/0143034396174003

Brown, E. C., Low, S., Smith, B. H., & Haggerty, K. P. (2011). Outcomes from a school-randomized controlled trial of Steps to Respect: A bullying prevention program. *School Psychology Review, 40*(3), 423 – 443. https://doi.org/10.1037/e734362011-045

Busch, V., Leeuw, R., & Schrijvers, A. J. P. (2013). Results of a multibehavioral health-promoting school intervention in a Dutch secondary school. *Journal of Adolescent Health, 52*(4), 400 – 406. https://doi.org/10.1016/j/jadohealth.2012.07.008

Bull, H. D., Schultze, M., & Scheithauer, H. (2009). School-based prevention of bullying and relational aggression: The fairplayer.manual. *European Journal of Developmental Science, 3*(3), 312 – 317. https://doi.org/10.3233/dev-2009-3310

Chaux, E., Velásquez, A. M., Schultze-Krumbholz, A., & Scheithauer, H. (2016). Effects of the cyberbullying prevention program Media Heroes (*Medienhelden*) on traditional bullying. *Aggressive Behavior, 42*(2), 157 – 165. https://doi.org/10.1002/ab.21637

Cissner, A. B., & Ayoub, L. H. (2014). *Building healthy teen relationships: An evaluation of the Fourth R curriculum with middle school students in the Bronx.* Center for Court Innovation, U.S.A.

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

Connolly, J., Josephson, W., Schnoll, J., Simkins-Strong, E., Pepler, D., MacPhearson, A., Weiser, J., Moran, M., & Jiang, D. P. (2015). Evaluation of a youth-led program for preventing bullying, sexual harassment, and dating aggression in middle schools. *Journal of Early Adolescence, 35*(3), 403 – 434. https://doi.org/10.1177/0272431614535090

Cross, D., Monks, H., Hall, M., Shaw, T., Pintabona, Y., Erceg, E., Hamilton, G., Roberts, C., Waters, S., & Lester, L. (2011). Three-year results of the Friendly School whole-of-school intervention on children's bullying behaviour. *British Educational Research Journal, 37*(1), 105 – 129. https://doi.org/10.1080/01411920903420024

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 & Adolescent Psychology, 34*(1), 140 – 150. https://doi.org/10.1207/s15374424jccp3401_13

Domino, M. (2013). Measuring the impact of an alternative approach to school bullying. *Journal of School Health, 83*(6), 430 – 437. https://doi.org/10.1111/josh.12047

Elledge, L. C., Cavell, T. A., Ogle, N. T., & Newgent, R. A. (2010). School-based mentoring as selective prevention for bullied children: A preliminary test. *Journal of Primary Prevention, 31*, 171 – 187. https://doi.org/10.1007/s10935-010-0215-7

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. https://doi.org/10.1080/01443410701309258

Espelage, D. L., Low, S., Polanin, J. R., & Brown, E. C. (2015). Clinical trial of Second Step© middle-school program: Impact on aggression & victimization. *Journal of Applied Developmental Psychology, 37,* 52 – 63. https://doi.org/10.1016/j.appdev.2014.11.007

Evers, K. E., Poskiparta, 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. https://doi.org/10.1080/00131880701717271

Fekkes, M., Pijpers, F. I. M., Verloove-Vanhorick, P. (2006). Effects of antibullying school program on bullying and health complaints. *Archives of Pediatrics and Adolescent Medicine, 160*(6), 638 – 644. https://doi.org/10.1001/archpedi.160.6.638

Fekkes, M., van de Sande, M. C. E., Gravesteijn, J. C., Pannebakker, F. D., Buijs, G. J., Diekstra, R. F. W., & Kocken, P. L. (2016). Effects of the Dutch Skills for Life program on the health behaviour, bullying, and suicidal ideation of secondary school students, *Health Education, 116*(1), 2 – 15. https://doi.org/10.1108/HE-05-2014-0068

Finn, K. O'K. (2009). *An evaluation of the Olweus Bullying Prevention Program.* (Unpublished doctoral dissertation). Hofstra University.

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. *The Journal of Child Psychology and Psychiatry, 50*(5), 607 – 616. https://doi.org/10.1111/j.1469-7610.2008.02025.x

Fox, C., & Boulton, M. (2003). Evaluating the effectiveness of a social skills training (SST) program for victims of bullying. *Educational Research, 45,* 231 – 247. https://doi.org/10.1080/0013188032000137238

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. https://doi.org/10.1037/0012-1649.41.3.479

Garaigordobil, M., & Martínez-Valderrey, V. (2015). Effects of Cyberprogram 2.0 on "face-to-face" bullying, cyberbullying and empathy. *Psciothema, 27*(1), 45 – 51. https://doi.org/10.7334/psciotherma2014.78

Gini, G., Belli, B., & Casagrande, M. (2003). Le prepotenze a scuola: Una esperienza di ricerca-intervento antibullisimo [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. https://doi.org/10.1024/1421-0185.65.2.125

Herrick, C. (2012). *An investigation into the effectiveness of an anti-bullying campaign*. (Unpublished doctoral dissertation). University of Nottingham.

Holen, S., Waaktaar, T., Lervåg, A., & Ystgaard, M. (2013). Implementing a universal stress management program for young school children: Are there classroom climate or academic effects? *Scandinavian Journal of Educational Research, 57*(4), 420 – 444. https://doi.org/10.1080/00313831.2012.656320

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. https://doi.org/10.1111/j.1475-3588.2006.00417.x

Jenson, J. M., Brisson, D., Bender, K. A., & Williford, A. P. (2013). Effects of the Youth Matters prevention program on patterns of bullying and victimization in elementary and middle school. *Social Work Research*, *37*(4), 361 – 372. https://doi.org/10.1093/swr/svt030

Joronen, K., Konu, A., Rankin, H. S., & Astedt-Kurki, P. (2011). An evaluation of a drama program to enhance social relationships and anti-bullying at elementary school: A controlled study. *Health Promotion International, 27*(1), 5 – 14. https://doi.org/10.1093/heapro/dar012

Ju, Y., Shuqiong, W., & Wenxin, Z. (2009). Intervention research on school bullying in primary schools. *Frontiers of Education in China, 4,* 111 – 122. https://doi.org/10.1007/s11516-009-0007-0

Kaljee, L., Zhang, L., Langhaug, L., Munjile, K., Tembo, S., Menon, A., Stanton, B., Li, X., & Malungo, J. (2017). A randomized-controlled trial for the teachers' diploma programme on psychosocial care, support and protection in Zambian government primary schools. *Psychology, Health, and Medicine, 22*(4), 381 – 392. https://doi.org/10.1080/13548503.2016.1153682

Kärnä, A., Voeten, M., Little, T. D., Alanen, E., Poskiparta, E., & Salmivalli, C. (2013). Effectiveness of the KiVa Antibullying Program: Grades 1 - 3 and 7 - 9. *Journal of Educational Psychology, 105*(2), 535 - 551. https://doi.org/10.1037/a0030417

Kärnä, A., Voeten, M., Little, T. D., Poskiparta, E., Alanen, E., & Salmivalli, C. (2011a). Going to scale: A nonrandomized nationwide trial of the KiVa antibullying Program for Grades 1 - 9. *Journal of Consulting and Clinical Psychology, 79*(6), 796 - 805. https://doi.org/10.1037/a0025740

Kärnä, A., Voeten, M., Little, T. D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011b). A large-scale evaluation of the KiVa antibullying program: Grades 4 - 6. *Child Development, 82*(1), 311 - 330. https://doi.org/10.1111/j.1467-8624.2010.01.557.x

Kimber, B., Sandell, R., & Bremberg, S. (2008). Social and emotional training in Swedish classrooms for the promotion of mental health: Results from an effectiveness study in Sweden. *Health Promotion International, 23*(2), 134 - 143. https://doi.org/10.1093/heapro/dam046

Knowler, C., & Frederickson, N. (2013). Effects of an emotional literacy intervention for students identified with bullying behaviour. *Educational Psychology, 33*(7), 862 – 883. https://doi.org/10.1080/01443410.2013.785052

Krueger, L. M. (2010). *The implementation of an anti-bullying program to reduce bullying behaviours on elementary school buses.* (Unpublisehd doctoral dissertation) D'Youville College.

Li, K., Washburn, I., DuBois, D. L., Vuchinich, S., Ji, P., Brechling, V., Day, J., Beets, M. W., Acock, A.C., Berbaum, M., Snyder, F., & Flay, B. R. (2011). Effects of the *Positive Action* programme on problem behaviours in elementary school students: A matched-pair randomised control trial in Chicago. *Psychology and Health, 26*(2), 187 - 204. https://doi.org/10.1080/08870446.2011.531574

Limber, S. P., Olweus, D., Wang, W., Masiello, M., & Breivik, K. (2018). Evaluation of the Olweus Bullying Prevention Program: A large scale study of U.S. students in grades 3 – 11, *Journal of School Psychology, 69,* 56 – 72. https://doi.org/10.1016/j.jsp.2018.04.004

Losey, R. A. (2009). *An evaluation of the Olweus Bullying Prevention Program's effectiveness in a high school setting*. (Unpublished doctoral dissertation). University of Cincinnati.

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.

McLaughlin, L. P. (2009) *The effect of cognitive behavioral therapy and cognitive behavioral therapy plus media on the reduction of bullying and victimization and the increase of empathy and bystander response in a bully prevention program for urban sixth-grade students.* (Unpublished doctoral dissertation). University of Toledo.

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. K. (2014). Evaluation of Bully-Proofing Your School as an elementary school antibullying intervention. *Journal of School Violence, 13*(2), 188 - 209. https://doi.org/10.1080/15388220.2013.840641

Menesini, E., Codescasa, 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. https://doi.org/10.1002/ab.80012

Menesini, E., Nocentini, A., & Palladino, B. E. (2012). Empowering students against bullying and cyberbullying: Evaluation of an Italian peer-led model. *International Journal of Conflict and Violence, 6*(2), 314 - 320. https://doi.org/10.1002/ab.80012

Meyer, N., & Lesch, E. (2000). An analysis of the limitations of a behavioural programme for bullying boys from a sub-economic environment. *Southern African Journal of Child* *and Adolescent Mental Health, 12*(1), 59 – 69. https://doi.org/10.1080/16826108.2000.9632368

Nocentini, A., & Menesini, E. (2016). KiVa antibullying program in Italy: Evidence of effectiveness in a randomized control trial. *Prevention Science, 17*(8), 1012 – 1023. https://doi.org/10.1007/s11121-016-0690-z

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). Sage.

Olweus, D. (1993). Bully/victim problems among school children: Long-term consequences and an effective intervention program. In S. Hodgins (Ed.), *Mental disorder and crime,* (pp. 317 – 349). 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. https://doi.org/10.1111/j.1469-7610.1994.tb01229.x

Olweus, D. (1994b). Bullying at school: Basic facts and an effective intervention programme. *Promotion and Education, 1,* 27 – 31. https://doi.org/10.1177/102538239400100414

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). Plenum.

Olweus, D. (1995). Peer abuse or bullying at school: Basic facts and a school-based intervention programme. *Prospects, 25*(1), 133 – 139. https://doi.org/10.1007/bf02334290

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). 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. https://doi.org/10.1111/j.1749-6632.1996.tb32527.x

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. https://doi.org/10.1080/03033910.1997.10558138

Olweus, D. (1997a). Bully/victim problems in school: Knowledge base and an effective intervention project. *Irish Journal of Psychology, 18,* 170 – 190. https://doi.org/10.1007/bf02195509

Olweus, D. (1997b). Bully/victim problems in schools: Facts and intervention. *European Journal of Psychology of Education, 12,* 495 – 510. https://doi.org/10.1007/bf03172807

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). 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 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. (2005a). A useful evaluation design, and the effects of the Olweus bullying prevention program. *Psychology, Crime and Law, 11,* 389 – 402. https://doi.org/10.1080/10683160500255471

O'Moore, A. M., & Minton, 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 University Press.

Ortega-Ruiz, R., Del Rey, R., & Casas, J. A. (2012). Knowing, building and living together on Internet and social networks: The ConRed cyberbullying prevention program. *International Journal of Conflict and Violence, 6*(2), 303 – 313. https://doi.org/10.4119/UNIBI/ijcv.250

Ostrov, J. M., Godleski, S. A., Kamper-DeMarco, K. E., Blakely-McClure, S. J., & Celenza, L. (2015). Replication and extension of the Early Childhood Friendship Project: Effects on Physical and Relational Bullying. *School Psychology Review, 44*(4), 445 – 463. https://doi.org/10.17105/spr-15.0048.1

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.

Palladino, B. E., Nocentini, A., & Menesini, E. (2012). Online and offline peer led models against bullying and cyberbullying. *Psicothema, 24*(4), 634 - 639. Retrieved from: https://www.redalyc.org/pdf/727/72723959020.pdf

Palladino, B. E., Nocentini, A., & Menesini, E. (2016). Evidence-based intervention against bullying and cyberbullying: Evaluation of the NoTrap! program in two independent trials. *Aggressive Behavior, 42*(2), 194 – 206. https://doi.org/10.1002/ab.21636

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 University Press.

Polanin, M. K. (2015). *Effects of cultural awareness training in conjunction with an established bullying prevention program.* (Unpublished doctoral dissertation). Loyola University Chicago.

Pryce, S., & Frederickson, N. (2013). Bullying behaviour, intentions and classroom ecology. *Learning Environment Research, 16,* 183 – 199. https://doi.org/10.1007/s10984-013-9137-7

Rahey, L., & Craig, W. M. (2002). Evaluation of an ecological program to reduce bullying in schools. *Canadian Journal of Counselling, 36,* 281 – 295. Retrieved from https://journalhosting.ucalgary.ca/index.php/rcc/article/view/58699

Rawana, J. S., Norwood, S. J., & Whitley, J. (2011). A mixed-method evaluation of a strength-based bullying prevention program. *Canadian Journal of School Psychology, 26*(4), 283 – 300. https://doi.org/10.1177/0829573511423741

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. https://doi.org/10.1111/j.1749-6632.1996.tb32554.x

Roland, E., Bru, E., Midthassel, U. V., & Vaaland, G.S. (2010). The Zero programme against bullying: Effects of the programme in the context of the Norwegian manifesto against bullying. *Social Psychology of Education, 13,* 41 – 55. https://doi.org/10.1007/s11218-009-9096-0

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 University Press.

Salmivalli, C., Kaukiainen, A., & Voeten, M. (2005). Anti-bullying intervention: Implementation and outcome. *British Journal of Educational Psychology, 75,* 465 – 487. https://doi.org/10.1348/000709905x26011

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 University Press.

Sapouna, M., Wolke, D., Vannini, N., Watson, S., Woods, S., Schneider, W., Enz, S., Hall, L., Paiva, A., André, E., Dautenhahn, K., & Aylett, R. (2010). Virtual learning intervention to reduce bullying victimization in primary school: A controlled trial. *Journal of Child Psychology and Psychiatry, 51*(1), 104 - 112. https://doi.org/10.1111/j.1469-7610.2009.02137.x

Silva, J.L. da, Oliveira, W.A. de, Braga, I.F., Farias, M.S., Silva Lizzi, E.A. da, Carvalho Gonçalves, M. F., Pereira, B.O., & Iossi Silva, M.A. (2016). The effects of a skill-based intervention for victims of bullying in Brazil. *International Journal of Environmental Research and Public Health, 13,* 1042. https://doi.org/10.3390/ijerph13111042

Solomontos-Kountouri, O., Gradinger, P., Yanagida, T., & Strohmeier, D. (2016). The implementation and evaluation of the ViSC program in Cyprus: Challenges of cross-national dissemination and evaluation results. *European Journal of Developmental Psychology, 13*(6), 737 – 755. https://doi.org/10.1080/17405629.2015.1136618

Spröber, 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: Evalation 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.

Stallard, P., Phillips, R., Montgomery, A. A., Spears, M., Anderson, R., Taylor, J., Araya, R., Lewis, G., Ukoumunne, O. C., Millings, A., Georgiou, L., Cook, E., & Sayal, K. (2013). A cluster randomised controlled trial to determine the clinical effectiveness and cost-effectiveness of classroom-based cognitive–behavioural therapy (CBT) in reducing symptoms of depression in high-risk adolescents. *Health Technology Assessment, 17*(47), https://doi.org/10.3310/hta17470

Strohmeier, D., Hoffmann, C., Schiller, E., Stefanek, E., & Spiel, C. (2012). ViSC Social Competence Program. *New Directions for Youth Development, 133,* 71 – 84. https://doi.org/10.1002/yd.20008

Sutherland, A. E. (2010). *The roles of school climate and peers in bullying.* (Unpublished doctoral dissertation). Queen's University, Canada.

Toner, B. K. (2010). *The implementation of the bully prevention program: Bully Proofing Your School and its effect on bullying and school climate on sixth grade suburban Students.* (Unpublished doctoral dissertation). Widener University.

Topper, L. R. (2011). *Bullying victimisation and alcohol-misuse in adolescence: Investigating the functional relationship and new prevention strategies.* (Unpublished doctoral dissertation) King's College London.

Trip, S., Bora, C., Sipos-Gug, S., Tocai, I., Gradinger, P., Yanagida, T., & Strohmeier, D. (2015). Bullying prevention in schools by targeting cognitions, emotions, and behavior: Evaluating the effectiveness of the REBE-ViSC program. *Journal of Counselling Psychology, 62*(4), 732 – 740. https://doi.org/10.1037/cou0000084

Tsiantis, A. C. J., Beratis, I. N., Syngelaki, E. M., Stefanakou, A., Asimopolous. C., Sideridis, G. D., & Tsiantis, J. (2013). The effects of a clinical prevention program on bullying, victimization, and attitudes toward school of elementary school students. *Behavioral Disorders, 38*(4), 243 – 257. https://10.1177/019874291303800406

Waasdorp, T. E., Bradshaw, C. P., & Leaf, P. J. (2012). The impact of schoolwide positive behavioural interventions and supports on bullying and peer rejection. *Archives of Pediatric and Adolescent Medicine, 166*(2), 149 – 156. https://doi.org/10.1001/archpediatrics.2011.755

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). Routledge.

Williams, J., Miller, S., Cutbush, S., Gibbs, D., Clinton-Sherrod, M., & Jones, S. (2015). A latent transition model of the effects of a teen dating violence prevention initiative. *Journal of Adolescent Health, 56,* S27 – S32. https://doi.org/10.1013.j,adohealth.2014.08.019

Wölfer, R., & Scheithauer, H. (2014). Social influence and bullying behavior: Intervention-based network dynamics of the fairplayer.manual bullying prevention program. *Aggressive Behavior, 40,* 309 – 319. https://doi.org/10.1002/ab.21524

Wong, D. S. W., Cheng, C. H. K., Ngan, R. M. H., & Ma, S. K. (2011). Program effectiveness of a restorative whole-school approach for tackling school bullying in Hong Kong. *International Journal of Offender Therapy and Comparative Criminology, 55*(6), 846 – 862. https://doi.org/10.1177/0306624X10374638

Yaakub, N. F., Haron, F., & Leong, G. C. (2010). Examining the efficacy of the Olweus prevention programme in reducing bullying: The Malaysian experience. *Procedia – Social and Behavioral Sciences, 5,* 595 – 598. https://doi.org/10.1016/j.sbspro.2010.07.148

[22]Yanagida, T., Strohmeier, D., & Spiel, C. (2019). Dynamic change of aggressive behavior and victimization among adolescents: Effectiveness of the ViSC program. *Journal of Clinical Child & Adolescent Psychology, 48,* S90 – S104. https://doi.org/10.1080.15374416.2016.1233498

[22] Published online in 2016.

12.2 Studies excluded from the school-bullying meta-analysis[23]

** = excluded due to missing information or overlapping samples*

*Ahtola, A., Haataja, A., Kärnä, A., Poskiparta, E., & Salmivalli, C. (2013). Implementation of anti-bullying lessons in primary classrooms: How important is head teacher support? *Educational Research, 55*(4), 376 – 392. https://doi.org/10.1080/00131881.2013.844941

*Ahtola, A., Haataja, A., Kärnä, A., Poskiparta, E., & Salmivalli, C. (2012). For children only? Effects of the KiVa antibullying program on teachers. *Teaching and Teacher Education, 28,* 851 – 859. https://doi.org/10.1016/j.tate.2012.03.006

Al-Samarrai, L. (2011). *Evaluation of Mythodrama intervention among middle school students.* (Unpublished doctoral dissertation). Institute for Clinical Social Work, Chicago.

Allen, K. P. (2010). A bullying intervention system in high school: A two-year school-wide follow-up. *Studies in Educational Evaluation, 36,* 83 – 92. https://doi.org/10.1016/j.stueduc.2011.01.002

*Alsaker, F. D. (2004). Bernese program against victimization to 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 University Press

Amundsen, E. J., & Ravndal, E. (2010). Does successful school-based prevention of bullying influence substance use among 13- to 16-year-olds? *Drugs: Education, Prevention and Policy, 17*(1), 42 – 54. https://doi.org/10.3109/09687630802095591

[23] References for cyberbullying intervention programs are not included here.

Avşar, F., & Alkaya, S. A. (2017). The effectiveness of assertiveness training for school-aged children on bullying and assertiveness level. *Journal of Pediatric Nursing, 36,* 186 – 190. https://doi.org/10.1016/j.pedn.2017.06.020

Azad, N. E., & Amiri, S. (2012). Effectiveness of Olweus Bullying Prevention Program on Iranian boys, *Iranian Journal of Psychiatry and Clinical Psychology, 18*(3), 175 – 183.

Beckman, L., & Svensson, M. (2015). The cost-effectiveness of the Olweus Bullying Prevention Program: Results from a modelling study. *Journal of Adolescence, 45,* 127 – 137. https://doi.org/10.1016/j.adolescence.2015.07.020

Beets, M. W., Flay, B. R., Vuchinich, S., Snyder, F. J., Acock, A., Burns, K., Washburn, I. J., & Durlak, J. (2009). Use of a social and character development program to prevent substance use, violent behaviors, and sexual activity among elementary-school students in Hawaii. *American Journal of Public Health, 99*(8), 1438 – 1445. https://doi.org/10.2105/AJPH.2008.142919

Beightol, J., Jevertson, J., Carter, S., Gray, S., & Gass, M. (2012). Adventure education and resilience enhancement. *Journal of Experiential Education, 35*(2), 307 – 325. https://doi.org/10.5193.JEE35.2.307

Beightol, J., Jevertson, J., Gray, S., Carter, S., & Gass, M. (2009). The effect of an experiential adventure-based "Anti-Bullying Initiative" on levels of resilience: A mixed methods study. *Journal of Experiential Education, 31*(3), 420 – 424. https://doi.org/10.1177/105382590803100312

Boulton, M. J. (2014). Teachers' self-efficacy, perceived effectiveness beliefs, and reported use of cognitive-behavioral approaches to bullying among pupils: Effects of in-service training with the I DECIDE program. *Behavior Therapy, 45,* 328 – 343. https://doi.org/10.1016/j.beth.2013.12.004

Bowes, D., Marquis, M., Young, W., Holowaty, P., & Isaac, W. (2009). Process evaluation of a school-based intervention to increase physical activity and reduce bullying. *Health Promotion Practice, 10*(3), 394 – 401. https://doi.org/10.1177/1524839907307886

Brenick, A., Shattuck, J., Donlan, A., Duh, S., & Zurbriggen, E. I. (2014). Empowering children with safety-skills: An evaluation of the Kidpower Everyday Safety-Skills program. *Children and Youth Services Review, 44,* 152 – 162. https://doi.org/10.1016/j.childyouth.2014.06.007

Bundy, A. C., Naughton, G., Tranter, P., Wyver, S., Baur, L., Schiller, W., Bauman, A., Engelen, L., Ragen, J., Luckett, T., Niehaues, A., Stewart, G., Jessup, G., & Brentnall, J. (2011). The Sydney playground project: popping the bubblewrap – unleashing the power of play: A cluster randomized controlled trial of a primary school playground-based intervention aiming to increase children's physical activity and social skills. *BMC Public Health, 11*(1), 680. https://doi.org/10.1186/1471-2458-11-680

Burkhart, K. M. (2012). *Parental factors contributing to bullying and the effects of the ACT Parents Raising Safe Kids program on the reduction of bullying.* (Unpublished doctoral dissertation). University of Toledo.

Burkhart, K. M., Knox, M., & Brockmeyer, J. (2013). Pilot evaluation of the ACT Raising Safe Kids program on children's bullying behavior. *Journal of Child and Family Studies, 22,* 942 – 951. https://doi.org/10.1007/s10826-012-9656-3

Cecil, H., & Molnar-Main, S. (2015). Olweus Bullying Prevention Program: Components implemented by elementary classroom and specialist teachers. *Journal of School Violence, 14*(4), 335 – 362. https://doi.org/10.1080/15388220.2014.912956

Černi Obrdalj, E., Zadro, K., Batić-Mujanović, O., & Zalihić, A. (2013). Satisfaction with the program of school bullying prevention and mental health promotion – Cross sectional study among primary school pupils in Mostar. *Acta Medica Academica, 43*(1), 35 – 39. https://doi.org/10.5644/ama2006-124.98

Chu, B. C., Hoffman, L., Johns, A., Reyes-Portillo, J., & Hansford, A. (2014). Transdiagnositc behavior therapy for bullying-related anxiety and depression: Initial development and pilot study. *Cognitive and Behavioral Practice,* https://doi.org/10.1016/j.cbpra.2014.06.007

Cobb, A. B. (2009). *Assessing the effectiveness of disciplinary alternative education programs for secondary students in the Dallas independent school district.* (Unpublished doctoral dissertation). University of Texas at Dallas.

Cooke, M. B., Ford, J., Levine, J., Bourke, C., Newell, L., & Lapidus, G. (2007). The effects of city-wide implementation of "Second Step" on elementary school students' prosocial and aggressive behaviors. *Journal of Primary Prevention, 28*(2), 93 – 115. https://doi.org/10.1007/s10935-007-0080-1

Cornell, D., Sheras, P., Gregory, A., & Fan, X. (2009). A retrospective study of school safety conditions in high schools using the Virginia Threat Assessment Guidelines versus alternative approaches. *School Psychology Quarterly, 24*(2), 119 – 129. https://doi.org/10.1037/a0016182

*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 University Press.

Cross, D., Waters, S., Pearce, N., Shaw, T., Hall, M., Erceg, E., Burns, S., Roberts, C., & Hamilton, G. (2012). The Friendly Schools Friendly Families programme: Three-year bullying behaviour outcomes in primary school children. *International Journal of Education Research, 53,* 394 – 406. https://doi.org/10.1016/j.ijer.2012.05.004

Daugherty, C. S. (2011). *Principal and teacher perceptions of the effectiveness of the Olweus Bullying Prevention Program.* (Unpublished doctoral dissertation). Lincoln Memorial University.

Davis, J. (2011). *Evaluating a social skills training protocol in a private setting: Outcomes and issues.* (Unpublished doctoral dissertation). University of Nevada.

DeNike, M. (2014). *No Bully Solution Team: Outcomes Evaluation 2012-14.* https://www.nobully.org/no-bully-system/research

*DeRosier, M. E. (2004). Building relationships and combating bullying: Effectiveness of a school-based social skills group intervention, *Journal of Clinical Child & Adolescent Psychology, 33*(1), 196 – 201. https://doi.org/10.1207/S15374424JCCP3301_18

*Domino, M. (2011). *The impact of Take the LEAD on school bullying among middle school youth.* (Unpublished doctoral dissertation). Walden University.

Drury, B. (2014). *An evaluation of the antibullying program at a public suburban high school.* (Unpublished doctoral dissertation). Walden University.

Earhart, J. A. Jr. (2011). *Promoting positive peer relationships among youths: A study examining the effects of a class-wide bullying prevention program.* (Unpublished doctoral dissertation). University of California, Santa Barbara.

Emfield, J. S. (2015). *Self-Defense as a bullying prevention strategy: An evaluation of the Gracie Bullyproof program.* (Unpublished doctoral dissertation). Northwest Nazarene University.

*Espelage, D. L., Low, S., Polanin, J. R., & Brown, E. C. (2013). The impact of a middle school program to reduce aggression, victimization, and sexual violence. *Journal of Adolescent Health, 53*(2), 180 – 186. https://doi.org/10.106/j.jadohealth.2013.02.021

Espelage, D. L., Rose, C. H., & Polanin, J. R. (2015). Social-emotional learning program to reduce bullying, fighting, and victimization among middle school students with disabilities. *Remedial and Special Education,* 1 – 13. https://doi.org/10.1177/0741932514564564

Farmer, T. W., Hall, C. M., Petrin, R., Hamm, J. V., & Dadisman, K. (2010). Evaluating the impact of a multicomponent intervention model on teachers' awareness of social networks at the beginning of middle school in rural communities. *School Psychology Quarterly, 25*(2), 94 – 106. https://doi.org/10.1037/a0020147

Farrell, A. D., Mehari, K., Mays, S., Sullivan, T. N., & Le, A. T. (2015). Participants' perceptions of a violence prevention curriculum for middle school students: Was it relevant and useful? *Journal of Primary Prevention, 36*(4), 227 – 246. https://doi.org/10.1007/s10935-015-0391-6

Fletcher, A., Fitzgerald-Yau, N., Wiggins, M., Viner, R. M., & Bonell, C. (2015). Involving young people in changing their school environment to make it safer. Findings from a process evaluation in English secondary schools. *Health Education, 115*(3-4), 322 – 338. https://doi.org/10.1108/HE-04-2014-0063

Frey, K. S., Hirschstein, M. K., Edstrom, L. V., & Snell, J. L. (2009). Observed reductions in school bullying, nonbullying aggression, and destructive bystander behavior: A longitudinal evaluation. *Journal of Educational Psychology, 101*(2), 466 – 481. https://doi.org/10.1037/a0013839

Frost, A. M. (2012). *Bully prevention, peer mediation and conflict resolution: Impact of prevention programs on reducing school suspensions*. (Unpublished doctoral dissertation). Kansas State University.

Fung, A. L. (2012). Group treatment of reactive aggressors by social workers in a Hong Kong school setting: A two-year longitudinal study adopting quantitative and qualitative approaches. *British Journal of Social Work, 42,* 1533 – 1555. https://doi.org/10.1093/bjsw/bcr156

*Garandeau, C. F., Lee, I. A., & Salmivalli, C. (2014). Differential effects of the KiVa anti-bullying program on popular and unpopular bullies. *Journal of Applied Developmental Psychology, 35*(1), 44 – 50. https://doi.org/10.1016/j.appdev.2013.10.004

*Garandeau, C. F., Poskiparta, E., & Salmivalli, C. (2014). Tackling acute cases of school bullying in the KiVa anti-bullying program: A comparison of two approaches. *Journal of Abnormal Child Psychology, 42*(6), 981 – 991. https://doi.org/10.1007/s10802-014-9861-1

Gibson, J. E., Flaspohler, P. D., & Watts, V. (2015). Engaging youth in bullying prevention through community-based participatory research. *Family and Community Health, 38*(1), 120 – 130. https://doi.org/10.1097/FCH.0000000000000048

Giesbrecht, G. F., Leadbeater, B. J., & MacDonald, S. W. S. (2011). Child and context characteristics in trajectories of physical and relational victimization among early elementary school children. *Development and Psychopathology, 23,* 239 – 252. https://doi.org/10.1017/S0954579410000763

Goncy, E. A., Sutherland, K. S., Farrell, A. D., Sullivan, T. N., & Doyle, S. T. (2015). Measuring teacher implementation in delivery of a bullying prevention program: The impact of instructional and procedural adherence and competence on student responsiveness. *Prevention Science, 16*(3), 440 – 450. https://doi.org/10.1007/s11121-014-0508-9

Good, C. P., McIntosh, K., & Gietz, C. (2011). Integrating bullying prevention into schoolwide positive behavior support. *Teaching Exceptional Children*, *44*(1), 48. https://doi.org/10.1177/004005991104400106

*Gradinger, P., Yanagida, T., Strohmeier, D., & Spiel, C. (2015). Prevention of cyberbullying and cyber victimization: Evaluation of the ViSC social competence program. *Journal of School Violence, 14*(1), 87 – 110. https://doi.org/10.1080/15388220.2014.963231

Gregus, S. J., Craig, J. T., Hernandez Rodriguez, J., Pastrana, F. A., & Cavell, T. A. (2015). Lunch buddy mentoring for children victimized by peers: Two pilot studies. *Journal of Applied School Psychology, 31*(2), 167 – 197. https://doi.org/10.1080/15377903.2015.1025172

Greytak, E. A., & Kosciw, J. G. (2010). Year one evaluation of the New York City Department of Education "Respect for All" training program. *Gay, Lesbian, and Straight Education Network,* http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED512335

Greytak, E. A., Kosciw, J. G., & Boesen, M. J. (2013). Educating the educator: Creating supportive school personnel through professional development. *Journal of School Violence, 12*(1), 80 – 97. https://doi.org/10.1080/15388220.2012.731586

Gyooyeong, C. (2013). Effects of recurrence prevention program of school violence for adolescents with the experience of bullying. *Journal of The Korean Data Analysis Society, 15*(6), 3293 – 3307.

*Haataja, A., Voeten, M., Boulton, A. J., Ahtola, A., Poskiparta, E., & Samlivalli, C. (2014). The KiVa antibullying curriculum and outcome: Does fidelity matter? *Journal of School Psychology, 52,* 479 – 493. https://doi.org/10.1016/j.jsp.2014.07.001

Hallam, S. (2009). An evaluation of the Social and Emotional Aspects of Learning (SEAL) programme: Promoting positive behaviour, effective learning and well-being in primary school children. *Oxford Review of Education, 35*(3), 313 – 330. https://doi.org/10.1080/03054980902934597

*Harpin, S. B. (2011). *Missingness in longitudinal research: Attrition analysis and imputation approaches in a school-based study of young adolescents.* (Unpublished doctoral dissertation). University of Minnesota.

Hatzenbuehler, M. L., & Keyes, K. M. (2013). Inclusive anti-bullying policies and reduced risk of suicide attempts in Lesbian and Gay youth. *Journal of Adolescent Health, 53,* S21 – S26. https://doi.org/10.1016/j.jadohealth.2012.08.010

Hawe, R., Bond, L., Ghali, L. M., Perry, R., Davison, C.M., Casey, D.M., Butler, H., Webster, C.M., & Scholz, B. (2015). Replication of a whole school ethos-changing intervention: Different context, similar effects, additional insights. *BMC Public Health, 15*, 265. https://doi.org/10.1186/s12889-015-1538-3

Hervey, L., & Kornblum, R. (2006). An evaluation of Kornblum's body-based violence prevention curriculum for children. *The Arts in Psychotherapy, 33,* 113 – 129. https://doi.org/10.1016/j.aip.2005.08.001

Hoglund, W. L. G., Hosan, N. E., & Leadbeater, B. J. (2012). Using your WITS: A 6-year follow-up of a peer victimization prevention program. *Research into Practice, 41*(2), 193 – 214. https://doi.org/10.1080/02796015.2012.12087520

Holden, B. C. (2015). *Student, teacher, and administrator perceptions of the effectiveness of the Olweus Bullying Prevention Program.* (Unpublished doctoral dissertation). Trevecca Nazarene University.

Hornblower, K. (2014). *Don't Stand By, Stand Up: A peer group anti-bullying intervention to increase pro-defending attitudes and behaviour in students that witness bullying.* (Unpublished doctoral dissertation). University of Exeter.

Huddleston, L. B., Varjas, K., Meyers, J., & Cadenhead, C. (2011). A case study with an identified bully: Policy and practice implications. *The Western Journal of Emergency Medicine, 12*(3), 316 – 323.

*Hutchings, J., & Clarkson, S. (2015). Introducing and piloting the KiVa bullying prevention programme in the UK. *Educational and Child Psychology, 32*(1), 49 – 61.

Isaacs, L. K. A. (2009). *Teachers' perceptions of the effectiveness of the Olweus Bullying Prevention program in hindering bullying behaviors.* (Unpublished doctoral dissertation). Widener University.

James, A. (2011). *The use and impact of peer support schemes in schools in the UK, and a comparison with use in Japan and South Korea.* (Unpublished doctoral dissertation) Goldsmiths, University of London.

James, D., Flynn, A., Lawlor, M., Courtney, P., Murphy, N., & Henry, B. (2011). A friend in deed? Can adolescent girls be taught to understand relational bullying? *Child Abuse Review, 20*(6), 439 – 454. https://doi.org/10.1002/car.1120

James, D., Lawlor, M., Murphy, N., & Flynn, A. (2013). 'It's a girl thing!' Do boys engage in relational aggression? Exploration of whether strategies to educate young people about relational aggression are relevant for boys. *Pastoral Care in Education, 31*(2), 156 – 172. https://doi.org/10.1080/02643944.2013.774042

*Jenson, J. M., & Dieterich, W. A. (2007). Effects of a skill-based prevention program on bullying and bully victimization among elementary school children. *Prevention Science, 8,* 285 – 296. https://doi.org/10.1007/s11121-007-0076-3

*Jenson, J. M., Dieterich, W. A., Brisson, D., Bender, K. A., & Powell, A. (2010). Preventing childhood bullying: Findings and lessons from the Denver Public Schools trial. *Research on Social Work Practice, 20*(5), 509 – 517. https://doi.org/10.1177/1049731509359186

Jeong-Lan, I., & Oh-Hyun, O. (2014). The effect of school violence prevention program for improvement of empathic ability on schoolchildren who are indifferent to school violence. *Korean Journal of Counselling, 15*(4), 1425 – 1440. https://doi.org/10.15703/kjc.15.4.201408.1425

Jiménez-Barbero, J. A., Ruiz-Hernández, J. A., Llor-Esteban, B., Llor-Zaragoza, & García, M. P. (2013). Efficacy of a brief intervention on attitudes to reduce school violence: A randomized clinical trial. *Children and Youth Services Review, 35,* 1313 – 1318. https://doi.org/10.1016/j.childyouth.2013.05.010

Knights, N. H. (2011). *Evaluating an intensive recovery programme for adolescents who have been bullied: A mixed methods study.* (Unpublished doctoral dissertation). University of Hertfordshire.

Konishi, C., Saewyc, E., Homma, Y., & Poon, C. (2013). Population-level evaluation of school-based interventions to prevent problem substance use among gay, lesbian and bisexual adolescents in Canada. *Preventive Medicine, 57,* 929 – 933. https://doi.org/10.1016/j.ypmed.2013.06.031

*Kyriakides, L., Creemers, B. P. M., Muijs, D., Rekers-Mombarg, L., Petegem, P. van, & Pearson, D. (2014). Using the dynamic model of educational effectiveness to design strategies and actions to face bullying. *School Effectiveness and School Improvement, 25*(1), 83 - 104. https://doi.org/10.1080/09243453.2013.771686

Langevin, M., & Prasad, N. G. N. (2012). A stuttering education and bullying awareness and prevention resource: A feasibility study. *Language, Speech & Hearing Services in Schools, 43*(3), 344 – 358A. https://doi.org/10.1044/0161-1461(2012/11-0031)

Layfield, B. H. (2014). *A multi-faceted approach to reduce incidents of student misbehaviour though prevention and response measures at Sussex Central High School.* (Unpublished doctoral dissertation). University of Delaware.

Leadbeater, B., & Sukhawatanakul, P. (2011). Multicomponent programs for reducing peer victimization in early elementary school: A longitudinal evaluation of the WITS primary program. *Journal of Community Psychology, 39*(5), 606 - 620. https://doi.org/10.1002/jcop.20447

Leff, S. S., Waasdorp, T. E., Paskewich, B., Gullan, R. L., Jawad, A. F., MacEvoy, J.P., Feinberg, B.E., & Power, T.J. (2010). The preventing relational aggression in schools everyday program: A preliminary evaluation of acceptability and impact. *School Psychology Review, 39*(4), 569 – 587. https://doi.org/10.1080/02796015.2010.12087742

*Lewis, K. M., Schure, M. B., Bavarian, N., DuBois, D. L., Day, J, Ji, P., Silverthorn, N., Acock, A., Vuchinich, S., & Flay, B.R. (2013). Problem behavior and urban, low-income youth: A randomized controlled trial of Positive Action in Chicago. *American Journal of Preventive Medicine, 44*(6), 622 – 630. https://doi.org/10.1016/j.amepre.2013.01.030

*Lishak, N. (2011). *Examination of bullying behaviours and the implementation of a social norms project in a middle school.* (Unpublished doctoral dissertation). Walden University.

*Low, S., & Van Ryzin, M. (2014). The moderating effects of school climate on bullying prevention efforts. *School Psychology Quarterly, 29*(3), 306 – 319. https://doi.org/10.1037/spq0000073

Low, S., Van Ryzin, M. J., Brown, E. C., Smith, B. H., & Haggerty, K. P. (2014). Engagement matters: Lessons from assessing classroom implementation of Steps to Respect: A Bullying Prevention Program over a one-year period. *Prevention Science, 15,* 165 - 176. https://doi.org/10.1007/s11121-012-0359-1

Lucassen, M. F. G., & Burford, J. (2015). Educating for diversity: An evaluation of a sexuality diversity workshop to address secondary school bullying. *Australasian Psychiatry, 23*(5), 544 - 549. https://doi.org/10.1177/1039856215592324

Macedo, E., Martins, F., Cainé, J., Macedo, J., & Novais, R. (2014). Bullying escolar y evaluación de un programa de intervención. *Revista Portuguesa de Enfermagem de Saúde Mental*, (SPE1), 15-20.

Malatino, K.W. (2012). *The impact of a systematic approach to student support on middle-childhood development for low-income, urban children.* (Unpublished doctoral dissertation). Boston College.

McElearney, A., Adamson, G., Shevlin, M., & Bunting, B. (2013). Impact evaluation of a school-based counselling intervention in Northern Ireland: Is it effective for pupils who have been bullied? *Child Care in Practice, 19*(1), 4 - 22. https://doi.org/10.1080/13575279.2012.732557

*Menard, S., Grotpeter, J., Gianola, D., & O'Neal, M. (2008). *Evaluation of Bully Proofing your school: Final report.* http://www.ncjrs.gov/pdffiles1/nij/grants/221078.pdf

Mendes, C. S. (2011). Preventing school violence: An evaluation of an intervention program. *Revista da Escola de Enfermagem da USP, 45*(3), 581 - 588.

Menesini, E., & Nocentini, A. (2012). Peer education intervention: Face-to-face versus online. In A. Costabile, & B. Spears (Eds.), *The impact of technology on relationships in educational settings.* (pp. 139 – 150). Routledge.

Migliaccio, T., & Raskauskas, J. (2013). Small-scale bullying prevention discussion video for classrooms: A preliminary evaluation. *Children and Schools, 35*(2), 71 – 81. https://doi.org/10.1093/cs/cdt003

Minton, S. J., O' Mahoney, M., & Conway-Walsh, R. (2013). A 'whole-school/community development' approach to preventing and countering bullying: The Erris Anti-Bullying Initiative (2009 - 2011). *Irish Educational Studies, 32*(2), 233 - 249. https://doi.org/10.1080/03323315.2013.784637

Miyairi, M. (2013). *Fat bias and weight-related teasing prevention among adolescents*. (Unpublished doctoral dissertation). University of Utah.

Nakamura, R., & Koshikawa, F. (2014). Evaluating a psycho-educational program for the prevention of bullying in junior high school. *Japanese Journal of Educational Psychology, 62*(2), 129 - 142. https://doi.org/10.5926/jjep.62.129

Nese, R. N. T., Horner, R. H., Dickey, C. R., Stiller, B., & Tomlanovich, A. (2014). Decreasing bullying behaviors in middle school: Expect Respect. *School Psychology Quarterly, 29*(3), 272 – 286. https://doi.og/10.1037/spq0000070

Newgent, R. A., Behrend, B. A., Lounsbery, K. L., Higgins, K. K., & Lo, W. -J. (2010). Psychosocial educational groups for students (PEGS): An evaluation of the treatment effectiveness of a school-based behavioral intervention program. *Counselling Outcome Research and Evaluation, 1*(2), 80 – 94. https://doi.org/10.1177/2150137810373919

Nixon, C. L., & Werner, N. E. (2010). Reducing adolescents' involvement with relational aggression: Evaluating the effectiveness of the creating a safe school (CASS) intervention. *Psychology in the Schools, 47*(6), 606 – 620. https://doi.org/10.1002/pits.20494

*Noland, B. (2011). *Effects of the KiVa anti-bullying program on adolescents' perceptions of peers, depression, and anxiety.* (Unpublished doctoral dissertation). University of Kansas.

Pack, C., White, A., Raczynski, K., & Wang, A. (2011). Evaluation of the Safe School Ambassadors program: A student-led approach to reducing mistreatment and bullying in schools. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 84*(4), 127 – 133. https://doi.org/10.1080/00098655.2011.564974

Park, J., Kim, J., & Wie, S. (2014). The expected effect of food-therapeutical leisure program on preventing school bullying. *Journal of Tourism Sciences, 38*(6), 57 – 78.

Peagram, K. (2013). *Bully prevention and intervention program: Bulldog solution intervention model.* (Unpublished doctoral dissertation). The Chicago School of Professional Psychology.

Pepler, D., & Craig, W. (2011). Promoting relationships and eliminating violence in Canada. *International Journal of Behavioral Development, 35*(5), 389 – 397. https://doi.org/10.1177/0165025411407455

Phillips, L. D. (2015). *Educators' perceptions of bullying before and after implementing a bullying prevention program.* (Unpublished doctoral dissertation). Walden University.

Pister, R. L. (2010). *Working against youth violence everywhere: Evaluating a peer-led approach to bullying prevention.* (Unpublished masters dissertation). Wilfrid Laurier University.

*Ploeg, R. van der, Steglich, C., & Veenstra, R. (2016). The support group approach in the Dutch KiVa anti-bullying programme: Effects on victimization, defending and well-being at school. *Educational Research, 58*(3), 221 – 236. https://doi.org/10.1080/00131881.2016.1884949

Ramírez, F. C., & Lacasa, C. S. (2013). Efficacy of the CIP program for improve the coexistence and preventing bullying in primary schoolchildren. *Apuntes de Psicología, 31*(2), 173 – 181.

Renshaw, T. L., & Jimerson, S. R. (2012). Enhancing student attitudes via a brief, universal-level bullying prevention curriculum. *School Mental Health, 4,* 115 – 128. https://doi.org/10.1007/s12310-011-9069-2

Rigby, K., & Griffiths, C. (2011). Addressing cases of bullying through the method of shared concern. *School Psychology International, 32*(3), 345 – 357. https://doi.org/10.1177/0143034311402148

Roberto, A. J., Eden, J., Savage, M. W., Ramos-Salazar, L., & Deiss, D. M. (2014). Outcome evaluation results of school-based cybersafety promotion and cyberbullying prevention intervention for middle school students. *Health Communication, 29*(10), 1029 – 1042. https://doi.org/10.1080/10410236.2013.831684

Ross, S. W. (2009). *Bully prevention in positive behavior support.* (Unpublished doctoral dissertation). University of Oregon.

Ross, S. W., & Horner, R. H. (2014). Bully prevention in positive behavior support: Preliminary evaluation of third-, fourth-, and fifth-grade attitudes toward bullying. *Journal of Emotional and Behavioral Disorders, 22*(4), 225 – 236. https://doi.org/10.1177/1063426613491429

Ross, S. W., & Horner, R. H. (2010). Bully prevention in positive behavior support. *Journal of Applied Behavior Analysis, 42,* 747 – 759. https://doi.org/10.1901/jaba.2009.42-747

Rubin-Vaughan, A., Pepler, D., Brown, S., & Craig, W. (2011). Quest for the Golden Rule: An effective social skills promotion and bullying prevention program. *Computers and Education, 56,* 166 - 175. https://doi.org/10.1016.j.compedu.2010.08.009

*Şahin, M. (2012). An investigation into the efficiency of empathy training program on preventing bullying in primary schools. *Children and Youth Services Review, 34,* 1325 – 1330. https://doi.org/10.1016/j.childyouth.2012.03.013

*Sainio, M., Veenstra, R., Huitsing, G., & Salmivalli, C. (2012). Same- and other-sex victimization: Are the risk factors similar? *Aggressive Behavior, 38,* 422 – 455. https://doi.org/10.1002/ab.21445

*Salmivalli, C., Kärnä, A., & Poskiparta, E. (2012). Counteracting bullying in Finland: The KiVa program and its effects on different forms of being bullied. *International Journal of Behavioral Development, 35*(5), 405 – 411. https://doi.org/10.1177/0165025411407457

Santos, R. G., Chartier, M. J., Whalen, J. C., Chateau, D., & Boyd, L. (2011). Effectiveness of school-based violence prevention for children and youth: A research report. *Healthcare Quarterly, 14,* 80 – 91. https://doi.org/10.12927/hcq.2011.22367

Saurini, L. (2011). *The effects of anger management on bullying incidences in the middle school system.* (Unpublished doctoral dissertation). Walden University.

Scheithauer, H., & Bull, H. D. (2010). Tuitional-based promotion of social competencies and prevention of bullying in adolescence – the fairplayer.manual: Results of a pilot evaluation study. *Praxis der Kinderpsychologie und Kinderpsychiatrie, 59*(4), 266 – 281. https://doi.org/10.13109/prkk.2010.59.4.266

Schroeder, B. A., Messina, A., Schroeder, D., Good, K., Barto, S., Saylor, J., & Masiello, M. (2012). The implementation of a state-wide bullying prevention program: Preliminary findings from the field and the importance of coalitions. *Health Promotion Practice, 13*(4), 49 – 495. https://doi.org/10.1177/1524839910386887

Shek, D. T. L., & Yu, L. (2013). Impact of the Project P.A.T.H.S. on adolescent risk behavior: A five-year longitudinal study. In D. T. L. Shek & R. C. F. Sun (Eds.), *Development and evaluation of positive adolescent training through holistic social programs,* pp. 291 – 304. Springer.

Spiel, C., Wagner, P., & Strohmeier, D. (2012). Violence prevention in Austrian schools: Implementation and evaluation of a national strategy. *International Journal of Conflict and Violence, 6*(2), 176 – 186. https://doi.org/10.1177/0165025411407458

Splett, J. D., Maras, M. A., & Brooks, C. M. (2015). GIRLSS: A randomized, pilot study of a multisystemic, school-based intervention to reduce relational aggression. *Journal of Child and Family Studies, 24*(8), 2250 – 2261. https://doi.org/10.1007/s10826-014-0027-0

Sreckovic, M. A. (2015). *Peer network interventions for secondary students with ASD: Effects on social interaction and bullying victimization.* (Unpublished doctoral dissertation). University of North Carolina.

Stallard, P., & Buck, R. (2013). Preventing depression and promoting resilience: Feasibility study of a school-based cognitive-behavioural intervention. *British Journal of Psychiatry, 202*(s54), s18 – s23. https://doi.org/10.1192/bjp.bp.112.119172

Steiger, A. D. (2010). *Solution Team: A program evaluation of an anti-bullying intervention*. (Unpublished doctoral dissertation). California Institute of Integral Studies, San Francisco.

Tokarick, K. (2015). *The effect of bullying prevention programs on female youth.* (Unpublished doctoral dissertation). Walden University.

Tomić-Latinac, M., & Nikčević-Milković, A. (2010). Procjena učinkovitosti UNICEF-Ovog programa prevencije vršnjačkog nasilja i zlostavljanja. *Ljetopis scoijalnog rada, 16*(3), 635 – 657.

Vannini, N., Enz, S., Sapouna, M., Wolke, D., Watson, S., Woods, S., Dautenhahn, K., Hall, L., André, E., Aylett, R., & Schneider, W. (2011). "FearNot!": A computer-based anti-bullying-programme designed to foster peer intervention. *European Journal of Psychology of Education, 26*(1), 21 – 44. https://doi.org/10.1007/s10212-010-0035-4

Velderman, C. J. (2015). *A preliminary evaluation of a team-based institute model of professional development in bullying prevention.* (Unpublished doctoral dissertation). University of New Jersey.

Watson, S. E. J., Vannini, N., Woods, S., Dautenhahn, K., Sapouna, M., Enz, S., Schneider, W., Wolke, D., Hall, L., Paiva, A., André, E., & Aylett, R. (2010). Inter-cultural differences in response to a computer-based anti-bullying intervention, *Educational Research, 52*(1), 61 – 80. https://doi.org/10.1080/00131881003588261

Westheimer, K., & Szalacha, L. A. (2015). Welcoming schools: Lesbian, gay, bisexual, transgender, and gender-inclusive bullying prevention in elementary schools. In P. Goldblum, D. L. Espelage, J. Chu, & B. Bongar (Eds.), *Youth suicide and bullying: Challenges and strategies for prevention and intervention,* pp. 231 – 245. Oxford University Press

*Williford, A., Boulton, A., Noland, B., Little, T. D., Kärnä, A., & Salmivalli, C. (2012). Effects of the KiVa anti-bullying program on adolescents' depression, anxiety, and perception of peers. *Journal of Abnormal Child Psychology, 40,* 289 – 300. https://doi.org/10.1007/s10802-011-9551-1

*Williford, A., Elledge, L. C., Boulton, A. J., DePaolis, K. J., Little, T. D., & Salmivalli, C. (2013). Effects of the KiVa anti-bullying program on cyberbullying and cybervictimization frequency among Finnish youth. *Journal of Clinical Child and Adolescent Psychology, 42*(6), 820 – 833. https://doi.org/10.1080/15374416.2013.787623

Wolfe, D. A., Crooks, C. V., Chiodo, D., Hughes, R., & Ellis, W. (2012). Observations of Adolescent peer resistance skills following a classroom-based healthy relationship program: A post-intervention comparison. *Prevention Science, 13*(2), 196 – 205. https://doi.org/10.1007/s11121-011-0256-z

Wood, B. F. (2012). *An evaluation of the implementation fidelity and outcomes of the Olweus Bullying Prevention Program in three elementary schools in Virginia.* (Unpublished doctoral dissertation). College of William and Mary.

Wright, W., Bailey, C. L., & Bergin, J. J. (2012). The effects of a bully intervention program on the relational aggressive behaviors of 5th grade girls. *Georgia Counsellors Association Journal, 19*(1), 65 – 91.

*Wurf, G. (2012). High school anti-bullying interventions: An evaluation of curriculum approaches and the method of Shared Concern in four Hong Kong international schools. *Australian Journal of Guidance and Counselling, 22*(1), 139 – 149. https://doi.org/10.1017/jgc.2012.2

Yamashiro, S. K. M. (2013). *Exploring heterosexist bias and oppressive bullying in anti-bullying prevention programs.* (Unpublished masters dissertation). University of Nevada.

*Yang, A., & Salmivalli, C. (2015). Effectiveness of the KiVa antibullying programme on bully-victims, bullies and victims. *Educational Research, 57*(1), 80 – 90. https://doi.org/10.1080/00131881.2014.983724

Young, A., Hardy, V., Hamilton, C., Biernesser, K, Sun, L., & Niebergall, S. (2009). Empowering students: Using data to transform a bullying prevention and intervention program. *Professional School Counselling, 12*(6), 413 – 420. https://doi.org/10.1177/2156759x0901200616

12.3 Studies included in the cyberbullying meta-analysis

Athanasiades, C., Kamariotis, H., Psalti, A., Baldry, A. C., & Sorrentino, A. (2015). Internet use and cyberbullying among adolescent students in Greece: The 'Tabby' project. *Hellenic Journal of Psychology, 12,* 14 – 39.

Chaux, E., Velásquez, A. M., Schultze-Krumbholz, A., & Scheithauer, H. (2016). Effects of the cyberbullying prevention program Media Heroes *(Medienhelden)* on traditional bullying. *Aggressive Behavior, 42,* 157 – 165. https://doi.org/10.1002/ab.21637

Cross, D., Shaw, T., Hadwen, K., Cardoso, P., Slee, P., Roberts, C., Thomas, L., & Barnes, A. (2016). Longitudinal impact of the Cyber Friendly Schools program on adolescents' cyberbullying behavior. *Aggressive Behavior, 42*(2), 166 – 180. https://doi.org/10.1002/ab.21609

Del Rey, R., Casas, J. A., & Ortega, R. (2012). The ConRed program: An evidence-based practice. *Communicar, 23*(45), 129 – 138. https://doi.org/10.3916/c39-2012-03-03

Del Rey, R., Casas, J. A., & Ortega, R. (2015). The impacts of the ConRed program on different cyberbullying roles. *Aggressive Behavior, 42*(2), 123 – 135. https://doi.org/10.1002/ab.21608

Del Rey, R., Casas, J. A., & Ortega, R. (2016). Impact of the ConRed program on different cyberbullying roles. *Aggressive Behavior, 42,* 123 – 135. https://doi.org/10.1002/ab.21608

DeSmet, A., Bastiaensens, S., Cleemput, K., van, Poels, K., Vandebosch, H., Deboutte, G., Herrewijn, L., Malliet, S., Pabian, S., Broeckhoven, F. van, Troyer, O. de, Deglorie, G., Hoecke, S. van, Samyn, K., & Bourdeaudhuij, I., de (2018). The efficacy of the Friendly Attac serious digital game to promote prosocial bystander behavior in cyberbullying among young adolescents: A cluster-randomized controlled trial.

Computers in Human Behavior, 78, 336 – 347. https://doi.org/10.1016/j.chb.2017.10.011

Espelage, D. L., Low, S., van Ryzin, M. J., & Polanin, J. R. (2015). Clinical trial of Second Step middle school program: Impact on bullying, cyberbullying, homophobic teasing, and sexual harassment perpetration. *School Psychology Review, 44*(4), 464 – 479. https://doi.org/10.17105/spr-15-0052.1

Fekkes, M., van de Sande, M. C. E., Gravesteijn, J. C., Pannebakker, F. D., Buijs, G. J., Diekstra, R. F. W., & Kocken, P. L. (2016). Effects of the Dutch Skills for Life program on the health behaviour, bullying, and suicidal ideation of secondary school students, *Health Education, 116*(1), 2 – 15. https://doi.org/10.1108/HE-05-2014-0068

Garaigordobil, M., & Martínez-Valderrey, V. (2015a). Effects of Cyberprogram 2.0 on “face-to-face” bullying, cyberbullying, and empathy. *Psicothema, 27*(1), 45 – 51. https://doi.org/10.7334/psicothema2014.78

Garaigordobil, M., & Martínez-Valderrey, V. (2016). Impact of Cyberprogram 2.0 on different types of school violence and aggressiveness. *Frontiers in Psychology, 7.* https://doi.org/10.3389/fpsyg.2016.00428

Gradinger, P., Yanagida, T., Strohmeier, D., & Spiel, C. (2015). Prevention of cyberbullying and cyber victimization: Evaluation of the ViSC social competence program. *Journal of School Violence, 14*(1), 87 – 110. https://doi.org/10.1080/15388220.2014.963231

Gradinger, P., Yanagida, T., Strohmeier, D., & Spiel, C. (2016). Effectiveness and sustainability of the ViSC social competence program to prevent cyberbullying and cyber-victimization: Class and individual level moderators. *Aggressive Behavior, 42,* 181 – 193. https://doi.org/10.1002/ab.21631

Menesini, E., Nocentini, A., & Palladino, B. E. (2012a). Empowering students against bullying and cyberbullying: Evaluation of an Italian peer-led model. *International Journal of Conflict and Violence, 6*(2), 314 – 321.

Nocentini, E., & Palladino, A. (2012). Peer education intervention: Face-to-face versus online. In A. Costabile, & B. Spears (Eds.), *The Impact of Technology in relationships in educational settings,* (pp. 139 – 150). Routledge, Taylor & Francis Group.

Ortega-Ruiz, R., Del Rey, R., & Casas, J. A. (2012). Knowing, building and living together on Internet and social networks: The ConRed cyberbullying prevention program. *International Journal of Conflict and Violence, 6*(2), 303 – 313. https://doi.org/10.4119/UNIBI/ijcv.250

Palladino, B. E., Nocentini, A., & Menesini, E. (2012). Online and offline peer led models against bullying and cyberbullying. *Psicothma, 24*(4), 634 – 639.

Palladino, B. E., Nocentini, A., & Menesini, E. (2016). Evidence-based intervention against bullying and cyberbullying: Evaluation of the NoTrap! program in two independent trials. *Aggressive Behavior, 42,* 194 – 206. https://doi.org/10.1002/ab.21636

Pieschl, S., Kourteva, P., & Stauf, L. (2017). Challenges in the evaluation of cyberbullying prevention: Insights from two case studies. *International Journal of Developmental Science, 11,* 45 – 54. https://doi.org/10.3233/DEV-160209

Pieschl, S., & Urbasik, S. (2013). Does the cyberbullying prevention program Surf-Fair work? An evaluation study. In R. Hanewald (Ed.), *From cyberbullying to cyber safety: Issues and approaches in education contexts,* (pp. 205 – 224). Nova Science Publishers.

Roberto, A. J., Eden, J., Savage, M. W., Ramos-Salazar, L., & Deiss, D. M. (2014). Outcome evaluation results of school-based cybersafety promotion and cyberbullying prevention intervention for middle school students. *Health Communication, 29*(10), 1029 – 1042. https://doi.org/10.1080/10410236.2013.831684

Schultze-Krumbholz, A., Schultze, M., Zagorscak, P., Wölfer, R., & Scheithauer, H. (2016). Feeling cybervictims' pain: The effect of empathy training on cyberbullying. *Aggressive Behavior, 42,* 147 – 156. https://doi.org/10.1002/ab.21613

Shaw, T., Cross, D., & Zubrick, S. R. (2015). Testing for response shift bias in evaluations of school antibullying programs. *Evaluation Review, 39*(6), 527 – 554. https://doi.org/10.1177/019384IXI6629863

Solomontos-Kountouri, O., Gradinger, P., Yanagida, T., & Strohmeier, D. (2016). The implementation and evaluation of the ViSC program in Cyprus: Challenges of cross-national dissemination and evaluation results. *European Journal of Developmental Psychology, 13*(6), 737 – 755. https://doi.org/10.1080/17405629.2015.1136618

Tanrikulu, T., Kinay, H., & Aricak, O. T. (2015). Sensibility development program against cyberbullying. *New Media and Society, 17*(5), 708 – 719. https://doi.org/10.1177/1461444813511923

Williford, A., Elledge, L. C., Boulton, A. J., DePaolis, K. J., Little, T. D., & Salmivalli, C. (2013). Effects of the KiVa antibullying program on cyberbullying and cybervictimization frequency among Finnish youth. *Journal of Clinical Child and Adolescent Psychology, 42*(6), 820 – 833. https://doi.org/10.1080/15374416.2013.787623

Wölfer, R., Schultze-Krumbholz, A., Zagorscak, P., Jäkel, A., Göbel, K., & Scheithauer, H. (2014). Prevention 2.0: Targeting cyberbullying @ school. *Prevention Science, 15,* 879 - 887. https://doi.org/10.1007/s11121-013-0438-y

12.4 Studies excluded from the cyberbullying meta-analysis[24]

** = excluded due to missing information, repeat publications or overlapping samples*

Appelqvist-Schmidlechner, K., Liski, A., Pankakoski, M., Solantaus, T., Ojala, T., Kampman, M., & Santalahti, P. (2016). Together at school intervention programme. A pilot study on the feasibility and perceived benefits of a programme focusing on improving socio-emotional skills among school children in Finland. *International Journal of Mental Health Promotion, 18*(3), 127 – 143. https://doi.org/10.1080/14623730.2016.1193763

Avşar, F., & Alkaya, S. A. (2017). The effectiveness of assertiveness training for school-aged children on bullying and assertiveness level. *Journal of Pediatric Nursing, 36,* 186 – 190. https://doi.org/10.1016/j.pedn.2016.060.020

Barkoukis, V., Lazuras, L., Ourda, D., & Tsorbatzoudis, H. (2016). Tackling psychosocial risk factors for adolescent cyberbullying: Evidence from a school-based intervention. *Aggressive Behavior, 42,* 114 – 122. https://doi.org/10.1002/ab.21625

Camelford, K. G., & Ebrahim, C. (2016). The cyberbullying virus: A psychoeducational intervention to define and discuss cyberbullying among high school females. *Journal of Creativity in Mental Health, 11*(3 – 4), 458 – 468. https://doi.org/10.1080/15401383.2016.1183545

Chaux, E., Barrera, M., Molano, A., Velásquez, A. M., Castellanos, M., Chaparro, M. P., & Bustamante, A. (2017). Classrooms in Peace within violent contexts: Field evaluation of *Aulas en Paz* in Colombia. *Prevention Science, 18,* 828 – 838. https://doi.org/10.1007/s11121-017-0754-8

[24] *Note* = studies included in the school-bullying meta-analysis are not in this list.

Clarkson, S., Axford, N., Berry, V., Edwards, R. T., Bjornstad, G., Wrigley, Z., Charles, J., Hoare, Z., Ukoumunne, O. C., Matthews, J., & Hutchings, J. (2016). Effectiveness and micro-costing of the KiVa school-based bullying prevention programme in Wales: Study protocol for a pragmatic definitive parallel group cluster randomised controlled trial. *BMC Public Health, 16,* 104. https://doi.org/10.1186/s12889-016-2746-1

Cleemput, K., van, Vandebosch, H., Poels, K., Bastiaensens, S., DeSmet, A., & Bourdeaudhuij, I., de (2016). The development of a serious game on cyberbullying. In T. Völlink, F. Dehue, & C. McGuckin (Eds.), *Cyberbullying: From theory to intervention,* (pp. 93 – 109). Routledge.

Dillon, K. P., & Bushman, B. J. (2015). Unresponsive or un-noticed? Cyber-bystander intervention in an experimental cyberbullying context. *Computers in Human Behavior, 45,* 144 – 150. https://doi.org/10.1016/j.chb.2014.12.009

Doane, A. N., Kelley, M. L., & Pearson, M. R. (2016). Reducing cyberbullying: A theory of reasoned action-based video prevention program for college students. *Aggressive Behavior, 42,* 136 – 146. https://doi.org/10.1002/ab.21610

Foshee, V. A., Benefield, T., Dixon, K. S., Chang, L. -Y., Senkomago, V., Ennett, S. T., Moracco, K. E., & Bowling, J. M. (2015). The effects of Moms and Teens for Safe Dates: A dating abuse prevention program for adolescents exposed to domestic violence. *Journal of Youth and Adolescence, 44,* 995 – 1010. https://doi.org/10.1007/s10964-015-0272-6

Garaigordobil, M., & Martínez-Valderrey, V. (2014). Effect of Cyberprogram 2.0 on reducing victimization and improving social competence in adolescence. *Revista de Psicodidáctica, 19*(2), 289 – 305. https://doi.org/10.1387/RevPsicodidact.10239

Garaigordobil, M., & Martínez-Valderrey, V. (2015b). The effectiveness of Cyberprogram 2.0 on conflict resolution strategies and self-esteem. *Journal of Adolescent Health, 57,* 229 – 234. https://doi.org/10.1016/j.jadohealth.2015.04.007

Garaigordobil, M., Martínez-Valderrey, V., & Machimbarrena, J. M. (2017). Intervention in bullying and cyberbullying: Assessment of Martin's case. *Revista de Psicologia clinica con ninos y adolescentes, 4*(1), 25 – 32.

Guo, S., Wu, Q., Smokowski, P. R., Bacallao, M., Evans, C. B. R., & Cotter, K. L. (2015). A longitudinal evaluation of the Positive Action program in a low-income, racially diverse, rural county: Effects on self-esteem, school hassles, aggression, and internalizing symptoms. *Journal of Youth and Adolescence, 44,* 2337 – 2358. https://doi.org/10.1007/s10964-015-0358-1

*Harshman, K.L. (2014). *Assessing effectiveness of age-appropriate curriculum on Internet safety education and cyberbullying prevention.* (Unpublished doctoral dissertation). Grand Canyon University.

Hicks, J. F., Le Clair, B., & Berry, S. (2016). Using solution-focused dramatic empathy training to eliminate cyber-bullying. *Journal of Creativity in Mental Health, 11*(3-4), 378 – 390. https://doi.org/10.1080/15401383.2016.1172533

Jacobs, N. C. L., Dehue, F., Völlink, T., & Lechner, L. (2016). Online Pestkoppenstoppen [Stop the bully online]: The systematic development of a web-based tailored intervention for adolescent cyberbully victims to prevent cyberbullying. In T. Völlink, F. Dehue, & C. McGuckin (Eds.), *Cyberbullying: From theory to intervention,* (pp. 110 – 155). Routledge.

Juvonen, J., Schacter, H. L., Sainio, M., & Salmivalli, C. (2016). Can a school-wide bullying prevention program improve the plight of victims? Evidence for risk x intervention effects. *Journal of Consulting and Clinical Psychology, 84*(4), 334 – 344. https://doi.org/10.1037/ccp0000078

Lee, M-. S., Zi-Pei, W., Svanström, L., & Dalal, K. (2013). Cyber bullying prevention: Intervention in Taiwan. *PLOS One, 8*(5), e64031. https://doi.org/10.1371/journal.pone.0064031

Leff, S. S., Waasdorp, T. E., & Paskewich, B. S. (2016). The broader impact of Friend to Friend (F2F): Effects on teacher-student relationships, prosocial behaviors and relationally and physically aggressive behaviors. *Behavior Modification, 40*(4), 589 – 610. https://doi.org/10.1177/0145445516650879

Leung, A. N. M., Fung, D. C. -L., & Farver, J. A. (2017). A cyberbullying intervention for Hong Kong Chinese college students. *Applied Research Quality Life,* https://doi.org/10.1007/s11482-017-9572-1

McCuddy, T., & Esbensen, F. -A. (2017). After the bell and into the night: The link between delinquency and traditional, cyber-, and dual-bullying victimization. *Journal of Research in Crime and Delinquency, 54*(3), 409 – 441. https://doi.org/10.1177/0022427816683515

McElearney, A., Roosmale-Cocq, S., Scott, J., & Stephenson, P. (2008). Exploring the anti-bullying role of a befriending peer support programme: A case study within the primary school setting in Northern Ireland. *Child Care in Practice, 14*(2), 109 – 130. https://doi.org/10.1080/13575270701868694

Midgett, A., Moody, S. J., Reilly, B., & Lyter, S. (2017). The phenomenological experience of student-advocates trained as defenders to stop school bullying. *Journal of Humanistic Counselling, 56,* 53 – 71. https://doi.org/10.1002/johc.12044

Ploeg, R. van der, Steglich, C., & Veenstra, R. (2016). The support group approach in the Dutch KiVa anti-bullying programme: Effects on victimisation, defending and well-being at school. *Educational Research, 58*(3), 221 – 236. https://doi.org/10.1080/00131881.2016.1184949

Roberto, A. J., Eden, J., Deiss, D. M., Savage, M. W., & Ramos-Salazar, L. (2017). The short-term effects of a cyberbullying prevention intervention for parents of middle school students. *International Journal of Environmental Research and Public Health, 14,* 1038. https://doi.org/10.3390/ijerph14091038

Savage, M. W., Deiss Jr., D. M., Roberto, A. J., & Aboujaoude, E. (2017). Theory-based formative research on an anti-cyberbullying victimization intervention message. *Journal of Health Communication: International Perspectives, 22*(2), 124 – 134. https://doi.org/10.1080/10810730.2016.1252818

Silva, J. L. da, Oliveria, W. A. de, Fallerios Braga, I., Farias, M. S., Silva Lizzi, E. A., da, Carvalho Gonçalves, M. F., Oliveira Pereira, B., & Iossi Silva, M. A. (2016). The effects of a skill-based intervention for victims of bullying in Brazil. *International Journal of Environmental Research and Public Health, 13,* 1042 – 1052. https://doi.org/10.3390/ijerph13111042

Sullivan, T. N., Sutherland, K. S., Farrell, A. D., Taylor, K. A., & Doyle, S. T. (2017). Evaluation of violence prevention approaches among early adolescents: Moderating effects of disability status and gender. *Journal of Child and Family Studies, 26,* 1151 – 1163. https://doi.org/10.1007/s10826-016-0629-9

Tangen, D., & Campbell, M. (2010). Cyberbullying prevention: One primary schools' approach. *Australian Journal of Guidance &and Counselling, 20*(2), 225 – 234. https://doi.org/10.1375/ajgc.20.2.225

Tanrikulu, T., Kinay, H., & Aricak, O. T. (2015). Sensibility development program against cyberbullying. *New Media and Society, 17*(5), 708 – 719. https://doi.org/10.1177/1461444813511923

Timmons-Mitchell, J., Levesque, D. A., Harris III, L. A., Flannery, D. J., & Falcone, T. (2016). Pilot test of StandUp, an online school-based bullying prevention program. *Children and Schools, 38*(2), 71 – 79. https://doi.org/10.1093/cs/cdw010

Toshack, T., & Colmar, S. (2012). A cyberbullying intervention with primary-aged students. *Australian Journal of Guidance and Counselling, 22*(2), 268 – 278. https://doi.org/10.1017/jgc.2012.31

Usó, I., Villanueva, L., & Adrián, J. E. (2016). Impact of peer mediation programmes to prevent bullying behaviours in secondary schools/ El impacto de los programas de mediación entre iguales para prevenir conductas de acoso escolar en los centros de secundaria. *Infancia y Aprendizaje, 39*(3), 499 – 527. https://doi.org/10.1080/02103702.2016.1189122

Wang, C., & Goldberg, T. S. (2017). Using children's literature to decrease moral disengagement and victimization among elementary school students. *Psychology in the Schools, 54,* 918 – 931. doi: 10.1002/pits.22042

Wexler, L., Poudel-Tandukar, K., Rataj, S., Trout, L., Poudel, K. C., Woods, M., & Chachamovich, E. (2017). Preliminary evaluation of a school-based youth leadership and prevention program in rural Alaska native communities. *School Mental Health, 9,* 172 – 183. https://doi.org/10.1007/s12310-016-9203-2

Williford, A., Boulton, A., Noland, B., Little, T. D., Kärnä, A., & Salmivalli, C. (2012). Effects of the KiVa anti-bullying program on adolescents' depression, anxiety, and perception of peers. *Journal of Abnormal Psychology, 40,* 289 – 300. https://doi.org/10.1007/s10802-011-9551-1

Ybarra, M. L., Prescott, T. L., & Espelage, D. L. (2016). Stepwise development of a text messaging-based bullying prevention program for middle school students

(BullyDown). *JMIR mHealth and uHealth, 4*(2), e60.

https://doi.org/10.2196/mhealth.4936

12.5 Other references

Ackers, M. J. (2012). Cyberbullying: Through the eyes of children and young people. *Educational Psychology in Practice, 28*(2), 141 – 157. https://doi.org/10.1080/02667363.2012.665356

Allen, N., Grigsby, B., & Peters, M. L. (2015). Does leadership matter? Examining the relationship among transformational leadership, school climate, and student achievement. *International Journal of Educational Leadership Preparation, 10*(2), 1 - 22.

Allison, K. R., & Bussey, K. (2017). Individual and collective moral influences on intervention in cyberbullying. *Computers in Human Behavior, 74,* 7 – 15. https://doi.org/10.1016/j.chb.2017.04.019

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders,* 4th edition. Author.

Ang, R. P. (2015). Adolescent cyberbullying: A review of characteristics, prevention and intervention strategies. *Aggression and Violent Behavior, 25,* 35 – 42. https://doi.org/10.1016/j.avb.2015.07.011

Armitage, P. (1985). Statistical considerations: Conclusion. In N.J. Wald & R. Doll (Eds.), *Interpretation of negative epidemiological evidence for carcinogenicity.* International Agency for Research on Cancer.

Arsenault, L., Bowes, L., & Shakoor, S. (2010). Bullying victimization in youths and mental health problems: 'Much ado about nothing'?. *Psychological Medicine*, *40*(5), 717-729. https://doi.org/10.1017/s0033291709991382

Asam, A. E., & Samara, M. (2016). Cyberbullying and the law: A review of psychological and legal challenges. *Computers in Human Behavior, 65,* 127 – 141. https://doi.org/10.1016/j.chb.2016.08.012

Ashktorab, Z., & Vitak, J. (2016). *Designing cyberbullying mitigation and prevention solutions through participatory design with teenagers*. Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1142/2858036.2858548

Attar-Schwartz, S., Mishna, F., & Khoury-Kassabri, M. (2019). The role of classmates' social support, peer victimization and gender in externalizing and internalizing behaviors among Canadian youth. *Journal of Child and Family Studies, 28*(9), 2335 – 2346, https://doi.org/10.1007/s10826-017-0852-z

Attrill, A. (2015). Age versus goal-directed Internet use. In A. Attrill (Ed.), *Cyberpsychology,* (pp. 29 – 48). Oxford University Press.

Ayieko, J., Abuogi, L., Simchowitz, B., Bukusi, E. A., Smith, A. H., & Reingold, A. (2014). Efficacy of isoniazid prophylactic therapy in prevention of tuberculosis in children: A meta-analysis. *BMC Infectious Diseases, 14*(1), 91. https://doi.org/10.1186/1471-2334-14-91

Baier, D., Sung Hong, J., Kilem, S., & Bergmann, M. C. (2019). Consequences of bullying on adolescents' mental health in Germany: Comparing face-to-face bullying and cyberbullying. *Journal of Child and Family Studies, 28,* 2347 – 2357. https://doi.org/10.1007/s10826-018-1181-6

Baldry, A., Blaya, C., & Farrington, D.P. (2018). *International perspectives on cyberbullying.* Palgrave McMillan.

Baldry, A. C., Farrington, D. P., & Sorrentino, A. (2015). "Am I at risk of cyberbullying?" A narrative review and conceptual framework for research on risk of cyberbullying and cybervictimization: The risk and needs assessment approach. *Aggression and Violent Behavior, 23,* 36 – 51. https://doi.org/10.1016/j.avb.2015.05.014

Baldry, A. C., Farrington, D. P., & Sorrentino, A. (2017). School bullying and cyberbullying among boys and girls: Roles and overlap. *Journal of Aggression, Maltreatment and Trauma, 26*(9), 937 – 951. https://doi.org/10.1080/10926771.2017.1330793

Baldry, A. C., Sorrentino, A., & Farrington, D. P. (2019). Post-traumatic stress symptoms among Italian preadolescents involved in school and cyber bullying and victimization. *Journal of Child and Family Studies, 28*(9), 2358 – 2364. https://doi.org/10.1007/s10826-018-1122-4

Barboza, G. E., Schiamberg, L. B., Oehmke, J., Korzeniewski, S. J., Post, L. A., & Heraux, C. G. (2009). Individual characteristics and the multiple contexts of adolescent bullying: An ecological perspective. *Journal of Youth and Adolescence, 38,* 101 – 121. https://doi.org/10.1007/s10964-008-9271-1

Barlett, C. P. (2017). From theory to practice: Cyberbullying theory and its application to intervention. *Computers in Human Behavior, 72,* 269 – 275. https://doi.org/10.1016/j.chb.2017.02.060

Bauman, S. (2013). Why it matters. In S. Bauman, D. Cross, & J. Walker (Eds.), *Principles of Cyberbullying Research: Definitions, Measures, and Methodology,* pp. 23 –26. Routledge.

BBC (Producer). (2017). *Love Island's Olivia: I've received death threats.* http://bbcnews.co.uk

Besemer, S., Ahmad, S. I., Hinshaw, S. P., & Farrington, D. P. (2017). A systematic review and meta-analysis of the intergenerational transmission of criminal behavior, *Aggression and Violent Behavior, 37,* 161 – 178. https://doi.org/10.1016/j,avb.2017.10.004

Bevilacqua, L., Shackleton, N., Hale, D., Allen, E., Bond, L., Christie, D., Elbourne, D., Fitzgerald-Yau, N., Fletcher, A., Jones, R., Miners, A., Scott, S., Wiggins, M., Bonell, C., & Viner, R.M. (2017). The role of family and school-level factors in bullying and cyberbullying: A cross-sectional study. *BMC Pediatrics, 17,* 160. https://doi.org/10.1186/s12887-017-0907-8

Betts, L.R. (2016). Cyberbullying: Approaches, consequences, and interventions. In J. Binder (Ed.), *Palgrave studies in Cyberpsychology,* Palgrave Macmillan.

Betts, L. R., & Spenser, K. A. (2017). "People think it's a harmless joke": Young people's understanding of the impact of technology, digital vulnerability and cyberbullying in the United Kingdom, *Journal of Children and Media, 11*(1), 20 – 35. https://doi.org/10.1080/17482798.2016.1233893

Bhatia, R., Lopipero, P., & Smith, A. H. (1998). Diesel exhaust exposure and lung cancer, *Epidemiology, 9,* 84 – 91.

Birkeland Nielsen, M., Tangen, T., Idsoe, T., Berge Matthiesen, S., & Magerøy, N. (2015). Post-traumatic stress disorder as a consequence of bullying at work and at school. A literature review and meta-analysis. *Aggression and Violent Behavior, 21,* 17 – 24. https://doi.org/10.1016/j.avb.2015.01.001

Bjornebekk, G., & Howard, R. (2012). Validation of a motivation-based typology of angry aggression among antisocial youths in Norway. *Behavioral Sciences and the Law, 30,* 167 – 180. https://doi.org/10.1002/bsl.2007

Bloom, H. S., Hill, C. J., & Riccio, J. A. (2003). Linking program implementation and effectiveness: Lessons from a pooled sample of welfare-to-work experiments. *Journal of Policy Analysis and Management*, *22*(4), 551-575. https://doi.org/10.1002/9781444307399.ch26

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

Boulton, M. J., Hardcastle, K., Down, J., Fowles, J., & Simmonds, J. A. (2014). A comparison of preservice teachers' responses to cyber versus traditional bullying scenarios: Similarities and differences and implications for practice. *Journal of Teacher Education, 65*(2), 145 – 155. https://doi.org/10.1177/0022487113511496

Bowes, L., Arseneault, L., Maughan, B., Taylor, A., Caspi, A., & Moffitt, T. E. (2009). School, neighbourhood, and family factors are associated with children's bullying involvement: A nationally representative longitudinal study. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(5), 545 – 553. https://doi.org/10.1097/CHI.0b013e31819cb017

Brendgen, M., & Poulin, F. (2018). Continued bullying victimization from childhood to young adulthood: A longitudinal study of mediating and protective factors. *Journal of Abnormal Child Psychology, 46,* 27 – 39. https://doi.org/10.1007/s10802-017-0314-5

Brewer, G., Kerslake, J. (2015). Cyberbullying, self-esteem, empathy and loneliness. *Computers in Human Behavior, 48,* 255 – 260. https://doi.org/10.1016/j.chb.2015.01.073

Brighi, A., Ortega, R., Pyzalski, J., Scheithauer, H., Smith, P. K., & Tsormpatzoudis, J. (2012). *European cyberbullying intervention project questionnaire – ECIPQ,* https://www.bullyingandcyber.net

Bronfenbrenner, U. (1979). *Ecology of human development: Experiments by nature and design.* Harvard University Press.

Brown, C. F., Demaray, M. K., & Secord, S. M. (2014). Cyber victimization in middle school and relation to social emotional outcomes. *Computers in Human Behavior, 35,* 12 – 21. https://doi.org/10.1016/j.chb.2014.02.014

Callaghan, M., Kelly, C., & Molcho, M. (2015). Exploring traditional and cyberbullying among Irish adolescents. *International Journal of Public Health, 60,* 199 – 206. https://doi.org/10.1007/s00038-014-0638-7

Calvete, E., Orue, I., & Gámez-Guadix, M. (2016). Cyberbullying victimization and depression in adolescents: The mediating role of body image and cognitive schemas in a one-year prospective study. *European Journal on Criminal Policy and Research, 22,* 271 -284. https://doi.org/10.1007/s10610-015-9292-8

Cantone, E., Piras, A. P., Vellante, M., Preti, A., Daníelsdóttir, S., D'Aloja, E., Lesinskiene, S., Angermeyer, M. C., Carta, M. G., & Bhugra, D. (2015). Interventions on bullying and cyberbullying in schools: A systematic review. *Clinical Practice and Epidemiology in Mental Health, 11*(Suppl 1: M4), 58 – 76. https://doi.org/10.2174/1745017901511010058

Carlos-Wallace, F. M., Zhang, L., Smith, M. T., Rader, G., & Steinmaus, C. (2016). Parental, in utero, and early-life exposure to benzene and the risk of childhood leukaemia: A meta-analysis. *American Journal of Epidemiology, 183*(1), 1 – 14. https://doi.org/10.1093/aje/kwv120

Castellví, P., Miranda-Mendizábal, A., Parés-Badell, O., Almenara, J., Alonso, I., Blasco, M.J., Cebrià, A., Gabilondo, A., Gili, M., Lagares, C., Piqueras, J. A., Roca, M., Rodríguez-Marín, J., Rodríguez-Jimenez, T., Soto-Sanz, V., & Alsono, J. (2016). Exposure to violence, a risk for suicide in youths and young adults. A meta-analysis of longitudinal studies. *Acta Psychiatrica Scandinavica,* 1 – 17. https://doi.org/10.1111/acps.12679

Centers for Disease Control and Prevention [CDC] (2014). *Bullying surveillance among school- aged children: Uniform definitions and recommended data elements.* Author.

Chaffee, B. W., & King, J. C. (2012). Effect of zinc supplementation on pregnancy and infant outcomes: A systematic review. *Paediatric and Perinatal Epidemiology, 26*(Suppl 1), 118 – 137. https://doi.org/10.1111/j.1365-3016.2012.01289.x

Chaffee, B. W., & Weston, S. J. (2010). Association between chronic periodontal disease and obesity: A systematic review and meta-analysis. *Journal of Periodontology, 81*(12), 1708 – 1724. https://doi.org/10.1902/jop.2010.100321

Chalamandaris, A., & Piette, D. (2015). School-based anti-bullying interventions: Systematic review of the methodology to assess their effectiveness. *Aggression and Violent Behavior, 24,* 131 – 174. https://doi.org/10.1016/j.avb.2015.04.004

Chan, H. C., & Wong, D. S. W. (2017). Coping with cyberbullying victimization: An exploratory study of Chinese adolescents in Hong Kong, *International Journal of Law, Crime and Justice, 50,* 71 – 82. https://doi.org/10.1016/j.ijlcj.2017.04.003

Chen, L., Ho, S. S., & Lwin, M. O. (2017). A meta-analysis of factors predicting cyberbullying perpetration and victimization: From the social cognitive and media effects approach. *New Media and Society, 19*(8), 1194 – 1213. https://doi.org/10.1177/1461444816634037

Chisholm, J. F. (2014). Review of the status of cyberbullying and cyberbullying prevention. *Journal of Information Systems Education, 25*(1), 77 – 87.

Clarke, P., Crawford, C., Steele, F., & Vignoles, A. (2013). Revisiting fixed- and random-effects models: Some considerations for policy relevant education research. *Education Economics, 23*(3), 259 – 277. https://doi.org/10.1080/09645292.2013.855705

Cleemput, K. van, DeSmet, A., Vandebosch, H., Bastianensens, S., Poels, K., & De Bourdeaudhuji, I. (2014). *A systematic review and meta-analysis of the efficacy of* *cyberbullying prevention programs.* Paper presented at 'Etmaal van de Communicatiewetenschap', Wageningen, Netherlands.

Cohen, J., & Freiberg, J. A. (2013). *School climate and bullying prevention. School climate practice brief.* National School Climate Center.

Cooke, B. J., & Farrington, D. P. (2016). The effectiveness of dog-training programs in prison: A systematic review and meta-analysis of the literature. *The Prison Journal, 96*(6), 854 – 876. https://doi.org/10.1177/0032885516671919

Copeland, W. E., Wolke, D., Angold, A., & Costello, E. J. (2013). Psychiatric outcomes of bullying and being bullied by peers in childhood and adolescence. *JAMA Psychiatry, 70*(4), 419 – 426. https://doi.org/10.1001/jamapsychiatry.2013.504

Cowie, H., Bauman, S., Coyne, I., Myers, C., Pörhöla, M., & Almeida, A. (2013). Cyberbullying amongst university students. In P.K. Smith & G. Steffgen (Eds.), *Cyberbullying through the new media: Findings from an international network,* 165 – 186. Psychology Press, Taylor & Francis Group.

Corcoran, L., Connolly, I., & O'Moore, M. (2012). Cyberbullying in Irish schools: An investigation of personality and self-concept. *Irish Journal of Psychology, 33*(4), 153 – 165. https://doi.org/10.1080/03033910.2012.67799

Cotter, P., & McGilloway, S. (2011). Living in an 'electronic age': Cyberbullying among Irish adolescents. *Irish Journal of Education, 39,* 44 – 56. https://www.jstor.org/stable/41548683

Cox, E., Leung, R., Baksheev, G., Day, A., Toumbourou, J. W., Miller, P., Kremer, P., & Walker, A. (2016). Violence prevention and intervention programmes for adolescents in Australia: A systematic review. *Australian Psychologist, 51,* 206 – 222. https://doi.org/10.1111/ap.12168

Criado Perez, C. (2019). *Invisible Women: Exposing data bias in a world designed for men.* Chatto & Windus.

Cross, D., Barnes, A., Papageorgiou, A., Hadwen, K., Hearn, L., & Lester, L. (2015). A social-ecological framework for understanding and reducing cyberbullying behaviours. *Aggression and Violent Behavior, 23,* 109 – 117. https://doi.org/10.1016/j.avb.2015.05.016

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science, 25*(1), 7 – 29. https://doi.org/10.1177/0956797613504966

Davis, J. P., Merrin, G. J., Ingram, K. M., Espelage, D. L., Valido, A., & El Sheikh, A. J. (2019). Examining pathways between bully victimization, depression, & school belonging among early adolescents. *Journal of Child and Family Studies, 28*(9), 2365 – 2378, https://doi.org/10.1007/s10826-019-01340-9

deLara, E. W. (2019). Consequences of childhood bullying on mental health and relationships for young adults. *Journal of Child and Family Studies, 28*(9), 2379 – 2389, https://doi.org/10.1007/s10826-018-1197-y

Del Rey, R., Casas, J. A., Ortega-Ruiz, R., Schultze-Krumbholz, A., Scheithauer, H., Smith, P., Thompson, F., Barkoukis, V., Tsorbatzoudis, H., Brighi, A., Guarini, A., Pyżalski, J., & Plichta, P. (2015). Structural validation and cross-cultural robustness of the European cyberbullying intervention project questionnaire. *Computers in Human Behavior, 50,* 141 – 147. https://doi.org/10.1016/j.chb.2015.03.065

Della Cioppa, V., O'Neil, A., & Craig, W. (2015). Learning from traditional bullying interventions: A review of research on cyberbullying and best practice. *Aggression and Violent Behavior, 23,* 61 – 68. https://doi.org/10.1016/j.avb.2015.05.009

Devine, P., & Lloyd, K. (2012). Internet use and psychological well-being among 10-year-old and 11-year-old children. *Child Care in Practice, 18*(1), 5 – 22. https://doi.org/10.1080/13575279.2011.621888

Dodd, V. (2017, August 21). CPS to crack down on social media hate crime, says Alison Saunders. *The Guardian,* http://www.theguardian.com

Donner, A., & Klar, N. (2002). Issues in meta-analysis of cluster randomized trials. *Statistics in Medicine, 21,* 2971 – 2980. https://doi.org/10.1002/sim.1301

Donner, A., Piaggio, G., & Villar, J. (2001). Statistical methods for the meta-analysis of cluster randomized trials. *Statistical Methods in Medical Research, 10,* 325 – 338. https://doi.org/10.1191/096228001680678322

Dorjee, K., Choden, T., Baxi, S. M., Steinmaus, C., & Reingold, A. L. (2018). Risk of cardiovascular disease associated with exposure to abacavir among individuals with HIV: A systematic review and meta-analyses of results from 17 epidemiologic studies. *International Journal of Antimicrobial Agents, 52,* 541 – 553. https://doi.org/10.1016/j.ijantimicag.2018.07.010

Duong, A., Steinmaus, C., McHale, C. M., Vaughan, C. P., & Zhang, L. (2011). Reproductive and developmental toxicity of formaldehyde: A systematic review. *Mutation Research/Reviews in Mutation Research, 728*(3), 118 – 138. https://doi.org/10.1016/j.mrrev.2011.07.003

Earnshaw, V. A., Reisner, S. L., Menino, D. D., Poteat, V. P., Bogart, L. M., Barnes, T. N., & Schuster, M. A. (2018). Stigma-based bullying interventions: A systematic review. *Developmental Review, 48,* 178 – 200. https://doi.org/10.1016/j.dr.2018.02.001

Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet, 337*(8746), 867 – 872. https://doi.org/10.1016/0140-6736(91)90201-Y

Eisner, M., & Humphreys, D. (2012). Measuring conflict of interest in prevention and intervention research: A feasibility study. In T. Bliesener, A. Beelmann, & M. Stemmler (Eds.), *Antisocial Behavior and Crime,* (pp. 165 – 180). Hogrefe.

Ellis, P. D. (2010). *The Essential Guide to Effect sizes: Statistical power, Meta-analysis, and the interpretation of research results.* Cambridge University Press.

Erren, T. C., Glende, C. B., Morfeld, P., & Piekarski, C. (2009). Is exposure to silica associated with lung cancer in the absence of silicosis? A meta-analytical approach to an important public health question. *International Archives of Occupational and Environmental Health, 82*(8), 997 – 1004. https://doi.org/10.1007/s00420-008-0381-0

Erreygers, S., Pabian, S., Vandebosch, H., & Baillien, E. (2016). Helping behavior among adolescent bystanders of cyberbullying: The role of impulsivity. *Learning and Individual Differences, 48,* 61 – 67. https://doi.org/10.1016/j.lindif.2016.03.003

Espelage, D.L., Hong, J.S., & Valido, A. (2019). Cyberbullying in the United States. In A. Baldry, C. Blaya, & D.P. Farrington (Eds). *International Perspectives on Cyberbullying,* (pp. 65 – 100). Palgrave McMillan: Palgrave Studies in Cybercrime and Cybersecurity Series.

Espelage, D. L., Polanin, J. R., & Low, S. K. (2014). Teacher and staff perceptions of school environment as predictors of student aggression, victimization, and willingness to intervene in bullying situations. *School Psychology Quarterly, 29*(3), 287 – 305. https://doi.org/10.1037/spq0000072

Evans, C. B. R., Fraser, M. W., & Cotter, K. L. (2014). The effectiveness of school-based bullying prevention programs: A systematic review. *Aggression and Violent Behavior, 19,* 532 – 544. https://doi.org/10.1016/j.avb.2014.07.004

Eysenck, H. J., & Eysenck, S. B. G. (1975). *Junior Eysenck Personality Questionnaire (JEPQ).* Hodder & Stoughton.

Fahimi, J., Singh, A., & Frazee, B. W. (2015). The role of adjunctive antibiotics in the treatment of skin and soft tissue abscesses: A systematic review and meta-analysis. *Canadian Journal of Emergency Medicine, 17*(4), 420 – 432. https://doi.org/10.1017/cem.2014.52

Farrington, D. P. (1983). Randomized experiments on crime and justice. *Crime and Justice, 4,* pp. 257 – 308. https://doi.org/10.1086/449091

Farrington, D. P. (1993). Understanding and preventing bullying. In M. Tonry (Ed.), *Crime and Justice, 17,* 381 – 458. https://www.jstor.org/stable/1147555

Farrington, D.P. (2003). Methodological quality standards for evaluation research. *The Annals of the American Academy of Political and Social Science*, 587 (1), 49 – 68.

Farrington, D. P., Gottfredson, D. C., Sherman, L. W., & Welsh, B. C. (2002). The Maryland Scientific Methods Scale. In L.W. Sherman, D. P. Farrington, B. C. Welsh, & D. Layton MacKenzie (Eds.), *Evidence-based crime prevention,* pp. 13 – 21. Routledge.

Farrington, D. P., & Loeber, R. (1989). Relative improvement over chance (RIOC) and phi as measures of predictive efficiency and strength of association in 2x2 tables. *Journal of Quantitative Criminology, 5,* 201 – 213. https://doi.org/10.1007/bf01062737

Farrington, D. P., Lösel, F., Braga, A. A., Mazerolle, L., Raine, A., Sherman, L. W., & Weisburd, D. (2020). Experimental criminology: Looking back and forward on the 20th anniversary of the Academy of Experimental Criminology, *Journal of Experimental Criminology, 16,* 649 – 673. https://doi.org/10.1007/s11292-019-09384-z

Farrington, D. P., Lösel, F., Ttofi, M. M., & Theodorakis, N. (2012). *School bullying, depression and offending behavior later in life: An updated systematic review of longitudinal studies.* Swedish National Council for Crime Prevention.

Farrington, D. P., & MacKenzie, D. L. (2013). Long-term follow-ups of experimental interventions. *Journal of Experimental Criminology, 9*(4), 385 – 388. https://doi.org/10.1007/s11292-013-9195-y

Farrington, D. P., & Petrosino, A. (2001). The Campbell Collaboration Crime and Justice Group. *ANNALS of the American Academy of Political and Social Science, 578*(1), 35 – 49. https://doi.org/10.1177/000271620157800103

Farrington, D. P., & Ttofi, M. M. (2009). School-based programs to reduce bullying and victimization, *Campbell Systematic Reviews, 6,* 1 – 148. https://doi.org/10.4073/csr.2009.6

Farrington, D. P., & Welsh, B. C. (2008). *Saving children from a life of crime: Early risk factors and effective interventions.* Oxford University Press.

Farrington, D. P., & Welsh, B. C. (2013). Measuring effect size in meta-analysis, with special reference to area-based crime prevention programs and the effects of close-circuit television on crime. In A. A. Kuhn, C. Schawarzenegger, P. Margot, A. Donatsch, M. Aebi, & D. Jositsch (Eds.), *Criminology, criminal policy and criminal law from an international perspective* (pp. 75 – 89). Stampfli.

Ferguson, C. J., San Miguel, C., Kilburn, J. C., & Sanchez, P. (2007). The effectiveness of school-based anti-bullying programs: A meta-analytic review. *Criminal Justice Review, 32,* 401 – 414. https://doi.org/10.1177/0734016807311712

Ferreira, P. C., Simão, A. M. V., Ferreira, A., Souza, S., & Francisco, S. (2016). Student bystander behavior and cultural issues in cyberbullying: When actions speak louder than words. *Computers in Human Behavior, 60,* 301 – 311. https://doi.org/10.1016/j.chb.2016.02.05

Fink, C. A., & Bates, M. N. (2005). Melanoma and ionizing radiation: Is there a causal relationship? *Radiation Research, 164,* 701 – 710. https://doi.org/10.1667/rr3447.1

Finkelhor, D., Mitchell, K. J., & Wolak, J. (2000). *Online victimization: A report on the nation's youth (NCMEC 6-00-020).* National Center for Missing & Exploited Children.

Flay, B. R., Biglan, A., Boruch, R. F., González Castro, F., Gottfredson, D., Kellam, S., Mościcki, E. K., Schinke, S., Valentine, J. C., & Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6,* 151 – 175. https://doi.org/10.1007/s11121-005-5553-y

Fletcher, A., Fitzgerald-Yau, N., Jones, R., Allen, E., Viner, R. M., & Bonell, C. (2014). Brief report: Cyberbullying perpetration and its associations with socio-demographics, aggressive behaviour at school, and mental health outcomes. *Journal of Adolescence, 37,* 1393 – 1398. https://doi.org/10.1016/j.adolescence.2014.10.005

Foody, M., Samara, M., & O'Higgins Norman, J. (2017). Bullying and cyberbullying studies in the school-aged population on the island of Ireland: A meta-analysis. *British Journal of Educational Psychology, 87,* 535 – 557. https://doi.org/10.1111/bjep.12163

Furlong, M. J., Sharkey, J.D., Felix, E.D., Tanigawa, D., & Greif Green, J. (2010). Bullying assessment: A call for increased precision of self-reporting procedures. In S.R. Jimerson, S.M. Swearer, & D.L. Espealage (Eds.), *Handbook of Bullying in Schools: An International Perspective,* (pp. 329 – 345). Routledge.

Gaertner, R. R. W., & Thériault, G. P. (2002). Risk of bladder cancer in foundry workers: A meta-analysis. *Occupational and Environmental Medicine, 59*(10), 655 – 663. https://doi.org/10.1136/oem.59.10.655

Gaffney, H., & Farrington, D. P. (2018). Cyberbullying in the United Kingdom and Ireland. In A. Baldry, C. Blaya, & D.P. Farrington (Eds), *International Perspectives on Cyberbullying,* (pp. 101 – 143). Palgrave McMillan: Palgrave Studies in Cybercrime and Cybersecurity Series.

Gaffney, H., & Farrington, D. P. (2020). Systematic reviews and meta-analyses of the effectiveness of school-based anti-bullying programmes: A review of reviews. In P. K. Smith & J. O'Higgins Norman (Eds.), *Handbook of Bullying, Volume 2: Development, coping and intervention.* Wiley-Blackwell, submitted for publication.

Gaffney, H., Farrington, D. P., Espelage, D. L., & Ttofi, M. M. (2019a). Are cyberbullying intervention and prevention programs effective? A systematic and meta-analytical review. *Aggression and Violent Behavior*, *45,* 134 – 153. https://doi.org/10.1016/j.avb.2018.07.002

Gaffney, H., Farrington, D. P., & Ttofi, M. M. (2019b). Examining the effectiveness of school-bullying intervention programs globally: A meta-analysis. *International Journal of Bullying Prevention, 1,* 14 – 31. https://doi.org/10.1007/s42380-019-0007-4

Gaffney, H., Ttofi, M. M., & Farrington, D. P. (2019c). Evaluating the effectiveness of school-bullying prevention programs: An updated meta-analytical review. *Aggression and Violent Behavior*, *45,* 111 – 133. https://doi.org/10.1016/j.avb.2018.07.001

Gaffney, H., Ttofi, M. M., & Farrington, D. P. (2020). What works in anti-bullying programs? Analysis of effective intervention components. *Journal of School Psychology,* accepted for publication.

Gage, N. A., Prykanowski, D. A., & Larson, A. (2014). School climate and bullying victimisation: A latent class growth model analysis. *School Psychology Quarterly, 29*(3), 256 – 271. https://doi.org/10.1037/spq0000064

Gámez-Guadix, M., Orue, I., Smith, P. K., & Calvete, E. (2013). Longitudinal and reciprocal relations of cyberbullying with depression, substance use, and problematic Internet use among adolescents. *Journal of Adolescent Health, 53,* 446 – 452. https://doi.org/10.1016/j.jadohealth.2013.03.030

Gardella, J. H., Fisher, B. W., & Teurbe-Tolon, A. R. (2017). A systematic review and meta-analysis of cyber-victimization and educational outcomes for adolescents. *Review of Educational Research, 87*(2), 283 – 308. https://doi.org/10.3102/0034654316689136

Gardner, F., Montgomery, R., & Knerr, W. (2016). Transporting evidence-based parenting programs for child problem behavior (age 3 – 10) between countries: Systematic review and meta-analysis. *Journal of Clinical Child & Adolescent Psychology, 45*(6), 749 – 762. https://doi.org/10.1080/15374416.2015.1015134

Gastic, B. (2008). School truancy and the disciplinary problems of bullying victims. *Educational Review, 60*(4), 391 – 404. https://doi.org/10.1080/00131910802393423

Geel, M. van, Goemans, A., & Vedder, P. H. (2016). The relation between peer-victimization and sleeping problems: A meta-analysis. *Sleep Medicine Reviews, 27,* 89 – 95. https://doi.org/10.1016/j.smrv.2015.05.004

Genta, M. L., Smith, P. K., Ortega, R., Brighi, A., Guarini, A., Thompson, F., Tippett, N., Mora-Merchán, J. A., & Calmaestra, J. (2012). Comparative aspects of cyberbullying in Italy, England, and Spain: Findings from a DAPHNE project. In Q. Li, D. Cross, & P. K. Smith, (Eds). *Cyberbullying in the global playground: Research from international perspectives* (pp. 15 – 31). Wiley-Blackwell.

Goodman, J., Madaris, J., Verity, K., & Hott, B. (2013). A synthesis of international school-based bullying interventions. *The Journal of Special Education Apprenticeship, 2*(2), Article 3. https://scholarworks.lib.csusb.edu/josea/vol2/iss2/3

Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews.* Sage.

Gradinger, P., Strohmeier, D., & Spiel, C. (2017). Parents' and teachers' opinions on bullying and cyberbullying prevention. *Zeitschrift für Psychologie, 225*(1), 76 – 84. https://doi.org/10.1027/2151-2604/a000278

Guo, S. (2016). A meta-analysis of the predictors of cyberbullying perpetration and victimization. *Psychology in the Schools, 53*(4), 432 – 453. https://doi.org/10.1002/pits.21914

Hase, C. N., Goldberg, S. B., Smith, D., Stuck, A., & Campain, J. (2015). Impacts of traditional bullying and cyberbullying on the mental health of middle school and high school students. *Psychology in the Schools, 52*(6), 607 – 617. https://doi.org/10.1002/pits

He, J., van de Vijver, F. J. R., Dominguez Espinosa, A., Abubakar, A., Dimitrova, R., Adams, B. G., Aydinli, A., Atitsogbe, K., Alonso-Arbiol, I., Bobowik, M., Fischer, R., Jordanov, V., Mastrotheodoros, S., Neto, F., Ponizovsky, Y. J., Reb, J., Sim, S., Sovet, L., Stefenel, D., … Villieux, A. (2015). Socially desirable responding: Enhancement and denial in 20 countries. *Cross-Cultural Research, 49*(3), 227 – 249. https://doi.org/10.1177/1069397114552781

Hedges, L. V. (1982a). Estimation and testing for differences in effect size: Comment on Hsu. *Psychological Bulletin, 91,* 391 – 393. https://doi.org/10.1037/0033-2909.91.3.691

Hedges, L.V. (1982b). Estimating effect size from a series of independent experiments. *Psychological Bulletin, 92,* 490 – 499. https://doi.org/10.1037/0033-2909.92.2.490

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Academic Press.

Henry, J. A., & Reingold, A. L. (2012). Prehospital trauma systems reduce mortality in developing countries: A systematic review and meta-analysis. *Journal of Trauma and Acute Care Surgery, 73*(1), 261 – 268. https://doi.org/10.1097/TA.0b013e31824bdele

Herrera-López, M., Casas, J. A., Romera, E. M., Ortega-Ruiz, R., & Del Rey, R. (2017). Validation of the European Cyberbullying Intervention Project Questionnaire for

Colombian adolescents. *Cyberpsychology, Behavior, and Social Networking, 20*(2), 117 – 125. https://doi.org/10.1089/cyber.2016.0414

Hickenbotham, A., Roorda, A., Steinmaus, C., & Glasser, A. (2012). Meta-analysis of sex differences in presbyopia. *Investigative Ophthalmology and Visual Science, 53*(6), 3215 – 3220. https://doi.org/10.1167/iovs.12-9791

Higgins, J. P. T., Deeks, J. J., & Altman, D. G. (2011). Special topics in statistics. In J.P.T. Higgins, & S. Green (Eds.), *Cochrane Handbook for Systematic Reviews of Interventions,* (pp. 481 – 530). John Wiley & Sons.

Higgins, J. P. T., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions.* John Wiley & Sons.

Hinduja, S., & Patchin, J. W. (2017). Cultivating youth resilience to prevent bullying and cyberbullying victimization. *Child Abuse and Neglect, 73,* 51 – 62. https://doi.org/10.1016/j.chiabu.2017.09.010

Holt, M. K., Vivolo-Kantor, A.M., Polanin, J.R., Holland, K.M., Degue, S., Matjasko, J.L., Wolfe, M., & Reid, G. (2015). Bullying and suicidal ideation and behaviors: A meta-analysis. *Pediatrics, 135*(2), e496 – e509. https://doi.org/10.1542/peds.2014-1864

Hong, J. S., & Espelage, D. L. (2012). A review of research on bullying and peer victimization in school: An ecological model. *Aggression and Violent Behavior, 17,* 311 – 322. https://doi.org/10.1016/j.avb.2012.03.003

Houchins, D. E., Peia Oakes, W., & Johnson, Z. G. (2016). Bullying and students with disabilities: A systematic literature review of intervention studies. *Remedial and Special Education, 37*(5), 259 – 273. https://doi.org/10.1177/0741932516648678

Howard, R. (2011). The quest for excitement: A missing link between personality disorder and violence? *Journal of Forensic Psychiatry and Psychology, 22,* 692. https://doi.org/10.1080/14789949.2011.617540

Hutson, E., Kelly, S., & Militello, L. K. (2017). Systematic review of cyberbullying interventions for youth and parents with implications for evidence-based practice. *Worldviews on Evidence-Based Nursing, 15*(1), 72 – 79. https://doi.org/10.1111/wvn.12257

Janssen, I., Craig, W. M., Boyce, W. F., & Pickett, W. (2004). Associations between overweight and obesity with bullying behaviors in school-aged children. *Pediatrics, 113,* 1187 – 1194. https://doi.org/10.1542/peds.113.5.1187

Jiménez Barbero, J. A., Ruiz-Hernández, J. A., Llor Esteban, B., & Pérez García, M. (2012). Effectiveness of antibullying school programmes: A systematic review by evidence levels. *Children and Youth Services Review, 34,* 1646 – 1658. https://doi.org/10.1016/j.childyouth.2012.04.025

Jiménez-Barbero, J. A., Ruiz-Hernández, J. A., Llor-Zaragoza, L., Pérez-García, M., & Llor-Esteban, B. (2016). Effectiveness of anti-bullying school programs: A meta-analysis. *Children and Youth Services Review, 61,* 165 – 175. https://doi.org/10.1016/j.childyouth.2015.12.015

John, A., Glendenning, A. C., Marchant, A., Montgomery, P., Stewart, A., Wood, S., Lloyd, K., & Hawton, K. (2018). Self-harm, suicidal behaviours, and cyberbullying in children and young people: Systematic review. *Journal of Medical Internet Research, 20*(4), e129. https://doi/.org/10.2196/jmir.9044

Kaminski, J. W., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology, 36,* 567 – 589. https://doi.org/10.1007/s10802-007-9201-9

Kowalski, R. M., & Limber, S. P. (2013). Psychological, physical, and academic correlates of cyberbullying and traditional bullying. *Journal of Adolescent Health, 53*(1), S13 – S20. https://doi.org/10.1016/j.jadohealth.2012.09.018

Kwan, M. L., Buffler, P. A., Abrams, B., & Kiley, V. A. (2004). Breastfeeding and the risk of childhood leukaemia: A meta-analysis. *Public Health Reports, 119*, 521 – 535. https://doi.org/10.1016/j.phr.2004.09.002

Langford, R., Bonell, C., Jones, H., Pouliou, T., Murphy, S., Waters, E., Komro, K., Gibbs, L., Magnus, D., & Campbell, R. (2015). The World Health Organization's Health Promoting schools framework: A Cochrane systematic review and meta-analysis. *BMC Public Health, 15,* 130. https://doi.org/10.1186/s12889-015-1360-y

Langos, C. (2012). Cyberbullying: The challenge to define. *Cyberpsychology, Behavior, and Social Networking, 15*(6), 285 – 289. https://doi.org/10.1089/cyber.2011.0588

Lasher, S., & Baker, C. (2015). *Bullying: Evidence from the Longitudinal Study of Young People in England 2, wave 2.* HM Department of Education. https://www.gov.uk/government/publications

Lee, M. -S., Zi-Pei, W., Svanström, L., & Dalal, K. (2013). Cyberbullying prevention: Intervention in Taiwan, *PLoS ONE, 8*(5), e64031. https://doi.org/10.1371/journal.pone.0064031

Lee, S., Kim, C-J., & Kim, D. H. (2015). A meta-analysis of the effect of school-based anti-bullying programs. *Journal of Child Health Care, 19*(2), 136 – 153. https://doi.org/10.1177/1367493515593581

Li, Q., Smith, P. K., & Cross, D. (2012). *Cyberbullying in the global playground: Research from international perspectives.* Wiley-Blackwell.

Lier, P. A. C. van, Vuijk, P., & Crijnen, A. A. M. (2005). Understanding mechanisms of change in the development of antisocial behavior: The impact of a universal intervention. *Journal of Abnormal Child Psychology, 33,* 521 – 535. https://doi.org/10.1007/s10802-005-6735-7

Lipsey, M. W. (2003). Those confounded moderators in meta-analysis: Good, bad, and ugly. *ANNALS of the American Academy of Political and Social Science, 587,* 69 – 81. https://doi.org/10.1177/0002716202250791

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

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

Liu, X., Huang, Y., & Liu, Y. (2018). Prevalence, distribution, and associated factors of suicide attempts in young adolescents: School-based data from 40 low-income and middle-income countries. *PLoS ONE, 13*(12), e0207823. https://doi.org/10.1371/journal.pone.0207823

Liu, R., Zhang, L., McHale, C. M., & Hammond, S. K. (2011). Paternal smoking and risk of childhood acute lymphoblastic leukaemia: Systematic review and meta-analysis. *Journal of Oncology,* https://doi.org/10.1155/2011/854584

Livingstone, S., Haddon, L., Görzig, A., & Ólafsson, K. (2011). *EU Kids Online: Final report.* EU Kids online, London School of Economics and Political Science.

Livingstone, S., & Smith, P. K. (2014). Annual research review: Harms experienced by child users of online and mobile technologies: The nature, prevalence and management of sexual and aggressive risks in the digital age. *Journal of Child Psychology and Psycyhiatry, 55*(6), 635 – 654. https://doi.org/10.1111/jcpp.12197

Machimbarrena, J. M., Alvarez-Bardon, A., Leon-Mejia, A., Gutierrez-Ortega, M., Casadiego-Cabrales, A., & Gonzalez-Cabrera, J. (2019). Loneliness and personality profiles involved in bullying victimization and aggressive behavior. *School Mental Health, 11*(4), 807 – 818. https://doi.org/10.1007/s12310-019-09328-y

Marczak, M., & Coyne, I. (2015). A focus on online bullying. In A. Attrill (Ed.), *Cyberpsychology* (pp. 145 – 163). Oxford University Press.

McAuley, L., Tugwell, P., & Moher, D. (2000). Does the inclusion of grey literature influence estimates of intervention effectiveness reported in meta-analysis? *The Lancet, 356,* (9237), 1228 – 1231. https://doi.org/10.1016/s0140-6736(00)02786-0

McClure Watters, R. S. (2011). *The nature and extent of pupil bullying in schools in the North of Ireland.* (Research Report no 56). Department of Education for Northern Ireland.

McGuckin, C., Cummins, P. K., & Lewis, C. A. (2010). f2f and cyberbullying among children in Northern Ireland: Data from the Kids Life and Times survey. *Psychology, Society, and Education, 2*(2), 83 – 96. https://doi.org/10.25115/psye.v2i2.437

Merrell, K. W., Isava, D. M., Gueldner, B. A., & Ross, S. W. (2008). How effective are school bullying intervention programs? A meta-analysis of intervention research. *School Psychology Quarterly, 23*(1), 26 – 42. https://doi.org/10.1037/1045-3830.23.1.26

Mishna, F., Cook, C., Saini, M., Wu, M-.J., & McFadden, R. (2011). Interventions to prevent and reduce cyber abuse of youth: A systematic review. *Research on Social Work Practice (21)*1, 5 – 14. https://doi.org/10.1177/1049731509351988

Modecki, K. L., Minchin, J., Harbaugh, A. G., Guerra, N. G., & Runions, K. C. (2014). Bullying prevalence across contexts: A meta-analysis measuring cyber and traditional bullying. *Journal of Adolescent Health, 55,* 602 – 611. https://doi.org/10.1016/j.adohealth.2014.06.007

Moher, D., Liberati, A., Tetzlaff, & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRIMSA statement. *Annals of Internal*

Medicine, 151(4), 264 – 269. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

Moore, A., & Enquobahrie, D. A. (2011). Paternal occupational exposure to pesticides and risk of neuroblastoma among children: A meta-analysis. *Cancer, Causes, & Control, 22,* 1529 – 1536. https://doi.org/10.1007/s10552-011-9829-1

Monks, C. P., Robinson, S., & Worlidge, P. (2012). The emergence of cyberbullying: A survey of primary school pupils' perceptions and experiences. *School Psychology International, 33*(5), 477 – 491. https://doi.org/10.1177/0143034312445242

Morita, Y. (1996). Bullying as a contemporary behaviour problem in the context of increasing 'social privatisation' in Japan. *International Bureau of Education, 26,* 311 – 329.

Morita, Y., Soeda, H., Soeda, K., & Taki, M. (1999). Japan. In P.K. Smith, Y. Morita, J. Junger-Tas, & D. Olweus (Eds.), *The nature of school bullying: A cross-national perspective,* pp. 309 – 323. Routledge.

Murray, D. M., & Blitstein, J. L. (2003). Methods to reduce the impact of intraclass correlation in group-randomized trials. *Evaluation Review, 27,* 79 – 103. https://doi.org/10.1177/0193841x02239019

Mynard, H., & Joseph, S. (1997). Bully/victim problems and their association with Eysenck's personality dimensions in 8 to 13-year-olds. *British Journal of Educational Psychology, 67*(1), 51 – 54. https://doi.org/10.1111/j.2044-8279.1997.tb01226.x

Nagin, D. S., & Sampson, R. J. (2019). The Real Gold Standard: Measuring counterfactual worlds that matter most to social science and policy. *Annual Review of Criminology, 2,* 123 – 145. https://doi.org/10.1146/annurev-criminol-011518-024838

Nocentini, A., Zambuto, V., & Menesini, E. (2015). Anti-bullying programs and Information and Communication technologies (ICTs): A systematic review. *Aggression and Violent Behavior, 23,* 52 – 60. https://doi.org/10.1016/j.avb.2015.05.012

Nuccitelli, M. (2012). Cyber bullying tactics. *Forensic Examiner, 21*(3), 24 – 27.

O'Higgins Norman, J., O'Moore, M., & McGuire, L. (2016). *Cyberbullying in Ireland: A survey of parents internet usage and knowledge.* ABC, Irish national Anti-bullying research and resource centre.

O'Moore, M. (2012). Cyber-bullying: The situation in Ireland. *Pastoral Care in Education, 30*(3), 209 – 223. https://doi.org/10.1080/02643944.2012.688065

Ofcom, (2017). *Communications Market Report.* United Kingdom: Ofcom. https://www.ofcom.org.uk/__data/assets/pdf_file/0017/105074/cmr-2017-uk.pdf

Olejnik, S., & Algina, J. (2000). Measures of effect size for comparative studies: Applications, interpretations, and limitations. *Contemporary Educational Psychology, 25*(3), 241 – 286. https://doi.org/10.1006/ceps.2000.1040

Oliver, C., & Candappa, M. (2003). *Tackling bullying: Listening to the views of children and young people.* (Research Report RR400). HM Department for Education and Skills.

Olweus, D. (1986). *The Olweus Bully/Victim Questionnaire.* University of Bergen.

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). Erlbaum.

Olweus, D. (1993). Bully/victim problems among school children: Long-term consequences and an effective intervention program. In S. Hodgins (Ed.), *Mental Disorder and Crime,* (pp. 317 – 349). Sage.

Olweus, D., & Limber, S. P. (2017). Some problems with cyberbullying research. *Current Opinion in Psychology, 19,* 139 – 143. https://doi.org/10.1016/j.copsyc.2017.04.012

Orue, I., & Calvete, E. (2019). Psychopathic traits and moral disengagement interact to predict bullying and cyberbullying among adolescents. *Journal of Interpersonal Violence, 34*(11), 2313 – 2332. https://doi.org/10.1177/0886260516660302

Ostrov, J. M., & Kamper, K. E. (2015). Future directions for research on the development of relational and physical peer victimization. *Journal of Clinical Child and Adolescent Psychology, 44,* 509 – 519. https://doi.org/10.1080/15374416.2015.1012733

Oxman, A.D., & Guyatt, G. H. (1992). A consumer's guide to subgroup analyses. *Annals of Internal Medicine, 116*(1), 78 – 84. https://doi.org/10.7326/0003-4819-116-1-78

Paul, S., Smith, P .K., & Blumberg, H. H. (2010). Addressing cyberbullying in school using the Quality Circle approach. *Australian Journal of Guidance and Counselling, 20*(2), 157 – 168. https://doi.org/10.1375/ajgc.20.2.157

Pearce, N., Cross, D., Monks, H., Waters, S., & Falconer, S. (2011). Current evidence of best practice in whole-school bullying intervention and its potential to inform cyberbullying interventions. *Australian Journal of Guidance and Counselling, 21*(1), 1 – 21. https://doi.org/10.1375/ajgc.21.1.1

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

Phillips, D., & Ertl, H. (Eds.)(2003). *Implementing European Union Education and Training Policy.* Kluwer Academic Publishers.

Piers, E. V., & Herzberg, D. S. (2002). *Piers-Harris 2: Piers-Harris children's self-concept scale second edition.* Western Psychological Services.

Plog, A. E., Epstein, L., Jens, K., & Porter, W. (2006). *Assessment of bullying: The Colorado school climate survey.* Unpublished manuscript, Creating Caring Communities.

Polanin, J. R., Espelage, D. L., & Pigott, D. L. (2012). A meta-analysis of school-based bullying prevention programs' effects on bystander intervention behavior. *School Psychology Review, 41*(1), 47 – 65. https://doi.org/10.1080/02796015.2012.12087375

Pornai, C. D., & Wood, J. (2010). Peer and cyber aggression in secondary school students: The role of moral disengagement, hostile attribution bias, and outcome expectancies. *Aggressive Behavior, 36,* 81 – 94. https://doi.org/10.1002/ab.20336

Portnoy, J., & Farrington, D. P. (2015). Resting heart rate and antisocial behavior: An updated systematic review and meta-analysis. *Aggression and Violent Behavior, 22,* 33 – 45. https://doi.org/10.1016/j.avb.2015.02.004

Priesman, E., Newman, R., & Ford, J. A. (2017). Bullying victimization, binge drinking, and marijuana use among adolescents: Results from the 2013 National Youth Risk Behavior Survey, *Journal of Psychoactive Drugs, 50*(2), 133 – 142. https://doi.org/10.1080.02791072.2017.1371362

Prochaska, J. O., Evers, K. E., Prochaska, J. M., van Marter, D., & Johnson, J. L. (2007). Efficacy and effectiveness trials: Examples from smoking cessation and bullying prevention. *Journal of Health Psychology, 12*(1), 170 – 178. https://doi.org/10.1177/1359105307071751

Purdy, N., & McGuckin, C. (2015). Cyberbullying, schools and the law: A comparative study in Northern Ireland and the Republic of Ireland. *Educational Research, 57*(4), 420 – 436. https://doi.org/10.1080/00131881.2015.1091203

Purdy, N., & Smith, P. K. (2016). A content analysis of school anti-bullying policies in Northern Ireland. *Educational Psychology in Practice, 32*(3), 281 – 295. https://doi.org/10.1080/02667363.2016.1161599

Purdy, N., & York, L. (2016). A critical investigation of the nature and extent of cyberbullying in two post-primary schools in Northern Ireland. *Pastoral Care in Education, 34*(1), 13 – 23. https://doi.org/10.1080/02643944.2015.1127989

Pyżalski, J. (2012). From cyberbullying to electronic aggression: Typology of the phenomenon. *Emotional and Behavioural Difficulties, 17*(3 – 4), 305- 317. https://doi.org/10.1080/13632752.2012.704319

Rawlings, J. R., & Stoddard, S. A. (2019). A critical review of anti-bullying programs in North American elementary schools. *Journal of School Health, 89,* 759 – 780. https://doi.org/10.1111/josh.12814

Riddell, J., Pepler, D., & Craig, W. (2018). Cyberbullying in Canada. In A. Baldry, C. Blaya, & D.P. Farrington (Eds). *International Perspectives on Cyberbullying,* (pp. 39 – 64). Palgrave McMillan.

Rigby, K. (1996). *Manual for the Peer Relations Questionnaire (the PRQ).* University of South Australia.

Rigby, K., & Johnson, B. (2006). Expressed readiness of Australian schoolchildren to act as bystanders in support of children who are being bullied. *Educational Psychology, 26*(3), 425 – 440. https://doi.org/10.1080/01443410500342047

Rivers, I., & Noret, N. (2010). "I h8 u": Findings from a five-year study of text and email bullying, *British Educational Research Journal, 36*(4), 643 – 671. https://doi.org/10.1080/01411920903071918

Rose, C. A., & Tynes, B. (2015). Longitudinal associations between cybervictimization and mental health among U.S. adolescents. *Journal of Adolescent Health, 57,* 305 – 312. https://doi.org/10.1016/j.jadohealth.2015.05.002

Rosenthal, R., & Rubin, D. B. (1982). Comparing effect sizes of independent studies. *Psychological Bulletin, 92,* 500 – 504. https://doi.org/10.1037/0033-2909.92.2.500

Runions, K. C., Bak, M., & Shaw, T. (2017). Disentangling functions of online aggression: The cyber-aggression typology questionnaire (CATQ). *Aggressive Behavior, 43,* 74 – 84. https://doi.org/10.1002/ab.21663

Runions, K.C., Salmivalli, C., Shaw, T., Burns, S., & Cross, D. (2018). Beyond the reactive-proactive dichotomy: Rage, revenge, reward, and recreational aggression predict early high school bully and bully/victim status. *Aggressive Behavior, 44,* 501 – 511. https://doi.org/10.1002/ab.21770

Saarento, S., Boulton, A. J., & Salmivalli, C. (2015). Reducing bullying and victimization: Student- and classroom-level mechanisms of change. *Journal of Abnormal Child Pscyhology, 43*(1), 61 – 76. https://doi.org/10.1007/s10802-013-9841-x

Sanders, C. E. (2019). Breakfast skipping, psychological distress, and involvement in bullying: Is there a connection. *International Journal of Bullying Prevention, 1*(2)*,* 147 – 157. https://doi.org/10.1007/s42380-019-00008-8

Salmivalli, C. (2010). Bullying and the peer group: A review. *Aggression and Violent Behavior, 15,* 112 – 120. https://doi.org/10.1016/j.avb.2009.08.007

Salmivalli, C., Lagerspetz, K., Björkqvist, K., Österman, K., & Kaukiainen, A. (1996). Bullying as a group process: Participant roles and their relations to social status within the group. *Aggressive Behavior, 22,* 1 – 15. https://doi.org/10.4324/9781351153683-13

Sampson, R. J. (1986). Crime in cities: The effects of formal and informal social control. *Crime and Justice, 8,* 271 – 311. https://doi.org/10.1086/449125

Sapouna, M., Enz, S., & Samaras, M., & Wolke, D. (2015). Learning how to cope with cyberbullying in the virtual world. In T. Völlink, F. Dehue, & C. McGuckin (Eds.), *Cyberbullying: From theory to intervention,* (pp. 79 – 92). Routledge.

Schultze-Krumbholz, A., Göbel, K., Scheithauer, H., Brighi, A., Guarini, A., Tsorbatzoudis, H., Barkoukis, V., Pyżalski, J., Plichta, P., Del Rey, R., Casas, J. A., Thompson, F., & Smith, P. K. (2015). A comparison of classification approaches for cyberbullying and traditional bullying using data from six European countries. *Journal of School Violence, 14*(1), 47 – 65. https://doi.org/10.1080/15388220.2014.961067

Schwartz, D., Dodge, K. A., Coie, J. D., Hubbard, J. A., Cillessen, A. H., Lemerise, E. A., & Bateman, H. (1998). Social-cognitive and behavioral correlates of aggression and victimization in boy's play groups. *Journal of Abnormal Child Psychology, 26,* 431 – 440. https://doi.org/10.1023/a:1022695901088

Schwilk, E., Zhang, L., Smith, M. T., Smith, A. H., & Steinmaus, C. (2010). Formaldehyde and leukaemia: An updated meta-analysis and evaluation of bias. *Journal of Occupational and Environmental Medicine, 52*(9), 878 – 886. https://doi.org/10.1097/JOM.0b013e3181ef7e31

Semprevivo, L. K., Agnich, L. E., & Peguero, A. A. (2020). Is victimization associated with carrying a weapon? Investigating the intersection of sex and race/ethnicity. *Journal of School Violence, 19*(1), 20 – 34. https://doi.org/10.1080/15388220.2019.1703714

Setia, M. S., Steinmaus, C., Ho, C. S., & Rutherford, G. W. (2006). The role of BCG in prevention of leprosy: A meta-analysis, *Lancet Infect, 6,* 162 – 170. https://doi.org/10.1016/s1473-3099(06)70412-1

Shore, R. E., Gardner, M. J., & Pannett, B. (1993). Ethylene oxide: An assessment of the epidemiological evidence on carcinogenicity. *British Journal of Industrial Medicine, 50,* 971 – 997. https://doi.org/10.1136/oem.50.11.971

Silva, J. L. da, Oliveira, W. A. de, Mello, F. C. M. de, Andrade, L. S. de, Bazon, M. R., & Silva, M. A. I. (2017). Anti-bullying interventions in schools: A systematic literature review. *Ciência & Saúde Coletiva, 22*(7), 2329 – 2340. https://doi.org/10.1590/1413-81232017227.16242015

Silver, E., & Miller, L. L. (2004). Sources of informal social control in Chicago neighborhoods. *Criminology, 42*(3), 551 – 584. https://doi.org/10.1111/j.1745-9125.2004.tb00529.x

Singh, M. M., Ng, P. J., Yap, K. M., Husin, M. H., & Malim, N. H. A. H. (2017). Cyberbullying and a mobile game app? An initial perspective on an alternative solution. *Journal of Information Processing Systems, 13*(3), 559 – 572. https://doi.org/10.3745/JIPS.04.0033

Sivaraman, B., Nye, E., & Bowes, L. (2019). School-based anti-bullying interventions for adolescents in low- and middle-income countries: A systematic review. *Aggression and Violent Behavior, 45,* 154 – 162. https://doi.org/10.1016/j.avb.2018.07.007

Slee, P. T., & Rigby, K. (1993). The relationship of Eysenck personality-factors and self-esteem to bully victim behavior in Australian schoolboys. *Personality and Individual Differences, 14*(2), 371 – 373. https://doi.org/10.1016/0191-8869(93)90136-Q

Slonje, R., Smith, P. K., & Frisén, A. (2013). The nature of cyberbullying, and strategies for prevention. *Computers in Human Behavior, 2,* 26 – 32. https://doi.org/10.1016/j.chb.2012.05.024

Smith, P. K., Ananiadou, K., & Cowie, H. (2003). Interventions to reduce school bullying. *The Canadian Journal of Psychiatry*, *48*(9), 591-599. https://doi.org/10.3390/books978-3-03921-360-3

Smith, P. K., Barrio, C. del, & Tokunaga, R. S. (2013). Definitions of bullying and cyberbullying: How useful are the terms? In S. Bauman, D. Cross, & J. Walker (Eds.), *Principles of Cyberbullying Research: Definitions, Measures, and Methodology,* (pp. 26 – 40). Routledge.

Smith, P. K., & Berkkun, F. (2017). How research on cyberbullying has developed. In C. McGuckin & L. Corcoran (Eds.), *Bullying and cyberbullying: Prevalence, psychological impacts and intervention strategies.* Nova Science.

Smith, P. K., Cowie, H., Olafsson, R. F., & Liefooghe, A. P. (2002). Definitions of bullying: A comparison of terms used, and age and gender differences, in a fourteen–country international comparison. *Child Development*, *73*(4), 1119-1133. https://doi.org/10.1111/1467-8624.00461

Smith, P.K., Kwak, K., & Toda, Y. (2016). *School bullying in different cultures: Eastern and Western perspectives,* Cambridge, UK: Cambridge University Press.

Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry, 49*(4), 376 – 385. https://doi.org/10.1111/j.1469-7610.2007.01846.x

Smith, P. K., Salmivalli, C., & Cowie, H. (2012). Effectiveness of school-based programs to reduce bullying: A commentary. *Journal of Experimental Criminology, 8,* 433 – 441. https://doi.org/10.1007/s11292-012-9142 -3

Smith, J. D., Schneider, B. H., Smith, P. K., & Ananiadou, K. (2004). The effectiveness of whole-school antibullying programs: A synthesis of evaluation research. *School Psychology Review*, 33, 547–560. https://doi.org/10.1080/02796015.2004.12086267

Solberg, M. E., & Olweus, D. (2003). Prevalence estimation of school bullying with the Olweus Bully/Victim Questionnaire. *Aggressive Behavior*, *29*(3), 239-268. https://doi.org/10.1002/ab.10047

Sourander, A., Klomek, A. B., Ikonen, M., Lindroos, J., Luntamo, T., Koskelainen, M., Ristkari, T., & Helenius, H. (2010). Psychological risk factors associated with cyberbullying among adolescents: A population-based study. *Archives of General Psychiatry, 67*(7), 720 – 728. https://doi.org/10.1001/archgenpsychiatry.2010.79

Ssekitoleko, R., Kamya, M. R., & Reingold, A. L. (2013). Primary prophylaxis for cryptococcal meningitis and impact on mortality in HIV: A systematic review and meta-analysis. *Future Virology, 8*(9), https://doi.org/10.2217/fvl.13.71

Steinmaus, C. M., Nuñez, S., & Smith, A. H. (2000). Diet and bladder cancer: A meta-analysis of six dietary variables. *American Journal of Epidemiology, 151,* 693 – 702. https://doi.org/10.1093/oxfordjournals.aje.a010264

Steinmaus, C., Smith, A. H., Jones, R. M., & Smith, M. T. (2008). Meta-analysis of benzene exposure and non-Hodgkin lymphoma: Biases could mask an important association. *Occupational & Environmental Medicine, 65*(6), 371 – 378. https://doi.org/10.1136/oem.2007.036913

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. https://doi.org/10.1348/000709900158056

Strøm, I. F., Thoresen, S., Wentzel-Larsen, T., & Dyb, G. (2013). Violence, bullying and academic achievement: A study of 15-year-old adolescents and their school environment. *Child Abuse & Neglect, 37*(4), 243 – 251. https://doi.org/10.1016/j.chiabu.2012.10.010

Sullivan, G. M., & Feinn, R. (2012). Using effect size – or why the *p* value is not enough. *Journal of Graduate Medical Education, 4*(3), 279 – 282. https://doi.org/10.4300/JGME-D-12-00156.1

Swearer, S., Siebecker, A. B., Johnsen-Frerichs, L. A., & Wang, C. (2010). Assessment of bullying/victimization: The problem of comparability across studies and methodologies. In S. R. Jimerson, S. M. Swearer, & D. L. Espelage (Eds.), *Handbook of Bullying in Schools: An International Perspective* (pp. 305–327). Routledge.

Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life Course Criminology, 2,* 85 – 112. https://doi.org/10.1007/s40865-016-0026-5

Tanrikulu, I. (2018). Cyberbullying prevention and intervention programs: A systematic review. *School Psychology International, 39*(1), 74 – 91. https://doi.org/10.1177/0143034317745721

Thompson, F., & Smith, P. K. (2012). Anti-bullying strategies in schools: What is done and what works. *British Journal of Educational Psychology, 11*(9), 154 – 173.

Tippett, N., & Wolke, D. (2014). Socioeconomic status and bullying: A meta-analysis. *American Journal of Public Health, 104*(6), E48 – E59. https://doi.org/10.2105/AJPH.2014.301960

Toda, Y. (2016). Bullying (*ijime*) and related problems in Japan: History and research. In P. K. Smith, K. Kwak, & Y. Toda (Eds.), *School Bullying in Different Cultures: Eastern and Western Perspectives,* (pp. 73 – 92). Cambridge University Press.

Totten, M., Quigley, P., & Morgan, M. (2004). *Safe schools study.* Department of Justice Canada and Canadian Public Health Association. https://tottenandassociates.ca/wp-content/uploads/2008/02/safeschlstudyresearch-report-jan-05-totten.pdf

Ttofi, M. M. (2015). Adolescent bullying linked to depression in early adulthood: Evidence supports early intervention. *British Medical Journal*, *350*, h2694. https://doi.org/10.1136/bmj.h2694

Ttofi, M. M., Eisner, M., & Bradshaw, C. P. (2014). Bullying prevention: Assessing existing meta-evaluations. In G. Bruinsma & D. Weisburd (Eds), *Encyclopedia of*

Criminology and Criminal Justice. New York: Springer. https://doi.org/10.1007/978-1-4614-5690-2

Ttofi, M. M. & Farrington, D. P. (2011). Effectiveness of school-based programs to reduce bullying: A systematic and meta-analytic review. *Journal of Experimental Criminology*, 7, 27 – 56. https://doi.org/10.1007/s11292-010-9109-1

Ttofi, M.M., Farrington, D.P., & Lösel, F. (2012). School bullying as a predictor of violence later in life: A systematic review and meta-analysis of prospective longitudinal studies. *Aggression and Violent Behavior, 17,* 405 – 418. https://doi.org/10.1016/j.avb.2012.05.002

Ttofi, M. M., Farrington, D. P., Lösel, F., & Loeber, R. (2011a). Do the victims of school bullies tend to become depressed later in life? A systematic review and meta-analysis of longitudinal studies. *Journal of Aggression, Conflict, and Peace Research, 3*(2), 63 – 73. https://doi.org/10.1108.17596591111132873

Ttofi, M. M., Farrington, D. P., Lösel, F., & Loeber, R. (2011b). The predictive efficiency of school bullying versus later offending: A systematic/meta-analytic review of longitudinal studies. *Criminal Behaviour and Mental Health, 21,* 80 – 89. https://doi.org/10.1002/cbm.808

Ttofi, M. M., Farrington, D. P., Piquero, A. R., Lösel, F., DeLisi, M., & Murray, J. (2016). Intelligence as a protective factor against offending: A meta-analytic review of prospective longitudinal studies. *Journal of Criminal Justice, 45,* 4 – 18. https://doi.or/10.1016/j.jcrimjus.2016.02.003

United National Educational, Scientific, and Cultural Organisation (2019). *School violence and bullying: Global status and trends, drivers and consequences.* UNESCO.

Valdebenito, S., Eisner, M., Farrington, D. P., Ttofi, M. M., & Sutherland, A. (2018). School-based interventions for reducing disciplinary school exclusion: A systematic review. *Campbell Systematic Reviews* (1). https://doi.org/10.4073/csr.2018.1

Valdebenito, S., Ttofi, M., & Eisner, M. (2015). Prevalence rates of drug use among school bullies and victims: A systematic review and meta-analysis of cross-sectional studies. *Aggression and Violent Behavior, 23,* 137 – 146. https://doi.org/10.1016/j.avb.2015.05.004

Valdebenito, S., Ttofi, M. M., Eisner, M., & Gaffney, H. (2017). Weapon carrying in and out of school among pure bullies, pure victims and bully-victims: A systematic review and meta-analysis of cross-sectional and longitudinal studies. *Aggression and Violent Behavior*, *33*, 62 - 77. https://doi.org/10.1016/j.avb.2017.01.004

Verseveld, M. D. A. van, Fukkink, R. G., Fekkes, M., & Oostdam, R. J. (2019). Effects of antibullying programs on teachers' interventions in bullying situations. A meta-analysis. *Psychology in the Schools.* https://doi.org/10.1002/pits.22283

Vinnikov, D., Blanc, P. D., & Steinmaus, C. (2016). Is smoking a predictor for acute mountain sickness? Findings from a meta-analysis. *Nicotine & Tobacco Research, 18*(6), 1509 – 1516. https://doi.org/10.1093/ntr/ntv218

Vivolo-Kantor, A.M., Martell, B. N., Holland, K. M., & Westby, R. (2014). A systematic review and content analysis of bullying and cyberbullying measurement strategies. *Aggression and Violent Behavior, 19,* 423 – 434. https://doi.org/10.1016/j.avb.2014.06.008

Volk, A. A., Veenstra, R., & Espelage, D. L. (2017). So you want to study bullying? Recommendations to enhance the validity, transparency, and comparability of bullying research. *Aggression and Violent Behavior*, 36, 34 – 43. https://doi.org/10.106/j.avb.2017.07.003

Völlink, T., Dehue, F., McGuckin, C., & Jacobs, N. C. L. (2016). An introduction to cyberbullying research. In T. Völlink, F. Dehue, & C. McGuckin, (Eds.), *Cyberbullying: From theory to intervention,* pp. 3 – 14. Routledge.

Waasdorp, T. E., & Bradshaw, C. P. (2015). The overlap between cyberbullying and traditional bullying. *Journal of Adolescent Health, 56,* 483 – 488. https://doi.org/10.1016/j.adohealth.2014.12.002

Walker, J., Craven, R. G., & Tokunaga, R. S. (2013). Introduction. In S. Bauman, D. Cross, & J. Walker (Eds.), *Principles of Cyberbullying Research: Definitions, Measures, and Methodology,* pp. 3 – 20. Routledge.

Wang, C., Berry, B., & Swearer, S. M. (2013). The critical role of school climate in effective bullying prevention. *Theory into Practice, 52*(4), 296 – 302. https://doi.org/10.1080/00405841.2013.829735

Webster-Stratton, C., Gaspar, M. F., & Seabra-Santos, M. J. (2012). Incredible years® parent, teachers, and children's series: Transportability to Portugal of early intervention programs for preventing conduct problems and promoting social and emotional competence. *Psychosocial Intervention, 21*(2), 157 – 169. https://doi.org/10.5093/in2012a15

Weisburd, D. (2003). Ethical practice and evaluation of interventions in crime and justice: The moral imperative for randomized trials. *Evaluation Review, 27*(3), 336 – 354. https://doi.org/10.1177/0193841X03252516

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. https://doi.org/10.4324/9781315089256-3

Welling, R., Beaumont, J. J., Petersen, S. J., Alexeeff, G. V., & Steinmaus, C. (2015). Chromium VI and stomach cancer: A meta-analysis of the current epidemiological evidence. *Occupational and Environmental Medicine, 72*(2), 151 – 159. https://doi.org/10.1136/oemed-2014-102178

Willard, N.E. (2006). *Cyberbullying and Cyberthreats: Responding to the challenge of online social cruelty, threats, and distress.* Centre for Safe and Responsible Internet Use https://www.internetsafetyproject.org/wiki/center-safe-and-responsible-internet-use

Williams, K. R., & Guerra, N. G. (2011). Perceptions of collective efficacy and bullying perpetration in schools. *Social Problems, 58*(1), 126 – 143. https://doi.org/10.1525/sp.2011.58.1.126

Wilson, D. B. (2010). Meta-analysis. In A. R. Piquero & D. Weisburd, *Handbook of Quantitative Criminology,* (pp. 181 – 208). Springer.

Wolke, D., Lee, K., & Guy, A. (2017). Cyberbullying: A storm in a teacup? *European Child and Adolescent Psychiatry, 26,* 899 – 908. https://doi.org/10.1007/s00787-017-0954-6

Woolf-King, S. E., Steinmaus, C.M., Reingold, A. L., & Hahn, J. A. (2013). An update on alcohol use and risk of HIV infection in sub-Saharan Africa: Meta-analysis and future research directions. *International Journal of Alcohol and Drug Research, 2*(1), 99 – 110. https://doi.org/10.7895/ijadr.v2i1.45

Ybarra, M. L., & Mitchell, K. J. (2004). Youth engaging in online harassment: Associations with caregiver-child relationships, Internet use, and personal characteristics, *Journal of Adolescence, 27*(3), 319 – 336. https://doi.org/10.1016/j.adolescence.2004.03.007

Yeager, D. S., Fong, C. J., Lee, H. Y., & Espelage, D. L. (2015). Declines in efficacy of anti-bullying programs among older adolescents: Theory and a three-level meta-analysis. *Journal of Applied Developmental Psychology, 37,* 36 – 51. https://doi.org/10.1016/j.appdev.2014.11.005

Yusuf, S., Wittes, J., Probstfield, J., & Tyroler, H. A. (1991). Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA, 266*(1), 93 – 98. https://doi.org/10.1001/jama.266.1.93

Zhang, L., Steinmaus, C., Eastmond, D. A., Xin, X. K., & Smith, M. T. (2009). Formaldehyde exposure and leukemia: A new meta-analysis and potential mechanisms. *Mutation Research/Reviews in Mutation Research, 681*(2-3), 150 – 168. https://doi.org/10.1016/j.mrrev.2008.07.002

Zhuo, H., Smith, A. H., & Steinmaus, C. (2004). Selenium and lung cancer: A quantitative analysis of heterogeneity in the current epidemiological literature. *Cancer Epidemiology, Biomarkers & Prevention, 13*(5), 771 – 778.

Zych, I., Baldry, A. C., Farrington, D. P., & Llorent, V. J. (2019). Are children involved in cyberbullying low on empathy? A systematic review and meta-analysis of research on empathy versus different cyberbullying roles. *Aggression and Violent Behavior, 45,* 83 – 97. https://doi.org/10.1016/j.avb.2018.03.004

Zych, I., & Farrington, D. P. (in press). Stability and change in bullying and cyberbullying throughout the school years. In P. K. Smith & J. O'Higgins Norman (Eds.), *Handbook on Bullying Volume 2: Development, coping, and intervention.* Wiley-Blackwell.

Zych, I., Farrington, D. P., Llorent, V. J., & Ttofi, M. M. (2017). Protecting children against bullying and its consequences. *Springer Briefs in Behavioral Criminology,* https://doi.org/10.1007/978-3-319-53028-4

Zych, I., Farrington, D. P., & Ttofi, M. M. (2019). Protective factors against bullying and cyberbullying: A systematic review of meta-analyses. *Aggression and Violent Behavior, 45,* 4 – 19. https://doi.org/10.1016/j.avb.2018.06.008

Zych, I., Ortega-Ruiz, R., & Del Rey, R. (2015). Systematic review of theoretical studies on bullying and cyberbullying: Facts, knowledge, prevention, and intervention.

Aggression and Violent Behavior, 23, 1 – 21. https://doi.org/10.1016/j.avb.2015.10.001

Zych, I., Ttofi, M.M., & Farrington, D.P. (2019). Empathy and callous-unemotional traits in different bullying roles: A systematic review and meta-analysis. *Trauma, Violence & Abuse, 20*(1), 3 – 21. https://doi.org/10.1177/1524838016683456

Zych, I., Viejo, C., Vila, E., & Farrington, D. P. (2019). School bullying and dating violence in adolescents: A systematic review and meta-analysis. *Trauma, Violence, & Abuse.* https://doi.org/10.1177/1524838019854460

Appendices

Appendix 1: Worked example of the Multiplicative variance adjustment (MVA) model

In the present meta-analysis, the summary effect size estimated for bullying perpetration was OR = 1.324 with 95% confidence intervals of 1.298 – 1.351 under a fixed effects model. The effect size in the MVA model is the same as the effect size in the fixed effects model. The variance of the effect size in the MVA model is calculated as follows:

$$MVA_{var} = FE_{var} \times \frac{Q}{df}$$

Therefore, in the above example of the summary effect size for bullying perpetration outcomes, the FE_{var} is 0.000104. Therefore, with Q = 458.555 and df = 109, the MVA adjustment for fixed effects is 0.02098, calculated as:

$$MVA_{var} = 0.000104 \times \frac{458.555}{109} = 0.000438$$

Therefore, the adjusted standard error is 0.0209. In this example thus, the MVA fixed effect is OR = 1.324, and the 95% confidence intervals are 1.271 to 1.380.

Appendix 2: Example of estimating an odds ratio for the before-after intervention effect

Williams et al. (2015) evaluated the effectiveness of the Start Strong program based on students' self-reported experiences of bullying victimization. The primary study found that, at baseline, 23% of participants in the experimental group (N = 717) reported bullying victimization, while 23% of participants in the control group (N = 800) also reported bullying victimization at baseline. Hence, the baseline OR was calculated as follows:

Table 36

Example of estimating an odds ratio; baseline

	Non-Victims	*Victims*	*N*
Experimental	552	165	717
Control	616	184	800

Thus, the OR_{before} = 0.999; Ln OR_{before} = -0.002; and *var* Ln OR_{before} = 0.015. Williams et al. (2015) report that after implementation of the Start Strong program, bullying victimization was reported by 28% of experimental participants and 34% of control participants. Accordingly, the post-test OR was calculated as follows:

Table 37

Example of estimating an odds ratio; post-intervention

	Non-Victims	*Victims*	*N*
Experimental	516	201	717
Control	526	272	800

Thus, the OR_{after} = 1.323; Ln OR_{after} = 0.28; and *var* Ln OR_{after} = 0.013. Employing these figures, the ln OR for the intervention effect of the Start Strong program was calculated as:

$$\text{Ln } OR_{change} = \text{Ln } OR_{after} - \text{Ln } OR_{before}$$

$$\text{Ln } OR_{change} = 0.28 - (-0.002) = 0.282$$

$$var \text{ Ln } OR_{change} = 0.75 \text{ x } (0.015 + 0.013) = 0.021$$

$$\text{SE of Ln } OR_{change} = \sqrt{0.021} = 0.145$$

The ln OR_{change} is computed as the difference between the before and after effect size and the variance of this new estimate is adjusted by multiplying the sum of the variances of before and after variances by 0.75. This is an approximation of the assumed correlation between before and after effect sizes. The ln OR_{change} and the SE of ln OR_{change} were then entered into CMA as an estimation of the intervention effect.

Appendix 3: Example of converting an odds ratio to a percentage.

The conversion from weighted mean odds ratio to percentage value is also described in the previous Campbell report (see Farrington & Ttofi, 2009). The formula involves assuming equal allocation of participants to experimental and control conditions and that the % of bullies and/or victims was lesser in the experimental condition than in the control condition (as supported by our overall positive mean effect size).

For example, if there are 200 participants in each experimental condition and approximately 30% of participants report bullying victimization in the control condition and 25% victims in the experimental condition, the numbers of victims and non-victims would be as follows:

Table 38

Example of converting an odds ratio to a percentage

	Non-Victims	*Victims*	*N*
Experimental	150	50	200
Control	140	60	200
Total	290	110	400

Therefore, using the previously described formula for estimating an odds ratio, the following data would correspond to an odds ratio of 1.286 (i.e., [150 x 60]/[140 x 50]). Moreover, the percentage decrease would be approximately 16.67% (i.e., (10/60) x 100).

Using this basic formula, we can manipulate the % and number of victims in each experimental condition in order to achieve a odds ratio that corresponds to our weighted mean effect size (i.e., MVA OR = 1.324 and RE OR = 1.309 for bullying perpetration; MVA

OR = 1.248 and RE OR = 1.242 for bullying victimization). Using the *n* values that give the closest possible mean effect size we can thus estimate the corresponding percentage reduction in either bullying perpetration or victimization outcomes.

Appendix 4: Studies excluded from the school-bullying meta-analysis

Table 39

Description of studies included in category 4 studies

Study	*Reasons for exclusion*
Ahtola et al. (2013)	Explore teachers' perceptions of support from schools' principals in the KiVa program, and whether this predicted implementation adherence. Did not compare bullying outcomes of program. *[Outcomes]*
Ahtola et al. (2012)	Examined the effects of the KiVa anti-bullying program on teacher perceptions of bullying, no outcome of bullying behaviors in students is included *[Outcomes]*
Al-Samarri (2011)	Evaluated the effectiveness of the 'Mythodrama' violence prevention program, on verbal and physical bullying, but did not employ a control group. *[No control group]*
Allen (2010)	Evaluated a whole-school bullying intervention initiative for the effectiveness in reducing bullying, however, did not employ a control group for comparison *[No control group]*
Amundsen & Ravndal (2010)	Assessed the effectiveness of the OBPP to reduce alcohol and substance use in adolescents, but no measure/outcome of bullying behaviors actually employed. *[Outcomes]*
Azad & Amiri (2012)	Carried out an evaluation of the Olweus Bullying Prevention Program in a randomized controlled trial with Iranian primary school boys, however only abstract was published in English and did not provide enough details for meta-analysis. *[Other: Language]*
Beckman & Svensson (2015)	Evaluates the cost effectiveness of the Olweus Bullying Prevention Program, not the effectiveness of the program to reduce bullying. *[Method]*
Beets et al. (2009)	Conducted and evaluated an intervention program for Hawaiian elementary-school students for a number of outcomes, including violent behaviors, but no outcomes relevant to school bullying. *[Outcomes]*
Beightol et al. (2012)	Re-publication of Beightol et al. (2009). This report evaluates treatment effects on participant goals, empathy, self-efficacy and resilience. Only qualitative data refers to bullying outcomes. Employed the 'Anti-bullying Initiative Survey' which does include 6 items regarding bullying behaviors, however did not administer this section. *[Outcomes]*
Beightol et al. (2009)	Evaluates the effectiveness of an adventure-based intervention, but main outcome is participants' 'resilience', implications for reducing bullying, but provide no empirical evaluation data. *[Outcomes]*
Boulton (2014)	Conducted an evaluation of the teacher-training component of the I DECIDE anti-bullying program, and its effectiveness at increasing teachers' perceived effectiveness, self-efficacy and implementation of the program. Implications for the impact of the program on bullying are discussed, however no direct evaluation is conducted. *[Outcomes]*

Bowes et al. (2009)	Conducted a process evaluation of the 'Peers Running Organized Play Stations (PROPS)' intervention program. Outcome of interest was the implementation rate of the program by teachers, not the effect of the program on bullying behaviors. *[Outcomes]*
Brenick et al. (2014)	Evaluation of a safety-skills program for elementary school children. Study did include a measure of victimization, however only the outcome 'safety skills knowledge' was analysed pre- and post-test as an indicator of the effectiveness of the program. Additionally, the victimization measure refers to "participants' perceptions of the regularity of bullying..." and not their actual experiences of being victimized. *[Outcomes]*
Bundy et al. (2011)	Evaluation of a program to develop physical and social skills in children who are overweight. Main aim of program was to increase physical activity levels of children, and authors suggest that such outcomes would decrease childhood obesity and as a result, bullying. However, do not employ any bullying-related outcome measures to assess the impact of the program on bullying experiences/ behaviors directly. *[Outcomes]*
Burkhart et al. (2013); Burkhart (2012)	Evaluation of a community-based family violence intervention and prevention program, that included parent-measures of early childhood bullying. However, was excluded because bullying measures were not specific enough to school bullying. *[Outcomes]*
Cecil & Molnar-Main (2015)	Explored the effect of implementer (e.g. teachers) characteristics, beliefs of self-efficacy, and perceptions and attitudes towards bullying on OBPP implementation and fidelity. *[Outcomes]*
Černi Obrdalj et al. (2014)	Conducted an evaluation of a violence prevention program which involved family physicians (GPs). Included a measure of 'frequency of experiencing violence at school', however did not employ a control group to compare effect. *[No control group]*
Chu et al. (2013)	Tailored intervention of victims of bullying suffering from anxiety and depressive disorders. Measures included a scale measuring impairment (on family/ peer relations and academic performance) that occurs as a result of bullying. Outcomes of effectiveness are changes in psychological clinical symptoms as a result of victimization, and participant satisfaction with the intervention. No change in victimization is reported. *[Outcomes]*
Cobb (2009)	Investigated the effectiveness of Disciplinary Alternative Education Programs (DAEPs) for improving academic performance of students who demonstrate challenging behaviors, for example, those that bully others. *[Outcomes]*
Cooke et al. (2007)	Examined the impact of the violence prevention program 'Second Step' on a number of outcomes, including bullying behaviors, measured by 4 items on the Modified Aggression Scale, did not employ experimental and control conditions. *[No control group]*
Cornell et al. (2009)	Explore differences between schools that implement a violence prevention set of guidelines on constructs such as bullying, but no pre- and post-test measures, is a 'nonexperimental' study. *[Method]*
Cross et al. (2012)	Report the results of a three-year evaluation study of the Friendly Schools, Friendly families program, however no control group is utilised as after the second year of implementation, many schools wished to implement the program. Authors compare the effectiveness of the program across three different levels of implementation, low, moderate, and high. *[No control group]*
Daugherty (2011)	Aimed to evaluate the effectiveness of the Olweus Bullying Prevention Program, but the main outcome of interest were teacher and school principals' perceptions of the

	effectiveness of the program. Survey does include an item referring to a decrease in bullying incidents, however, this is related to teacher and principal perceptions and opinions about whether or not bullying decreased, rather than actual records indicating they did. *[Outcomes]*
Davis (2011)	Abstract outlines that the study evaluated the effectiveness of a social skills treatment program for children displaying problems behaviors such as bullying, aggression, and poor social skills. However, do not evaluate the program's effectiveness of altering these problem behaviors. Instead, assess the change in variables such as empathy, social skills, and motivation. *[Outcomes]*
DeNike (2014)	Abstract outlines that the report evaluated the effectiveness of just one part of the 'No Bully System' anti-bullying intervention, the Solution Team, limited information is available, but do not refer to a comparison group in graphical representation of findings. *[No control group]*
Drury (2014)	Investigated whether an anti-bullying program reduced "HIB" incidents (i.e. harassment, intimidation and bullying). Do not compare effect of intervention with a control group. *[No control group]*
Earhart (2011)	Investigated the effect of implementing the 'Promoting Positive Peer Relationships' program, however excluded as effectiveness of the program was measured using attitudinal outcomes of bullying rather than bullying behaviors. *[Outcomes]*
Emfield (2015)	Evaluated the experiences of participants in an anti-bullying self-defence training program. Qualitative data only about the participants' opinions and thoughts on the program, no quantitative measure of bullying outcomes *[Outcomes; Method; No control group]*
Espelage et al. (2015)	Randomized clinical trial of the Second Step: Student Success Through Prevention program in middle schools to reduce bullying. However, excluded from present review as sample utilised were disabled. *[Sample]*
Farmer et al. (2010)	Conducted an evaluation of the 'Rural Early Adolescent Learning Program (Project REAL), to explore the impact of the program on teachers' abilities to identify peer groups amongst their students and also identify the incidents of bullying occurring in peer groups *[Outcomes]*
Farrell et al. (2015)	Qualitatively explored participants in the 'Second Step' violence prevention programs' implementation and perceptions of the skills they learnt during the program. No measure of actual bullying behaviors or victimization is utilised. *[Outcomes; Method]*
Fletcher et al. (2015)	A qualitative study evaluating the implementation of an anti-bullying program, specifically, how young were involved and young peoples' experiences of the program. *[Method]*
Frost (2012)	Examined the prevalence of school programs implemented in Kansas, including, bullying prevention, conflict resolution and peer mediation programs. Compare official records of school suspension for violence in relation to the type of program implemented. However, do not use any indicator of specific school bullying perpetration or victimization. *[Outcomes]*
Fung (2012)	Tested the effects of an intervention with high-risk reactive aggressors (i.e. bullies) over 5 time-points in one year, however no control group was utilised. *[No control group]*
Garandeau, Poskiparta & Salmivalli (2014)	Using data from a previous evaluation study of the KiVa anti-bullying intervention program, the authors compared the impact of the 'Confronting' and 'Non-Confronting' approaches on bullying victimization. Thus, compare intervention participants according to which arm they were assigned to, but do not compare either with control group.

	[No control group]
Gibson et al. (2015)	Evaluates the outcomes of a bullying-focused program, refer to outcomes such as fear of bullying and peer/teacher interventions in bullying. *[Outcomes]*
Giesbrecht et al. (2011)	WITS violence prevention program, reduced levels of physical and relational victimization. Excluded because outcome variables are not specific enough to school bullying. *[Outcomes]*
Goncy et al. (2015)	Investigates the influence of several aspects of teacher implementation of the OBPP, such as: *adherence; competence; and student responsiveness,* on student engagement with the intervention, not any change in their bullying behaviors as a result of the program. *[Outcomes]*
Good et al. (2011)	Report presents a case study example of a school in Canada that implemented the 'School Wide Positive Behavior Support' Program, using discipline referrals for bullying as an effectiveness indicator. However, do not employ a comparison school as a control. *[No control group]*
Gregus et al. (2015)	Describe two separate studies that tested the effects of a Lunch Buddy mentoring program. First study was with victimized elementary school children, and the second was with bully-victim children. Excluded due to lack of control group. *[No control group]*
Greytak & Kosciw (2010)	Present the results of a one-year training program 'Respect for All' for secondary school teachers in order to increase their abilities to intervene and be aware of LGBT bullying in their schools. Evaluated the effectiveness of the program for teachers' attitudes towards LGBT students and various variables relating to their self-efficacy beliefs to intervene, but not on actual bullying behaviors of their students. *[Outcomes]*
Greytak et al. (2013)	Evaluate a professional development program for teachers that aims to help them to develop better strategies and attitudes towards LGBT youth and prevent bullying. Do not evaluate the outcomes of this program in relation to actual bullying incidents in schools. Focus instead on teacher-related outcomes, similar to Greytak & Kosciw (2010). *[Outcomes]*
Gyooyeong (2013)	Evaluated the effectiveness of a program designed for victimized adolescents. Looked at changes in ego-resiliency, self-esteem, somatic symptoms, aggression and social withdrawal in intervention and control group, but change in bullying behaviors/experiences was not an outcome. *[Other; Language]*
Haataja et al. (2014)	This study evaluates the link between implementation fidelity of the KiVa anti-bullying program and its outcomes, do not actually explore the effectiveness of the program as a whole *[No control group]*
Hallam (2009)	Qualitative aspect of the evaluation of school staffs' (i.e. teachers, principals and non-teaching staff) perceptions of the effectiveness of the Social and Emotional Aspects of Learning program (SEAL) on a range of outcomes, including bullying. Quantitative student measures include measures of emotional and behavioral skills, perceptions of classroom and school ethos and their attitudes towards school, but not bullying behaviors. *[Method (Teacher-report); Outcomes (Student-report)]*
Hatzenbuehler & Keyes (2013)	Evaluated the impact of anti-bullying policies that incorporate an anti-homophobic element on suicide and attempted suicide in homosexual adolescents. However, do not explore the impact of these policies on reported bullying behaviors. *[Outcomes]*

Hawe et al. (2015)	Replicated the Gatehouse project intervention in Canadian schools and investigated the effects of program on a series of health risk behaviors, including bullying victimization. Excluded due to lack of inclusion of a control group *[No control group]*
Hervey & Kornblum (2006)	Evaluation of a violence prevention program, 'Disarming the Playground', on a variety of different outcomes. The behavioural measure included does include some aggressive items, but these are not specified as being related to bullying behaviors. *[Outcomes]*
Hoglund et al. (2012)	Evaluated effectiveness of a community-based, whole-school prevention program 'WITS Primary Program' for peer victimization. However, victimization measures are not specifically related to school bullying, thus, excluded from the current review. *[Outcomes]*
Holden (2015)	Evaluated the effectiveness of the Olweus Bullying Prevention Program. However, excluded from the present meta-analysis as did not include a control group for comparison. *[No control group]*
Hornblower (2014)	Evaluated an anti-bullying program implemented in an English secondary school but did not include a control condition. *[No control group]*
Huddleston et al. (2011)	Describe the implementation and evaluation of an individualized intervention for one adolescent middle school bully and investigated the impact on their bully behaviors, however no control student/group. *[No control group]*
Hutchings & Clarkson (2015)	Presents results from the pilot implementation of the KiVa anti-bullying program in the UK. However, do not employ any control condition in order to evaluate the significance of any results. *[No control group]*
Isaacs (2009)	Examined the impact of the OBPP in U.S. middle schools, however, conduct a 'single school' study, and thus, did not include a control school *[No control group]*
James (2011)	Conducted cross-cultural comparisons of the effect of peer support approaches to bullying prevention. In two studies conducted in UK, compare quantitative measures of bullying as a result of program. Excluded on the basis that no control condition was employed. *[No control group; Method]*
James et al. (2011)	Evaluation of an educational program to raise awareness of relational aggression/bullying in teenage girls, however, knowledge and attitudes of relational bullying and change in these constructs were the primary outcome of interest *[Outcomes]*
James et al. (2013)	Evaluated the applicability of the relational aggression educational program implemented by James et al. (2011), for boys, but main focus is knowledge and attitudes towards relational bullying. *[Outcomes]*
Jeong-Lan & Oh-Hyun (2014)	Evaluated a school violence prevention program and its effectiveness to increase levels of empathy in school children. Do not refer to any bullying-related outcomes. Full text only available in Korean. *[Outcomes; Other: Language]*
Jiménez-Barbero et al. (2013)	Explored the effects of a school violence prevention program on a range of outcomes, such as attitudes towards violence and perceived violent victimization. Imply modifying attitudes towards violence can reduce prevalence of bullying, but no bullying measure. *[Outcomes]*

Knights (2011)	Conducted an evaluation of the impact specialized schools for highly victimized adolescents, "Red Balloon Learner Centers". However, the evaluation outcomes are clinical and academic-related constructs, such as levels of anxiety/ depression in RBLC participants and victimized children from Local Authority comparison schools. The only bullying-related measure is concerned with establishing retrospective bullying experiences, and the severity of past bullying experiences. *[Outcomes]*
Konishi et al. (2013)	Explored the association between schools implementing anti-homophobic bullying policies and LGBT youths' alcohol and drug use, however, do not investigate the effect of these program on bullying/victimization experiences. *[Outcomes]*
Langevin et al. (2012)	Examined the effects of an anti-bullying program specifically targeting bullying of children who have a speech impediment. Assess change in attitudes towards and knowledge of this type of bullying. Authors did conduct a measure of bullying behaviors, but only at pre-test baseline. Thus, the effect of the intervention on bullying behaviors cannot be assessed. *[Method]*
Layfield (2014)	An exploratory case study of one school's implementation and methods for reducing problem behaviors, such as bullying. No control school utilised. Dissertation, only preview available *[No control group]*
Leadbeater & Sukhawatanakul (2011)	Evaluated the effect of the WITs program on elementary school children to reduce peer victimization trajectories. However, victimization outcomes do not relate to school bullying. *[Outcomes]*
Leff et al. (2010)	Evaluates a program designed to reduce relational aggression in schools, discuss implications for bullying prevention in text, but main outcome is aggression. *[Outcomes]*
Low et al. (2014)	Using data from a previous evaluation of the *Steps to Respect* program (Brown et al., 2011), this study assessed the predictors of implementation factors such as: engagement and adherence. Bullying victimization and perpetration are included as possible indicators, but the study does not compare these measures in relation to the effectiveness of the intervention. *[Outcomes]*
Lucassen & Burford (2015)	Evaluated a sexuality diversity workshop in secondary schools and its potential impact to reduce school bullying. The effect of the program is primarily assessed through changes in participants valuing and understanding of sexually diverse individuals, no actual measure of bullying experiences utilised. *[Outcomes]*
Macedo et al. (2014)	Implemented an evaluated the program 'We are the Others' in a group of Portuguese students, did not employ a control group. *[No control group]*
Malatino (2012)	Conducted an evaluation of the program 'City Connects' on a range of social development outcomes, including bullying behaviors. However, no true control group is utilised. All participants had been exposed to the intervention, just at different 'dosage' levels, i.e. for longer/shorter periods of time. *[No control group]*
McElearney et al. (2013)	Examined the effectiveness of a school counselling intervention in improving peer relationships in children identified as victims of bullying. Measures included the Strengths and Difficulties Questionnaire, and the Peer Problems subscale, but no direct measure of bullying behaviors/experiences utilised. *[Outcomes]*
Mendes (2011)	Examined the effects of an anti-violence school program on the levels of bullying in a Lisbon school, however, do not include a control condition

	[No control group]
Menesini & Nocentini (2012)	Conducted an evaluation of the efficacy of a peer-led intervention program to reduce cyberbullying perpetration and victimization. Authors do not include any measures of traditional/offline bullying. *[Outcomes; Cyberbullying]*
Migliaccio & Raskauskas (2013)	Evaluated a small-scale video-based bullying awareness program, but the main outcomes were changes in knowledge about and attitudes towards bullying behaviors and no measure of actual bullying behaviors was employed. *[Outcomes]*
Minton et al. (2013)	Implemented and evaluated an anti-bullying intervention described as a 'whole school/community development' program in Ireland primary and post primary schools on self-reported involvement and experiences of bullying. Excluded due to lack of control condition. *[No control group]*
Miyari (2013)	Implemented and evaluated a weight-related 'teasing' (or bullying) prevention program but did not employ any control group. *[No control group]*
Nakamura & Koshikawa (2014)	Conducted an evaluation of a social skills training and psychoeducational program for preventing bullying in Japan, however the full text was not available in English. *[Other: Language]*
Nese et al. (2014)	Evaluated the Expect Respect intervention program, using a non-concurrent multiple baseline design. All participants received the intervention; thus, no control group was used for analysis. *[No control group]*
Newgent et al. (2010)	Carried out an evaluation of a psychoeducational program in order to determine the effect on several outcomes, including bullying behaviors. However, comparison groups were formed on the basis of pre-test clinical symptoms, and all students received the intervention, thus, no true control group employed. *[No control group]*
Nixon & Werner (2010)	Evaluation of the intervention program 'Creating a Safe School' (The Ophelia Project) to reduce relation aggression and victimization in children. Thus, 'relational aggression' and 'relational victimization' are the primary outcomes, not specifically related to bullying. *[Outcomes]*
Pack et al. (2011)	Conducted an evaluation of the Safe School Ambassadors program me, however outcomes of interest are participants' perceptions of the impact of the project. Did not employ a direct measure of actual bullying experiences.
Park et al. (2014)	Effects of a 'food-therapy' program on bullying/school violence (crossover between terms used in Abstract). Full text in Korean. *[Other: Language]*
Peagram (2013)	Evaluated the impact of the Bulldog Solution Intervention Model as a way to reduce bullying and aggression and increase empathy, and self-esteem. However, measure of bullying is inadequate, student measure relates to be a bystander or witness to bullying ("I have seen bullying"). *[Outcomes]*
Pepler & Craig (2011)	Do not directly evaluate the effectiveness of a specific anti-bullying intervention or prevention program. Authors examine the effects that the establishment and work of the 'Promoting Relationships and Eliminating Violence Network (PREVNet)' Canadian research network has had on research on bullying and participation in anti-bullying initiatives.

	[Method]
Phillips (2015)	Implemented a bullying prevention program in order to ascertain its effectiveness in changing educators' perceptions of bullying, thus, the main outcome evaluated was not bullying behaviors by students. Additionally, did not employ a control group. Dissertation, only preview available *[No control group; Outcomes]*
Pister (2010)	Evaluated the 'Working against Youth Violence Everywhere' program to prevent bullying and violence in schools, however unable to obtain full text. *[Other: Unavailable]*
Ramierz & Lacasa (2013)	Conducted an evaluation of an anti-bullying program in Spanish primary schools but did not employ a control group. Full text in Spanish *[No control group]*
Renshaw & Jimerson (2012)	Examined the impact of a bullying prevention curriculum for middle school students, however, effectiveness outcomes do not refer to bullying behaviors, but attitudes towards bullying and perceptions of bullying-related support services within the school. *[Outcomes]*
Rigby & Griffiths (2011)	Qualitative evaluation data from interviews with students and practitioners involved in the anti-bullying initiative 'Method of Shared Concern' are reported, but there was no quantitative evaluation of effectiveness of program *[Method]*
Roberto et al. (2014)	Evaluated the effects of the 'Arizona Attorney General's Social Networking Safety Promotion and Cyberbullying Prevention' presentation on cyberbullying perpetration and victimization. No measures of traditional bullying are employed. *[Outcomes; Cyberbullying]*
Ross (2009); Ross & Horner (2009)	Evaluated the single-subject program Bully Prevention in Positive Behavior Support to reduce bullying behaviors. However, do not employ a control group. *[No control group]*
Ross & Horner (2014)	Investigated the effect of the 'School-Wide Positive Behavior Interventions and Supports, and measures employed did include 9 items that refer to bullying perpetration and victimisation but did not employ a control group. *[No control group]*
Rubin-Vaughan et al. (2011)	Evaluated the effect of the 'Quest for the Golden Rule' e-learning anti-bullying program, but outcomes were attitudes and knowledge of bullying issues and effective intervention and coping strategies. *[Outcomes]*
Santos et al. (2011)	Investigated the impact of a school violence prevention program widely implemented in Canada, 'Roots of Empathy', but targeted outcomes are mental health or generic aggression/violence related and not specified to refer to bullying. *[Outcomes]*
Saurini (2011)	Explored the effect of a psychoeducational anger management program on bullying behaviors, but do not utilise a control condition. *[No control group]*
Scheithauer & Bull (2010)	Imply that text presents the results of a pilot evaluation of the 'fairplayer.manual' school bullying preventative intervention program on prevalence of bullying; however, no control group was employed. *[No control group]*
Shek & Yu (2013)	Evaluation of the Project P.A.T.H.S, an intervention program in Hong Kong for adolescent males' risky behaviors. School bullying is not an outcome. *[Outcomes]*

Spiel et al. (2012)	Qualitative evaluation study of Austria's national school violence prevention program. *[Methods]*
Splett et al. (2015)	Describes evaluation of intervention program for reducing relational aggression, not specific to bullying. *[Outcomes]*
Srekovic (2015)	Effectiveness of a social intervention program for students with Autism Spectrum Disorder who were identified as being bullied, or at risk of being bullied. Conducted a peer network intervention, however, did not employ any control or comparison group. *[No control group]*
Stallard & Buck (2013)	Evaluated an intervention program where the main outcome was reducing depression in participants, thus, bullying experiences and behaviors were not the primary outcomes. Qualitative focus groups conducted after the interview did review participants' perceptions of bullying issues covered in the intervention. *[Outcomes]*
Steiger (2010)	Assessed the effectiveness of the 'Solution Team' anti-bullying program for primary school children identifying as victims of bullying, but do not employ a control group for comparison. *[No control group]*
Tokarick (2015)	Evaluated the effect of bullying prevention program on adolescent females' perceptions of bullying, thus, not actual bullying behaviors *[Outcomes]*
Tomic-Latinac & Nikcevic-Milkovic (2010)	Evaluated the efficacy of the UNICEF bullying prevention program in high school students. However, full text is published in Croatian. *[Other: Language]*
Vannini et al. (2011)	Investigated the impact of the 'FearNot!' virtual anti-bullying program in UK and German schools on participants' 'defender' status. Thus, indicator of effectiveness was an increase in peer-reported bystander intervention, not decreases in reports of bullying behaviors. *[Outcomes]*
Velderman (2015)	Evaluation of a professional development program for teachers, and the impact the development program had on their knowledge of bullying related issues and implementation of anti-bullying plans. Do not however, evaluate the effectiveness in reducing bullying behaviors amongst their students. *[Outcomes]*
Watson et al. (2010)	Examined the efficacy of the FearNot! bullying prevention program in UK and German schools, comparison is done cross-nationally. However, effectiveness outcome is coping strategy knowledge in relation to bullying victimization, not actual reports of being bullied. *[Outcomes]*
Westheimer & Szalacha (2015)	Chapter outlining the Welcoming School program for LGBT anti-bullying. Do outline an evaluation study, but none of the outcomes relate to bullying perpetration or victimization. *[Outcomes]*
Wolfe et al. (2012)	Evaluated the classroom-based intervention program, the 'Fourth R program' which aims to decrease abusive and health-risk behaviors in adolescents. No outcome of bullying is included, 'peer resistance skills', i.e. ability to withstand peer pressure is the primary targeted outcome. During intervention, one of the pressures adolescents are pressed to comply with is a bullying scenario. *[Outcomes]*
Wood (2012)	Evaluate the 'implementation fidelity' of the Olweus Bullying Prevention Program, but do not employ a control comparison group. *[No control group]*
Wright et al. (2012)	Investigated the effectiveness of a bullying intervention program, The Ophelia Project, but outcome measure was relational aggression, not bullying behaviors.

	[Outcomes]
Yamashiro (2013)	Qualitative evaluation using semi-structured interviews with participants in the Anti-Bullying Prevention Pilot Program (ABPPP) *[Methods]*
Young et al. (2009)	Appears to evaluate a bullying prevention approach adopted by school counsellors in one school. Effectiveness is measured using discipline referral rates; however no control group was employed. *[No control group]*

Note. (1) Studies excluded from the school-bullying meta-analysis because they reported cyberbullying outcomes are not included in this table; (2) Studies that were excluded because no full-text was available are also excluded from this table. The majority of these were unpublished dissertations.

Appendix 5: Intervention components coding for school-bullying meta-analysis

Table 40

Intervention components coding results: School-bullying meta-analysis

			School			Classroom		Teacher		Parent		Peer			Individual			Intervention				
Study	**Intervention**	*Rich_Score*	*WSA*	*SUP*	*ABP*	*CRule*	*CManage*	*TInfo*	*TTrain*	*PInfo*	*PInvolve*	*Peer1*	*Peer2*	*Peer3*	*BULL*	*VIC*	*CoOp*	*Curriculum*	*SESkills*	*MH*	*Punitive*	*Non-punitive*
										Randomised Controlled Trials												
Baldry2004	Bulli & Pupe	5						✓	✓			✓					✓	✓				
Beran2005	Project Ploughshares Puppets for Peace	1																✓				
Berry2009[b]	Confident Kids	5									✓					✓	✓	✓		✓		
Bonnell2015[b]	INCLUSIVE	10	✓		✓		✓	✓	✓					✓			✓	✓	✓			✓
Boulton1996[a]	Short Video ABP	1										✓										
Brown2011	Steps to Respect	11	✓	✓	✓		✓	✓	✓	✓		✓	✓					✓			✓	
Chaux2016	MediaHeroes	8	✓					✓	✓		✓	✓	✓	✓				✓				
Cissner2014	Fourth R	3						✓	✓									✓				
Connolly2015[b]	Youth-led	3												✓			✓	✓				
Cross2011	Friendly Schools	9		✓			✓	✓	✓	✓	✓							✓	✓			✓
DeRosier2005	S.S. GRIN	7										✓			✓	✓	✓	✓	✓	✓		
Domino2013	Take the LEAD	7						✓	✓	✓	✓	✓						✓	✓			
Espelage2015	Second Step	7				✓		✓	✓			✓	✓					✓	✓			
Fekkes2006	Skills for Life	9	✓	✓	✓	✓	✓	✓	✓	✓								✓				

Study	**Intervention**	*Rich_Score*	*WSA*	*SUP*	*ABP*	*CRule*	*CManage*	*TInfo*	*TTrain*	*PInfo*	*PInvolve*	*Peer1*	*Peer2*	*Peer3*	*BULL*	*VIC*	*CoOp*	*Curriculum*	*SESkills*	*MH*	*Punitive*	*Non-punitive*
Fekkes2016	Skills for Life	8	✓					✓	✓			✓					✓	✓	✓	✓		
Fonagy2009	SPC + CAPSLE	9	✓					✓	✓			✓	✓		✓	✓	✓	✓				
Frey2005	Steps to Respect	11	✓		✓	✓	✓	✓	✓	✓		✓	✓					✓	✓			
Garaigordobil2015	Cyberprogram 2.0	3										✓					✓	✓				
Holen2013[a]	Zippy's Friends	8	✓					✓	✓			✓					✓	✓	✓	✓		
Hunt2007	Australian ABP	5						✓		✓	✓	✓							✓			
Jenson2013	Youth Matters	6	✓			✓						✓					✓	✓	✓			
Ju2009[b]	Chinese ABP	6						✓	✓		✓	✓			✓	✓						
Kaljee2017	Teacher Diploma	3						✓	✓										✓			
Kärnä2011b[c]	KiVa	15	✓	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓	✓				✓
Krueger2010[a]	School-bus ABP	4			✓							✓	✓					✓				
Li2011[a]	Positive Action	5	✓				✓	✓		✓								✓				
McLaughlin2009	CBT + Media	6			✓								✓				✓	✓	✓	✓		
Meyer2000[a]	"Bullying Boys"	4													✓		✓	✓				✓
Nocentini2016	KiVa	14	✓	✓	✓	✓		✓	✓			✓	✓	✓	✓	✓		✓			✓	✓
Ostrov2015[a]	Early Childhood Friendship	4										✓					✓	✓	✓			
Polanin2015	Second Step	7	✓					✓	✓	✓		✓						✓	✓			
Rosenbluth2004	Expect Respect	11	✓		✓			✓	✓	✓	✓	✓	✓			✓		✓				

Study	**Intervention**	*Rich_Score*	*WSA*	*SUP*	*ABP*	*CRule*	*CManage*	*TInfo*	*TTrain*	*PInfo*	*PInvolve*	*Peer1*	*Peer2*	*Peer3*	*BULL*	*VIC*	*CoOp*	*Curriculum*	*SESkills*	*MH*	*Punitive*	*Non-punitive*
Sprober2006	Pro-ACT+E	10				✓	✓	✓	✓	✓	✓							✓		✓	✓	✓
Topper2011[b]	Adventure	4						✓	✓								✓			✓		
Stallard2013[a]	Resourceful Adolescent	5										✓					✓	✓	✓	✓		
Trip2015	ViSC + REBE	7	✓					✓	✓			✓					✓	✓	✓			
Tsiantis2013	Greek ABP (1)	8				✓		✓	✓	✓	✓	✓					✓	✓				
Waasdorp2012[a, c]	SWPBIS	11	✓	✓	✓			✓	✓						✓	✓	✓		✓			✓
Wölfer2014[a]	fairplayer.manual	9	✓			✓		✓	✓		✓	✓	✓				✓	✓				
Yanagida2016	ViSC	7		✓				✓	✓			✓	✓					✓	✓			
									Before-After/Quasi-experimental designs													
Alsaker2001	Be-Prox	10				✓	✓	✓	✓		✓		✓		✓	✓	✓	✓				
Andreou2007	Greek ABP (2)	5				✓		✓	✓			✓						✓				
Battey2009[b]	BPCCC	3							✓								✓		✓			
Bauer2007[b]	OBPP	14	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓			✓	
Beran2004[b]	Bully Proofing Your School	11	✓		✓			✓	✓	✓	✓	✓			✓	✓		✓				✓
Bergen 2/Olweus	OBPP	14	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓			✓	
Bull2009	fairplayer.manual	9	✓			✓		✓	✓		✓	✓					✓	✓				
Ciucci1998	Progetto Pontassieve	8						✓	✓	✓		✓		✓			✓		✓	✓		
Elledge2010[b]	Lunch Buddy	2														✓	✓					

Study	**Intervention**	*Rich_Score*	*WSA*	*SUP*	*ABP*	*CRule*	*CManage*	*TInfo*	*TTrain*	*PInfo*	*PInvolve*	*Peer1*	*Peer2*	*Peer3*	*BULL*	*VIC*	*CoOp*	*Curriculum*	*SESkills*	*MH*	*Punitive*	*Non-punitive*
Evers2007	Transtheroetcial ABP	3						✓		✓								✓				
Finn2009	OBPP	14	✓	✓		✓		✓	✓	✓	✓	✓			✓	✓		✓			✓	✓
Fox2003[b]	Social Skills training	4														✓	✓	✓	✓			
Gini2003	Stare bene a scuola	5				✓		✓	✓								✓	✓				
Gollwitzer2006	ViSC	6				✓		✓	✓	✓		✓	✓				✓	✓				
Joronen2011	Drama program	5						✓	✓		✓	✓							✓			
Kimber2008[b]	Socio-emotional training	6						✓	✓		✓	✓						✓	✓			
Losey2009	OBPP	13	✓	✓			✓	✓	✓	✓	✓	✓			✓	✓	✓	✓			✓	
Martin2005	Granada ABP	10				✓		✓		✓	✓	✓	✓		✓	✓	✓	✓				
Melton1998	OBPP	14	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓			✓	
Menard2014	Bully-Proofing Your School	11				✓	✓	✓	✓	✓	✓	✓	✓					✓			✓	✓
Menesini2003	Befriending intervention	7						✓	✓	✓		✓	✓	✓		✓	✓					
Menesini2012	NoTrap!	2												✓				✓				
Ortega-Ruiz2012	ConRed	7	✓		✓			✓	✓		✓						✓	✓				
Palladino2012	NoTrap!	5						✓				✓		✓		✓		✓				
Palladino2016	NoTrap!	6						✓				✓	✓	✓		✓		✓				
Pepler2004	Toronto ABP	12	✓	✓	✓	✓		✓	✓	✓	✓	✓			✓	✓		✓				

Study	**Intervention**	*Rich_Score*	*WSA*	*SUP*	*ABP*	*CRule*	*CManage*	*TInfo*	*TTrain*	*PInfo*	*PInvolve*	*Peer1*	*Peer2*	*Peer3*	*BULL*	*VIC*	*CoOp*	*Curriculum*	*SESkills*	*MH*	*Punitive*	*Non-punitive*
Pryce2013	Anti-bullying Pledge Scheme	4	✓		✓			✓		✓												
Rahey2002	Ecological ABP	9	✓					✓				✓		✓	✓	✓	✓	✓	✓			
Rawana2011	Strengths in Motion	6	✓								✓				✓	✓	✓		✓			
Rican1996	Short intensive ABP	12	✓	✓	✓	✓		✓	✓	✓		✓			✓	✓		✓			✓	
Sapouna2010	FearNot	2						✓										✓				
Silva2016	Social-skills training	7										✓	✓			✓	✓	✓	✓	✓		
Sismani2014	Daphne III	1																✓				
Solomontos-Kountouri2016	ViSC	10	✓	✓				✓	✓			✓	✓		✓	✓		✓	✓			
Sutherland2010	Beyond the Hurt	3	✓											✓					✓			
Toner2010	Bully-Proofing Your School	7	✓				✓	✓	✓	✓		✓						✓				
Williams2015[b]	Start Strong	1																✓				
Wong2011	Restorative Whole-school approach	5	✓		✓										✓	✓						✓
Yaakub2010	OBPP	14	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓			✓	
											Age Cohort designs											
Busch2013	Healthy Schools	7	✓		✓			✓	✓		✓							✓	✓			
Ertesvåg2004	Respect	6	✓			✓	✓	✓	✓	✓												
Kärnä2011a[c]	KiVa	15	✓	✓	✓	✓		✓	✓	✓		✓	✓	✓	✓	✓	✓	✓				✓
Limber2018[c]	OBPP	14	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓			✓	

Study	**Intervention**	*Rich_Score*	*WSA*	*SUP*	*ABP*	*CRule*	*CManage*	*TInfo*	*TTrain*	*PInfo*	*PInvolve*	*Peer1*	*Peer2*	*Peer3*	*BULL*	*VIC*	*CoOp*	*Curriculum*	*SESkills*	*MH*	*Punitive*	*Non-punitive*
Olweus/Bergen 1	OBPP	14	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓			✓	
Olweus/New National [c]	OBPP	14	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓			✓	
Olweus/Oslo 1	OBPP	14	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓			✓	
Olweus/Oslo 2	OBPP	14	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓			✓	
O'Moore2004	Donegal ABP	11	✓		✓		✓	✓	✓	✓	✓	✓	✓					✓				✓
Pagliocca2007	OBPP	14	✓	✓	✓	✓	✓	✓	✓	✓	✓				✓	✓		✓			✓	
Purugulla2011	OBPP	14	✓	✓	✓	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓			✓	
Roland2010[c]	Zero Program	10		✓		✓	✓	✓	✓		✓				✓	✓			✓		✓	
Salmivalli2005	Finnish ABP	10	✓		✓	✓	✓	✓	✓						✓	✓		✓				✓
Whitney1994	Sheffield ABP	13	✓	✓	✓	✓	✓	✓		✓		✓		✓	✓	✓		✓				✓

Note. ✓ = component present; ABP = Anti-bullying program; OBPP = Olweus Bullying Prevention Program; BPCCC = Bully Prevention Challenge Course Curriculum; Rich_Score = sum total number of components included in intervention; WSA = whole school approach; SUP = increased supervision; ABP = anti-bullying policy; CRule = classroom rules; CManage = classroom management; TInfo = Information for teachers; TTrain = Teacher training; PInfo = Information for parents; PInvolve = Parental involvement; Peer1 = informal peer involvement; Peer2 = Encouraging bystanders; Peer3 = Formal peer involvement; BULL = Work with individual bullies; VIC = Work with individual victims; CoOp = Co-operative group work; Curriculum = Set intervention curriculum materials; SESkills = Socio-emotional skills; MH = Mental health; Punitive = Punitive disciplinary methods; Non-punitive = Non-punitive disciplinary methods.

a. Studies only reported effectiveness in reducing bullying perpetration outcomes
b. Studies only reported effectiveness in reducing bullying victimization outcomes
c. Studies were deemed 'over-powered' and thus removed from the model for the purpose of intervention component analyses

Appendix 6: Estimation of between-group variance in subgroup analyses

This example uses the AgeCat1 variable from the cyberbullying meta-analysis and cyberbullying perpetration outcomes to illustrate how to estimate Q_B in subgroup analyses. In this comparison, 9 studies that included 'younger' participants (i.e., group X) are compared with 9 studies that included 'older' participants (i.e., group Y).

The overall Q value for the meta-analysis on cyberbullying perpetration outcomes was 67.49 (df = 17). As reported in Table 33, the Q values for these groups are as follows:

$$Q_X = 29.282, \text{ df} = 8$$

$$Q_Y = 30.899, \text{ df} = 8$$

Adopting the formula outlined by Borenstein et al. (2009, p. 157), the QW is calculated as the sum of QX and QY, and QB = Q – QW. Therefore, the QB is calculated as follows:

$$QW = 29.282 + 30.899 = 60.181$$

$$QB = 67.49 - 60.181 = 7.309$$

The significance of QB is estimated using the chi-square distribution function in excel [CHIDIST(Q,df)]. Degrees of freedom are calculated as the number of groups minus 1.