

## Preventing prejudice and improving intergroup attitudes: A meta-analysis of child and adolescent training programs ☆☆☆



Andreas Beelmann\*, Kim Sarah Heinemann

Department of Research Synthesis, Intervention, and Evaluation, Institute of Psychology, Friedrich-Schiller-University of Jena, Germany

### ARTICLE INFO

#### Article history:

Received 29 August 2012

Received in revised form 1 November 2013

Accepted 7 November 2013

Available online 4 January 2014

#### Keywords:

Prejudice  
Intergroup attitudes  
Meta-analysis  
Intervention  
Children  
Adolescents

### ABSTRACT

This article reports the results of a meta-analysis of 81 research reports containing 122 intervention–control comparisons of structured programs to reduce prejudice or promote positive intergroup attitudes in children and adolescents. Overall, the analysis revealed a mean effect size of  $d = 0.30$ , indicating low to moderate intervention effects. From the great variety of different approaches, interventions based on direct contact experiences along with social-cognitive training programs designed to promote empathy and perspective taking showed the strongest effect sizes. In addition, effects varied according to the program participant's social status (higher effects for majority groups), the target out-group (lower effect sizes for ethnic vs. disabled and aged out-groups), and the type of outcome assessment (higher effects for cognitive vs. affective and behavioral measures of intergroup attitudes). The discussion considers several limitations including the lack of implementation and follow-up research as well as future direction of research on promoting intergroup relations.

© 2013 Elsevier Inc. All rights reserved.

Prejudice and other forms of negative intergroup attitudes cause serious social problems in many societies throughout the world. Even in their mildest form, the consequences for human beings may include social exclusion and segregation, health and behavioral problems, poorer chances on the labor market, and even a generally more negative quality of life (see, e.g., Paradies, 2006; Williams, Neighbors, & Jackson, 2003). Previous social-developmental research has indicated clearly that prejudice and other forms of biased intergroup attitudes are not just restricted to adulthood, but that children start to favor their own social group as soon as the basic processes of social categorization and identification emerge in early childhood (see Levy & Killen, 2008; Raabe & Beelmann, 2011).

Although definitions of prejudice have changed over the last decades (see Duckitt, 2010), it can generally be viewed as “any attitude, emotion, or behavior toward members of a group, which directly or indirectly implies some negativity or antipathy toward that group” (Brown, 2010, p. 7). According to this definition, prejudice may have different manifestations and is multifaceted in nature. On a global level, one can first differentiate between in-group preference (e.g., ascribing more positive characteristics to our own social group) and out-group derogation (ascribing more negative characteristics to the social group we do not belong to). These seem to be distinct constructs, although they have

common negative consequences for out-group members (Brewer, 1999). In addition, negative attitudes also have a multifaceted character and are normally divided into a cognitive component (attributing negative characteristics such as being mean or aggressive to members of the out-group), an emotional component (e.g., liking them less), and a behavioral component (e.g., exhibiting negative behavior such as social exclusion). In line with the broad range of different operationalizations, recent social-developmental research has studied prejudice within a broader framework of intergroup attitude development. This has led to the identification of important individual and social developmental factors and processes. For example, it has been proposed that cognitive and social-cognitive abilities such as classification skills and perspective taking (Aboud, 1988; Bigler & Liben, 2007), the development of social identity (Nesdale, 2004), or moral decision making (Killen & Rutland, 2011) are crucial contributors to intergroup attitude development in children. Other research has focused more on social variables such as intergroup contact and cross-group friendships (Davies, Tropp, Aron, Pettigrew, & Wright, 2011), social norms (Rutland, 2004), and intergroup threat (Bar-Tal & Teichman, 2005; Riek, Mania, & Gaertner, 2006).

However, although all this research documents major scientific progress in understanding the emergence of prejudiced attitudes, there is an ongoing debate over which are the most important individual and social factors that need to be addressed in systematic interventions designed to prevent negative intergroup attitudes and associated problems of intolerance and discrimination in intergroup relations (Killen, Rutland, & Ruck, 2011; Tropp & Mallett, 2011). As a result, rather diverse intervention programs have been developed and tested during the last 30 years (e.g., Oskamp, 2000; Stephan & Stephan, 2001). The present meta-analysis summarizes research on

☆ This research was supported by the German Research Foundation Grant BE 3731/2-1.

☆☆ We thank Jonathan Harrow for native-speaker advice.

\* Corresponding author at: Department of Research Synthesis, Intervention, and Evaluation, University of Jena, Humboldtstrasse 26, DE-07743 Jena, Germany. Tel.: +49 3641 945901; fax: +49 3641 945902.

E-mail addresses: andreas.beelmann@uni-jena.de (A. Beelmann), kim.heinemann@uni-jena.de (K.S. Heinemann).

the effectiveness of standardized psychological and educational intervention programs to prevent or reduce prejudice or otherwise promote positive intergroup attitudes and relations in childhood and adolescence. It studies the general effectiveness as well as specific effects indicating which program will be the best alternative for whom and on what outcome measure.

Interventions aiming to promote intergroup attitudes and relations have a long history and can be traced back to the integrative school systems movement in the United States (Brown v. Board of Education, 1954) and to Allport's (1954) classic work on the psychological nature of prejudice. Both initiatives generated a great deal of empirical research (see Pettigrew & Tropp, 2006; Schofield, 1995; Schofield & Hausmann, 2004), particularly in the social context of schools and universities, and this has led to a variety of intervention concepts. Other approaches place more emphasis on education in political values (e.g., democracy), multicultural diversity, morality, and values such as equality, respect, and tolerance (Stephan & Vogt, 2004). Despite their different histories and disciplinary roots, all these approaches are based on the assumption that prejudice and negative intergroup attitudes are associated with a number of social problems and phenomena within our societies such as the social exclusion and discrimination of ethnic minorities, immigrants, the handicapped, or other groups; racism, hate crimes, and right-wing extremism; or even international conflicts. Therefore, preventing prejudice and negative intergroup attitudes is assumed to be an appropriate tool for avoiding or at least reducing these problems and generally promoting more positive intergroup relations, tolerance, and justice within our societies.

In an earlier review of this research with children and adolescents, Aboud and Levy (2000) distinguished between five groups of interventions ranging from integrated schooling, bilingual education, multicultural education, promoting social-cognitive skills, to training in role-playing and empathy. They classified these as being based on three different theoretical concepts about intergroup attitudes (see also Killen et al., 2011). The first most prominent theoretical concept is founded on the *intergroup contact hypothesis* (Allport, 1954; Pettigrew & Tropp, 2011). This states that contact between members of different social groups leads to less prejudiced attitudes and generally promotes positive intergroup relations. This effect of intergroup contact is even more pronounced when the situation complies with the criteria of positive contact, that is, an equal status of members of both social groups within the situation, support for the contact from authorities, and cooperation in achieving a joint goal. The intergroup contact hypothesis is one of the best supported theories in social psychology with hundreds of studies showing that the effects apparently hold for a variety of different situations, settings, and samples (see Paluck & Green, 2009; Pettigrew & Tropp, 2006, for reviews). Well-known examples of this intervention type are integrated schooling, cooperative learning techniques, or even public campaigns with high-profile individuals from the social out-group. In addition, recent extensions of intergroup contact theory have led to programs based on the idea of extended, that is, indirect contact—for example, when someone gets to know an in-group member who is in contact with an out-group member (see Wright, Aron, McLaughlin-Volpe, & Ropp, 1997).

The second theoretical construct refers to *general socialization theory* and *social learning theory*. This assumes that social experiences, knowledge, and information about the social world and social out-groups as well as knowledge about relevant concepts in the context of intergroup attitudes (e.g., information about prejudice, democratic values, and cultural diversity) lead to more positive intergroup relations. Multicultural or antibias training programs are good examples for these knowledge- and information-based interventions.

Finally, the third theoretical construct is *social-cognitive developmental theory*. In general, it assumes that children's intergroup attitudes often reflect their stage in sociocognitive development and that the increasing development of distinct sociocognitive abilities leads to a decreasing tendency to have biased attitudes. This forms the basis for

trainings in cognitive and social-cognitive skills (e.g., classification skills, social categorization, perspective taking, conflict resolution, moral decision making) that interrelate with the development of intergroup attitudes.

Aboud and Levy's (2000) classification of interventions into three different theoretical concepts is, of course, neither independent nor exhaustive. Because they are grounded in diverse scientific disciplines such as social and developmental psychology, multicultural education, sociology of migration and diversity, criminal justice, or politics, there are several ways to classify interventions in this field. For example, Paluck and Green (2009) recently presented an extended review of prejudice reduction interventions that distinguished between categories such as cooperative learning, entertainment strategies using books or films, discussion and peer influence, and instruction. Other authors such as Killen et al. (2011) have differentiated between interventions to promote intergroup attitudes by influencing peer relations, adult-child interaction, and social-cognitive judgments. In a recently published article, Aboud et al. (2012) differentiated between contact-based, media/instruction-based, and antibias/multicultural interventions. These and other examples illustrate that interventions in this field could possibly be best characterized as a mixture of programs with diverse underlying theories, goals, contents, strategic concepts, and intervention methods.

Perhaps as a consequence of this diversity in classifying intergroup interventions, the outcomes of systematic evaluations in this field have revealed no clear empirical consensus on what should be viewed as the best way to prevent prejudice and promote positive intergroup attitudes among children and adolescents (Aboud & Levy, 2000; Paluck & Green, 2009). However, several meta-analyses and systematic reviews point to some promising approaches. For example, Pettigrew and Tropp's (2006) comprehensive meta-analysis on the *intergroup contact hypothesis* showed that programs based on contact between members of different social groups seem to be of central value in influencing intergroup attitudes independent of age, type of attitudes, cultures, and social conditions. Aboud et al. (2012) confirmed these results even for young children under the age of 8. Likewise, several educational techniques seem to offer promising ways of reducing prejudice. For example, intensive studies on the effects of *cooperative learning techniques* (Johnson & Johnson, 1989, 2000; Roseth, Johnson, & Johnson, 2008) have shown that they markedly increase not only school and academic achievement but also interpersonal attraction and relations between members of different ethnic groups in the classroom compared with individual and competitively oriented learning strategies.

Other approaches such as *multicultural, diversity, and antiracism training programs* have a somewhat more limited empirical foundation, but also seem to produce low to moderate, but nonetheless positive, effect sizes on prejudice—at least in adult populations (Paluck & Green, 2009; Stephan, Renfro, & Stephan, 2004; Verkuyten, 2010). Empirical evaluations of *cognitive and social-cognitive programs* are also basically positive, showing that a reduction in biased judgment processes and prejudice can be achieved (Aboud & Levy, 2000; Paluck & Green, 2009). Nonetheless, especially in this field, any final conclusion is difficult to reach, because this label covers a wide variety of different approaches such as programs to improve cognitive skills (e.g., Bigler & Liben, 1992), social perspective taking and empathy (e.g., bystander interventions, see Stephan & Finlay, 1999), moral development (e.g., value self-confrontation, see Grube, Mayton, & Ball-Rokeach, 1994), and conflict resolution (Johnson & Johnson, 1996).

In sum, a variety of measures have been applied to improve intergroup relations and prevent prejudice. However, most of the aforementioned and other reviews (see Beelmann, 2009) do not focus specifically on psychological and educational programs with children and adolescents (Paluck & Green, 2009; Pettigrew & Tropp, 2006), are limited to ages 8 and below and do not use a meta-analytic approach (Aboud et al., 2012), are narrative in nature, or somewhat outdated (Aboud &

Levy, 2000; Pfeifer, Spears Brown, & Juvonen, 2007). Therefore, the present meta-analysis aims to integrate existing research on the effectiveness of structured intervention programs designed to promote positive intergroup attitudes and prevent prejudice in children and adolescents—*independent of the underlying theoretical concept (e.g., contact- or knowledge/information-based, promoting individual social-cognitive competencies).*

Alongside the question of the general effectiveness of the selected programs, there are also several differential research questions. First, because different intervention concepts exist, we are naturally interested in the specific effects of different types of programs such as multicultural trainings or training in moral reasoning, different intervention components such as information about the out-group versus teaching empathy, and the influence of implementation parameters such as the number of sessions or the type of administrator. Second, because Raabe and Beelmann (2011) found a characteristic developmental course of ethnic prejudice along with significant moderators of this course, we shall analyze several possible moderators of effectiveness including age, social status (majority vs. minority), and target out-group (ethnicity, disabilities, etc.). Finally, because prejudice and intergroup attitudes are clearly multifaceted constructs (Rutland, Killen, & Abrams, 2010) with divergent operationalization and assessment strategies, we shall also analyze the effectiveness of the programs in terms of different measures of intergroup attitudes and prejudice.

## Method

### Study selection

Primary studies were selected according to the following eligibility criteria: First, the study had to evaluate a standardized psychological or educational intervention program aimed at reducing prejudice or otherwise improving intergroup attitudes and relations on the basis of intergroup contact, information/knowledge acquisition, or promotion of individual social-cognitive competencies. The main feature for selecting programs was their standardization, that is, there had to be at least a recognizable stepwise order of different units. Such units do not necessarily mean different program sessions, but rather different components or methodologies used in a structured manner. For this reason, we did not include interventions based solely on contact between members of an in-group and out-group with no further components (such as structured discussions about contact experience). Likewise, we excluded research on broadband educational approaches (such as integrated schooling) or special learning methods (e.g., cooperative learning techniques), because they are not constructed as standardized and structured intervention programs. In addition, contact-based interventions and cooperative learning techniques have already been investigated and summarized sufficiently well in recent years (see Pettigrew & Tropp, 2006; Roseth et al., 2008). Second, the study had to contain an intervention group and at least one untrained control group. Third, the age of the youngsters had to be below 18 years. Fourth, the outcome assessment had to contain some kind of intergroup attitude measure (e.g., prejudice, in-group favoritism, attitudes toward discriminating behavior, or tolerance). Fifth, results had to be reported in a manner that permits a reliable calculation of effect sizes. Finally, sixth, research reports had to be written in English, German, French, Spanish, Italian, or another major European language.

### Literature search

We applied several strategies to identify relevant evaluation studies. First, we examined the reference lists of existing reviews and meta-analyses on the modification of intergroup attitudes and on intervention programs to reduce or prevent prejudice in children and adolescents (Aboud & Levy, 2000; Graves, 1999; Levy, 1999; Levy & Killen, 2008; McGregor, 1993; Paluck & Green, 2009; Pettigrew & Tropp,

2006; Pfeifer et al., 2007; Stephan & Finlay, 1999; Stephan et al., 2004). Second, we intensively scanned electronic databases such as PsycINFO, ERIC, PSYNDEX, and ProQuest Dissertations and Theses with the combination of three groups of keywords with and without truncations (\*): (a) intervention/prevention/evaluation/effect\*/training, (b) child\*/adolescen\*, and (c) intergroup attitud\*/prejudic\*/ingroup bias/discriminat\*/tolera\*. Third, we analyzed the references given in previously identified primary studies for further relevant publications (cross-referencing). Finally, we conducted an Internet search to look for additional promising evaluation studies.

In the first round, these strategies yielded a total of 214 research reports that we considered to be relevant. We excluded 71 reports (33.1%) because they did not meet our first eligibility criterion on the program character. We dropped further 31 reports (15.4%) because a control group was missing (second eligibility criterion); and 10 reports (4.7%), because the sample was older than 18 years of age (third eligibility criterion). Finally, we excluded 15 reports (7.0%) because of missing data for the effect-size calculation and another 6 reports (2.8%) that could not be retrieved. Thus, we ended up with 81 research reports that met our eligibility criteria. Of these 81 reports, only slightly more than one-half (56%) had already been integrated in the above-mentioned reviews and meta-analyses with the highest coverage rates being 56% (Paluck & Green, 2009) and 24% (Pettigrew & Tropp, 2006).

Some of the research reports contained more than one study and/or more than one intervention or control group, therefore permitting multiple intervention–control group comparisons. As a result, our final database (and unit of analysis) for this meta-analysis was 122 intervention–control group comparisons.

### Coding and computation of effect sizes

The coding scheme contained characteristics of the publications (e.g., author, year of publication), intervention programs (e.g., target out-group, content component, program parameter, duration), sample and participants (e.g., age, gender, social status), methods (e.g., design, sample size), outcome assessment (e.g., content and type of dependent measures), and, finally, study results (see Tables 1–4 for a detailed list). The second author and a trained advanced student used this scheme to code these reports and comparisons. They met periodically to discuss and clarify critical coding decisions. In addition, 30 intervention–control group comparisons (24.6%) were randomly selected and coded by an independent rater to calculate the interrater reliability of the coding sheet.

**Table 1**  
Description of the 81 research reports.

Characteristics	Coding	n	%
Publication year	1958–1970	4	4.9
	1971–1980	23	28.4
	1981–1990	18	22.2
	1991–2000	20	24.7
	2001–2010	16	19.8
Publication type	Journal article	51	63.0
	Dissertation	26	32.0
	Book chapter	2	2.5
	Conference paper	2	2.5
Publication language	English	78	96.3
	German	2	2.5
	French	1	1.2
Country of intervention	United States	60	74.1
	Germany	6	7.4
	Canada	4	4.9
	Great Britain	4	4.9
	Israel	3	3.7
	Australia	1	1.2
	Finland	1	1.2
	Ireland	1	1.2
	Portugal	1	1.2

Note. n = number of research reports.

**Table 2**  
Description of the 122 intervention–control group comparisons.

Characteristics	Coding	<i>k</i>	%	<i>d</i>	<i>d<sub>w</sub></i>
Program					
Target out-group based on...	Ethnicity	75	61.5	0.26	0.25 <sup>a</sup>
	Disability	30	24.6	0.45	0.41 <sup>a</sup>
	Age (elderly)	12	9.8	0.41	0.38
	Combination of out-groups	5	4.1	0.43	0.46
Content components	Knowledge acquisition/information on...	93	77.2	0.33	0.32
	Out-group	81	66.4	0.30	0.29
	Prejudice and intergroup attitudes	26	21.3	0.26	0.27
	Values and norms	14	11.5	0.33	0.30
	Intergroup contact	66	54.1	0.33	0.31
	Indirect contact	44	36.1	0.26	0.26
	Direct contact	22	18.0	0.48	0.43
	Social-cognitive training in...	45	36.9	0.40	0.33
	Perspective taking/empathy	24	19.7	0.50	0.44
	Classification/social categorization	18	14.8	0.16	0.14 <sup>a</sup>
	Problem-solving skills	11	9.0	0.20	0.18 <sup>a</sup>
	Moral development	9	7.4	0.36	0.32
	Social skills	7	5.7	0.50	0.39
	Other cognitive skills	4	3.3	0.22	0.21 <sup>a</sup>
Program parameter	Active trainer	104	85.2	0.35	0.31
	Printed materials	91	74.9	0.29	0.27
	Group discussions	86	70.5	0.31	0.28
	Curriculum implementation	62	50.8	0.35	0.34
	TV material	46	37.7	0.32	0.31
	Role taking/role play	33	27.0	0.42	0.35
	Manual used	20	16.4	0.21	0.23
Duration	≤4 weeks	39	32.0	0.39	0.27
	5–8 weeks	32	26.2	0.25	0.21 <sup>ab</sup>
	9–12 weeks	11	9.0	0.31	0.31
	≥13 weeks	25	20.5	0.35	0.32
	Not specified	15	12.3	0.33	0.39 <sup>b</sup>
Number of sessions	≤5	36	29.5	0.38	0.37 <sup>a</sup>
	6–10	35	28.7	0.26	0.25 <sup>a</sup>
	11–15	7	5.7	0.23	0.19 <sup>ab</sup>
	16–20	10	8.3	0.36	0.26
	>20	8	6.6	0.56	0.52 <sup>b</sup>
	Not specified	26	21.3	0.30	0.29
Intensity <sup>b</sup>	Low (<5 h)	50	41.0	0.29	0.28
	Medium (5–10 h)	19	15.6	0.37	0.33
	High (>10 h)	27	22.1	0.41	0.34
	Not specified	26	21.3	0.30	0.29
Trainer	Teachers	66	54.1	0.30	0.30
	Study authors	22	18.0	0.24	0.20 <sup>a</sup>
	Research associates/supervised students	20	16.4	0.56	0.39 <sup>a</sup>
	Not specified	14	11.5	0.29	0.32
Subjects					
Mean age (years; months)	3; 6–7; 11	35	28.7	0.30	0.28
	8–9; 11	29	23.8	0.42	0.35
	10–13; 11	28	23.0	0.30	0.30
	14–18	30	24.6	0.31	0.30
Gender (% male)	0–49	34	27.9	0.28	0.26
	50	54	44.2	0.38	0.35
	51–73	34	27.9	0.29	0.27
Ingroup	Majority	85	69.7	0.34	0.32
	Minority	4	3.3	0.15	0.17 <sup>a</sup>
	Mixed/not specified	33	27.0	0.33	0.29
Methodology					
Design quality	Post-control	8	6.6	0.40	0.40
	Pre-post-control	35	28.7	0.33	0.29
	Pre-post-control with demonstrated comparability	70	57.4	0.32	0.29
	Pre-post-control with randomization and demonstrated comparability	9	7.4	0.37	0.32
Sample size	<50	42	34.4	0.42	0.39 <sup>a</sup>
	50–99	37	30.3	0.26	0.25 <sup>a</sup>
	100–300	31	25.4	0.31	0.31
	>300	12	9.8	0.27	0.27

Note. *k* = number of comparisons. *d* = unweighted effect size. *d<sub>w</sub>* = weighted effect size based on random effects model. Weighted mean effect sizes within the same characteristic category sharing a common subscript differed significantly from each other. A positive effect size indicates a successful intervention (i.e., reduced prejudice or improved intergroup attitudes). In cases of a missing mean age, the age group was specified according to the school grade (e.g., 1st grade = 5 to 6 years) or the mean of a specified age range (e.g., 7 to 11 years = mean age of 9.5).

<sup>a</sup> Effect size did not differ significantly from zero ( $p > .10$ ).

<sup>b</sup> Intensity = session duration × number of sessions.

Overall, this interrater reliability was satisfactory ( $M = 93.2\%$ ), ranging from 78.3% (program parameter) to 100% (e.g., country of conduct, mean age of the sample).

We used Cohen's (1988) *d* to compute unified effect sizes. When relevant data were available, we computed the effect sizes as the difference between the pre- and posttest scores in the intervention and

**Table 3**  
Coding of outcome assessment with examples of measurement instruments.

Characteristics	Categories	Definition and example
Intergroup attitudes dimension	Cognitive	Stereotypes or characteristics of the out-group (e.g., PRAM II, MRA)
	Affective	Emotional indicators such as sympathy, fear, or threat (e.g., Liking board)
	Behavioral	Intergroup behavior (discrimination) or behavior intention (e.g., PATHS, IBM)
Type of intergroup attitude measure	Prejudice	Negative dimension of intergroup attitudes (e.g., PRAM II)
	In-group favoritism	Variables favoring the in-group on a positive dimension (e.g., Liking board)
	Tolerance	Tolerant attitudes toward out-group members (e.g. F-Scale)
	Social distance	Variables measuring the distance to social interaction (e.g. PATHS)
	Knowledge	Knowledge about the out-group including stereotypes (e.g., KMH);

Note. PRAM = Preschool Racial Attitude Measure II (Williams, Best, Boswell, Mattson, & Graves, 1975); MRA = Multi-racial attitude measure (Doyle & Aboud, 1995); Liking board (Aboud, 1988); F-Scale (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950); PATHS = Peer Attitudes Toward the Handicapped Scale (Bagley & Greene, 1981); IBM = Intended Behavior Measure (Cameron et al., 2006); KMH = Knowledge of Mental Handicap (McConkey, McCormack, & Naughton, 1983).

the control group divided by the pooled standard deviation of the intervention and control group at the pretest. When only posttest scores were available, we used these to compute effect sizes. When no means and standard deviations had been reported, we used recomputation and effect-size estimation techniques (see Lipsey & Wilson, 2001). When the research reports mentioned nonsignificant results with no further details, we counted these as zero effects. By using this quite conservative procedure, we avoided any overestimation of intervention effects. Overall, we computed 268 individual effect sizes of which 134 (50.0%) were computed via means and standard deviations, 110 (41.0%) were recomputed via test statistics, and 24 (9.0%) were coded as zero effects because results were not significant.

*Integration and statistical analysis*

Based on these 268 individual effect sizes, we computed the intervention–control group (comparison) effect size in a two-step

**Table 4**  
Outcome measure characteristics on the dependent variable level ( $n_{ES} = 266$ ) and the comparison level ( $k = 122$ ).

Characteristics	Categories	Dependent variable level				Comparison level			
		$n_{ES}$	%	$d$	$d_w$	$k$	%	$d$	$d_w$
Intergroup attitudes dimension <sup>a</sup>	Cognitive	162	60.9	0.32	0.29 <sup>a</sup>	107	87.7	0.38	0.33 <sup>a</sup>
	Affective	31	11.7	0.05	0.05 <sup>ac</sup>	14	11.5	0.06	0.03 <sup>ac</sup>
	Behavioral	66	24.8	0.23	0.19 <sup>a</sup>	43	35.2	0.27	0.20 <sup>a</sup>
	Not specified <sup>b</sup>	7	2.6	0.21	0.28 <sup>c</sup>				
Type of intergroup attitude <sup>a</sup>	Prejudice	140	52.6	0.36	0.31 <sup>a</sup>	94	77.0	0.38	0.35 <sup>ab</sup>
	In-group favoritism	47	17.7	0.01	0.03 <sup>abc</sup>	25	20.5	0.09	0.10 <sup>acc</sup>
	Tolerance	21	7.9	0.23	0.19 <sup>b</sup>	14	11.5	0.33	0.28 <sup>c</sup>
	Social distance	36	13.5	0.21	0.22 <sup>c</sup>	24	19.7	0.20	0.20 <sup>b</sup>
	Knowledge	9	3.4	0.48	0.54 <sup>abcd</sup>	9	7.7	0.44	0.46
	Not specified <sup>b</sup>	13	4.9	0.24	0.17 <sup>cd</sup>				
Type of instrument	Questionnaires	149	56.0	0.36	0.32 <sup>a</sup>	82	67.2	0.38	0.35 <sup>a</sup>
	Tests	99	37.2	0.15	0.13 <sup>a</sup>	33	27.0	0.24	0.20 <sup>a</sup>
	Interview	16	6.0	0.16	0.16 <sup>c</sup>	8	6.6	0.17	0.16 <sup>c</sup>
	Observation	2	0.8	0.00	0.00 <sup>f</sup>	2	1.6	0.00	0.00 <sup>ac</sup>
Stimulus material for category activation	Lexical	151	56.8	0.24	0.22 <sup>a</sup>	82	67.2	0.26	0.25
	Visual	60	22.6	0.24	0.22 <sup>b</sup>	28	23.0	0.41	0.35
	Lexical and visual	20	7.5	0.35	0.29	12	9.8	0.41	0.33
	Not specified <sup>b</sup>	35	12.8	0.38	0.34 <sup>ab</sup>				

Note.  $n_{ES}$  = number of effect sizes.  $k$  = number of comparisons.  $d$  = unweighted effect size.  $d_w$  = weighted effect size based on the random effects model. Weighted mean effect sizes within the same characteristic category sharing a common subscript differed significantly from each other. A positive effect size indicates a successful intervention (i.e., reduced prejudice or improved intergroup attitudes).

<sup>a</sup> Explanations on coding categories are explained in Table 3.

<sup>b</sup> Not specified measures were not integrated on comparison level.

<sup>c</sup> Effect size did not differ significantly from zero ( $p > .10$ ).

procedure: First, we calculated each individual effect size within studies by identifying the mean effect for each outcome measure used. Second, we calculated the mean comparison effect by averaging effect sizes across the different outcome measures within each comparison. The mean effect sizes across intervention–control group comparisons were calculated according to methods proposed by Hedges and Olkin (1985). These included weighting effect sizes by the inverse of the sampling error and performing subsequent homogeneity analyses in order to analyze effect-size variance (i.e., whether it variance exceeded sampling error). When effect sizes were homogeneous, we applied the fixed-effects model. When the distribution of the effect sizes remained heterogeneous, we used the random-effects model (see Lipsey & Wilson, 2001, for details).

**Results**

*Description of research reports and comparisons*

The 81 research reports were published between 1958 and 2010 and mostly written in English ( $k = 78, 96.3\%$ ). The majority of studies was conducted in the United States ( $k = 60, 74.1\%$ ) and published in scientific journals ( $k = 51, 63.0\%$ ). A further description of the 81 research reports is given in Table 1.

The 122 intervention–control group comparisons contained a great diversity of intervention programs and samples (see Table 2 for details). The majority of interventions addressed prejudice against and intergroup attitudes toward persons of other ethnicities ( $k = 75, 61.5\%$ ) followed by persons with disabilities ( $k = 30, 24.6\%$ ) and the aged ( $k = 12, 9.8\%$ ). Only five interventions (4.1%) addressed different out-groups at the same time (two interventions targeted ethnicity and disability, three targeted all three out-groups). When coding the intervention programs, prior research has indicated that it might be necessary not only to describe programs on a single dimension but also to distinguish between different program components and parameters in order to gain an appropriate description of the interventions. Therefore, we coded studies according to the primary *theoretical orientation* of the program (i.e., socialization/knowledge acquisition, intergroup contact theory, social-cognitive development theory). Within these category groups, we then coded the programs according to their specific *content* (e.g., information on the out-group, on prejudice and intergroup attitudes, or on norms and values; direct or indirect contact; different

kinds of social-cognitive skill promotion). In addition, we distinguished the programs according to several *program parameters* (e.g., methods such as group discussion and role play, materials used, use of intervention manuals, or implementation within a regular curriculum). Table 2 shows that 93 of the 122 interventions (76.2%) were based theoretically on socialization/knowledge acquisition (by using mostly information on some social out-group), 66 (54.1%) on intergroup contact theory (indirect and direct contact), and 45 (36.9%) on social-cognitive developmental theory (mostly with perspective taking/empathy training as content). A total of 74 interventions (60.7%) used a combination of these approaches—the most frequent being a combination of socialization/knowledge acquisition and intergroup contact ( $k = 42$ , 36.1%) followed by a combination of socialization/knowledge acquisition and social-cognitive development ( $k = 18$ , 14.8%). The combination of intergroup contact and social-cognitive development was used in six interventions (4.9%). Eight studies tested the results of interventions combining all three elements (6.6%). Of the *implementation components*, the most frequently used methods were group discussion ( $k = 86$ , 70.5%) and printed materials ( $k = 91$ , 74.6%).

The program duration ranged from a single meeting to 44 weeks of intervention with the majority lasting 8 weeks or less ( $k = 71$ , 58.2%). The number of sessions also showed a broad range from one single session ( $k = 9$ , 7.4%) to 44 sessions ( $k = 2$ , 1.6%), although most programs contained 10 sessions or less ( $k = 71$ , 58.2%). Combined with the duration of each session, the majority of programs were of only low to medium intensity ( $k = 69$  or 56% of programs lasted up to 10 h) with a relatively high number of studies with an unspecified dosage.

The mean age of the participants varied from 3 years 6 months to 17 years 3 months ( $M = 10.2$  years,  $SD = 3.67$ ). However, there was a clear focus on elementary school children aged 7 to 10 years (see Table 2). A mixed sample of boys and girls was used in 44.2% of the comparisons ( $k = 54$ ), and in 69.7% of the comparisons, the participants were majority children and adolescents ( $k = 85$ ), whereas only four comparisons addressed minority groups ( $k = 33$  comparisons included mixed samples or delivered no information on the social status of the sample).

The study design quality ranged from post-control-only studies to randomized experiments (see Table 2), with the highest frequency for pre-post-control studies with demonstrated comparability between experimental and control group ( $k = 70$ , 57.4%). Sample sizes ranged from 16 to 1,480 subjects with 64.8% of the samples containing 100 subjects or less ( $k = 79$ ).

#### Overall intervention effects

Post-intervention comparison effect sizes ( $k = 122$ ) ranged between  $d = -0.78$  and  $d = 2.14$  with 8.2% of the comparisons having negative effect sizes ( $k = 10$ ). The unweighted overall mean comparison effect size was  $d = 0.33$  ( $k = 122$ ). When the fixed-effects model (FEM) was applied, the overall effect became slightly smaller ( $d = 0.27$ ). Because the FEM revealed significant heterogeneity,  $Q(df = 121) = 345.39$ ,  $p < .001$ , the random-effects model (REM) seemed to be more appropriate for our data, especially in view of the diversity of the integrated interventions. According to the REM, the overall effect was  $d = 0.30$  ( $k = 122$ ). However, the mean overall effect differed significantly from zero ( $p < .001$ ) for both the FEM and the REM.

#### Publication bias

A potential threat to the interpretation and validity of any meta-analytic result is the problem of publication bias (Beelmann & Lipsey, *in press*; Rothstein, 2008). Publication bias refers to the tendency for only studies reporting a significant result to actually get published. The reasons for such a bias may be manifold (e.g., publication guidelines of journals, researchers' opinions about "good," i.e., significant data), but

the effect is the same: an over-representation of significant results in published studies and therefore an overestimation of effect sizes.

To address the issue of potential publication bias, we used the funnel plot technique to display the distribution of unweighted contrast effect sizes by sample size (Light & Pillemer, 1984). Fig. 1 shows the distribution of the effect sizes by their standard errors that serve as an indicator for the precision of the effect-size estimation. Because the funnel plot was not fully symmetrical, as indicated by the 95% confidence interval around the mean weighted effect size, we conducted a regression analysis using the sample size as the independent variable and the unweighted effect sizes as the dependent variable (Macaskill, Walter, & Irwig, 2001). This allowed an informed testing of the asymmetry of the funnel plot. The regression slope ( $\beta = -.10$ ,  $p = .25$ ) did not differ significantly from zero, indicating that there was no significant prediction from sample size to effect size. This cast doubt on the assumption that the observed mean effect size in this meta-analysis was distorted due to publication bias—possibly because of the relatively high number of unpublished reports (e.g., dissertations, conference paper) in this dataset (see Table 1). Indeed, a comparison between results based on published versus unpublished reports revealed at least a statistical trend in favor of published results,  $d = 0.34$ ,  $k = 82$  for published, and  $d = 0.23$ ,  $k = 40$  for unpublished reports;  $Q(df = 1) = 2.66$ ,  $p = .10$ , using the REM.

#### Effects on different intergroup attitude measures

The studies used a great variety of outcome measures to assess intervention effects. Table 3 reports the definitions and most important codings of the outcome measures according to the intergroup attitudes dimension and the type of intergroup attitude measure. Table 4 presents an overview of the effect sizes based on these and additional codings of the outcome assessment on the dependent variable level ( $n_{ES} = 266$ ) and comparison level ( $k = 122$ ) on which effect sizes of the same category were averaged *within* comparisons before being integrated between comparisons. Studies focused mostly on the cognitive dimension of intergroup attitudes ( $n_{ES} = 162$ ,  $k = 107$ ; 60.9 and 87.7%, respectively), and prejudice measures were the main type of outcome ( $n_{ES} = 140$ ,  $k = 94$ ; 52.6 and 77.0%, respectively). Effects were assessed mainly by questionnaires ( $n_{ES} = 149$ ,  $k = 82$ , 50.6 and 67.2%, respectively), and the stimulus material addressing the social out-group was activated most often lexically ( $n_{ES} = 151$ ,  $k = 82$ ; 56.8 and 67.2%, respectively).

Because categories were, of course, not independent on the comparison level (i.e., the same comparison was usually represented in more than one category), we ran the moderator analyses at the dependent variable level. For all moderator analyses, we selected the REM of effect-size integration because of the aforementioned heterogeneity of the overall effect size. These analyses revealed that the *intergroup*

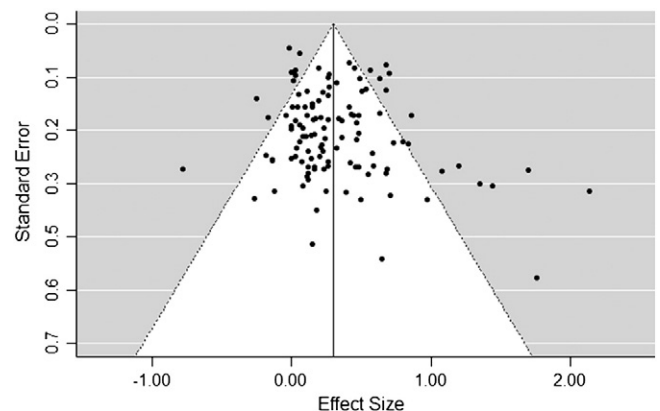


Fig. 1. Funnel plot of study effect sizes by standard error (with 95% confidence interval around the mean weighted effect size).

*attitudes dimension* was a significant moderator,  $Q(df = 3) = 11.41$ ,  $p = .009$ , with outcome measures focusing on the cognitive dimension of intergroup attitudes yielding significantly higher effect sizes than outcome measures focusing on the behavioral and affective dimensions. In particular, outcomes on the affective dimension of intergroup attitudes showed effect sizes that did not differ significantly from zero on both analytic levels. Also, the *type of intergroup attitude* was a significant moderator,  $Q(df = 5) = 33.25$ ,  $p < .001$ , with knowledge about the out-group and prejudice measures yielding the highest, and measures for in-group bias the lowest and nonsignificant effect sizes (see Table 4). Finally, the *type of instrument* used for the outcome measure accounted significantly for effect-size variability,  $Q(df = 4) = 19.55$ ,  $p < .001$ , with questionnaires yielding significantly higher effect sizes than test scores. The *stimulus material* used for the activation of the out-group category was not a significant moderator,  $Q(df = 3) = 3.70$ ,  $p = .30$ .

#### Moderator analyses on program and sample characteristics

Before analyzing the influence of program and sample characteristics on effectiveness, we first looked at the influence of the *study design quality* and the *sample size* to control for the influence of methodological variables on effect-size variability. Although it seemed as if poor study design (only post-control design) and low sample size (below 50) correlated negatively with effect sizes (see Table 2), neither the study design quality nor the sample size were significant moderators,  $Q(df = 3) = 0.78$ ,  $p = .86$  and  $Q(df = 3) = 2.80$ ,  $p = .42$ , respectively.

Turning to program characteristics, the analysis of the *target out-group* revealed at least a tendency for the target out-group to be a potential moderator,  $Q(df = 2) = 5.33$ ,  $p = .07$ . Programs addressing intergroup attitudes toward persons with disabilities yielded the highest effect size and were even significantly more effective than programs addressing intergroup attitudes toward members of other ethnicities (see Table 2).

Effect sizes for the three different *theoretical orientations* (socialization/knowledge acquisition, intergroup contact, social-cognitive development) were almost identical, with  $d$  ranging from 0.31 to 0.33. However, results differed for the various content components, especially in the area of intergroup contact and social-cognitive development interventions (see Table 2). The highest effect sizes were found for interventions with *direct contact* elements ( $d = 0.43$ ,  $k = 22$ ) and training in *perspective taking/empathy* ( $d = 0.44$ ,  $k = 24$ ), whereas the lowest and nonsignificant effect sizes emerged for programs with training components in classification/social categorization ( $d = 0.14$ ,  $k = 18$ ) and social problem-solving skills ( $d = 0.18$ ,  $k = 11$ ). Of course, as already mentioned, the majority of programs used a combination of contents (see above). Therefore, we compared interventions according to their complexity and found a nearly significant difference between unimodal (e.g., only one content aspect;  $d = 0.24$ ,  $k = 48$ ) and multimodal programs, (e.g., more than one content component,  $d = 0.35$ ,  $k = 74$ ;  $Q(df = 1) = 3.41$ ,  $p = .06$ ). However, there were no differences between several types of content combinations (e.g., knowledge acquisition/intergroup contact vs. knowledge acquisition/social-cognitive development).

In addition to these analyses, we tested whether interventions differed in their effects on various outcome measures and compared the programs according to different intergroup attitudes dimensions. In general, there were only very few differential effects. The most notable were, first, the relatively high effect of programs with a social-cognitive developmental foundation on cognitive measures of intergroup attitudes ( $d = 0.47$ ,  $k = 40$ ). Second, was the above-average effect sizes of programs with a *direct contact* component and programs covering training in *perspective taking/empathy* on both cognitive measures ( $d = 0.45$ ,  $k = 17$  and  $d = 0.64$ ,  $k = 21$ , respectively) and behavioral measures of intergroup attitudes ( $d = 0.38$ ,  $k = 5$  and  $d = 0.45$ ,

$k = 13$ , respectively). Finally, this pattern could be found generally in multimodal programs that had systematically higher effect sizes on each intergroup attitude dimension, especially on behavioral measures, on which they differed significantly from unimodal programs ( $d = .29$ ,  $k = 25$  vs.  $d = 0.12$ ,  $k = 18$ ),  $Q(df = 1) = 4.63$ ,  $p = .03$ .

Effect sizes for the different *program parameters* varied between  $d = 0.23$  and  $d = 0.35$ . Interventions using role taking/role play as a method yielded the highest effect sizes, but effect sizes in general did not differ very much from each other (see Table 2). However, because the program components and program parameters were usually confounded, we conducted a meta-regression controlling for the mutual influence of these program variables (see Table 5). Overall, this regression model was significant,  $Q(df = 16) = 38.90$ ,  $p < .001$ , leaving no proportions of variance unexplained,  $Q(df = 105) = 110.33$ ,  $p = .34$ . In detail, the  $\beta$  coefficients for the program components *direct intergroup contact* and *perspective taking* and the program parameters *active trainer* and *implementation within the curriculum* revealed (at least as a tendency) that these variables were significant positive moderators; in other words, the inclusion of these components and parameters yielded higher effect sizes. In contrast, programs using *group discussion* and *printed materials* yielded lower effect sizes than programs not including these components (see Table 4).

No further characteristics of the program (*duration*, *number of sessions*, *intensity rating*, *trainer*) or the subjects (*age group*, *gender*, *in-group*) accounted significantly for effect-size variability (all  $p > .15$ , see Table 2). Thus, none of these characteristics were significant moderators, although some single categories differed significantly from each other (e.g., programs implemented by research assistants/supervised students had higher effect sizes than programs implemented by study authors, see Table 2). In addition, effect sizes for some single categories did not differ significantly from zero (i.e., if the target in-group was minority children,  $d = 0.17$ ), indicating that there was no significant intervention effect for these groups.

#### Description of research reports including follow-up data

In addition to posttest data, 10 of the 81 research reports included follow-up results. These reports contained 13 intervention–control group comparisons of which 5 addressed intergroup attitudes toward persons of other ethnicities, 5 toward persons with disabilities, and 3 toward the aged (see Table 6 for further details). The follow-up assessment

**Table 5**  
Metaregression with program components and parameters.

Program components	$\beta$	$p$	$B$	95% CI
Socialization/knowledge acquisition				
Information on groups	-.15	.135	-0.11	[-0.25, 0.03]
Information on prejudice	-.09	.432	-0.07	[-0.26, 0.11]
Information on values and norms	.07	.604	0.07	[-0.21, 0.35]
Intergroup contact				
Indirect contact	.17	.105	0.12	[-0.03, 0.27]
Direct contact	.25	.012	0.24	[0.05, 0.43]
Social-cognitive skills				
Classification/social categorization	-.22	.029	-0.21	[-0.40, -0.02]
Perspective taking/empathy	.36	.001	0.33	[0.14, 0.53]
Social problem solving	-.18	.096	-0.22	[-0.47, 0.04]
Moral development	-.01	.953	-0.01	[-0.34, 0.32]
Program parameter				
Group discussion	-.20	.040	-0.15	[-0.29, -0.01]
Role-taking/role playing	-.03	.795	-0.02	[-0.18, 0.14]
Printed materials	-.32	.003	-0.26	[-0.42, -0.09]
TV-material	-.04	.644	-0.03	[-0.17, 0.11]
Manual used	-.02	.862	-0.02	[-0.20, 0.17]
Active trainer	.20	.055	0.20	[-0.00, 0.40]
Curriculum implementation	.17	.108	0.12	[-0.03, 0.26]

Note.  $B$  = unstandardized regression coefficient. CI = confidence interval. Positive  $\beta$  = inclusion of component increases mean effect size, exclusion decreases it. Negative  $\beta$  = inclusion of component decreases the mean effect size, exclusion increases it.

**Table 6**  
Description of the 10 research reports including follow-up data.

Author	Country	Ingroup	Intergroup attitudes toward	N		Age (in months)		Intervention components			Effect size <i>d</i>		
				IG	CG	M	Range	Content <sup>a</sup>	Implementation <sup>b</sup>	No. of sessions	Follow-up (in month)	Post	Follow-up
Aday, Sims, and Evans (1991)	USA	Children	Aged	24	25	NR	108–120	1, 2b	1, 3, 7, 8	8	12	1.20	–1.11
Beelmann et al. (2010)	D	German children	Other ethnicities	219	211	120	96–108	1, 2a, 3	1, 3, 5, 7, 8	15	16	0.19	0.27
Clunies-Ross, and O'Meara, (1989)	AUS	Nondisabled	Disabilities	15	15	108	108–120	1, 2b	1, 2, 3, 4, 7, 8	4	3	0.39	0.47
		Nondisabled	Disabilities	15	15	108	108–120	1, 2b	1, 2, 3, 4, 7, 8	4	3	0.97	1.15
Cross (1990)	USA	Children	Aged	51	24	NR	120–132	1, 2a	2	3	1	0.36	0.24
Glass and Trent (1980)	USA	Adolescents	Aged	224	227	168	168–180	1	1, 3, 5, 7, 8	10	4 to 6	0.45	0.45
Hillis (1987)	USA	Nondisabled	Disabilities	39	39	120	108–132	1	1, 2, 3, 7	6	2	0.08	–0.17
Krahé and Altvasser (2006)	D	Nondisabled	Disabilities	22	24	178	NR	1, 2a	1, 2, 3, 7	2	3	0.00	0.00
		Nondisabled	Disabilities	24	24	178	NR	1, 2b	1, 2, 3, 7	2	3	0.20	0.20
Küchel and Beelmann (2008)	D	German children	Other ethnicities	36	40	109	84–120	2a	1, 3, 7	5	6	0.01	0.01
Lovelace (1975)	USA	Whites	Other ethnicities	41	49	171	168–180	1, 3	1, 2, 3, 4, 5, 8	15	0.5	0.03	0.31
		Whites	Other ethnicities	47	49	171	168–180	1, 3	1, 2, 3, 4, 5, 8	16	0.5	–0.17	0.00
Weiner and Wright (1973)	USA	Whites	Other ethnicities	27	27	96	96–108	3	4, 7, 8	2	0.5	1.70	2.30

Note. NR = not reported; IG = intervention group, CG = control group. *d* = unweighted effect size. A positive *d* indicates a successful intervention (i.e., reduced prejudice or improved intergroup attitudes).

<sup>a</sup> Coding: 1 = knowledge acquisition, 2a = indirect contact, 2b = direct contact, 3 = promotion of cognitive abilities.

<sup>b</sup> Coding: 1 = printed materials, 2 = tv material, 3 = active participation, 4 = first-hand experience of discrimination, 5 = manual used, 6 = out-group member as trainer, 7 = actively involved trainer, 8 = curriculum implementation.

varied from 2 weeks to 16 months with a mean of about 4 months. The follow-up effect sizes ( $k = 13$ ) ranged between  $d = -1.11$  and  $d = 2.30$  with two negative effect sizes (see Table 6). The unweighted overall mean comparison effect size was  $d = 0.32$ . Following the FEM, the overall follow-up effect decreased slightly,  $d = 0.29$ ,  $p < .001$ , and again, showed significant heterogeneity,  $Q(df = 12) = 74.44$ ,  $p < .001$ . Therefore, the random-effects model seemed to be more appropriate for our follow-up data, too. According to this model, the overall follow-up effect was  $d = 0.29$ ,  $p = .05$ .

## Discussion

The present meta-analysis integrated the effects of standardized intervention programs designed to prevent and reduce prejudice or otherwise improve intergroup attitudes in children and adolescents. The interventions proved to be generally effective with a mean effect size of around  $d = 0.30$ . These outcomes are largely comparable to those effects ascertained within other prevention fields such as antisocial behavior or the outcomes for social and emotional training programs in general (see, e.g., Beelmann & Raabe, 2009; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011). According to the binomial effect-size display, this could be considered to be an improvement in intergroup attitudes of around 15% (Lipsey & Wilson, 2001). Although this effect size is viewed typically as a small to medium effect (Cohen, 1988), one should recognize that most of the programs under review are of low intensity and cost, and even small effects could have impressive implications and practical significance—for example, if slightly improved intergroup attitudes lead to less discrimination or prevent violent behavior toward out-group members (see Ellis, 2010).

However, we have to consider at least two restrictions to our results. First, the mean effect sizes must be evaluated against the background that a large proportion of the studies summarized in the present meta-analysis used proximal outcome assessments. Such assessments usually have strong ties to the content of programs and generally result in larger effect sizes (Beelmann, 2011). Therefore, the present results reflect more of the upper limit of program effectiveness. Second, the majority of primary studies assessed only immediate effects on intergroup attitudes, and only relatively few studies evaluated long-term developments or the stability of outcomes. Nonetheless, these few

long-term evaluations confirmed that durable positive effects are not only possible but sometimes even higher than short-term changes (e.g., Beelmann, Saur, Ziegler, Diener, & Noack, 2010). The problem of insufficient investigations into the long-term effects of programs is well known in intervention research in general. It could perhaps be viewed as the most pressing issue when it comes to justifying prevention measures, because, in most cases, long-lasting effects (e.g., on tolerance attitudes, hate crimes, or right-wing extremism) are the intention (Beelmann, 2011). However, apart from this open question regarding effect stability and long-term effects, the present results can be viewed as reliable estimators of effectiveness, because neither publication bias nor significant confounds with methodological variables seem to restrict the validity of the data.

Therefore, the results of this meta-analysis imply first of all, that prejudice prevention and the promotion of intergroup attitudes are possible via psychological and educational training programs in children and adolescents. Nonetheless, the moderate effect sizes suggest that this approach is in need of further development and should be viewed as one way of dealing with problems of prejudice and negative intergroup relations (Aboud et al., 2012; Killen et al., 2011). Second, as in many intervention fields, program outcomes are moderated by several variables involving the assessment method, program components, and target groups. This makes it necessary to take a closer look at what works, for whom, and under which conditions.

### Effects on different outcome dimensions

Many meta-analyses have shown that characteristics of the outcome measures are important effect-size moderators (Lipsey & Wilson, 2001). This could be endorsed here as well, especially with regard to the *intergroup attitude dimension*. Apparently, the cognitive and behavioral dimensions of intergroup attitudes are easier to influence than the emotional dimension (e.g., sympathy). This result holds for all types of intervention, and it contradicts some previous findings reporting that emotional indicators tend to show stronger effects than cognitive indicators—at least for contact-based interventions (Tropp & Pettigrew, 2005).

However, our findings should be qualified by the fact that almost all emotional indicators tap positive instead of negative aspects of intergroup attitudes (e.g., liking and sympathy and not disliking, anger,



fear, or intergroup threat), and this positive dimension does not seem to be sensitive to intended changes. Indeed, contact-based interventions primarily promote intergroup attitudes by changing negative emotions such as fear or intergroup threat (Pettigrew & Tropp, 2008). Unfortunately, the studies integrated here hardly ever used these measures. Moreover—at least for interventions focusing on modifying negative facets of intergroup attitudes—an increase in sympathy and liking is not necessarily a primary objective. In addition, several intergroup attitude constructs such as tolerance or respect do not inevitably imply sympathy, even if they are not independent from each other (see van Quaquebeke, Henrich, & Eckloff, 2007).

In sum, the finding that the integrated studies neglect negative emotional ratings almost completely as outcome measures is highly unsatisfactory. This is particularly because negative emotions seem to play an important role in the development of serious prejudice and discrimination. For example, intergroup threat and negative feelings toward an out-group are one of the best confirmed risk factors that will lower intergroup contact and impede intergroup friendships (e.g., Riek et al., 2006). Therefore, future intervention programs for children and adolescents should focus more intensively on intergroup emotions, and future evaluation studies should pay more attention to measuring especially the negative part of the emotional dimension of prejudice and intergroup attitudes.

Effect-size differences in outcome assessment have also been ascertained between the categories representing the *type of intergroup attitude* in which measures of in-group favoritism (e.g., higher rating of in-group members on a positive characteristic such as intelligence) had an effect size close to zero. In contrast, prejudice measures, that is, measures representing some kind of bias on a *negative* dimension (e.g., being mean), showed higher and significant effect sizes. At first glance, this pattern of results appears to correspond with the central aims of the interventions, because changing negative evaluations seems to be a priority goal of many intervention programs. However, even small differences in in-group favoritism could also have considerable negative effects for out-group members (e.g., Brewer, 1999), although the probability of severe consequences from negative evaluations is perhaps higher. In addition, Brewer (1999) argues that in-group favoritism is close to a more subtle kind of prejudice that may well be harder to change via interventions (see also Dixon, Levine, Reicher, & Durrheim, 2012). Again, future programs should focus more on applying several measures and indicators of intergroup attitudes at the same time in order to investigate effects on divergent dimensions and facets.

#### Results on program components

This meta-analysis also provided some insights into which programs or program elements can possibly be recommended. However, as in other reviews, the high range of intervention programs makes it virtually impossible to state which is the most promising single approach for altering prejudice and improving intergroup attitudes. Because of this diversity, we coded the intervention programs separately on the most highly distinguishing program components and parameters. The advantage of this coding procedure was that we could also enter the diversity of the program components into our meta-regressions in order to rule out confounds. This at least allowed us to examine whether any of these components were superior to others, and it should deliver indications for future developments in prejudice prevention programs.

This detailed analysis revealed that direct contact and training in empathy and perspective taking are the most promising *program content components*. The contact effect is a well-known result that has been confirmed in several reviews during the last decade (e.g., Aboud et al., 2012; Paluck & Green, 2009; Pettigrew & Tropp, 2006). In addition, our analyses revealed that direct contact leads to better results than any kind of vicarious, extended, or—more generally—indirect contact.

However, although the latter shows lower effects, these are still significant. The relatively small effect produced by indirect contact is somewhat surprising in any case, because it has been recommended strongly by some authors (e.g., Wright et al., 1997). However, previous reviews have also reported that some moderating and mediating factors have to be considered when looking at contact effects regardless of whether they are direct or indirect (Pettigrew, Tropp, Wagner, & Christ, 2011). For example, Schofield's (1995; Schofield & Hausmann, 2004) analysis of *school desegregation* produced rather mixed findings, revealing that, within this context, contact alone is not a sufficient condition for positive effects. Success seems to depend critically on whether the programs succeed in initiating personal relationships and friendships between members of different ethnic groups (see also Aboud & Levy, 2000; Tausch, Schmid, & Hewstone, 2010). In addition, recent meta-analytic results (e.g., Davies et al., 2011) confirm the central role of personal friendships, and these are impossible or at least difficult to promote in an indirect contact situation. Therefore, program developers should integrate direct contact elements in order to acknowledge the necessary or at least beneficial contribution of the contact situation (see above; Tropp & Prenovost, 2008). However, direct contact is frequently impossible in some social contexts, because members of the out-group are physically not present. In such cases, some kind of indirect contact has to be implemented (e.g., Beelmann et al., 2010; Cameron, Rutland, Brown, & Douch, 2006). Nonetheless, our results indicate the benefits of actively including at least some element of direct contact (e.g., some joint play activities with visiting children or member of the out-group as trainer) that would at least support and complement indirect contact through personal interaction experience.

In addition to the direct contact effects, training in empathy and perspective taking showed high potentials for promoting intergroup attitudes via social-cognitive abilities. The beneficial influence of empathy and perspective taking has long been discussed in the developmental literature on intergroup attitudes (Aboud & Fenwick, 1999; Bigler & Liben, 2007), and a recent meta-analysis has confirmed its importance for prejudice development, especially in elementary school age children (Heinemann & Beelmann, 2011). Although moral or value education was not significant in our meta-regression analysis, our categorical analyses indicate that it seems to deliver some promising results. Promoting moral development is most likely to be confounded with perspective taking and empathy, and any analysis of it may suffer from the relatively small number of relevant training–control comparisons. Within the field of moral development, we clearly need more high-standard evaluation studies, especially because recent studies suggest that it plays a central role in prejudice and intergroup attitude development (Killen, Kelly, Richardson, & Jampol, 2010; Rutland et al., 2010).

Training in classification skills and reorganizing social categorization has been recommended in a series of studies, especially within the social psychological literature (Gaertner & Dovidio, 2000; Gaertner et al., 2008). However, the results of our meta-analysis fail to confirm these recommendations—at least for children and adolescents. We did not find any significant positive changes, and sometimes even negative effects of training classification skills or restructuring the processing of social categorization. There might be different reasons for this: First, classification skills may not be as important as several authors and social cognitive developmental theories assume. For example, in a meta-analysis on individual prejudice predictors in childhood and adolescence, Heinemann and Beelmann (2011) found that these skills did not correlate significantly with prejudice at any stage of development. Thus, it seems less useful to train these abilities in order to promote intergroup attitudes—possibly because the trainings are restricted mainly to classification skills for nonsocial characteristics.

Second, because social categorization and stereotyping form the psychological basis of intergroup attitudes (Brown, 2010), modifying social categorizations might also have negative effects when training sensitizes members of the target group to social categories that they

were not aware of before the start of the program (i.e., they serve as a root for a subsequent prejudice). Therefore—and in contrast to adult target groups—intervention programs for children and adolescents should probably work on social categorization only in cases in which social categories are already known and used for producing biased social attitudes.

A third possible reason might lie in the type of implementation of social categorization. Several studies have shown that how social categorization takes place seems to be important. For example, Cameron et al. (2006) reported that only reconstruction following the dual identity model (i.e., invoking a superordinate social category but simultaneously retaining the subgroup identity) had positive effects. In contrast, the common in-group model (i.e., summarizing in-group and out-group within the same superordinate social group) and decategorization (i.e., teaching that the corresponding social category is not important) were less effective—at least for the age group under investigation (preschool and elementary school children). In general, results on program content indicate clearly that more research is needed on how to structure social categorization within training programs in order to promote positive intergroup relations in children and adolescents. Simple strategies such as “these categories are not important” or “we all belong to one social group of human beings” do not seem to be enough and are, at least in some cases, even counterproductive.

Besides these substantial considerations, we also identified diverse program parameters as important factors for successful program conduct and implementation. For example, interventions with a trainer who was actively involved in the program administration yielded significantly higher effect sizes than interventions without this element—independent of the person adopting that role (e.g., teacher, program author, or trained student). This result clearly corresponds to the contact literature in which supervision (or taking an active role) by an authoritative person is assumed to be one of the beneficial conditions during intergroup contact (Tropp & Prenovost, 2008). This interpretation probably also fits in with the surprisingly negative influence of the use of printed materials and group discussions on the mean effect size. Printed materials are frequently used when children teach themselves about prejudice and intergroup attitudes, and group discussions are probably less structured intervention methods. Thus, our findings indicate that training concepts receiving only low support from the teacher or program administrator are less likely to deliver significant effects on improving intergroup attitudes.

The results on content and program parameters clearly have some further implications for the prevention of prejudice and for the promotion of intergroup relations in general. First, future program concepts should maybe focus more on combined and multitheoretical models. Because there are divergent aspects of intergroup attitudes and, apparently, several promising ways to prevent prejudice and to promote intergroup attitudes and relations, it would seem to be worth developing new integrative models of intervention. In addition, we clearly find that multimodal programs have higher and more generalized effects on different outcomes. Therefore, the development of new programs would profit clearly from a systematic combination of different intergroup approaches. At the same time, however, it should not neglect the most effective elements such as structured contact experiences and training in perspective taking. For example, Beelmann et al. (2010) combined contact experiences with training in different sociocognitive skills (including perspective taking) and multicultural information and found relatively high long-term effects on different measures of intergroup attitudes. Nonetheless, these approaches mark perhaps only the beginning of a new era characterized by a more comprehensive approach to prejudice prevention research combining multiple theories, methods, contents, and outcome assessments.

Second, because an active trainer is apparently a crucial factor in promoting intergroup attitudes, intervention concepts should not be

restricted to didactic materials alone. Programs need to integrate an interaction component involving adults or competent partners. At the same time, we could assume (even though this was only a statistical trend) that interventions should be integrated into daily school activities, because implementing a program within the curriculum seems to be favorable. In sum, findings on program parameters indicated that interactive, structured, well-planned, and well-implemented programs seem to be more valuable and successful—an outcome that has been confirmed frequently in prevention research with children and adolescents in general (Durlak & DuPre, 2008; Durlak et al., 2011).

#### *Results on different target groups*

In addition to the results on content and program parameters, we identified two further variables that correlated systematically with effect sizes. On the one hand, higher effects were found when the target out-group consisted of persons who are handicapped compared to ethnic out-groups. The fact that attitudes toward the disabled are apparently easier to improve is probably due to several factors. First, although integrated schooling has been implemented broadly in several countries, contact with the handicapped is perhaps less probable in everyday life than that with ethnic groups—at least in multiethnic countries such as the United States. Moreover, the majority of studies addressing intergroup attitudes toward the handicapped were performed outside integrative schools. As a result, children may have less experience in interacting with children who are disabled and thus have more potential to learn something about this group and promote their attitudes toward them—at least in information-based and contact-based interventions. Although we have no comparable data on the preintervention level of prejudice, it is a well-known result in intervention research that groups showing higher problems or less competence profit more from those interventions than less burdened groups (Beelmann & Raabe, 2009; Durlak et al., 2011) simply due to ceiling or floor effects in outcome measures.

Second, the intergroup threat is probably higher in intergroup contexts with ethnic out-groups compared to intergroup contexts with people who are disabled, especially in high conflict areas or multiethnic countries. This may hold for elderly people as a target group as well, for whom we also found relatively high effects. This higher mean level of intergroup threat may well be responsible for the significantly lower effect sizes in ethnic prejudice and intergroup attitudes, because of more difficulties in concrete implementation of programs. However, we found virtually no intervention studies addressing this issue.

Finally, western societies may well have stronger antiprejudice and antidiscrimination norms for the handicapped and for elderly people than for ethnic minorities, probably again because of different level of intergroup threat. Therefore, program goals with these out-groups are either easier to attain or less sensitive to interfering norms in the social context. However, despite this relatively unfavorable result for ethnic out-groups, it should not be forgotten that there are also significant effect sizes for improving intergroup attitudes toward other ethnic groups. Future research should focus more on the concrete social context and the ethnic conflicts underlying attitudes in order to optimize intervention efforts in this area.

The second variable showing a systematic correlation with effect sizes was the in-group participating in the intervention. Studies with minority groups (e.g., in which they are the in-group and the intervention addresses prevention of prejudice against majority children) revealed nonsignificant effect sizes and were far less effective than interventions with majority groups (e.g., addressing prejudice against minority children). One reason for this result is probably the poor fit between interventions and the development of intergroup attitudes in minority groups (see also Aboud et al., 2012). A recent meta-analysis has shown that the developmental course of intergroup attitudes differs for minority compared to majority groups (Raabe & Beelmann, 2011).

Majority (high-status) children showed an increase in the level of prejudice against minority children between ages 3 and 5 to 7 and a subsequent decrease up to ages 8 to 10. Minority (low-status) children, in contrast, start to increase their prejudice against majority children at a later age and fail to show the decrease found in majority children. These different developmental trajectories are probably due to different developmental processes, indicating the need for different kinds of interventions at different developmental stages. For example, the decrease in majority children could probably be ascribed to emerging sociocognitive abilities (e.g., [Aboud, 2008](#); [Bigler & Liben, 2007](#)). If so, this indicates the benefits of training in these abilities at this age to promote natural development. In minority children in contrast, the later increase probably relates to specific identification processes (e.g., after experiencing first discrimination through majority children). This indicates the benefits of more training in self-efficacy, coping with negative experiences, or establishing specific contact opportunities to majority members (e.g., [Banaji, Baron, Dunham, & Olson, 2008](#)). Therefore, developmental knowledge on different developmental courses should lead to different prevention concepts. However, the four studies addressing minority children ([Aboud & Fenwick, 1999](#); [Avci-Werning, 2004](#); [Feinman, 1982](#); [Walker, 1971](#)) used essentially the same intervention as that used for majority children.

In addition, although concrete intervention planning should depend on the specific in-group–out-group constellation in a given social environment, the aforementioned developmental knowledge leads us to assume that minority groups should not simply be integrated within programs for majority children and adolescents. For example, information on and contact to the out-group is perhaps redundant for minority children, because they are exposed to this out-group constantly in their everyday lives. Therefore, it seems advisable to apply different antibias interventions for majority and minority members. However, the present comparison of majority versus minority children and adolescents and the conclusions made are based on a very small number of studies for minorities, indicating a need for further research to derive more valid conclusions on this important topic.

One final word should address the age of the participants as effect moderator. We failed to find differences in effectiveness of programs according to the age group treated. On the one hand, this is surprising, because different phases of intergroup attitude development would lead us to expect age-related differences in the amenability to intervention programs ([Raabe & Beelmann, 2011](#)). In addition, prevention research is largely based on the assumption that “earlier is better,” suggesting that prevention measures should be more effective in younger age groups.

On the other hand, there may be several reasons why age could not be identified as a significant moderator. First, although nonsignificant, we nevertheless found a slightly higher effect size for the age group between 8 and 10 years, which corresponds to highly distinctive developmental changes—at least for majority children (see above). Therefore, results may indicate small age-related differences that could not be confirmed because of low statistical power. Second, age is normally confounded with other variables (e.g., intervention concept, level of prejudice). As in many meta-analyses, these confounds are mostly neither easy to control (e.g., in order to consider the level of prejudice, we need comparable pretest scores) nor easy to take into account systematically when interpreting results. For example, we found no systematic relation between the content and program parameters of the intervention and age group apart from a higher percentage of socio-cognitive training programs in older children from age 8 on. This confound is probably responsible for some bias when comparing the results for age groups, but cannot explain entirely why they remain largely comparable over the course of development. Therefore, the lack of age-specific results suggests that even effective prevention of prejudice and promotion of positive intergroup attitudes is possible at all ages as long as the underlying intervention concepts are developmentally appropriate (see also [Aboud et al., 2012](#)).

### *Limitations and future research on promoting intergroup relations*

This meta-analysis also has its limitations that simultaneously deliver indications for further research on promoting intergroup attitudes and relations. First, the great diversity of intervention programs and approaches makes it almost impossible to identify clearly defined programs that are more effective and therefore to be recommended. In addition, we can make no systematic statement about the appropriateness of the programs within their specific social context and how they could be tailored to these contextual requirements. In general, programs seem to have clear deficits in relation to developmental theories and existing empirical knowledge on intergroup relations. For example, only a few intervention concepts analyzed in this meta-analysis derive their content and further program parameters such as age at start from modern developmental theories on intergroup attitudes (e.g., [Bigler & Liben, 2007](#); [Nesdale, 2004](#)). In addition, although empirical research has revealed that several risk factors are involved in prejudice development, only a few programs have multicomponent concepts (e.g., [Beelmann et al., 2010](#)). However, this recommendation for closer links between developmental research and intervention concepts also indicates the need for further developmental studies. For example, up to now, it is relatively unclear how stable intergroup attitudes actually are in childhood and adolescence, and which prejudice thresholds deliver a clear negative developmental prognosis and are linked to negative intergroup behavior in the future. A few longitudinal results indicate some strong continuities over the course of development as well as a linkage to behavior problems (see [Beelmann & Heinemann, 2011](#); [Raabe & Beelmann, 2011](#)). This is important information for constructing intervention programs and deciding at what age they should be applied. Therefore, future social psychological and developmental research should focus more on applying longitudinal designs and developing closer links to prevention and intervention programs in this field. This, in turn, should make these interventions more effective (see also [Beelmann, 2011](#); [Killen et al., 2011](#)).

Second, in light of the only moderate effect sizes found here, it seems advisable to look for alternative intervention approaches in the future such as programs with parents, teachers, and peers in order to promote their function in socializing toward more positive social norms (see [Killen et al., 2011](#)). This strand of intervention research is quite neglected despite the strong ties between, for example, parent and child intergroup attitudes ([Degner & Dalege, 2013](#)).

Third, most of the systematic intervention research has been performed in North America and Europe. However, problems in intergroup relations are universal phenomenon that needs to be addressed in many countries—especially in those with massive conflicts between ethnicities or other social groups. The few studies coming from Israel (e.g., [Slone, Tarrasch, & Hallis, 2000](#)) have taught us that principles of intergroup contact and promoting empathy and perspective taking may also hold for these social contexts. However, especially in the context of strong and long-lasting historical conflicts such as those in the Middle East or some regions of Africa, it is in some way questionable whether these kinds of psychological programs for children and youth alone will suffice. Alternatively, we should think about more multidimensional and systemic interventions including, for example, new legislation, peace education within society, and reducing intergroup threat by promoting reconciliation between conflict parties ([Bar-Tal & Rosen, 2009](#); [Staub, 2006](#)). Research on this integrative approach to intergroup conflicts is only just beginning ([Nadler, Malloy, & Fisher, 2008](#); [Tropp, 2012](#)), but is much needed to apply existing psychological knowledge to solve problems in these adverse social contexts.

Fourth, turning to the type of outcome assessment, only a small variety of different outcome measures was used, and these do not represent the multifaceted character of intergroup attitudes. Most studies used cognitive intergroup attitude indicators to measure prejudice and applied questionnaires to assess it. More data are needed, especially to estimate the effect on the negative emotional facets of intergroup

attitudes (intergroup threat, fear, or anger), because these indicators exert a tremendous influence on intergroup behavior. Moreover, despite the long history of measuring social attitudes in social psychology, there is still a need for more age-appropriate, reliable, and valid strategies for assessing intergroup attitudes in children and adolescents in general. The lack of comparable and scientifically acceptable assessment tools for intergroup attitudes is also one reason why evaluation studies are so difficult to compare. For example, when selecting target groups, it would be of great value if we could compare the preintervention level of biased intergroup attitudes across studies in order to decide whether universal or targeted prevention programs promise to be more effective.

In addition, implicit measures of prejudice and intergroup attitudes are seldom used for evaluation purposes. This may have important consequences, because implicit measures point to a different developmental pathway compared to explicit measures—perhaps because age increases the problem of social desirability in assessment (Banaji et al., 2008; Degner & Wentura, 2010; Raabe & Beelmann, 2011). However, social desirability is not just an assessment error in this field of intervention. It can also be viewed as one sign of learning social (antibias) norms. Nonetheless, it would be helpful to measure effects in a more indirect and implicit way in order to control social desirability and to extend our knowledge of changes in intergroup attitudes via intervention programs.

Fifth, in light of the prevention of more serious societal problems such as social exclusion, discriminative behavior, racism, or right-wing extremism, one main limitation of the prevention approaches in this field is the restriction of many studies to the assessment of attitudes alone. Direct behavior measures and more long-term effects have hardly ever been assessed, making it scarcely possible to appropriately evaluate any positive long-term effect on these problems. Although we do possess some data indicating the effectiveness of prejudice prevention programs in the long run, future primary studies will have to investigate and report follow-up data on their interventions before any reliable conclusions can be made.

Finally, from the standpoint of intervention research, the studies do not permit any reliable conclusions about their implementation quality. But implementation is important, as the conduct and dissemination of prevention programs clearly shows (Durlak & DuPre, 2008). This could be expected to be a particular problem in the field of prejudice interventions in which implementation in the social context is nearly always a difficult task, because the aims of programs probably interfere with the social norms, social interests, and identities of those involved and may well even trigger political resistance. Therefore, as in other prevention fields, the implementation perspective also needs to be included in intervention research on intergroup attitudes.

In sum and taking account of the aforementioned limitations, the results of this meta-analysis indicate that it is not only possible but also worthwhile to promote positive intergroup attitudes and prevent prejudice with child and youth programs. However, at the moment, no one can say whether this line of research will lead to more tolerance and less discrimination within our societies in the long run. But promising first results should lead to the refinement and improvement of intervention strategies in the near future. When facing such phenomena as rapid globalization, the need for cooperation between social groups, and the worldwide risk of conflict, there is a clear need for effective programs and strategies to promote positive intergroup attitudes and relations in perhaps all social contexts. However, educational and psychological prevention programs are only one way to address these challenging goals. Other ways include new legislation, mass media campaigns, work with parents and teachers, and integrated schooling (see Oskamp, 2000; Stephan & Stephan, 2001). They should all share the common mission of not only reducing prejudice and discrimination but also enabling children, adolescents, and finally adults to develop a greater sense of equity, tolerance, and justice (Killen et al., 2011) and to build up positive intergroup relations so that they may

benefit rather than suffer from social diversity (Dixon et al., 2012; Tropp & Mallett, 2011).

## References<sup>1</sup>

- Aboud, F. E. (1988). *Children and prejudice*. Oxford, England: Blackwell.
- Aboud, F. F. (2008). A social-cognitive theory of prejudice. In J. F. Dovidio, P. Glick, & A. Rudman (Eds.), *On the nature of prejudice: Fifty years after Allport* (pp. 310–326). Malden, MA: Blackwell.
- Aboud, F., & Fenwick, V. (1999). Exploring and evaluating school-based interventions to reduce prejudice. *Journal of Social Issues*, 55, 767–786. <http://dx.doi.org/10.1111/0022-4537.00146>.
- Aboud, F. E., & Levy, S. R. (2000). Interventions to reduce prejudice and discrimination in children and adolescents. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 269–293). Mahwah, NJ: Lawrence Erlbaum.
- Aboud, F. E., Tredoux, C., Tropp, L. R., Brown, C. S., Niens, U., Noor, N. M., et al. (2012). Interventions to reduce prejudice and enhance inclusion and respect for ethnic differences in early childhood: A systematic review. *Developmental Review*, 307–336. <http://dx.doi.org/10.1016/j.dr.2012.05.001>.
- Adorno, T. W., Frenkel-Brunswick, E., Levinson, D. J., & Sanford, R. N. (1950). *The authoritarian personality*. New York, NY: Harper.
- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison-Wesley.
- Bagley, M. T., & Greene, J. F. (1981). Peer attitudes toward the Handicapped Scale. Austin, TX: Pro-Ed.
- Banaji, M. R., Baron, A. S., Dunham, Y., & Olson, K. (2008). The development of intergroup social cognition: Early emergence, implicit nature, and sensitivity to group status. In S. R. Levy, & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 197–236). New York, NY: Oxford University Press.
- Bar-Tal, D., & Rosen, Y. (2009). Peace education in societies involved in intractable conflicts: Direct and indirect models. *Review of Educational Research*, 79, 557–575. <http://dx.doi.org/10.3102/0034654308330969>.
- Bar-Tal, D., & Teichman, Y. (2005). *Stereotypes and prejudice in conflict: Representations of Arabs in Israeli Jewish society*. New York, NY: Cambridge University Press.
- Beelmann, A. (2009). Prevention of right-wing extremism: European and international research results. In M.A. Niggli (Ed.), *Right-wing extremism in Switzerland: National and international perspectives* (pp. 252–264). Baden-Baden, Germany: Nomos.
- Beelmann, A. (2011). The scientific foundation of prevention: The status quo and future challenges for developmental crime prevention. In T. Bliesener, A. Beelmann, & M. Stemmler (Eds.), *Antisocial behavior and crime: Contributions of developmental and evaluation research to prevention and intervention* (pp. 137–164). Cambridge, MA: Hogrefe Publishing.
- Beelmann, A., & Heinemann, K. S. (2011). Intergrupale soziale Einstellungen und der Zusammenhang zum Sozialverhalten von Grundschulern [Intergroup attitudes and their relation to the social behavior of elementary school children]. *Zeitschrift für Grundschulforschung*, 4, 100–112.
- Beelmann, A., & Lipsey, M. W. (2013). Meta-analysis of effect estimates from multiple studies. In M. W. Lipsey, & D. S. Cordray (Eds.), *Field experimentation: Methods for evaluating what works, for whom, under what circumstances, how, and why*. Thousand Oaks, CA: Sage (in press).
- Beelmann, A., & Raabe, T. (2009). The effects of preventing antisocial behavior and crime in childhood and adolescence: Results and implications of research reviews and meta-analyses. *European Journal of Developmental Science*, 3, 260–281.
- Bigler, R. S., & Liben, L. S. (1992). Cognitive mechanisms in children's gender stereotyping: Theoretical and educational implications of a cognitive-based intervention. *Child Development*, 63, 1351–1363. <http://dx.doi.org/10.2307/1131561>.
- Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and reducing children's social stereotyping and prejudice. *Current Directions in Psychological Science*, 16, 162–166. <http://dx.doi.org/10.1111/j.1467-8721.2007.00496.x>.
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate. *Journal of Social Issues*, 55, 429–444.
- Brown, R. (2010). *Prejudice: Its social psychology* (2nd ed.). Chichester, England: Wiley.
- Brown v. Board of Education, 347 U.S. 483 (1954).
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Davies, K., Tropp, L. R., Aron, A., Pettigrew, T. F., & Wright, S.C. (2011). Cross-group friendships and intergroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, 15, 332–351. <http://dx.doi.org/10.1177/1088868311411103>.
- Degner, J., & Dalege, J. (2013). The apple does not fall far from the tree, or does it: A meta-analysis of parent-child similarity in intergroup attitudes. *Psychological Bulletin*. <http://dx.doi.org/10.1037/a0031436> (Advance online publication).
- Degner, J., & Wentura, D. (2010). Automatic prejudice in childhood and early adolescence. *Journal of Personality and Social Psychology*, 98, 356–374. <http://dx.doi.org/10.1037/a0017993>.
- Dixon, J., Levine, M., Reicher, S., & Durrheim, K. (2012). Beyond prejudice: Are negative evaluations the problem and is getting us to like one another more the solution? *Behavioral and Brain Sciences*, 35, 411–466. <http://dx.doi.org/10.1017/S0140525X11002214>.
- Doyle, A.B., & Aboud, F. E. (1995). A longitudinal study of White children's racial prejudice as a social-cognitive development. *Merrill-Palmer Quarterly*, 41, 209–228.
- Duckitt, J. (2010). Historical overview. In J. Dovidio, M. Hewstone, P. Glick, & V. M. Esses (Eds.), *The SAGE handbook of prejudice, stereotyping and discrimination* (pp. 29–44). Los Angeles, CA: Sage.

<sup>1</sup> References marked with an asterisk indicate research reports included in the meta-analysis. The in-text citations to studies selected for meta-analysis are not preceded by asterisks.

- Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology*, 41, 327–350. <http://dx.doi.org/10.1007/s10464-008-9165-0>.
- Durlak, J. A., Weissberg, R. P., Dymnicki, A.B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, 82, 405–432. <http://dx.doi.org/10.1111/j.1467-8624.2010.01564.x>.
- Ellis, P. D. (2010). *The essential guide to effect sizes. Statistical power, meta-analysis, and the interpretation of research results*. Cambridge, England: Cambridge University Press.
- Gaertner, S. L., & Dovidio, J. F. (2000). *Reducing intergroup bias: The common ingroup identity model*. New York, NY: Psychology Press.
- Gaertner, S. L., Dovidio, J. F., Guerra, R., Rebelo, M., Monteiro, M. B., Riek, B.M., et al. (2008). The common in-group identity model: Applications to children and adults. In S. R. Levy, & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 204–219). New York, NY: Oxford University Press.
- Graves, S. (1999). Television and prejudice reduction: When does television as a vicarious experience make a difference? *Journal of Social Issues*, 55, 707–725. <http://dx.doi.org/10.1111/0022-4537.00143>.
- Grube, J. W., Mayton, D.M., & Ball-Rokeach, S. J. (1994). Inducing change in values, attitudes, and behaviors: Belief system theory and the method of value self-confrontation. *Journal of Social Issues*, 50, 153–173. <http://dx.doi.org/10.1111/j.15404560.1994.tb01202.x>.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. New York, NY: Academic Press.
- Heinemann, K. S., & Beelmann, A. (2011). Individual factors of prejudice development in childhood and adolescence. A meta-analysis. *Poster presented at the biennial meeting of the Society for Research in Child Development, Montréal, Canada* (April).
- Hillis, S. R. (1987). Modification of the attitudes of non-handicapped children toward the handicapped through information-based and contact-based intervention. Doctoral dissertation. Retrieved from ProQuest Dissertations and Theses database (UMI No. 8615637).
- Johnson, D. W., & Johnson, R. T. (1989). *Cooperation and competition: Theory and research*. Edina, MI: Interaction Book Co.
- Johnson, D. W., & Johnson, R. T. (1996). Conflict resolution and peer mediation programs in elementary and secondary schools: A review of the research. *Review of Educational Research*, 66, 459–506 (Retrieved from <http://www.jstor.org/stable/1170651>)
- Johnson, D. W., & Johnson, R. T. (2000). The three Cs of reducing prejudice and discrimination. In S. Oskamp (Ed.), *Reducing prejudice and discrimination* (pp. 239–268). Mahwah, NJ: Lawrence Erlbaum.
- Killen, M., Kelly, M. C., Richardson, C., & Jampol, N. S. (2010). Attributions of intentions and fairness judgments regarding interracial peer encounters. *Developmental Psychology*, 46, 1206–1213. <http://dx.doi.org/10.1037/a0019660>.
- Killen, M., & Rutland, A. (2011). *Children and social exclusion: Morality, prejudice, and group identity*. Chichester, England: Wiley-Blackwell.
- Killen, M., Rutland, A., & Ruck, M. (2011). Promoting equity, tolerance, and justice in childhood. *Social Policy Report*, 25, 1–33 (Retrieved from [http://www.srcd.org/index.php?option=com\\_content&task=view&id=232&Itemid=550](http://www.srcd.org/index.php?option=com_content&task=view&id=232&Itemid=550))
- Levy, S. R. (1999). Reducing prejudice: Lessons from social-cognitive factors underlying perceiver differences in prejudice. *Journal of Social Issues*, 55, 745–765. <http://dx.doi.org/10.1111/0022-4537.00145>.
- Levy, S. R., & Killen, M. (2008). *Intergroup attitudes and relations in childhood through adulthood*. Oxford, England: Oxford University Press.
- Light, R., & Pillemer, D. (1984). *Summing up: The science of reviewing research*. Cambridge, MA: Harvard University Press.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Macaskill, P., Walter, S. D., & Irwig, L. (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in Medicine*, 20, 641–654. <http://dx.doi.org/10.1002/sim.698>.
- McGregor, J. (1993). Effectiveness of role playing and antiracist teaching in reducing student prejudice. *Journal of Educational Research*, 86, 217–226 (Retrieved from <http://pao.chadwyck.co.uk/PDF/1278672123673.pdf>)
- Nadler, A., Malloy, T. E., & Fisher, J.D. (Eds.). (2008). *The social psychology of intergroup reconciliation*. New York, NY: Oxford University Press.
- Nesdale, D. (2004). Social identity processes and children's ethnic prejudice. In M. Bennett, & F. Sani (Eds.), *The development of the social self* (pp. 219–246). New York, NY: Psychology Press.
- Oskamp, S. (Ed.). (2000). *Reducing prejudice and discrimination*. Mahwah, NJ: Lawrence Erlbaum.
- Paluck, E. L., & Green, D. P. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *The Annual Review of Psychology*, 60, 339–367. <http://dx.doi.org/10.1146/annurev.psych.60.110707.163607>.
- Paradies, Y. (2006). Ethnicity and health. A systematic review of empirical research of self-reported racism and health. *International Journal of Epidemiology*, 35, 888–901. <http://dx.doi.org/10.1093/ije/dyl056>.
- Pettigrew, T., & Tropp, L. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*, 90, 751–783. <http://dx.doi.org/10.1037/0022-3514.90.5.751>.
- Pettigrew, T. F., & Tropp, L. R. (2008). How does intergroup contact reduce prejudice? Meta-analytic tests of three mediators. *European Journal of Social Psychology*, 38, 922–934. <http://dx.doi.org/10.1002/ejsp.504>.
- Pettigrew, T. F., & Tropp, L. R. (2011). *When groups meet: The dynamics of intergroup contact*. Philadelphia, PA: Psychology Press.
- Pettigrew, T. F., Tropp, L. R., Wagner, U., & Christ, O. (2011). Recent advances in intergroup contact theory. *International Journal of Interpersonal Relations*, 35, 271–280. <http://dx.doi.org/10.1016/j.ijintrel.2011.03.001>.
- Pfeifer, J. H., Spears Brown, C., & Juvonen, J. (2007). Prejudice reduction in schools: Teaching tolerance in schools: Lessons learned since Brown v. Board of Education about the development and reduction of children's prejudice. *Social Policy Report*, 21, 1–24 (Retrieved from [http://www.srcd.org/index.php?option=com\\_content&task=view&id=232&Itemid=550](http://www.srcd.org/index.php?option=com_content&task=view&id=232&Itemid=550))
- Raabe, T., & Beelmann, A. (2011). Development of ethnic, racial, and national prejudice in childhood and adolescence: A multinational meta-analysis of age differences. *Child Development*, 82, 1715–1737. <http://dx.doi.org/10.1111/j.1467-8624.2011.01668.x>.
- Riek, B.M., Mania, E. W., & Gaertner, S. L. (2006). Intergroup threat and outgroup attitudes: A meta-analytic review. *Personality and Social Psychology Review*, 10, 336–353. [http://dx.doi.org/10.1207/s15327957pspr1004\\_4](http://dx.doi.org/10.1207/s15327957pspr1004_4).
- Roseth, C. J., Johnson, D. W., & Johnson, R. T. (2008). Promoting early adolescents' achievement and peer relationships: The effects of cooperative, competitive, and individualistic goal structures. *Psychological Bulletin*, 134, 223–246. <http://dx.doi.org/10.1037/0033-2909.134.2.223>.
- Rothstein, H. R. (2008). Publication bias as a threat to the validity of meta-analytic results. *Journal of Experimental Criminology*, 4, 61–81. <http://dx.doi.org/10.1007/s11292-007-9046-9>.
- Rutland, A. (2004). The development and self-regulation of intergroup attitudes in children. In M. Bennett, & F. Sani (Eds.), *The development of the social self* (pp. 247–265). New York, NY: Psychology Press.
- Rutland, A., Killen, M., & Abrams, D. (2010). A new social-cognitive developmental perspective on prejudice: The interplay between morality and group identity. *Perspectives on Psychological Science*, 5, 279–291. <http://dx.doi.org/10.1177/1745691610369468>.
- Schofield, J. W. (1995). Promoting positive intergroup relations in school settings. In W. D. Hawley, & A. W. Jackson (Eds.), *Toward a common destiny: Improving race and ethnic relations in America* (pp. 257–289). San Francisco, CA: Jossey-Bass.
- Schofield, J., & Hausmann, L. R. M. (2004). School desegregation and social science research. *American Psychologist*, 59, 538–546. <http://dx.doi.org/10.1037/0003-066X.59.6.538>.
- Staub, E. (2006). Reconciliation after genocide, mass killing, or intractable conflict: Understanding the roots of violence, psychological recovery, and steps toward a general theory. *Political Psychology*, 27, 867–894.
- Stephan, W. G., & Finlay, K. (1999). The role of empathy in improving intergroup relations. *Journal of Social Issues*, 55, 729–743. <http://dx.doi.org/10.1111/0022-4537.00144>.
- Stephan, C. W., Renfro, L., & Stephan, W. G. (2004). The evaluation of multicultural education programs: Techniques and a meta-analysis. In W. G. Stephan, & W. P. Vogt (Eds.), *Education programs for improving intergroup relations* (pp. 227–242). New York, NY: Teachers College Press.
- Stephan, W. G., & Stephan, C. W. (2001). *Improving intergroup relations*. Thousand Oaks, CA: Sage.
- Stephan, W. G., & Vogt, W. G. (Eds.). (2004). *Education programs for improving intergroup relations*. New York, NY: Teachers College Press.
- Tausch, N., Schmid, K., & Hewstone, M. (2010). The social psychology of intergroup relations. In G. Salomon, & E. Cairns (Eds.), *Handbook on peace education* (pp. 75–86). New York, NY: Psychology Press.
- Tropp, L. R. (Ed.). (2012). *The Oxford handbook of intergroup conflict*. New York: Oxford University Press.
- Tropp, L. R., & Mallett, R. K. (2011). *Moving beyond prejudice reduction: Pathways to positive intergroup relations*. Washington, DC: APA.
- Tropp, L. R., & Pettigrew, T. F. (2005). Differential relationships between intergroup contact and affective and cognitive dimensions of prejudice. *Personality and Social Psychology Bulletin*, 31, 1145–1158 (Retrieved from <http://psp.sagepub.com/content/31/8/1145>)
- Tropp, L. R., & Prenovost, M.A. (2008). The role of intergroup contact in predicting children's interethnic attitudes: Evidence from meta-analytic and field studies. In S. R. Levy, & M. Killen (Eds.), *Intergroup attitudes and relations in childhood through adulthood* (pp. 236–248). New York, NY: Oxford University Press.
- van Quaquebeke, N., Henrich, D., & Eckloff, T. (2007). "It's not tolerance I'm asking for, it's respect!" A conceptual framework to differentiate between tolerance, acceptance and (two types of) respect. *Gruppendynamik und Organisationsberatung*, 38, 185–200. <http://dx.doi.org/10.1007/s11612-007-0015-6>.
- Verkuyten, M. (2010). Multiculturalism and tolerance: An intergroup perspective. In R. Crisp (Ed.), *The psychology of social and cultural diversity* (pp. 147–170). Chichester, England: Wiley-Blackwell.
- Walker, P. A. (1971). *The effects of hearing selected children's stories that portray Blacks in a favorable manner on the racial attitudes of groups of Black and White kindergarten children*. (Doctoral dissertation). (Retrieved from ProQuest Dissertations and Theses database. (UMI No. 7209425)).
- Williams, J. E., Best, D. L., Boswell, D. A., Mattson, L. A., & Graves, D. J. (1975). Preschool racial attitude measure II. *Educational and Psychological Measurement*, 35, 3–18. <http://dx.doi.org/10.1177/001316447503500101>.
- Williams, D. R., Neighbors, H. W., & Jackson, J. S. (2003). Racial/ethnic discrimination and health: Findings from community studies. *American Journal of Public Health*, 93, 200–208. <http://dx.doi.org/10.2105/AJPH.93.2.200>.
- Wright, S.C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, 73, 73–90. <http://dx.doi.org/10.1037/0022-3514.73.1.73>.

## FURTHER READING

- \*Aday, R. H., Sims, C. R., & Evans, E. (1991). Youth's attitudes toward the elderly: The impact of intergenerational partners. *Journal of Applied Gerontology*, 10, 372–384. <http://dx.doi.org/10.1177/073346489101000310>.
- \*Alexander, R. C. (1977). *A moral education curriculum on prejudice*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 7711349).

- \*Avci-Werning, M. (2004). Prävention ethnischer Konflikte in der Schule: Ein Unterrichtsprogramm zur Verbesserung interkultureller Beziehungen [Preventing ethnic conflicts in schools: A teaching program to improve intercultural relations]. In U. Wagner (Ed.), *Texte zur Sozialpsychologie, Vol. 9*, Münster, Germany: Waxmann.
- \*Bagley, C., & Verma, G. K. (1972). Some effects of teaching designed to promote understanding of racial issues in adolescence. *Journal of Moral Education, 1*, 231–238. <http://dx.doi.org/10.1080/0305724720010307>.
- \*Beelmann, A., Saur, M., Ziegler, P., Diener, K., & Noack, P. (2010). The PARTS-Project: The evaluation of a multimodal intervention program to prevent prejudice and promote intergroup relations in elementary school children. *Paper presented at the 18th annual meeting of the Society on Prevention Research, Denver, CO, USA*.
- \*Berg, M., & Wolleat, P. (1973). A comparison of the effects of information and the effects of contact on children's attitudes toward other national groups. *California Journal of Educational Research, 24*, 200–208.
- \*Bernstein, J., Schindler Zimmerman, T., Werner-Wilson, R. J., & Vosburg, J. (2000). Pre-school children's classification skills and a multicultural education intervention to promote acceptance of ethnic diversity. *Journal of Research in Childhood Education, 14*, 181–192. <http://dx.doi.org/10.1080/02568540009594762>.
- \*Blunk, E. M., & Williams, S. W. (1997). The effects of a curriculum on preschool children's perception of the elderly. *Educational Gerontology, 23*, 233–241. <http://dx.doi.org/10.1080/0360127970230303>.
- \*Bronzo, A. F. (1970). *Changes in prejudice and cognitive behavior in high school students as a function of instruction*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 7024351).
- \*Bucich-Naylor, D. A. (1978). *The comparative effectiveness of a direct contact program and a didactic program in effecting changes in attitude of nondisabled children toward disabled children*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 7820694).
- \*Buege, C. (1993). The effect of mainstreaming on attitude and self-concept using creative drama and social skills training. *Youth Theatre Journal, 7*, 19–22.
- \*Cameron, L., Rutland, A., & Brown, R. (2007). Promoting children's positive intergroup attitudes towards stigmatized groups: Extended contact and multiple classification skills training. *International Journal of Behavioral Development, 31*, 454–466. <http://dx.doi.org/10.1177/0165025407081474>.
- \*Cameron, L., Rutland, A., Brown, R., & Douch, R. (2006). Changing children's intergroup attitudes toward refugees: Testing different models of extended contact. *Child Development, 77*, 1208–1219. <http://dx.doi.org/10.1111/j.1467-8624.2006.00929.x>.
- \*Clunies-Ross, G., & O'Meara, K. (1989). Changing the attitudes of students towards peers with disabilities. *Australian Psychologist, 24*, 273–284. <http://dx.doi.org/10.1080/0005006908259566>.
- \*Cross, C. W. (1990). *The influence of a positive portrayal of specific elderly individuals in a film series on the cognitive, affective, and behavioral components of children's attitudes toward the elderly*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 9012450).
- \*Dellmann-Jenkins, M., Lambert, D., Fruit, D., & Dinero, T. (1986). Old and young together: Effect of an educational program on preschoolers' attitudes toward older people. *Childhood Education, 62*, 206–212.
- \*Feinman, G. (1982). *The impact of humanistic education instruction on self-concept, racial prejudice, school attendance, school performance, and discipline*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 8223597).
- \*Finazzo, D. A. (1992). *The effects of a multicultural curriculum on the attitudes of first-grade students*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 9221127).
- \*Fisher, F. L. (1968). Influences of reading and discussion on the attitudes of fifth graders toward American Indians. *Journal of Educational Research, 62*, 130–134 (Retrieved from <http://pao.chadwyck.co.uk/PDF/1278599335431.pdf>)
- \*Florian, V., & Kehat, D. (1987). Changing high school students' attitudes toward disabled people. *Health & Social Work, 12*, 57–63 (Retrieved from <http://pao.chadwyck.co.uk/PDF/1278599508248.pdf>)
- \*Frech, W. P. (1975). The effect of cognitive training in anthropology on ethnocentric attitudes. *Psychology in the Schools, 12*, 364–370 doi:10.1002/1520-6807(197507)12:3<364::AID-PITS2310120324>3.0.CO;2-8.
- \*Gilbert, M.A. (1999). *Effects of exposure to well-elderly on children's negative stereotypical attitudes*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 9933932).
- \*Gimmestad, B. J., & de Chiara, E. (1982). Dramatic plays: A vehicle for prejudice reduction in the elementary school. *Journal of Educational Research, 76*, 45–49 (Retrieved from <http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=36&hid=111&sid=24c1d358-a3e2-462d-a748-1bddc6dab16c%40sessionmgr112>)
- \*Glass, J. C., Jr., & Trent, C. (1980). Changing ninth-graders' attitudes toward older persons: Possibility and persistence through education. *Research on Aging, 2*, 499–512. <http://dx.doi.org/10.1177/016402758024007>.
- \*Gorn, G. J., Goldberg, M. E., & Kanungo, R. N. (1976). The role of educational television in changing the intergroup attitudes of children. *Child Development, 47*, 277–280. <http://dx.doi.org/10.2307/1128313>.
- \*Gronberg, G. W. (1983). *Attitude responses of non-handicapped elementary students to specific information and contact with the handicapped*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 8305232).
- \*Guyton, J. (1987). *A developmental counseling approach to alter self-concept and racial prejudice in elementary children*. (Doctoral "dissertation"). Retrieved from ProQuest Dissertations and Theses database (UMI No. 8718814).
- \*Gwinn, C. B. (1998). The effect of multicultural literature on the attitudes of second grade students. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 9823817).
- \*Harte, T. F. (1981). The relationship between self-concept, empathy training, and attitudes toward the physically handicapped in ten to eleven year old boys and girls. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 8110520).
- \*Hébert, M., Voyer, J., & Valois, D. (2000). Evaluation du programme 'Les préjugés ... J'connais pas!' Auprès des jeunes de secondaire V. [Evaluation of the program "Prejudices... I don't know of any!" among youth in senior high school]. *Canadian Journal of Community Mental Health, 19*, 105–126.
- \*Hohn, R. L. (1973). Perceptual training and its effect on racial preferences of kindergarten children. *Psychological Reports, 32*, 435–441.
- \*Houlette, M.A., Gaertner, S. M., Johnson, K. M., Banker, B.S., Riek, B.M., & Dovidio, J. F. (2004). Developing a more inclusive identity: An elementary school intervention. *Journal of Social Issues, 60*, 35–55. <http://dx.doi.org/10.1111/j.0022-4537.2004.00098.x>.
- \*Houser, B. B. (1978). An examination of the use of audiovisual media in reducing prejudice. *Psychology in the Schools, 15*, 116–122. [http://dx.doi.org/10.1002/1520-6807\(197801\)15:1<116::AID-PITS2310150122>3.0.CO;2-G](http://dx.doi.org/10.1002/1520-6807(197801)15:1<116::AID-PITS2310150122>3.0.CO;2-G).
- \*Howell, M. (1974). A study of the effects of reading upon the attitudes of fifth graders toward Mexican Americans. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 7406211).
- \*Jugert, G., Kabak, S., & Notz, P. (2006). Fit for Differences: Training interkultureller und sozialer Kompetenz für Jugendliche [Fit for differences: A training in intercultural and social competence for adolescents]. Weinheim, Germany: Juventa.
- \*Katz, P. A. (1973). Stimulus predifferentiation and modification of children's racial attitudes. *Child Development, 44*, 232–237. <http://dx.doi.org/10.2307/1128041>.
- \*Kehoe, J. W., & Rogers, T. R. (1978). The effects of principle-testing discussions on student attitudes towards selected groups subjected to discrimination. *Canadian Journal of Education/Revue canadienne de l'éducation, 3*, 73–80 Retrieved from <http://www.jstor.org/stable/1494686>
- \*Kimoto, C. K. (1974). *The effects of a juvenile literature based program on majority group attitudes toward Black Americans*. (Doctoral dissertation). (Retrieved from ProQuest Dissertations and Theses database. (UMI No. 7416377)).
- \*Kowalski, K. (1998). The impact of vicarious exposure to diversity on preschooler's emerging ethnic/racial attitudes. *Early Child Development and Care, 146*, 41–51. <http://dx.doi.org/10.1080/0300443981460105>.
- \*Krahé, B., & Altwasser, C. (2006). Changing negative attitudes towards persons with physical disabilities: An experimental intervention. *Journal of Community & Applied Social Psychology, 16*, 59–69. <http://dx.doi.org/10.1002/casp.849>.
- \*Küchel, J., & Beelmann, A. (2008). Promoting positive intergroup relations in elementary school children: Effects of an extended contact intervention. *Poster presented at the 29th international congress of psychology, Berlin, Germany July*.
- \*Lazar, A. L., Gensley, J. T., & Orpet, R. E. (1971). Changing attitudes of young mentally gifted children toward handicapped persons. *Exceptional Children, 37*, 600–602.
- \*Lessing, E. E., & Clark, C. C. (1976). An attempt to reduce ethnic prejudice and assess its correlates in a junior high school sample. *Educational Research Quarterly, 1*(2), 3–16.
- \*Leyser, Y., Cumbalad, C., & Strickman, D. (1986). Direct intervention to modify attitudes toward the handicapped by community volunteers: The learning about handicaps programme. *Educational Review, 38*, 229–236. <http://dx.doi.org/10.1080/0013191860380303>.
- \*Leyser, Y., & Price, S. (1985). Improving attitudes of gifted children toward the handicapped. *Education, 105*, 432–437 (Retrieved from <http://web.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=5&hid=108&sid=25a978b5-baeb-4d20-8487-33e973e468ae%40sessionmgr112>)
- \*Liebkind, K., & McAlister, A. (1999). Extended contact through peer modeling to promote tolerance in Finland. *European Journal of Social Psychology, 29*, 765–780 doi: 10.1002/(SICI)1099-0992(199908/09)29:5<765::AID-EJSP958>3.0.CO;2-J.
- \*Litcher, J. H., & Johnson, D. W. (1969). Changes in attitudes toward Negroes of White elementary school students after use of multiethnic readers. *Journal of Educational Psychology, 60*, 148–152. <http://dx.doi.org/10.1037/h0027081>.
- \*Litcher, J. H., Johnson, D. W., & Ryan, F. L. (1973). Use of pictures of multiethnic interaction to change attitudes of White elementary school students toward Blacks. *Psychological Reports, 33*, 367–372.
- \*McAlister, A., Ama, E., Barroso, C., Peters, R. J., & Kelder, S. (2000). Promoting tolerance and moral engagement through peer modeling. *Cultural Diversity and Ethnic Minority Psychology, 6*, 363–373. <http://dx.doi.org/10.1037/1099-9809.6.4.363>.
- \*McConkey, R., McCormack, B., & Naughton, M. (1983). Changing young people's perceptions of mentally handicapped adults. *Journal of Mental Deficiency Research, 27*, 279–290.
- \*Meshel, D. S., & McGlynn, R. P. (2004). Intergenerational contact, attitudes, and stereotypes of adolescents and older people. *Educational Gerontology, 30*, 457–479. <http://dx.doi.org/10.1080/03601270490445078>.
- \*Mitnick, L. L., & McGinnies, E. (1958). Influencing ethnocentrism in small discussion groups through a film communication. *Journal of Abnormal and Social Psychology, 56*, 82–90. <http://dx.doi.org/10.1037/h0045323>.
- \*Peacock, B.D. (1992). The effects of an ethnic music program on attitudes of racial acceptance. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database (UMI No. 1348380).
- \*Persson, A., & Musher-Eizenman, D. R. (2003). The impact of a prejudice-prevention television program on young children's ideas about race. *Early Childhood Research Quarterly, 18*, 530–546. <http://dx.doi.org/10.1016/j.ecresq.2003.09.010>.
- \*Rich, P., Myrick, R. D., & Campbell, C. (1983). Changing children's perceptions of the elderly. *Educational Gerontology, 9*, 483–491. <http://dx.doi.org/10.1080/0380127830090512>.
- \*Schon, I., Hopkins, K. D., & Vojir, C. (1982). The effects of special curricular study of Mexican culture on Anglo and Mexican-American students' perceptions of Mexican-Americans. *Journal of Experimental Education, 50*, 215–218 (Retrieved from <http://pao.chadwyck.co.uk/PDF/1278685826218.pdf>)
- \*Schulze, B., Richter-Werling, M., Matschinger, H., & Angermeyer, M. C. (2003). Crazy? So what! Effects of a school project on students' attitudes towards people with schizophrenia. *Acta Psychiatrica Scandinavica, 107*, 142–150. <http://dx.doi.org/10.1034/j.1600-0447.2003.02444.x>.

- \*Slone, M., Tarrasch, R., & Hallis, D. (2000). Ethnic stereotypic attitudes among Israeli children: Two intervention programs. *Merrill-Palmer Quarterly*, 46, 370–390.
- \*Sousa, M., Neto, F., & Mullet, E. (2005). Can music change ethnic attitudes among children? *Psychology of Music*, 33, 304–316. <http://dx.doi.org/10.1177/0305735605053735>.
- \*Trent, C., Glass, J. C., & Crockett, J. (1979). Changing adolescent 4-H Club members' attitudes toward the aged. *Educational Gerontology*, 4, 33–48. <http://dx.doi.org/10.1080/0360127790040104>.
- \*Turner, R. N., & Brown, R. (2008). Improving children's attitudes toward refugees: An evaluation of a school-based multicultural curriculum and an anti-racist intervention. *Journal of Applied Social Psychology*, 38, 1295–1328. <http://dx.doi.org/10.1111/j.1559-1816.2008.00349.x>.
- \*Weiner, M. J., & Wright, F. E. (1973). Effects of undergoing arbitrary discrimination upon subsequent attitudes toward a minority group. *Journal of Applied Social Psychology*, 3, 94–102. <http://dx.doi.org/10.1111/j.1559-1816.1973.tb01298.x>.
- \*Wham, M.A., Barnhart, J., & Cook, G. (1996). Enhancing multicultural awareness through the storybook reading experience. *Journal of Research & Development in Education*, 30, 1–9.
- \*Winters, J. (1994). *The effect of a literature intervention program on children's attitudes toward the elderly*. (Doctoral dissertation). (Retrieved from ProQuest Dissertations and Theses database. (UMI No. 1361227)).
- \*Yawkey, T. D. (1973). Attitudes toward black Americans held by rural and urban white early childhood subjects based upon multi-ethnic social studies materials. *Journal of Negro Education*, 42, 164–169. <http://dx.doi.org/10.2307/2967011>.
- \*Zakay, D. (1985). The influence of information and daily contact on children's attitudes towards aphasic children. *British Journal of Educational Psychology*, 55, 1–10. <http://dx.doi.org/10.1111/j.2044-8279.1985.tb02600.x>.
- Dellmann-Jenkins, M., Lambert, D., & Fruit, D. (1991). Fostering preschoolers' prosocial behaviors toward the elderly: The effect of an intergenerational program. *Educational Gerontology*, 17, 21–32. <http://dx.doi.org/10.1080/0360127820170103>.
- Mathisen, J. H. (2000). *Stigma busting: Does strategic contact with individuals with severe mental illness reduce negative attitudes in an adolescent population*. (Doctoral dissertation). (Retrieved from ProQuest Dissertations and Theses database. (UMI No. 9937659)).
- Melloy, K. J. (1991). *Attitudes and behavior of nondisabled elementary-aged children toward their peers with disabilities in integrated settings: An examination of the effects of treatment on quality of attitude, social status and critical social skills*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 9103242).
- Robinson, J. F. (1981). *The effect of empathy training on the modification of attitudes and interpersonal effectiveness of learning disabled and nonhandicapped*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 8120370).
- Schultz, L. H., Barr, D. J., & Selman, R. L. (2001). The value of a developmental approach to evaluating character development programmes: An outcome study of Facing History and Ourselves. *Journal of Moral Education*, 30, 3–27. <http://dx.doi.org/10.1080/03057240120033785>.
- Shirley, O. B. (1988). *The impact of multicultural education on the self-concept, racial attitude, and student achievement of Black and White fifth and sixth graders*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 8818156).
- Simpson, B. V. (2007). *Exploring the influences of educational television and parent-child discussions on improving children's racial attitudes*. (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3268375).
- Simpson, R. L., Parrish, N. E., & Cook, J. J. (1976). Modification of attitudes of regular class children towards the handicapped for the purpose of achieving integration. *Contemporary Educational Psychology*, 46–51. [http://dx.doi.org/10.1016/0361-476X\(76\)90006-0](http://dx.doi.org/10.1016/0361-476X(76)90006-0).
- Thorman, A. E. (2003). *An evaluation of the impact of anti-bias curriculum practices on prejudicial racial attitudes among children attending Head Start*. (Doctoral dissertation). (Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3059649)).
- Westphal, R. C. (1974). *The effects of primary grade level interethnic curriculum on racial prejudice*. (Doctoral dissertation). (Retrieved from ProQuest Dissertations and Theses database. (UMI No. 7428468)).
- Whetstone, L. M. (1992). *An evaluation of prejudice reduction program for children*. (Doctoral dissertation). (Retrieved from ProQuest Dissertations and Theses database. (UMI No. 9135396)).