Children and Youth Services Review 61 (2016) 165–175

Contents lists available at ScienceDirect

Children and Youth Services Review

journal homepage: www.elsevier.com/locate/childyouth

Effectiveness of anti-bullying school programs: A meta-analysis

José Antonio Jiménez-Barbero [a,*], José Antonio Ruiz-Hernández [a], Laura Llor-Zaragoza [b], María Pérez-García [c], Bartolomé Llor-Esteban [a]

[a] University of Murcia, Faculty of Psychology, Espinardo, Murcia CP: 30100, Spain
[b] Catholic University San Antonio, Faculty of Social Sciences and Communication, Guadalupe, Murcia CP: 30107, Spain
[c] Mental Health Center of Cieza, Murcia CP: 30530, Spain

ARTICLE INFO

Article history:
Received 9 November 2015
Received in revised form 15 December 2015
Accepted 15 December 2015
Available online 28 December 2015

Keywords:
Bullying
Violence
Adolescents
School program
Effectiveness
Meta-analysis

ABSTRACT

The large number of studies published in recent years aimed at evaluating the effectiveness of school-based anti-bullying programs recommends research to focus on synthesizing the evidence found in this regard.
This study presents a meta-analysis of random clinical trials that assess the efficacy of 14 anti-bullying school programs. Sample size was of 30,934 adolescents aged between 7 and 16 years of whom 16,243 made up the Intervention Groups and 14,691 made up the Control Groups. Meta-analysis was conducted for each outcome measure, as well as heterogeneity analysis. Analysis of subgroups was performed when necessary, as well as analysis of publication bias.
Results show moderate effect sizes for the outcome measures *Bullying Frequency* and *Victimization Frequency*, *Attitudes* and *School Climate*. Greater impact was observed in interventions of less than one school year duration, as well as those targeting children younger than 10 years. Subgroup analysis confirmed greater heterogeneity in studies evaluating complex interventions.
In general, our results indicate that bullying and violence prevention programs in school settings are obtaining beneficial, albeit discrete, results in the outcome measures evaluated.

1. Introduction

The first studies of school bullying were performed by Heinemann (1972) and Olweus (1978, 1993, 1996, 2005). These authors refer to bullying, as a repeated behavior of persecution and physical, psychological, or moral aggression carried out by one student or group of students against another, with an inequality of power. School violence includes behaviors than can cause physical or emotional harm, ranging from verbal aggression to humiliation, social exclusion, physical harm, and the destruction of property (Allen, 2009; Benbenishty & Astor, 2005), including various categories, such as disruption in the classroom, disciplinary problems, and maltreatment among classmates (Olweus, 1993; Walker, 1995). The causes of this phenomenon are complex and multifactorial, involving psychological aspects (Jiménez-Barbero, Ruiz-Hernández, Llor-Esteban, & Waschgler, 2014), as well as variables related to family functioning, either directly or indirectly, through the choice of deviant friends (Olalla-Cutrín, Gómez-Fraguela, & Luengo, 2015).

We are facing a phenomenon that, although it has been the object of attention and social alarm in the past few decades, has probably always been present in our schools. Diverse studies have analyzed its prevalence in countries like the USA, Spain, Australia, the UK, or Germany, indicating that between 20 and 30% of the students were involved in episodes of violence ranging from simple verbal intimidation to severe forms of physical or sexual aggression (Currie et al., 2008; Defensor del Pueblo [Ombudsman], 2007; Department of Health and Human Services & Center for Disease Control and Prevention, 2006), which can foment the development of early behavior problems and social alienation among adolescents (Rudolph et al., 2013), causing important relational problems during adulthood (Ruiz-Hernández, García-Jiménez, Llor-Esteban, & Godoy-Fernández, 2015).

In recent years, the awareness of this problem has increased, leading to the proliferation of school prevention programs (Jiménez-Barbero, Ruiz-Hernández, Llor-Esteban, & Pérez-García, 2012; Ttofi & Farrington, 2011) as well as studies aimed at assessing their efficacy. In this sense, although several systematic reviews and meta-analyses in this regard have been published (Merrel, Gueldner, Ross, & Isava, 2008; Mytton, DiGuiseppi, Gough, Taylor, & Logan, 2006; Park-Higgerson,

* Corresponding author at: Plaza Vistabella, 17, 2L, CP: 30820 Alcantarilla, Murcia, Spain.
E-mail addresses: barbero49@hotmail.com (J.A. Jiménez-Barbero), jaruiz@um.es (J.A. Ruiz-Hernández), llorzaragoza@ucam.edu (L. Llor-Zaragoza), mariapg78@hotmail.com (M. Pérez-García), bllor@um.es (B. Llor-Esteban).

http://dx.doi.org/10.1016/j.childyouth.2015.12.015
0190-7409/

Perumean-Chaney, Bartolucci, Grimley, & Singh, 2008; Vreeman & Carroll, 2007), the capacity to generalize the results has been limited due to the inclusion criteria: for example, due to the date constraints imposed on the bibliographic searches, the use of evaluations that combine primary and secondary prevention interventions, or because the authors used random clinical trials with non-randomized methodologies.

The goal of our study was to assess the efficacy of anti-bullying school programs. For this purpose, we conducted a meta-analysis of random clinical trials (RCTs), contributing the following improvements with respect to the meta-analyses already carried out on this topic. Firstly, the meta-analyzed studies had to present an experimental design in order to ensure the internal validity of the effect size estimations. Secondly, the range of inclusion years of the studies was updated, extending it to June 2015, taking into account that the most recent meta-analyses of the efficacy of school programs for the prevention of bullying included up to May 2009 (Ttofi & Farrington, 2011). However, statistical random effects models were applied, which are currently considered more appropriate for the integration of the results of empirical studies, due to the variability they usually present (Rosa-Alcázar, Sánchez-Meca, Gómez-Conesa, & Marín-Martínez, 2008).

2. Method

2.1. Inclusion criteria of the studies

The studies were included in the review if they met the following criteria: (a) their goal should be the evaluation of the effectiveness of an intervention program to prevent violence or bullying in the school setting; (b) the assessed interventions had to target the study population directly (students of Primary or Secondary schools), not teachers or parents; (c) the studies should use an experimental design; (d) the studies had to provide the necessary statistical information from the control group (CG) and the intervention group (IG), at least at posttest (means and standard deviation); and (e) the studies should be published between 2000 and 2015.

Only comparisons between outcome measures assessed through standardized tests were included. The outcome measures included in the study were:

1. Bullying or school violence frequency. Including direct (physical or verbal) and indirect aggression.
2. School victimization frequency. Direct or indirect aggression would be taken into account.
3. Favorable attitudes toward bullying or school violence.
4. Attitudes against bullying or school violence.
5. School climate. Students' perceptions of school climate (trust, cooperation, and willingness to help…).

2.2. Search strategy

A systematic search in the following electronic databases was conducted: Medline, Tripdatabase, Cochrane, Academy Search Premier, PsycINFO, ERIC, and PsycARTICLES. The keywords and terms used were: "bullying" OR "*school violence*" AND "*attitudes toward violence* AND "*adolescents*" AND "*intervention*" OR "*prevention program*" AND "*self-esteem*" AND "*empathy*" AND "*school climate*" AND "*victimization*". The search was restricted to works published between 01/01/2000 and 31/05/2015. In all cases, the titles and abstracts were examined, rejecting the works that did not meet the previously defined inclusion criteria. The complete texts of the accepted articles were carefully read, and their lists of bibliographic references were examined in order to identify possible relevant articles that had not been located in initial search.

2.3. Rating criteria of the methodological quality

The studies were selected by two independent reviewers who assessed the methodological quality of the selected studies, observing a high interjudge reliability in the application of the assessment criteria, according to the Pearson correlational analysis of the obtained scores ($r = .86$). The criteria used by the reviewers to carry the assessment were the following:

(A) Score equal to or higher than 6 on a scale of methodological quality elaborated ad hoc from the guide published by the York University in 2001 for the elaboration of systematic reviews (NHS Centre for Reviews and Dissemination, 2001). This scale comprised 10 items: (1) operational definition of the constructs and terms used in the study; (2) adequate sample selection method; (3) adequate sample size; (4) prior distinction of the subgroups or use of appropriate clustering techniques; (5) validity of the assessment (information collected directly by the researchers); (6) reliability of the assessment (use of a validated instrument and/or with a high level of internal consistency to assess the intervention); (7) follow-up of the results; (8) use of outcome measures matching the object of study; (9) adequate statistical analyses; and (10) adequate presentation of the results by means of tables, figures, or similar.

(B) Score equal to or higher than 3 on the scale elaborated by Jadad et al. (1996) for randomized clinical trials. Studies that did not meet either of these two conditions were excluded.

2.4. Tabulation and data analysis

The eligible studies were directly coded on an Excel database data by the first author. The coding was reviewed by the second and third authors, and doubts were resolved through discussion among all the authors. Subsequently, a summary table was created in which the data of each selected study was recorded according to the following categories: date and country of study, goal of the investigation, name of the prevention or intervention program assessed, size and age of the sample used, and significant findings (Table 1).

2.5. Statistical analysis

The effects of the characteristics of 14 school programs to prevent or reduce violence or bullying in the school setting were assessed by means of comparison of the results of the IG and those of the CG at posttest, due to the fact that various studies included in the meta-analysis did not report the pretest results. The variable randomization (individual, class, or school) was not taken into account due to the small number of available studies for the comparisons. The total size of the effect was calculated for each comparison by means of RevMan 5.2 (Cochrane Centre). Due to the variability foreseen in the individual effect sizes, a random effects model was used. The differences between the IG and the IC for each comparison were grouped to obtain the estimation of the total effect. The grouped results were expressed as standardized mean differences (SMD) with a 95% confidence interval. As the studies included in the meta-analysis used different instruments to assess the outcome measures, we used SMD to standardize all results to a common scale, measured in units of standard deviation. As meta-analyses summarize the results through studies with different methods, topics, and results, it is important to include a measure of heterogeneity to determine whether the variation observed among the results in the studies is greater than that which could be expected by chance, so we used chi-square test with a significance level of .05 and calculated the I^2 index. For cases with significant heterogeneity, we conducted an analysis of subgroups according to the following moderating variables: (a) *year of publication of the study* (studies published between 2000 and 2007 vs. studies published after 2007); (b) *Sample size*