



DET PSYKOLOGISKE FAKULTETET



## **Exploring the Efficacy of Empathy-Raising Interventions in School Settings**

HOVEDOPPGAVE

*Profesjonsstudiet i Psykologi*

**Yannick Herrles**

Vår 2023

Veileder: Sebastian Jentschke

**Table of Contents**

Abstract.....	3
Sammendrag.....	4
Introduction.....	5
Empathy: Definition and Theoretical Conceptualization.....	5
The Importance of Fostering Empathy.....	7
Fostering Empathy at Home and at School.....	8
The Development of Empathy and the Impact of Genetics.....	11
Can Empathy be Raised for Anyone, and Anywhere?.....	13
Empathy's Role in Bullying Interventions.....	13
Possibilities of Promoting Empathy through Related Constructs.....	15
Global versus Specific Levels of Empathy.....	18
Recent Meta-Analytic Summaries of the Field.....	19
Methods.....	21
Search Procedure.....	21
Results.....	24
Description of the Individual Studies Included in this Literature Review.....	24
Study Characteristics and How They Affected the Observed Effect Sizes.....	32
Discussion.....	45
Overall Effects of Empathy Interventions.....	45
Effects specific to particular study characteristics.....	46
Limitations of the Review.....	57
Directions and Implications for Future Research.....	57
Concluding Remarks.....	61
References.....	62
Appendix.....	72

**Abstract**

This literature review explores the efficacy of randomized control studies in raising empathy in children in schools, as well as explores how empathy gains are affected by and interact with different study characteristics. The literature review identified effect sizes for empathy gains that generally did not exceed small effect sizes, and where none reached a large effect size, casting doubt on the efficacy of interventions aimed at raising empathy in children in schools. Empathy gains did not vary substantially across study characteristics. Exploring the relation between raising empathy and the effects on bullying reduction, effects were highly variable, and results suggested that affective empathy may be more effective in promoting this prosocial outcome. Implications for future research are discussed, noting a need to further explore variables that may affect interventions intending to raise empathy, and the distinct possibilities and effects of promoting affective versus cognitive empathy.

*Keywords:* Empathy, Children, School, Interventions, Literature Review

### **Sammendrag**

Denne litteraturoversikten utforsker effekten av intervensjoner rettet mot å øke empati i barn i skoler, samt hvordan studiekarakteristikk påvirker og interagerer med intervensjonens effekt. Litteraturoversikten identifiserte effektstørrelser som generelt sett ikke var større enn liten, og ingen av studiene rapporterte en stor effektstørrelse. Det ble ikke funnet betydelig variasjon i effektstørrelser på tvers av studiekarakteristikk. I utforskningen av relasjonen mellom økninger i empati og reduksjoner i mobbing fant litteraturoversikten svært varierende effektstørrelser, og affektiv empati ble identifisert som potensielt mer effektiv i å føre til reduksjoner i mobbing. Implikasjoner for framtidig forskning ble diskutert, og det ble vektlagt et behov for videre forskning på karakteristikk som påvirker intervensjoner rettet mot å øke empati i barn i skoler, samt et behov for forskning på distinkte effekter av affektiv versus kognitiv empati.

*Nøkkelord:* Empati, Barn, Skole, Intervensjoner, Litteraturoversikt

### **Exploring the Efficacy of Empathy-Raising Interventions in School Settings**

As society transitions into a more harmonious existence than ever before (e.g., Pinker, 2012), the eradication of conflict still seems an unreachable utopia. Impelled by ubiquitous human pain and suffering, the capacity for empathy has been at the forefront of the minds of philosophers and scientists for centuries. Psychology has identified empathy as a potent antidote against self-serving brutality, fostering the innate human capacity to care for those around us.

This paper seeks to explore whether empathy can be systematically promoted for children in schools. Its goal is to evaluate the contents and effects of empathy raising interventions for children in schools through a literature review. The review limits itself to: a) articles published between 2013 and the time point of the literature search (in February 2023; i.e., the past ten years), b) studies utilizing a randomized control trial design, c) the inclusion of a post-test measure of empathy, and d) only children with typical development (excluding, e.g., populations with neurological disorders such as autism).

The introduction aims to provide an overview over theoretical conceptualizations of empathy and its importance, its course of development, ways to foster empathy, connections with prosocial behavior, its limitations, and current standings and research in the field. The results aim to investigate effect sizes of empathy gains in school children due to interventions, how and to what degree they are affected by intervention characteristics, and relations between empathy and bullying. The discussion reflects on the results and on potential explanatory mechanisms, then denotes limitations of the literature review, and finally highlights avenues for future research.

#### **Empathy: Definition and Theoretical Conceptualization**

Malti et al. (2016) defined empathy as “requiring (and primarily involving) experiencing the same or a similar emotion as the other” (p. 719). A subdivision is made

between cognitive and affective empathy, both of which are core to moral development and behavior. Cognitive empathy is defined as “the ability to intellectually take the role or perspective of another person, involving the ability to decode and label emotions and their situational cues” (Thompson & Melchior, 2019, p. 200). Affective empathy is defined as “an affective response, more appropriate to or congruent with someone else’s situation than to one’s own situation” (Thompson & Melchior, 2019, p. 200). Thus, cognitive empathy allows us to apprehend and understand other’s emotions on a cognitive level, while affective empathy encapsulates the personal emotional activation that comes in response to another’s emotion. Their distinct nature is exemplified by high levels of one empathy being able to coexist with low levels of the other (i.e., low cognitive empathy but normal levels of affective empathy in autism, or low affective empathy but normal levels of cognitive empathy in psychopathy; Smith, 2006).

Sympathy represents a dimension of empathy, and refers to feeling sorry or concern for the distressed or needy other, rather than (only) feeling the same emotion as the other person is experiencing or is expected to experience (Spinrad & Gal, 2018). Sympathy can thus be described as stemming from the affective part of empathy (Malti et al., 2016). Core to affective empathy being transformed into sympathy is an understanding that the emotional experience is related not to the self but to the other (Malti et al., 2016). Another conceptually similar term often used in the literature is empathic concern, which is defined as “an emotional response of compassion and concern caused by witnessing someone else in need” (Niezink et al., 2012, p. 1). Many of the items in empathy scales essentially measure sympathy and empathic concern rather than pure affective empathy, and the use of empathy and affective empathy in this paper generally encompasses these dimensions.

Experiencing the same emotion as the other person without the consequent activation of sympathy can result in personal distress. Personal distress involves an unregulated affective

apprehension of the other's emotion, often leading to a protective shut-down of one's willingness to take in another's emotional state, and can thus be counteractive to empathy (Malti et al., 2016). The regulation of personal distress is therefore crucial to activating sympathy (Malti et al., 2016). Whereas personal distress induces feeling an emotion similar to the other, sympathy at least partially transforms this response into feeling concern (Spinrad & Gal, 2018). Emotion understanding, as well as emotion regulation and expression skills, all combine to make this possible (Malti et al., 2016).

### **The Importance of Fostering Empathy**

The importance of fostering empathy in children lies in its links with prosocial behavior. Empathy has repeatedly been linked to the promotion of children's other-oriented, prosocial behavior and reductions in antisocial maladaptive behavior (Malti et al., 2016). Empathy positively predicted coming to the defense of victimized peers, a core goal of many anti-bullying programs (Garandeanu & Salmivalli, 2018). Healthy empathy development was associated with lower levels of aggression, being more inclusive to others, and increased perspective taking in children (Thompson & Melchior, 2019). When sympathy emerges from empathy, altruistic motivation follows (Spinrad & Gal, 2018). In young children, sympathy and sometimes empathy have been linked to helping others (Spinrad & Gal, 2018).

Recent research based on functional decoding analysis, which describes "methods which attempt to predict mental states from neuroimaging data" (NiMARE, n.d.), revealed that emotion regulation and empathy were related through similar internally oriented and executive neural processes (Morawetz et al., 2022), and emotion regulation itself is related to many positive outcomes (Keefer et al., 2018). For instance, youth were less likely to behave aggressively when they learned to cope with stress and to regulate emotion (Keefer et al., 2018), which may minimize the interference of personal distress with empathy.

It can also be theorized that fostering empathy in schools may have echoing effects on other domains in life, for instance engendering more empathic sibling treatment or more prosocial views held by one's peer group. These are likely to reinforce other domains and areas in life, leading to cycles of reciprocal reinforcement.

### **Fostering Empathy at Home and at School**

Much research on the promotion of empathy was conducted in home settings, and it reveals important insights for interventions directed at raising empathy in schools.

#### ***Raising Empathy at Home***

The review *Fostering prosocial behavior and empathy in young children* by Spinrad and Gal (2018) provides a good overview of raising empathy in the home context, and a summary of it is provided here. Emotion socialization appears to play a key role in fostering empathy within the home setting. When mothers encouraged the expression of emotions at 18 months of age, toddlers were perceived as more empathic at age 2. Children's sympathy and prosocial behavior was also impacted by parent's use of strategies that model and down-regulate the child's personal distress. Frequent talk about emotions, as well as mothers' greater understanding of their child's feelings and needs, related to higher levels of empathy. The increased emotions' knowledge was likely to relate to an increase in prosocial behavior. A salient mechanism appeared to be the drawing of children's attention to others around them in order to stimulate empathy.

The same review (Spinrad & Gal, 2018) further indicated that a child's emotional knowledge was positively associated with parental warmth and responsiveness. This form of parenting involved more time spent explaining emotional content, while authoritarian parenting may have evoked feelings of fear that impede the internalization of moral scripts. Furthermore, rewarding children with material goods was linked with decreased prosocial behavior. Nonmaterial reinforcements such as encouragement of the behavior and verbal



praise were more conducive to fostering internal scripts and motivation for prosocial behavior. Finally, toddlers who helped their mothers with everyday chores also tended to show more prosocial interactions with the experimenter, indicating that the participation in prosocial behavior may be a reinforcing mechanism.

These findings are of great relevance for interventions directed at raising empathy in schools. Such interventions can seek to mimic natural positive developments found within child parent interactions, and integrate these into empathy-raising designs. Studies reviewed in this paper make use of similar methods and reported significant effects.

### ***Raising Empathy at School***

Schools represent an important arena for the raising of empathy. Failing to promote empathy or passively condoning hostility may lead to reduced expressions of empathy and accompanying personality changes through epigenetic interactions with one's environment (Holbrook & Hahn Holbrook, 2022). For instance, persons who had experienced hostile life events tended to see their environment as more hostile (Bernstein et al., 2018), which can often turn into a self-reinforcing cycle and may potentially undermine empathy.

Within the bounds of raising empathy in schools, programs differed in their focus on socio-affective versus socio-cognitive skills (Malti et al., 2016). Programs that focused on socio-affective aspects frequently combined experience and practice based learning, for instance through reading that activates emotional empathy, as well as through engagement in prosocial activity. Other interventions emphasized socio-cognitive dimensions of empathy, such as perspective taking. These served in order to increase the recognition of the contents of others' needs and perspectives. While the relative focus may vary, most interventions combined these two dimensions of empathy to create a more robust intervention protocol (Malti et al., 2016).

Social and emotional learning programs have long been at the forefront of promoting socio-emotional competencies in schools, and include the promotion of empathy in their core curriculum (Malti et al., 2016). Two of the five competencies promoted by social and emotional learning programs are closely related to empathy development. The competency of self-awareness describes “the ability to accurately recognize one’s emotions and thoughts and their influence on behavior” (Malti et al., 2016, p. 720), whereas the competency of social awareness defines “the ability to take the perspective of and empathize with others from diverse backgrounds and cultures, to understand social and ethical norms for behavior, and to recognize family, school, and community resources” (Malti et al., 2016, p. 720). While the former encapsulates the understanding and recognition of one’s own emotions, thoughts and their impact on behavior, the latter refers to other oriented processes such as perspective taking and affective empathy (Malti et al., 2016).

In a meta-analytic evaluation of 213 school-based social and emotional learning programs by Durlak et al. (2011), a number of commonly employed methods for promoting socio-emotional competencies were reported. Among these, freestanding lessons provided explicit instruction to teach social and emotional competencies across the five competency areas of social awareness, self-awareness, decision-making, self-management, and relationship skills. Topics ranged from setting goals and developing empathy and compassion, to resolving conflict, and making responsible decisions. Other learning techniques included discussions, role-play, and small group work. Opportunities were included for practicing skills beyond classroom or other connected academic lessons. An example for an elementary school classroom is students working together to resolve interpersonal problems (Salay, 2018).

### **The Development of Empathy and the Impact of Genetics**

Empathy skills appear to develop at different time points, which is important to consider in empathy raising interventions. At the age between nine and twelve years, children turn less egocentric: They increasingly see the feelings and perspectives of others, and develop a sense of wrong and right (Eisenberg et al., 2006). Longitudinal and cross-sectional research indicated that increases in concern for others occurred from middle childhood to early adolescence, though these may have limited themselves to cognitive empathy through adolescence (Malti et al., 2016). Thus, it may be easier to stimulate cognitive empathy in adolescents compared to affective empathy (Garandeanu & Salmivalli, 2018). Affective empathy appears to stabilize in adolescence, potentially making its promotion more difficult (Malti et al., 2016). However, despite this apparent negligible increase in affective empathy through development reported in longitudinal studies, children's self-regulatory capacities improved (Malti et al., 2016). As mentioned, targeting this capacity may work to regulate personal distress and convert affective empathy into sympathy.

The development of empathy across age appears to be influenced by the environment. A research analysis dissertation by Duong (2014) found that empathy in middle and late childhood remained stable for the majority of youth (60.8%), but also identified an early decreasing group (21.6%, moderate decline in empathy between grade 3 and grade 5) and a late decreasing group (17.6%, sharp decline from grade 4 to the end of grade 5). Interestingly, ecological context significantly differed between the trajectories, with the stable group coming from more well-off homes and community contexts. This underscores the potential importance of the environment, as it can be hypothesized that parental and community resources positively impact empathy. Possibly, similar findings could emerge on empathy gains according to the socio-economic resources of school students.

In accordance with differential developmental timings, the efficacy of promoting socio-emotional outcomes and also likely empathy varies across ages and grade levels. A meta-analysis by Boncu et al. (2017) on the efficiency of social and emotional learning programs reported on youth, with small effect sizes, but not on adolescents: The largest effect size was observed for youth aged seven to twelve ( $p < .01$ , Hedges'  $g = 0.38$ ), a little smaller effect was observed in youth aged three to six ( $p < .01$ , Hedges'  $g = 0.31$ ), whereas there was no significant difference related to the intervention in adolescents aged 13 to 18 ( $p > .05$ ).

Anti-bullying interventions, many aimed at raising empathy, have often been proven to be more effective in primary than in secondary school age (Garandeau & Salmivalli, 2018). Anti-bullying programs generally experienced a substantial drop in effectiveness from around grade 7, potentially in connection with the clash between adolescent drive for autonomy and their perception of intervention curricula as too controlling (Ryzin & Roseth, 2019).

Overall, a meta-analysis by Teding van Berkhout & Malouff (2016) identified few studies on empathy training in adolescent populations, and those identified did not show a significant overall effect. Thus, there appears to be a need for more research with younger and adolescent populations' potential for empathy training to reliably identify key timings.

Genetically conditioned differences in baseline levels of empathy also appear to exist, which may have impacted the efficacy of interventions. In terms of heritability, estimates for affective empathy ranged from 30% (e.g., Davis et al., 1994) among high school seniors to 68% (e.g., Rushton et al., 1986) among adults. For cognitive empathy, estimates were slightly lower, at around 28% (e.g., Warrier et al., 2018). Nonetheless, while biological factors appear to play a role, environmental influence is considerable, supporting the importance of interventions to raise empathy in various settings. Studies with toddlers revealed even lower estimates, indicating greater malleability of empathy at an earlier age (e.g., Knafo et al., 2008). Thus, interventions that aim to raise empathy in specific grade levels need to be

cognizant of biological intra-classroom variations in empathy and its consequences on individual receptivity to the intervention (Malti et al., 2016).

### **Can Empathy be Raised for Anyone, and Anywhere?**

Much debate has been around whether empathy can be raised across differing contexts and individual characteristics, and findings have largely supported this possibility in conventionally high risk groups such as bullies in anti-bullying interventions. Bullying perpetrators tended to be deficient in affective empathy (Garandau et al., 2021), opening up the question whether bullies can raise their levels of empathy just as effectively, or even more effectively, than uninvolved peers.

In support of the potential to raise prosocial outcomes related to empathy, van Baardewijk et al. (2009) reported that young bullies and children with psychopathic traits chose to be less aggressive in a computer game when they became aware of the distress of their virtual opponent, expressing fear via a message. Children high in callous unemotional traits increased in affective empathy and had fewer conduct problems in response to emotion recognition training, and bullying perpetrators in grade 6 increased in empathy after an empathy training program (Garandau et al., 2021). Interestingly, Raskauskas and Stoltz (2007) reported that almost 40% of cyberbullies engaged in their behavior for fun. This sheds light on a potential lack of empathy for a specific group that does not stem directly from more commonly assumed reasons such as retaliation or direct conflict. In a sense, for a sizable group of cyberbullying perpetrators, it points to a lack of awareness rather than a direct desire to hurt, which may potentially be ameliorated through fostering the capacity of empathy.

### **Empathy's Role in Bullying Interventions**

Of great interest to empathy interventions is whether empathy is related to prosocial societal outcomes such as reductions in bullying, and whether raising empathy in interventions affects bullying behaviors and attitudes. In support of this association, findings

suggest that core to successful anti bullying programs is a combination of (a) motivating children and adolescents that themselves engage in bullying to care about the feelings of others, and (b) empowering classmates by teaching them perspective taking and bullying intervention abilities (Henry & Ruffman, 2022). Defenders of victims often had empathy towards those in need and insight into others perspectives (Henry & Ruffman, 2022).

However, the link between empathy and bullying is not always clear-cut, and the literature diverges on the association between bullying and levels of either affective or cognitive empathy. In general, a negative relationship between empathy and aggressive behavior, bullying included, was found in a variety of studies (e.g., Zych et al., 2019). Most research explored affective empathy, finding mainly negative associations with bullying (e.g., van Noorden et al., 2014). Cognitive empathy had little to no association with bullying in some studies (e.g., Jolliffe & Farrington, 2010), while other research even identified negative correlations between cognitive empathy and bullying behaviors (e.g., van Noorden et al., 2014). Yet other studies suggest that bullying perpetrators actually surpassed their peers in cognitive empathy, which could have facilitated their strategic, abusive behavior (Garandau et al., 2021).

Findings additionally diverge for the categories of involvement in (stopping) bullying, apart from being a direct perpetrator of bullying. In defenders of bullying victims, affective empathy may make a stronger contribution to eliciting defending behavior, even though some such links have also been found for cognitive empathy (Garandau et al., 2021). These findings indicate that intervention contents may have to be tailored with a focus on affective empathy to achieve reductions, because the cognitive ability to understand the emotions of victims may not be sufficient to inhibit bullying in certain situations. It has been theorized that students able to experience the others' negative emotional state through affective empathy will be deterred from bullying behavior, while the lack of vicariously experiencing the others'

emotion in cognitive empathy may be insufficient to produce this behavior (Garandeau et al., 2021). However, even higher levels of affective empathy may not always be guaranteed to reduce levels of bullying: Salavera et al.(2021) did not find significant differences between affective and cognitive empathy when comparing victims and aggressors. These findings raise questions as to whether increasing cognitive or affective empathy in interventions would most effectively lead to prosocial outcomes such as reductions in bullying, whether effects would differ between groups, and what magnitude these effects would achieve. Based on the theoretical and neural distinctions between cognitive and affective empathy (Moore et al., 2015), it is likely that they relate differentially to prosocial outcomes and differ in terms of their effects on target groups.

In recent years, anti-bullying interventions have increasingly integrated prosocial, empathy-based approaches to reduce bullying, supplanting more traditional authoritarian interventions. These changes to bullying interventions were echoed in the findings of a dissertation by Roubin (2014), where representatives from the Los Angeles Unified School District voiced (in discussions with Roubin) “a desire to decrease use of traditional anti-bullying intervention strategies in preference of more prosocial, empathy-based approaches to bullying” (p. 28). Core to these approaches were empathy based approaches that prioritized prosocial skills training, empowering victims and bystanders to intervene in peer bullying through positive reinforcement (Roubin, 2014).

### **Possibilities of Promoting Empathy through Related Constructs**

Interventions that target negative and prejudiced beliefs may also induce positive changes in empathy levels, and empathy raising interventions could incorporate the deconstruction of such beliefs through the raising of knowledge. In support of this potential theoretical association, research on the contact theory revealed the power of interpersonal contact in reducing biased and prejudiced beliefs, finding that gains in empathy mediated

these effects (Pettigrew & Tropp, 2008). An intervention utilizing Harry Potter books aimed at reducing prejudiced beliefs likewise led to increases in empathy (Vezzali et al., 2014). A potential mechanism to explain this could be that the promotion of accurate knowledge by the intervention could counteract negative and prejudiced beliefs that otherwise would have reinforced negative and hostile views of an out-group, thus increasing empathy. This indicates that reductions in prejudiced beliefs thereby may lead to increases in empathy.

Specific sets of negative false beliefs often underlie the hostile attribution error, and an intervention that disproves such beliefs through raising knowledge could potentially raise empathy. The hostile attribution error is defined as an “interpretation bias in which individuals were more likely to interpret ambiguous situations as hostile than benign” (Wang et al., 2018, p. 2). When attributions to exclude a target from a group were experimentally manipulated, internal attributions (blaming the target for being excluded) decreased desire for siding with the target compared to external or ambiguous attributions (Bernstein et al., 2018). Internal attributions consequently seem to undermine empathy for the target. Challenging the veracity of prejudiced beliefs could therefore improve empathy.

Fitzgibbons (1986) found that anger caused by offensive life events promoted the formation of a psychological defense mechanism, which reduced the individual’s compassion and empathy for the offender and promoted the generation of hostile thinking toward others. Individuals with particularly high levels of trait anger were especially prone to negative perceived bias in the face of threat-related information, and were more sensitive to hostile social cues (Bernstein et al., 2018). Thus, certain negative life events or biological factors may engender a greater predisposition towards making hostile attributions for the reasons behind others behaviors, decreasing empathy. Interventions could target beliefs underlying these mechanisms.



Empathy is strongly interrelated and moderately correlated to moral disengagement and attitudes (e.g. Barkoukis et al., 2015), which could likewise be targeted in interventions aiming to raise empathy. A study by Influs et al. (2018) confirmed that such attitudes were held by participants: The belief that justice is solely on one's side may have represented a cognitive prejudiced belief that led to lower measures of behavioral empathy in one-on-one interactions with out-group members. In the same study, (negative) moral beliefs and attitudes seemed to be at play in reducing the perception of pain experienced by out-groups, leading to reduced activation in brain areas processing pain when observing pain in other races versus one's own race (Influs et al., 2018). While the brain responses were initially the same, responses directed towards the other race were blunted after 500 ms, seemingly shut down by top-down cognitive control mechanisms (Influs et al., 2018). In bullying perpetration and cyberbullying acts, aspects of moral disengagement have also been implicated (Cross et al., 2015). Thus, fostering accepting attitudes, disproving beliefs through knowledge and increasing perspective-taking may help reverse the shut-down of the brain's empathic response.

A multitude of interventions report on the efficacy of raising prosocial behavior in schools without the explicit mention of empathy (Malti et al., 2016). Prosocial behavior is defined as "behavior through which people benefit others, including helping, cooperating, comforting, sharing, and donating" (Ding et al., 2018, p. 1). Empathy and prosocial behavior are related, as empathy "mediates prosocial behaviour when sensitivity to others' distress is paired with a drive towards their welfare" (Decety et al., 2016, p. 1). Scales that are assessing prosocial behavior often include items that specifically reflect levels of sympathy, and empathy scales likewise often include prosocial behavior items (Malti et al., 2016). Prosocial behavior as an outcome measure was for instance taken into account in the review by Malti et al. (2016). However, in this review, prosocial behavior was not explicitly included in the

search terms, even though its promotion may have been a secondary goal in empathy interventions and though interventions aimed at prosocial behaviour likewise may have led to empathy gains.

While the aforementioned mechanisms are not the direct focus of this paper due to limitations of scope, it is likely that they directly or indirectly contribute to increases in empathy, and that they are promoted by many of the empathy raising initiatives discussed. Instructional lessons, group work and discussions all likely increase knowledge of another's situation, carrying the potential of counteracting and disproving prejudiced beliefs and moral disengagement, and thus likely have the potential of raising empathy. Given these possible associations, it may be important for future research to examine how these areas are intertwined.

### **Global versus Specific Levels of Empathy**

Individuals may differ in their global levels of empathy compared with their empathy towards a specific group or individual. For example, interventions can aim to improve a specific aspect of empathy: Studies suggested that violence perpetrators had specifically lowered empathy towards their victims only (e.g., Fernandez & Marshall, 2003), and that the use of violence in intimate relationships did not necessarily predict interpersonal violence in other settings (Holtzworth-Munroe & Meehan, 2004). On the one hand, a general global capacity for empathy does appear to exist in the majority of people, which may reflect positively on the potential of bringing one's specific and lowered levels of empathy towards certain individuals or groups up to one's global empathy levels. On the other hand, it may be necessary to structure the intervention curriculum in order to produce changes regarding specific aspects of empathy to ensure that these successfully target, e.g., a specific out-group. Ameliorative processes for increasing low empathy towards a specific group have been

successful in studies with youth where prejudice and lack of empathy abounded, such as with youth involved in the Palestine-Israel conflict (e.g., Influss et al., 2018).

### **Recent Meta-Analytic Summaries of the Field**

A search in PsychInfo, PubMed, Web of Science, Google Scholar and Google identified no meta-analyses specifically focused on interventions at a school level that aimed to promote empathy. Instead, the studies described below mainly reported results from social and emotional learning programs (that generally had a broader scope directed at social and emotional skills and outcomes).

A meta-analysis by Malti et al. (2016) reported stronger effects on the social outcomes of conduct problems and academic functioning when interventions incorporated more empathy-raising skills. This suggests that overall improvements in social outcomes may have been due to intervention efficacy in causing underlying changes in empathy. Drawing from the 2013 CASEL (Collaborative for Academic, Social, and Emotional Learning) guide and the Life course Interventions to Nurture Kids Successfully (LINKS) database, Malti et al. (2016) evaluated nineteen school-based social and emotional learning programs focused on empathy-related skills. The authors included only methodologically rigorous studies, namely only empirically backed and widely used programs which investigated at least one empathy-related construct and appeared in major intervention databases. Overall effect sizes of the social and emotional learning programs reported by Malti et al. (2016) were small, with the following effect sizes (Cohen's *d*) for socio-emotional competencies: 23 studies (77%) had a negligible to small effect sizes (i.e., 0.04 to 0.49), three studies (10%) had moderate effect sizes (i.e., 0.65 to 0.73), and four studies (13%) showed large effect sizes (i.e., > 0.80). The meta-analysis spanned classes from kindergarten to grade 8, and found intervention effects to be stronger when implemented at an earlier age. However, the scarcity of studies with adolescent age groups warranted their exclusion from the analysis. A meta-analysis by Teding

van Berkhout and Malouff (2016) on empathy raising initiatives also reported few studies with adolescents, and did not find a significant effect within this group.

Several meta-analyses focused on the efficacy of social and emotional learning interventions, which often directly employ empathy raising as a core intervention component. According to a recent meta-analysis (Mahoney et al., 2018), average effect sizes for school-based social and emotional learning interventions, spanning kindergarten to grade 12, ranged from Hedge's  $g = 0.19$  to  $0.74$  at post-test and  $0.23$  at follow-up. Durlak et al. (2011) reported a small effect size (Hedges'  $g = 0.31$ ) of social and emotional interventions. They found that the largest effect sizes with regard to the impact on behavior problems, academic performance and positive social behavior arose for programs that, among other variables, integrated empathy and emotion recognition (Hedges'  $g = 0.69$ , medium effect size). This suggests that underlying improvements in empathy, even though not directly measured, may have been partially responsible for the positive outcomes.

A meta-analytic study by Boncu et al. (2017) investigating the efficiency of social and emotional learning programs on the development of children and adolescents reported an effect size of Hedges'  $g = 0.36$ , a small effect size, for social and emotional skills in social and emotional learning programs. Specifically, prosocial behavior and attitudes toward self and others, concepts theorized to be highly interrelated with empathy, had negligible to small effect sizes (Hedges'  $g = 0.20$  for prosocial behavior, and Hedges'  $g = 0.19$  for attitudes).

Effects on socio-emotional skills were only observed for youth but not adolescent age groups: The effect size was largest for the seven- to twelve-year-olds (Hedges'  $g = .38$ ), followed by the three- to six-year-olds (Hedges'  $g = 0.31$ ), whereas the age group of 13- to 18-year-olds reported no significant effect. As a broad meta-analysis of social and emotional learning programs, the study failed to determine which study characteristics would produce more potent effects for individual age groups or dimensions of socio-emotional competencies.

Fewer social and emotional learning programs are implemented in secondary compared to elementary school, perhaps because they appear somewhat less effective (Yeager, 2017). Given the dearth of research assessing the efficacy of interventions aimed at raising empathy in school-aged children, this literature review hopes to add new insights.

## Methods

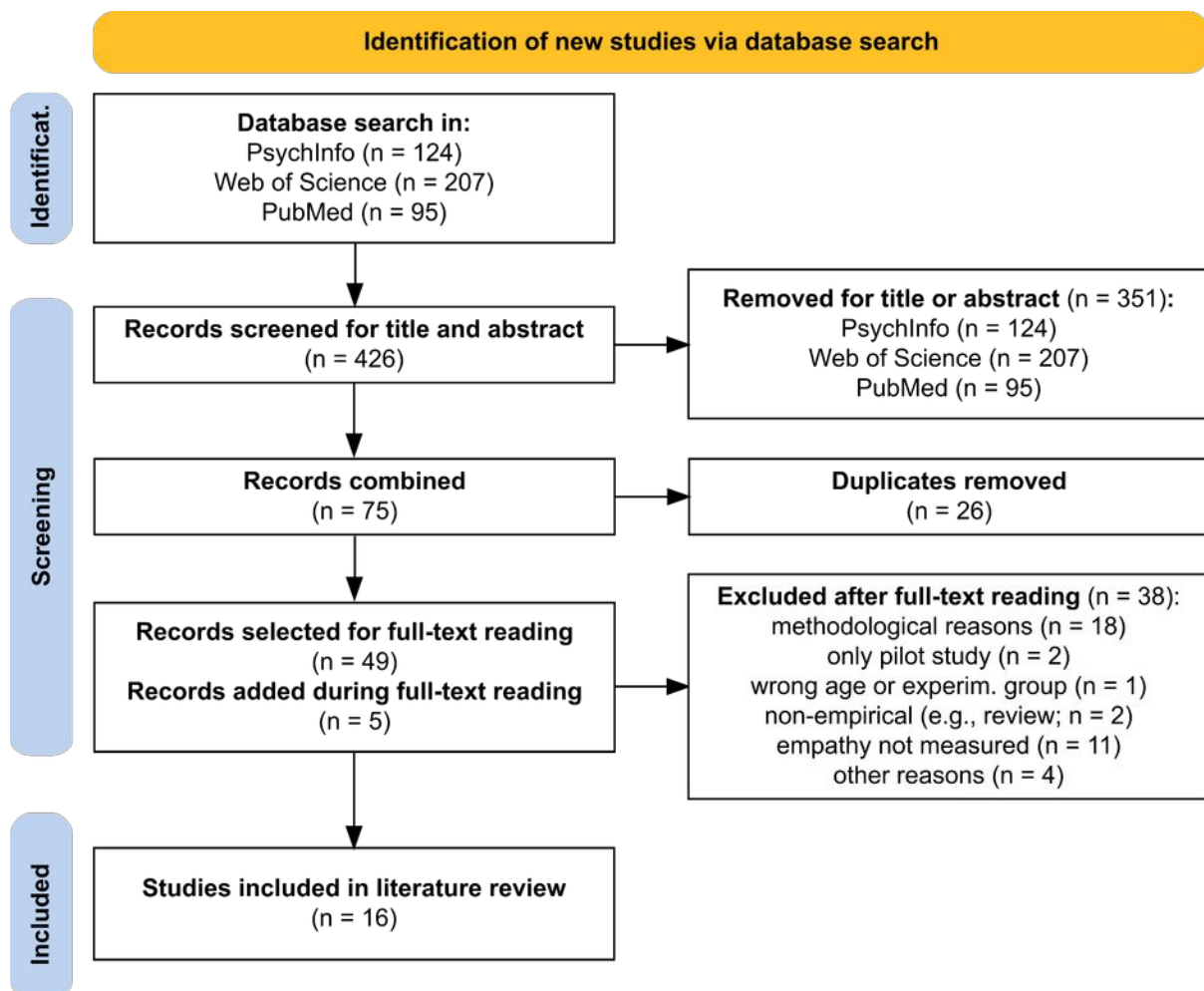
### Search Procedure

Searches were conducted in three databases: PsychInfo, Web of Science and PubMed. The search terms for all three searches were the same, and included the following keywords: (1) *selecting interventions (or related concepts)*: intervention, trial, program or evaluation; (2) *defining the target of such an intervention*: empathy, emotional intelligence, perspective taking or compassion; (3) *defining the context of the intervention*: school; (4) *excluding clinical conditions*: not autism and not disorder; and (5) *excluding where intervention targets (e.g., empathy or compassion) were directed towards the own person*: not self-compassion. The final search term that was common to all three searches was therefore: (intervention OR trial OR program OR evaluation) AND (empathy OR "emotional intelligence" OR "perspective taking" OR compassion) AND school NOT autism NOT disorder NOT "self-compassion". There were two slight changes, adjusting the final search term for some of the three searches: (1) Whereas PubMed provides a separate filter for Randomized Clinical Trials, the searches in PsychInfo and in Web of Science needed to implement such a filter via extending the search term by adding: AND (random\* OR rct or "random\* control\* trial"). (2) Web of Science has a much broader scope than the two other databases, and therefore the search was limited to the following research areas: "Psychiatry & Psychology", "Social Psychology", and "Education & Educational research". All three searches were conducted on the same date (February 27th, 2023), and the results were limited to the ten years previous to the search date, i.e., studies that were published between 2013 and the search date.

An overview over the whole process of searching and selecting the relevant articles is given in the PRISMA Flow Diagram in Figure 1, and described in more detail below.

**Figure 1**

*PRISMA Flow Diagram Depicting the Search Process and Selection Process*



The search in PsychInfo gave 124 results, the search in Web of Science gave 207 results, and the search in PubMed gave 95 results. After the search in each database, the abstracts and the titles of the articles returned by each search were screened, in order to identify studies that fulfilled the inclusion criteria of the literature review (as defined above).

(1) PsychInfo: 86 of the originally 124 studies were removed, leaving 38 studies. (2) Web of

Science: 183 of the originally 207 studies were removed, leaving 24 studies. (3) PubMed: 82 of the originally 95 studies were removed, leaving 13 studies.

The studies that were not excluded based upon their title or their abstract were combined afterwards: 38 studies came from PsychInfo, 24 studies came from Web of Science, and 13 studies came from PubMed, giving a total of 75 studies. Of those 75 studies, 26 were removed because they were duplicate entries found by more than one search, leaving 49 studies to be evaluated.

Those 49 studies were selected for full-text reading, in order to determine which studies should be included in this review. Given the scope of the review, studies were required to have a stringent randomization along all stages of the process (i.e., randomly selecting a representative sample of schools from a larger area) to allow for greater generalizability of the results. Out of the 49 studies selected for full text reading, 14 were deemed eligible to be included in the review based upon those criteria. During the full-text reading, five additional studies were identified for potential inclusion in the review (as they were referenced in a relevant section of a study). These were also full-text-read, and two were deemed eligible for inclusion. Together with the 14 studies already selected, 16 studies were finally included in the review. Supplementary Table 1 in the Appendix lists all studies that were selected for full-text reading, and gives information on why studies were excluded after reading. Altogether, 38 studies were excluded: 18 for methodological reasons, two for being only a pilot study, one for wrong age or experimental group characteristics, two for being non-empirical (e.g., meta-analyses), eleven because they did not include the relevant outcome measure (i.e., empathy or a related concept), and four for other reasons (e.g., being outside the defined time range for this review).

## Results

The results section is organized in two parts. The first part briefly describes each study that was included in the literature review. It is organized by grade level, with the divisions of: (a) “primary school age” (grades 1 to 5), reporting four studies; (b) “primary and secondary school age” (grades 1 to 9), with three studies; (c) “secondary school age” (grades 6 to 9), reporting four studies; (d) “high school age” (grades 9 to 12), with four studies; and (e) “too broad to classify” with the remaining study. Within each division, the studies are arranged alphabetically after the name of the first author. Generally, the studies were quite heterogeneous in terms of the country where they were conducted, the number of participants included in the study, and intervention type and length.

The second part of the results presents central characteristics of the studies included in the literature review, and an overview of these is provided by Table 1. For the characteristics listed in Table 1, it is then evaluated whether these characteristics affect the reported effect sizes.

### **Description of the Individual Studies Included in this Literature Review**

#### ***Primary School Age (Grades 1 to 5)***

**Building social cohesion in ethnically mixed Schools: An intervention on Perspective Taking (Alan et al., 2021).** The intervention curriculum contained written and animated class activities, making perspective taking a core concept. The reading and visual material stimulated students to understand and experience the emotions of others. For example, characters in the videos experienced events such as falling down, which is followed by the event happening to another character. These events were randomly distributed across ethnic groups, thus emphasizing similarities between different people in terms of the effects of the same adverse event. Other activities were reading diary extracts of an imagined new student at a school compared with the same event from another country, and an emotional



guessing game of friends' mental states. Perspective taking was explicitly fostered by stimulating students to think about others thoughts and feelings when witnessing an act of malfeasance towards animals.

**A cluster randomized trial of the Social Skills Improvement System-Classwide Intervention Program (SSIS-CIP) in first grade (DiPerna et al., 2018).** The universal intervention of Social Skills Improvement System-Classwide Intervention Program (SSIS-CIP) employed instructional strategies such as reinforcement, modeling, role play and problem-solving that are rooted in established student learning theories (i.e., operant, social learning). One skill was targeted by each of the ten units, for instance listening to others, asking for assistance, and getting along with others. Two of the units included self-control skills, and one focused on empathy.

**Promoting social-emotional competence: An evaluation of the elementary version of Second Step® (Low et al., 2015).** The intervention used the Second Step program, a skills-focused social and emotional learning curriculum which uses direct instruction to foster the ability to learn, empathy, emotion management, and problem-solving. In addition to direct instruction, its underlying model emphasized providing opportunities to practice skills and receiving reinforcement for behaviors. The recent 4th edition of the program added focus on three aspects of self-regulation: attention, working memory, and inhibitory control. Empathy was one of the four core units of the 22 lessons. The range of contents covered in these lessons were being respectful learners, planning to learn, identifying others feelings, showing compassion, making friends, and managing disappointment. Program components included to convey these skills were lesson cards, posters picturing learned skills; DVDs that demonstrate particular skills and brain builder games.

**Effects of a Collaborative Board Game on Bullying Intervention: A Group-Randomized Controlled Trial (Nieh & Wu, 2018).** The game used in this intervention was

called the galaxy rescuer game, and featured a storyline where the player chooses to be one of six characters. Points were earned in the first stage by answering questions about bullying. Then, players matched different characters such as bullies and victims to their roles in a bullying incident. The third stage featured a collaborative game where players collaborated to defend victims through methods such as talking to adults or accompanying the victim as they encountered bullying scenarios. They were faced with a choice between continuing their own assignment in the game and helping victims, preventing the victims' mood scale from dropping to zero.

### *Primary and Secondary School Age (Grades 1 to 9)*

**School bullies' intention to change behavior following teacher interventions: Effects of empathy arousal, condemning of bullying, and blaming of the perpetrator (Garandeau et al., 2016).** An anti-bullying team consisting of three school personnel adults was formed in each school in order to set in motion a five-step process. First, a screening was conducted to determine whether the conflict at hand was actually bullying. This was followed by a discussion with the victim and a discussion with the bullies, which were conducted separately for each part. Finally, a follow-up discussion was conducted with the victim and with the bullies, separately. Prosocial classmates were encouraged to support the victimized peer at the request of the classroom teacher. Discussions with the bullies were held as soon as possible and were organized to come as a surprise to the bullies, discouraging prior discussions between the bullies and their peers on the subject. The intervention explored the differences in using a confronting vs a non-confronting approach, where the former consists of making it clear to the student that the behavior of bullying is not accepted and thus condemning the act, while the latter emphasized increasing the bullies concern for the victim through reaching an agreement on the negative feelings experienced by the bullied student. Bullies were consulted on suggestions on improving the victim's situation.

**Effects of the KiVa Anti-Bullying Program on Affective and Cognitive Empathy in Children and Adolescents (Garandeau et al., 2021) and KiVa anti-bullying program in Italy: Evidence of effectiveness in a randomized control trial (Nocentini & Menesini, 2016).** The two studies implemented the same KiVa intervention program in two different countries and with similar age groups, Nocentini and Menesini (2016) did this in an Italian setting, Garandeau et al. (2021) in Finland. The rationale of the KiVa program was that positively changing behaviors of bystanders can minimize rewards obtained by bullies and thus alter their motivation to bully. To do so, KiVa included class activities that aimed to increase empathy, self-efficacy and anti bullying attitudes held by bystanders. Students learned to interpret subtle nonverbal cues of emotions through their enactment by peers, vicariously experienced emotions and put themselves in the position of others through using empty chairs to build a bullying scenario, and watched and discussed films of previously bullied adults to build empathic concern. The intervention sought to create a school climate where bullying is unacceptable. It also aimed to provide victims with support and to confront bullies on their actions. KiVa employed both universal and indicated methods. Student lessons were given by classroom teachers on the bystander role, to elicit empathy towards the victim, and lessons were followed by the adoption of classroom norms. KiVa used the methods of discussion, group work, role-play exercises, and short films about bullying. Computer games were used to reinforce knowledge and skills, where students had access to the thoughts and the emotions of characters. Other components were recess supervisors with vests, posters on KiVa content and guides to parents about KiVa content. When acute cases of bullying were identified, a process was set in motion involving discussions with the victims and the bullies.

***Secondary School Age (Grades 6 to 9)***

**Promoting Social and Emotional Learning in Middle School: Intervention Effects of Facing History and Ourselves (Domitrovich et al., 2022).** Facing history included a unit especially targeted at reducing prejudice and fostering empathy called The Choices in Little Rock. The focus were historical periods where intergroup conflict and racism abound. Through the creation of a democratic classroom environment, the free expression of opinions was theorized to lead to better discussions about society's challenges and changes in own beliefs and behavioral choices. Instructional practice was combined with fostering student reflection on the self, others and society, and with student cooperation and interaction. Critical social analysis aimed to transform citizenship beliefs and create anger at social injustice.

**Perspective taking and decision-making in educational game play: A mixed-methods study (Hilliard et al., 2016).** Quandary is an online free game that utilized storylines and accompanying characters to teach youth ethical decision-making and taking the perspectives of others. As the captain of a space colony, the player had to make helpful decisions for the colony. While there were no clear solutions, the player had to decide which consequences to face. Upon the player's decision, the solution implemented either boosted the colony's morale or decreased it. An example of an episode included was Fashion faction, where modifications to colony uniforms made by the colony tailor caused certain colonists to feel excluded.

**Effects of cooperative learning on peer relations, empathy, and bullying in middle school (Ryzin & Roseth, 2019).** This intervention applied the Johnsons' approach to cooperative learning. Under this approach, methods such as reciprocal teaching (i.e., Jigsaw), peer tutoring, collaborative reading and further methods consisting of peers helping one another to learn in small groups under positive interdependence-conditions were used. Important to the intervention was individual accountability, explicit coaching for social skills

in collaboration, high degrees of face to face interaction, and guided analysis of group performance. Cooperative learning represented a conceptual framework which could be used by teachers to design and implement their own group based activities based on existing curricula.

**Deeper engagement with live theater increases middle school students' empathy and social perspective taking (Troxler et al., 2022).** The intervention consisted of a live theater visit experimentally manipulated to include pre- and post-performance engagement experiences. While one group attended the theater only, others also received opportunities to familiarize themselves with the content in beforehand (through the addition of a performance guide) and discussed the contents afterward with an external expert (through the addition of a facilitated discussion). The theater performance was specifically focused on perspective taking. The play used was a modernized re-imagining of Mark Twain's *The Prince and the Pauper* in a hip hop style. In addition, the intervention also explored the effect of reading the book in beforehand.

### ***High School Age (Grades 9 to 12)***

**Tackling psychosocial risk factors for adolescent cyberbullying: Evidence from a school-based intervention (Barkoukis et al., 2015).** The intervention integrated empathy raising activities in an intervention aimed at reducing cyberbullying and decreasing moral disengagement. It employed awareness raising activities promoting policies and norms in the school, highlighted negative effects of cyberbullying on victims, denoted cyberbullying as a stigmatized behavior, and encouraged whole school efforts to prevent cyberbullying by galvanizing parents and teachers to take a stand against it. The messages were communicated through lectures, interactive talks with groups of students, and printed material such as posters and leaflets placed and distributed in student halls and classrooms. Posters illustrated examples of social media posts that could transform into cyberbullying and attached

suggestions for solutions and responsible behavior. The leaflets contained information on the negative effects of cyberbullying and intended to stimulate empathy through for instance stimulating reflection on feelings of regret when hurting someone through cyberbullying. Interactive discussions focused on the material presented, encouraging the reporting of cyberbullying incidents, and increasing negative attitudes and norms toward cyberbullying. One classroom based discussion following a lecturer was held by cyberbullying research experts, presenting the effects on victims. Students also watched documentary videos with cyberbullying victims in the intervention.

**Evaluation of a bystander-focused interpersonal violence prevention program with high school students (Edwards et al., 2019).** The intervention Bringing in the Bystander-High School Curriculum (BITB-HSC) used lectures, large and small group discussions, experiential and hands on exercises, skills building activities and videos to educate students on sexual harassment, sexual assault, dating violence and stalking. The emphasis lied on how these transgressions negatively impacted communities. The modules taught students how to recognize interpersonal violence, their role in fostering a healthy community, and how to effectively intervene as bystanders. School personnel were also trained in bystander behavior and in reinforcing information and skills learned by students.

**Effects of Cyberprogram 2.0 on “face-to-face” bullying, cyberbullying (Garaigordobil & Martínez-Valderrey, 2015).** The intervention aimed to improve abilities in analyzing the consequences of bullying and cyberbullying for victims, aggressors and observers, to increase condemnation of such actions, to develop coping strategies for acts of bullying and cyberbullying, and to improve other relevant skills (empathy included). At the start, group members sat in a circle listening to the adult explain the activity and its purpose. After the activity was finished, a guided reflection phase and discussion took place through non-judgmental critical questions formulated by the teacher. The program drew on the

techniques of role playing, brainstorming, case studies and guided discussions facilitated through well-thought-out questions.

**Training the social-emotional skills of youth school students in physical education classes (Malinauskas & Malinauskiene, 2021).** Employing the model of cooperative learning, physical education classes were used to implement socio-emotional skills training. The goals of the intervention were the promotion of socio-emotional skills, namely empathy, cooperation, assertion, self-control, optimism, understanding and analyzing emotions, appraisal, and utilizing emotions. To train the skills, five components were used: (1) skill description; (2) skill demonstration; (3) skill practice; (4) feedback and (5) reinforcement of trained skill. The training methods applied by the Social-Emotional Skills Training Program included impulse control (autogenic training), post-activity discussion about the shared experiences, group learning (cooperative learning), role-play scenarios, watching the videos, and written worksheets of the students. For example, relaxation and impulse control methods were used to improve emotional control, and real-life videos demonstrating difficulties experienced by students improved emotion understanding. Questions were directed at students after the video, for instance: “If your friend is sad, and you would like to say something to your friend about their mood or feeling during physical education classes, what would you say?”. Group learning was used to practice emotional understanding, followed by discussion about emotional experiences.

### ***Studies with a Broader Age Range***

**Prosocial video game as an intimate partner violence prevention tool among youth: A randomized controlled trial (Boduszek et al., 2019).** The intervention made use of a culturally sensitive prosocial video game called Jesse, aimed at increasing affective and cognitive empathy towards gender-based violence victims in a domestic context. The player role-played as a number of characters in different settings, exploring topics such as “gender

inequality, female economic dependency, intergenerational violence, the impact on primary and secondary victims, the physical and emotional impact of domestic violence and escalatory patterns, help-seeking behavior and empowerment-based interventions, the role of alcohol as a contributory factor, impact on children's behavior, the role of professionals, positive male role models as agents of change, and perpetrator accountability" (p. 262). For instance, role playing as Jesse's teacher, players asked Jesse questions aimed at discovering why Jesse's behavior and grades were decreasing. In the fourth level of the game, Rondell and his best friend Hayden had a conversation regarding Rondell's violent behavior toward Diana.

### **Study Characteristics and How They Affected the Observed Effect Sizes**

Table 1 presents an overview over the study characteristics and the effect sizes the interventions had on empathy measures and on reducing bullying behavior. Beneath the table is a more thorough summary of the results, with a specific focus on how different characteristics of the studies relate to the (main) effect sizes of the studies with these characteristics.

For the purpose of the comparison of study characteristics and effect sizes, effect sizes were classified into small, medium and large effects, following the classification developed by Cohen (e.g., Cohen, 1992). They differed in their respective ranges based on the type of effect size used in the study. An overview over those effect size categories (i.e., small, medium, and large) for the most common effect sizes (University of Cambridge, n.d.), recommends the following thresholds: For a small effect sizes, the lower thresholds would be 0.20 for Cohen's  $d$  and Hedges'  $g$ , 0.10 for measures derived from or equivalent to correlations (e.g.,  $r$ ,  $\beta$ ), and 0.01 for effect sizes from ANOVAs (i.e.,  $\eta^2$  or  $\eta^2_p$ ). All effect sizes that did not reach those lower thresholds were denoted as "negligible". For medium effect sizes, the lower thresholds would be 0.50 for Cohen's  $d$  and Hedges'  $g$ , 0.30 for measures derived from correlations, and 0.06 for  $\eta^2$ -effect-sizes. For large effect sizes, the



lower thresholds were 0.80 for Cohen's  $d$  and Hedges'  $g$ , 0.50 for measures derived from correlations, and 0.14 for  $\eta^2$ -effect-sizes. Some studies reported a positive effect but did not report an effect size, and are described accordingly.

Most study characteristics reported in the first columns in Table 1 were evaluated with respect to the (main) effect sizes of empathy gains reported in the studies (second-last column in Table 1) and some with respect to effect sizes reported for interventions focused on reducing bullying perpetration. Some further study characteristics were not consistently investigated or reported across all studies, and are therefore not presented in Table 1.

### ***Effect sizes for all interventions***

Overall, eleven out of 16 of the studies reported both significant results of the respective interventions aimed at raising empathy in schools, and the respective effect size. Of the remaining five studies, two did not obtain a significant result and three did not report an effect size. Of the studies without a significant main effect, Low et al. (2015) later identified a positive interaction effect with a proactive classroom environment. Barkoukis et al. (2015) did not find a significant effect and did not further explore interaction effects. Hilliard et al. (2016) found only a qualitative effect in terms of greater depth of moral responses. Two studies did not report a numeric effect size (Garandeanu et al., 2016; Nieh and Wu, 2018).

**Table 1**

*Overview over the characteristics of the studies reported in this literature review: The table contains the study reference (first column), the country in which the study was conducted (second column), the number of participants (third column) and attrition (fourth column, see table not for explanation), and the age and grade the participants were in (fifth column); which empathy measures were used (sixth column) and whether those were G = global or S = specific (seventh column); the length of the intervention (and how many sessions were administered; eighth column); and the effect sizes the intervention had on empathy measures and reductions in bullying perpetration (ninth and tenth column)*

<b>Study</b>	<b>Country</b>	<b>Participants</b>	<b>Age / grade</b>	<b>Empathy measures</b>	<b>Intervent. length</b>	<b>Empathy</b>	<b>Bully. reduct.</b>
Alan et al. (2021)	Turkey	N > 65000	0 primary, grades 3 and 4, ages 8-12	five-item Empathic Concerns Subscale	G 7 months; 3h / week	persp. taking: $d = 0.27$ (host children) $d = 0.28$ (refugee childr.)	violent event red.: $d > 1.23$
DiPerna et al. (2018)	USA	T: N = 373 C: N = 393	2 primary, grades 1 and 2	Social Skills scale (teacher observation)	G 12 weeks	$g = 0.31$	no effect
Low et al. (2015)	USA	N = 4891 T: 31 schools, C: 30 schools	2 primary, grades 0 to 2	Devereux Student Strengths Assessment (teacher rating)	G 4 months; 25 to 40 min / week	no effect	not assessed
Nieh and Wu (2018) <sup>a</sup>	Taiwan	T1: N = 116 T2: N = 125 C: N = 87	0 primary, grade 5	ten-item empathy scale	G 7 weeks (gaming: week 2 to 5); 40 min / week	not reported	not reported

Study	Country	Participants	Age / grade	Empathy measures	Intervent. length	Empathy	Bully. reduct.
Garandeau et al. (2016)	Finland	188 pairs (prim.) 153 pairs (sec.) 341 victims, 341 bullies	0 primary and secondary, grades 4 to 6	eleven-item KiVa scale	S one meeting and follow-up meeting (for both groups: victim and bully)	not reported	not reported
Garandeau et al. (2021)	Finland	N = 15403 T: 462 classes C: 399 classes	5 primary and secondary, Mage = 13.4; 51.5% girls	seven-item KiVa scale	S 9 months	$d = 0.07$	<i>peer report:</i> $d = 0.10$ <i>self report:</i> $d = 0.03$ to $d = 0.18$
Nocentini and Menesini (2016)	Italy	T: n= 1039 C: n= 1003	0 primary and secondary, grades 4 and 6	seven-item KiVa scale	S ~ 8 months	$d = 0.40$ (primary sch.) $d = 0.08$ (second. sch.)	$d = 0.24$ (primary sch.) $d = 0.21$ (second. sch.)
Domitrovich et al. (2022)	USA	T: N = 437 C: N = 257	2 secondary, grades 7 and 8	six-item Empathic Concerns Subscale	G 3 months; five sections with multiple lessons each	$\beta = 0.16$	no effect
Hilliard et al. (2016) <sup>b</sup>	USA	T1: N = 63 T2: N = 49 C: N = 51	3 secondary, grades 6 to 8	Davis Individual Reactivity Index	G three weeks, sessions in week 1 and 3 (20 min)	not reported	not assessed

Study	Country	Participants	Age / grade	Empathy measures	Intervent. length	Empathy	Bully. reduct.
Ryzin and Roseth (2019)	USA	N = 1890	2 secondary, grades 7 and 8	Basic Empathy Scale	G 2 years	$\beta = 0.45$ (cognitive) $\beta = 0.26$ (affective)	$\beta = 0.30$ (effect of coop. learn.)
Troxler et al. (2022) <sup>c</sup>	USA	T1: N = 275 T1 + T2: N = 386 T1 + T3: N = 402 T1 – T3: N = 368 C: N = 958	1 secondary, age 10 – 15	Social Perspective Taking-Motivation Survey  Basic Empathy Scale	G performance guide: 11 to 30 min (varied between teachers)  facilitated disc.: 5 to 12 min	$\beta = .06$ (facilitated discussion)	not assessed
Barkoukis et al. (2015)	Greece	N = 355	0 high school, age 13-17	Basic Empathy Scale	G 8 weeks interactive discussion (10 to 15 min) + lecture (45 min) / week	no effect	not reported
Edwards et al. (2019)	USA	T: N = 1081 C: N = 1322	4 high school, grades 9 to 12	Victim Empathy Scale	S 7 session (45 min)	$d = 0.13$	not reported
Garaigordobil and Martínez-Valderrey (2015)	Spain	T: N = 93 C: N = 83	0 high school, age 13 to 15	Index of Empathy for Children and Adolescents	S 9 months; 19 sessions (1h)	$d = 0.28$	bullying: $d = 0.75$ cyber-bully.: $d = 0.69$

Study	Country	Participants	Age / grade	Empathy measures	Intervent. length	Empathy	Bully. reduct.
Malinauskas and Malinauskiene (2021)	Lithuania	T: N = 49 C: N = 55	0 high school, Mage of 16.91, SD 1.12	Schutte Self-Report Inventory Social Skills Rating System	G 6 months; 48 sessions (15 min; total: 12 h)	$\eta^2_p = 0.08$	not assessed
Boduszek et al. (2019)	Barbados	T: N = 86 C: N = 86	0 primary to tertiary (prof. training)	None in Three Victim Responsiveness Assessment	S five days, 45 min / day	$d=0.56$	not assessed

Note. Abbreviations used in “Participants”: C = control group, T = treatment group[s], if there was more than one treatment condition, these are numbered, i.e., T1, T2, ...

The column after “Participants” encodes whether the study reported attrition: 0 = not reported, 1 = under 10%, 2 = 10 to 20%, 3 = 20 to 30%, 4 = longitudinal study (T2: 90%, T3: 77%; T4: 35% of the original sample), 5 = longitudinal study (T2: 82%, T4: 60% of the original sample).

<sup>a</sup> The conditions in the study of Nieh and Wu (2018) were: T1: game-only; T1 + T2: game-with-debriefing.

<sup>b</sup> The conditions in the study of Hilliard et al. (2016) were: T1: Quandary; T2: Quandary and Facilitation.

<sup>c</sup> The conditions in the study of Troxler et al. (2022) were: T1: Theater; T2: Performance guide; T3: Facilitated discussion.

Among the studies that reported numeric effect sizes, three studies provided multiple effect sizes (e.g., based on grade level, the distinction between affective and cognitive empathy, or intervention or target group). For cases, where several effect sizes were reported within one study, the effect sizes were summarized into a range (e.g., for Nocentini and Menesini, 2016, reporting Cohen's  $d = 0.40$ , a small effect, for primary school age and Cohen's  $d = 0.08$ , a negligible effect, for secondary school age, resulting in the classification "negligible to small"). Four effect sizes were regarded as negligible (three reported a Cohen's  $d < 0.20$ , one reported a coefficient  $\beta < 0.10$ ). One was negligible to small (Nocentini and Menesini, 2016, mentioned above), and three effect sizes were small (two reported Cohen's  $d$  or Hedges'  $g$  between 0.20 to 0.50, one a coefficient  $\beta$  between 0.10 to 0.29). Two effect sizes were of medium size (one reporting a Cohen's  $d = 0.56$ , one a  $\eta^2_p = 0.08$ ). Finally, one study reported effect sizes that were small to medium ( $\beta = 0.26$  for affective and  $\beta = 0.45$  for cognitive empathy). No large effect was reported by any of the studies.

### ***How Do Age and Grade Affect the Effect Sizes?***

For the purpose of exploring how age affects the effect sizes of empathy raising in the studies, the following divisions of age ranges were made: primary school age (grades 1 to 5), secondary school age (grades 6 to 9) and high school age (grades 9 to 12). Out of the 16 studies, four explored empathy raising in primary school grades, another three explored a broader range of ages that spanned both primary and secondary school grades, four explored secondary school grades, and another four explored high school grades. Boduszek et al. (2019) was not categorized into school grades as it explored too broad of an age group (spanning primary to tertiary education). This mix of ages is encouraging, given that several studies noted a dearth of research in upper grade levels.

Out of the seven studies which explored primary grades, one indicated a negligible effect size, and three reported small effect sizes. Low et al. (2015) reported no significant

main effect, but discovered an interaction effect with proactive classroom environment. Nieh and Wu (2018) and Garandau et al. (2016) did not report a numeric effect size.

Out of the seven studies exploring secondary grades, two reported a negligible effect. Two reported a small effect size. Ryzin and Roseth (2019) reported small to medium effect sizes. Hilliard et al. (2016) did not find a significant effect (and could not give an effect size), but reported qualitative findings of increased empathy. Garandau et al. (2016) did not provide an effect size.

Out of the four studies conducted with high school populations, Edwards et al. (2019) reported a negligible effect size. Garaigordobil and Martínez-Valderrey (2015) reported a small effect size. Malinauskas and Malinauskiene (2021) reported a medium effect size. Barkoukis et al. (2015) did not find a significant effect.

Further, five studies explicitly explored how empathy gains were affected by age or grade level. For grade 1 and 2, Low et al. (2015) did not observe grade level to affect empathy gains. Likewise, DiPerna et al. (2018) did not find empathy gains affected by grade level for grade 1 and 2. However, grade levels may lead to different degrees of empathy gains in different age groups. Garandau et al. (2016) noted a higher intention of bullies to change their behavior in secondary compared with primary schools, grade 4 vs grade 6. Nocentini and Menesini (2016) found effect sizes for empathy levels to differ recognizably between primary and secondary school age (grade 4 vs grade 6). However, Garandau et al. (2021) did not find secondary students to be less responsive to raising empathy through KiVa than primary school students, though the effect sizes of the intervention were negligible.

### ***Did the Cultural or Country Background Affect the Effect Sizes?***

Geographically, six studies were conducted in Europe, seven studies were conducted in the USA, and the remaining studies came from Taiwan, Barbados and Turkey. Out of the effect sizes originating from studies in Europe, one effect size was negligible, one was

negligible to small, one was small, one was medium, one had no significant effect and one did not report an effect size. Of the seven effect sizes from the studies conducted in the USA, two were negligible, two were small in effect size, one was small to medium, one reported no main effect, and one found no effect. The study from Turkey reported a small effect size, while the study from Barbados reported a medium effect size. The study from Taiwan did not provide an effect size.

### *Are the Effect Sizes Affected by Sample Size?*

Nine studies reached a sample size between 100 and 1000, five studies reached a sample size over 1000, and two even recruited over 10000 participants. For RCTs, sample sizes of over 100 participants are generally regarded to have good statistical power, and sample sizes over 1000 are expected to achieve over 90% power (Chan, 2003). Out of the nine studies with a sample size between 100 and 1000, three reported a small effect size, two reported a medium effect size, three did not report a numeric effect size, and one found no effect. Out of the seven studies reporting a sample size over 1000, three reported a negligible effect size (one of which reported an interaction effect), one reported a negligible to small effect size, one reported a small effect size, one reported a small to medium effect size, and one reported no significant main effect but an interaction effect.

### *Empathy Effect Sizes in Bullying Interventions versus Non-bullying Interventions*

As several studies aim to raise empathy in order to reduce bullying, it is of interest to explore whether effect sizes of empathy gains differ between studies directed at bullying versus those without this outcome measure. Out of the eleven studies that included an outcome measure related to bullying perpetration, two reported negligible effect sizes for empathy (one of which is for affective empathy only), one reported negligible to small effect sizes, four reported small effect sizes, one reported small to medium effect sizes, two did not report an effect size, and one had no significant effect. Out of the remaining five studies that



did not target bullying, one reported a negligible main effect size for empathy, two reported a medium effect size, one reported no main effects but interaction effects, and one found no effect.

**Empathy Effect Sizes versus Bullying Effect Sizes.** Additionally, it is of interest to explore how effect sizes of empathy gains related to or differed from effect sizes of bullying reductions, shedding light on whether one may impact the other. Seven studies targeted (cyber-)bullying or included it as an outcome measure, and reported an effect size for empathy. Four studies reported a small effect size for empathy, and bullying perpetration reductions ranged from no effect size to a large effect size. One study reported a negligible (primary school) and small (secondary school) effect size increase in empathy, but small effect size decreases in bullying in either grade. A second study reported small to medium empathy effect sizes, and a medium effect size decrease in bullying. A third study reported a negligible increase for affective empathy and a non-significant effect for cognitive empathy, and mostly negligible effect sizes for bullying.

**Relations Between Cognitive and Affective Empathy and Bullying Reduction.**

Only Ryzin and Roseth (2019) measured and explored both affective empathy and cognitive empathy changes and their relation to bullying reductions. They found that some of the effect on bullying outcomes is transmitted via affective empathy gains, a medium effect. However, while the intervention also increased cognitive empathy, it was not found to interact with bullying outcomes. No other study explored the relation of empathy as a whole, or the relation of cognitive and affective empathy, to bullying.

***How Does the Length of the Intervention Affect Effect Sizes?***

All 16 studies reported the length of the interventions employed. Two of the studies reported an intervention length of less than one week. One study (Garandau et al., 2016) reported two separate meetings with the bully/victim on two different occasions. Five of the

studies reported an intervention length between three and twelve weeks. Five reported an intervention length of one school term, and two others of approximately the same length, 22 weeks and nine months respectively. Finally, Ryzin and Roseth (2019) reported an intervention length of two years.

Of the three interventions with less than one week length, Troxler et al. (2022) reported a negligible effect size, while Boduszek et al. (2019) reported a medium effect size. Garandeau et al. (2016) did not report a numeric effect size. Of the five interventions employing lengths of three weeks to twelve weeks, Edwards et al. (2019) reported a negligible main effect size, DiPerna et al. (2018) reported a small effect size, Nieh and Wu (2018) did not report a numeric effect size, and Hilliard et al. (2016) and Barkoukis et al. (2015) did not observe a significant intervention effect. Of the seven interventions of over twelve weeks to a year in length, one reported a negligible effect size, Nocentini and Menesini (2016) reported negligible to small effect sizes, three reported small effect sizes, Malinauskas and Malinauskiene (2021) reported a medium effect size, and Low et al. (2015) did not find a significant main effect. The intervention that lasted more than a year (Ryzin & Roseth, 2019) reported small to medium effect sizes.

### ***Stability of the Intervention Effects over Time***

Out of the 16 studies, three reported an effect of the intervention at a follow-up measurement, shedding light on whether initial gains in empathy can be maintained. The shortest follow-up is seven days in the study by Boduszek et al. (2019; a medium effect size, Cohen's  $d = 0.56$ ), where empathy gains were maintained. Nieh and Wu (2018) reported a follow-up at two weeks (effect size not reported), but gains in empathy were not maintained. In the study by Edwards et al. (2019; with a negligible effect size, Cohen's  $d = 0.13$ ), empathy levels regressed to the levels of the control group at one year follow-up. Furthermore, while not defined as a follow-up, Domitrovich et al. (2022) reported conducting

a post-test three months after the unit designed to activate empathy was completed, reporting significant gains in empathy that remained stable over time (with a negligible effect size, Cohen's  $d = 0.16$ ).

### ***Effect Sizes by Cognitive versus Affective Empathy***

Six studies explicitly investigated the effect of the intervention (denoted as main effect in the description below) on affective versus cognitive empathy (described as interaction below). Alan et al. (2021) and Hilliard et al. (2016) specifically assess empathic concern, which forms part of affective empathy and, in contrast with personal distress, is related with accurate emotion recognition (Israelashvili et al., 2020). The distinction between affective and cognitive empathy has gained traction in the literature, in part due to new findings on the underlying neural mechanisms and the paths of development (Moore et al., 2015).

For *cognitive empathy*, Ryzin and Roseth (2019) found a small main effect, Alan et al. (2021) found a small main effect (for both refugee and host students), and Troxler et al. (2022) reported interaction effects, of which one had a small and three had a medium effect size. Garandean et al. (2021), Boduszek et al. (2019) and Hilliard et al. (2016) did not observe significant effects on cognitive empathy. For *affective empathy*, Garandean et al. (2021) reported a negligible effect size, Troxler et al. (2022) reported a negligible main effect and a small interaction effect, Ryzin and Roseth (2019) reported a small main effect, Boduszek et al. (2019) reported a medium effect size, and Alan et al. (2021) found an interaction effect (significant effect in the refugee group only, negligible). Hilliard et al. (2016) found no significant effect for empathic concern.

### ***Global vs Specific Level Instruments for Empathy***

The reviewed studies differed widely with regard to the instruments they used to measure empathy (see the column "Empathy measures" in Table 1). Generally, measurement

instruments for empathy can be distinguished into those that are focusing on specific aspects of empathy when assessing student's reactions to the intervention content, while others employed more global measures of empathy. There was little overlap in which instruments were used, apart from the Basic Empathy Scale and the Empathic Concerns Scale, both assessing empathy globally and used two times each, and the KiVa scale, assessing specific aspects of empathy and used three times. Other instruments, although they may have also measured specific aspects of or global levels of empathy, were either used only once or were even self-developed by the study authors, making comparing them difficult.

Specific versus global assessment did not seem to affect the obtained effect sizes to a large degree: Six studies use a context specific empathy measure, where two effect sizes were negligible, one reported negligible to small effect sizes, one was small, one was medium, and one did not report an effect size. Out of the ten studies that used instruments with a global measure of empathy, effect sizes similarly ranged from negligible to medium, with the majority being small.

### ***Control Group Empathy Changes***

Nine of the 16 studies reported changes in control group empathy. Three of these explicitly comment on changes in control group empathy levels in their analysis. Of these, Garaigordobil and Martínez-Valderrey (2015) reported maintained levels in control group empathy (high school). Nocentini and Menesini (2016) reported control group increases over time for grade 4 (primary), but a high likelihood for maintained or decreased levels of control group empathy for grade 6 (secondary). Garandean et al. (2021) reported decreases in empathy levels in the control group (primary and secondary).

Out of the six studies that reported pre- and post-test empathy level means without commenting on the implications of the findings, three reported mean decreases in post-test control group empathy, while the remaining three reported slight increases in mean post-test

control group empathy levels, however always lower than increases in the experimental group. In total, two control groups decreased, and two control groups increased in empathy levels in the primary schools, three control groups decreased in empathy levels in the secondary schools, and one control group decreased, two control groups increased, and one control group maintained empathy levels of empathy in high schools.

### **Discussion**

The discussion interprets the results from the studies in the literature review, especially in regard to intervention efficacy of raising empathy across different characteristics of the studies, and notes potential mechanisms to explain the outcomes. Based on the findings, future directions for research are discussed, as well as limitations of the review.

### **Overall Effects of Empathy Interventions**

The majority of the reviewed studies (14 out of 16) reported significant intervention effects, with the majority of studies reporting effect sizes that did not exceed small effect sizes, and with no study reporting a large effect size. A key characteristic of the studies included in this review was their considerable heterogeneity in terms of intervention design, how outcomes were measured, and the size of the study population, to name just a few characteristics along which studies differed from one another. This complicates a concise summary of their results, and presents challenges to providing clear advice regarding how future evidence-based interventions directed at the promotion of empathy in school-aged children best should be implemented, along with what potential these interventions possibly may have.

The rather small effect sizes revealed by the review are in accordance with existing literature, noting that effects of universal interventions often do not produce constant and sizeable effects across members of the entire population, and thus the studies included may still have lead to population level impacts (Greenberg & Abenavoli, 2016).

The small effect sizes may suggest that increasing empathy for victimized peers in school based programs are, at least to some degree, dependent on factors that are difficult to modify (e.g., genes, peer norms, or parenting style; cf. Garandean et al., 2021). Moreover, studies often targeted multiple outcomes and did not specifically structure the entire curriculum around empathy, potentially leading them to not reach possible effect sizes. Given the varying effect sizes and the heterogeneity in the combination of intervention techniques aimed at raising empathy, it is likewise difficult to give clear advice on how to develop future interventions that may eventually produce larger effect sizes (and that do so more consistently).

Consequently, the generally small magnitude of effect sizes, the heterogeneity in the intervention implementation as well as the small number of studies discussed for each of the following study characteristics should be kept in mind when reading the following discussion.

### **Effects specific to particular study characteristics**

#### ***Effect Sizes Across Age Groups or Grade Levels***

It is encouraging to see positive effect sizes having emerged across all age groups, lending support to the efficacy of empathy interventions beyond lower grade levels. On average, effect sizes identified in this literature review were slightly lower in secondary and high school than primary school. Nocentini and Menesini (2016), exploring how the size of the empathy gains caused by the intervention are affected by grade level, found larger effect sizes in primary schools than secondary schools. This supports findings that social and emotional learning programs incorporating empathy-related components tend to be more effective when implemented at an earlier age (Malti et al., 2016). However, Garandean et al. (2016) found a higher intention of bullies to change their behavior due to empathy arousal in secondary compared with primary schools, challenging this conclusion. It may be the case

that empathy gains were part of more complex interactions with other variables, given that Garandeau et al. (2016) intervention only targeted bullies, while Nocentini and Menesini (2016) provided measures of empathy gains for all students.

The development of different aspects of empathy (particularly affective and cognitive empathy) on different timescales may have contributed to interactions observed in control groups, where empathy decreases or increases variably in different studies. Findings of decreases in empathy in the control groups highlight the importance of implementing interventions to counteract these developments. Several studies noted that intervention effects, especially in secondary schools, were accompanied by levels of (affective) empathy decreasing significantly in control groups and less so in the experimental group, while they increased in both primary grade control and experimental groups. Thus, the intervention conditions often had a stabilizing effect on secondary school experimental group empathy levels. For instance, Garandeau et al. (2021) noted that “the positive effect of KiVa on affective empathy is due to levels of affective empathy decreasing significantly less in KiVa classrooms compared to control classrooms” (p. 524). However, it should be noted that decreases in control group empathy were also reported for primary grades in several studies. Further, while all control groups decreased in secondary school, findings were more mixed for high school control groups. More research is needed to ascertain the path of control group empathy changes for different grade levels and age groups.

### ***Effect Sizes by Sample Size***

Given that the majority of effect sizes reported were small or close to small, Cohen’s guidelines for small effect sizes indicating that an N of 393 is required to achieve a significant result at  $\alpha = 0.05$  (with a statistical power  $1 - \beta = 0.80$ ) can be used (Cohen, 1992). Six studies had under 393 participants, and three of these did not find a significant effect, possibly indicating the need for a larger sample size to obtain a significant effect.

Moreover, given that many effect sizes were negligible, an even larger N may be required in many of the reviewed studies in order to observe significant results at  $\alpha = 0.05$  (with a statistical power  $1 - \beta = 0.80$ ). The study of Garandean et al. (2021), employing a large N of 15,403 students, provides a good example: For affective empathy, they obtained a significant intervention effect, combined with only a negligible effect size (reflecting that with large sample sizes, even numerically very small differences can become significant). For this reason, the weighting of results should favor studies that achieved both significant effects and effect sizes that were at least small (and therefore at least to some degree practically meaningful).

In terms of study size, studies with a sample size over 1000 had a larger proportion of negligible effect sizes than studies with a sample size between 100 and 1000. However, the general pattern of results largely remained the same, regardless of sample size. While the studies with a sample sizes over 100 generally have good statistical power, weighing the conclusions of the literature analysis towards studies with over 1000 participants would effect somewhat lower average effect sizes and intervention efficacy in terms of promoting empathy in school children.

### ***Effect Sizes in Bullying versus Non-bullying Interventions***

No clear patterns emerged when comparing empathy effect sizes in interventions that direct themselves at reductions in bullying with empathy effect sizes in those that did not. Results not exceeding small effect sizes in either category may mean that both intervention content in anti-bullying interventions and in interventions not directed at bullying had difficulties in substantially raising empathy.

**Empathy Effect Sizes versus Bullying Effect Sizes.** Given the lack of overlap in effect size magnitudes between empathy gains and bullying reductions, empathy is likely not largely or solely responsible for increases in bullying. This conclusion is supported by the



similarity in effect sizes for bullying reduction in primary and secondary school, despite empathy gains being highly different in grade 4 versus grade 6 as observed by the study of Nocentini and Menesini (2016). Barkoukis et al. (2015) reported a lack of effect on empathy gains, but improvements in bullying attitudes. Ryzin and Roseth (2019) found an indirect effect of cooperative learning on bullying via peer-relatedness (though of negligible size,  $\beta = -0.04, p < .05$ ), indicating that bullying behavior may be influenced by other aspects of peer relations besides effects of interventions aiming to develop empathy.

However, empathy gains may more effectively lead to reductions in bullying in combination with other contextual factors. For instance, in the study by Alan et al. (2021), the likelihood of being bullied decreased only for the refugee group, even though refugee and host children experienced similar gains in perspective taking. A possible explanation could be a need for the rate of bullying to be “high” enough in a specific group in order for empathy gains to have an effect on bullying, or other characteristics pertaining specifically to either of the groups could account for this effect.

**The Relation of Cognitive and Affective Empathy and Bullying Reduction.** The finding of Ryzin and Roseth (2019) that only affective empathy reduced bullying is in line with previous literature reporting affective empathy to be more strongly linked to bullying reductions than cognitive empathy (e.g., van Noorden et al., 2014). Possibly supporting the role of affective empathy in reducing bullying, Garandeau et al. (2021) reported a negligible effect on affective empathy but none on cognitive empathy, drawing their data for analysis from the anti-bullying KiVa study done by Kärnä et al. (2011) that achieved reductions in bullying. However, the negligible effect size of affective empathy makes it unlikely that affective empathy accounts for much of this reduction.

However, the findings by Alan et al. (2021) of host students improving significantly on perspective taking, with no improvements occurring along the other dimensions of

behavioral norms, ethnic bias, impulsivity, empathic concern and eyes test, may have related to the central role of perspective taking in reducing bullying. Whereas the host students experienced no change in being bullied, the reductions in bullying of the refugee student group suggests that host students may have been responsible for this change. Further, Garandau et al. (2021) note that “bullying was negatively associated with both affective and cognitive empathy for victimized children” (p. 524) in their sample. They further postulate that “this implies that those who bullied showed a lack of understanding for victims that the program could not modify” (p. 524).

### *Affective and Cognitive Empathy*

Affective and cognitive empathy not only seem to affect bullying reduction differently, but also themselves appear to be differently affected by interventions aiming to increase empathy in school settings (these differences may be at least partially related to intervention length, which is discussed further below). While five out of six studies found an effect for affective empathy, only three found an effect for cognitive empathy. Several possible explanations can perhaps account for this difference. Boduszek et al. (2019) noted that cognitive empathy may take longer than their short intervention timespan of five days to consolidate. Additionally, they recommend the use of a facilitated discussion condition to make for a more intensive intervention capable of targeting cognitive empathy. However, the study by Troxler et al. (2022), which had an even shorter intervention time, did find significant changes in cognitive empathy. Possibly, the content of the intervention in terms of its ability to target cognitive empathy could be more important than intervention length.

While Ryzin and Roseth (2019) found significant effects for both affective and cognitive empathy over a 2-year period, the study by Garandau et al. (2021), conducted over nine months, reported no significant effects for cognitive empathy (although effect sizes for affective empathy were also very small, Cohen’s  $d = 0.07$ ). To explain the findings, the

authors noted that initial levels of cognitive empathy were higher than levels of affective empathy, potentially making it less likely for cognitive empathy to increase due to the intervention. Moreover, they suspect elements of the KiVa program to have been more focused on affective empathy (i.e., stories of past victimization), and less on better understanding the experience of bullied classmates. This ties in with the aforementioned importance of designing an intervention curriculum with specificity towards the target outcome. In sum, though caution is advised given the small number of studies, findings support the theoretical divergence of affective and cognitive empathy and the importance of ensuring that interventions explicitly target each construct.

Three studies did not observe significant effects for cognitive empathy. On the one hand, it may be argued that it is sufficient to feel sympathy in the face of others' pain due to affective empathy to be motivated to help, without the direct need to understand the reasons behind other's feelings (Boduszek et al., 2019). An argument could be that positive changes in children's behavior (such as a reduction in bullying) can be elicited by affective empathy alone. On the other hand, it could be argued that a more accurate understanding of other's emotions will lead to more helpful behaviors towards the person in need. Diverging effect sizes and findings for cognitive and affective empathy support research showing that the ability to empathize is independent of the capacity to mentalize, both neurally and behaviorally (Kanske et al., 2016).

The findings suggest the importance of affective empathy over cognitive empathy in reducing bullying. In support, Ryzin and Roseth (2019) reported on interactions between cognitive and affective empathy and bullying, and found that only affective empathy significantly interacted with bullying. Garandeau et al. (2021) only found a significant effect for affective empathy, and indicated reduced bullying levels found in the study by Kärnä et al. (2011) that their data is drawn from (both studies base themselves on data from the

evaluation of the KiVa anti-bullying program in Finland in 2007–2009). To support their findings theoretically, Ryzin and Roseth (2019) note that “students with high affective empathy will be better able to experience the negative emotional reaction of a bullied classmate, and will thus be less inclined to continue bullying or to bully others in the future. Cognitive empathy, on the other hand, only suggests the ability to understand another’s emotions, but not to experience them vicariously, and thus may be insufficient to inhibit bullying behavior” (p. 644).

### ***The Impact of the Length of the Intervention and Time-Stability of Intervention Effects***

In terms of attempting to identify an optimal or minimum time span for interventions to effect on empathy, all studies conducted within a timeframe of less than a week had a positive effect. While this is encouraging, the effects vary in size, with one effect size not being reported, and there is a risk for effects being short-lived. The seven-day follow-up provided by Boduszek et al. (2019) reveals that empathy levels were maintained in the experimental group (Cohen’s  $d = 0.56$ ). Nonetheless, a longer follow-up is needed to ascertain benefits over longer periods of time.

All seven studies ranging between twelve weeks to approximately a year indicated a positive intervention effect, potentially supporting the use of this intervention time frame. However, the observed effect sizes were mostly small, and none of the studies within this range include follow-up measurements. While not defined as a follow-up, Domitrovich et al. (2022) administered their post-test three months after the intervention meant to stimulate empathy was completed, reporting potentially maintained gains in empathy. This potentially indicates lasting gains in empathy, though empathy gains may have been maintained by other components of the curriculum.

Only three studies employed a follow-up measurement, making it difficult to draw firm conclusions. While empathy being maintained in a seven-day-follow-up by Boduszek et

al. (2019) is promising, this is a short time period. Nieh and Wu (2018) follow-up at two weeks did not maintain gains in empathy (however, they provided no numeric effect size). The one-year follow-up by Edwards et al. (2019), the intervention having been conducted over seven sessions, was unable to maintain gains in empathy. However, the study had few sessions and produced a negligible effect size, possibly indicating a lack of efficacy of the intervention curriculum in raising empathy.

### ***Global versus Specific Level Instruments for Empathy***

No clear differences in effect sizes emerged between instruments measuring empathy globally or specifically. For instance, Nocentini and Menesini (2016) used an instrument from the original KiVa study in Finland that assessed empathy specifically in bullying situations. In contrast, Troxler et al. (2022) used the Basic Empathy Scale instead of designing an instrument specific to the theater play attended by students. In order to detect empathy changes on a specific level as opposed to a global level, which may have remained the same, studies may have needed to use scales that assess context specific empathy to the group they promoted empathy towards. For instance, the more subtle differences picked up through the qualitative interviews in the study by Hilliard et al. (2016) could reflect changes in specific levels of empathy that were unable to be detected by the measurement instrument.

### ***Effect sizes by intervention technique***

The reviewed studies used differing intervention methods that can be largely categorized into computerized / computer-supported interventions (relying on, e.g., games and simulations) and behavioral interventions. An analysis of empathy effect sizes by intervention technique was not conducted, given that intervention curricula often highly overlapped in techniques used across studies, making use of a broad range of strategies simultaneously.

An example of a *computerized / computer-supported intervention* is the study of Boduszek et al. (2019), employing a computer game. This may be an especially interesting avenue for future interventions to explore. For instance, interventions utilizing virtual reality have attracted significant attention in the research community. Specifically, the immersive experiences these provide may especially enhance affective empathy (Martingano, 2021). While potentially not quite as immersive as virtual reality, Boduszek et al. (2019) likewise only find an effect for affective empathy. Such interventions may thus require and be potentiated by additional research exploring the promotion of cognitive empathy in immersive environments.

Among *behavioral interventions*, one possible approach is cooperative learning. Cooperative learning may be more effective for certain age groups than others, as it provides older children with greater autonomy (Yeager et al., 2015). This has been observed in anti bullying programs, which experience a drop in effectiveness in middle school, potentially due to adolescents objecting to highly structured lessons (Yeager et al., 2015). It would be of interest to explore whether interventions directed at raising empathy in school children produce similar results.

Another *behavioral intervention* approach is aiming at reducing prejudice and negative out-group beliefs. The explicit focus of the study by Domitrovich et al. (2022) on targeting moral beliefs around prejudice and social justice as a way to promote empathy supports the relation between these two concepts, though it is difficult to ascertain whether one could promote the other. For instance, a change in beliefs around the extent to which America is democratic may be theorized to foster greater beliefs in personal self-efficacy and the potential of the impact of personal actions, culminating in more prosocial behavior, as well as an increased belief in the meaningfulness of attempting to be empathic. Promoting empathy via a partial focus on moral constructs might even be more effective in upper grades,

as adolescents increase their interest in identity and moral commitments (Domitrovich et al., 2022).

However, not all interventions aimed at the reduction of prejudice also succeed in increasing empathy. A study by Barkoukis et al. (2015) identified no significant intervention effect on empathy despite lower scores in distortion of consequences (i.e., downplaying the detrimental effects of the misconduct on victims), and attribution of blame (i.e., blaming the victim), even though moral disengagement and attitudes were found to be moderately correlated with empathy in the same study. Thus, decreases in moral disengagement may not automatically foster a higher affective emotional activation or better emotional understanding of others. It may thus be important to specifically integrate empathy promotion skills within interventions, and not to simply assume that theoretically related concepts will lead to gains in empathy.

#### ***Potential Explanations for Non-Significant Effects***

In terms of the non-significant effects found, studies noted several possible explanations. Hilliard et al. (2016) noted that intervention length and time between intervention and outcome measurement may have been insufficient for the empathy scale to pick up changes in trait empathy, which often take longer to consolidate. Failure of interventions in raising empathy may therefore potentially be attributed to the little number of sessions. In accordance with this, Hilliard et al. (2016) describe a time of three sessions over three weeks as potentially insufficient at producing changes in an empathy scale measuring trait empathy. In terms of the lack of a significant effect on cognitive empathy, Garandeau et al. (2021) write that “it should be noted that average levels of cognitive empathy for victimized children were initially higher than average levels of affective empathy, which could imply that cognitive empathy was less likely to benefit from an intervention” (p. 524).

While finding no main effect, Low et al. (2015) found benefits for the intervention in general to be greater for students with lower baseline competencies in terms of positive socio-emotional skills and positive behaviors. Given the interaction effect identified between proactive classroom environment and empathy in the same study, it would be of interest to explore whether lower baseline competences interacted with a more proactive classroom environment to produce greater gains in empathy.

Barkoukis et al. (2015) attribute the non-significant findings to a lack of focus on empathy skills and empathic understanding in the intervention curriculum. While the feelings of cyberbullying victims were made explicit, there was no focused training on empathy skills. Given that the intervention led to reductions in moral disengagement, it appears important to specifically integrate empathy promotion skills within interventions, and not to simply assume that theoretically related concepts will lead to gains in empathy. Additionally, the intervention curriculum may have been too specific towards raising empathy for specific groups to obtain significant effects. Previous research indicated that violence accepting attitudes can be specific to certain groups, for instance in the sense that men committing acts of violence in intimate relationships often did not engage in interpersonal violence in other domains of life (Holtzworth-Munroe & Meehan, 2004). Thus, raising empathy in general may not be enough to induce changes in empathy or in negative/prejudiced beliefs towards a subgroup of the population (e.g. women that were victims of intimate partner violence). Hilliard et al. (2016) made a similar claim regarding the game episodes that were used in the study intervention, positing that they may not have enough personal relevance to the students lives to elicit greater personal effort in problem-solving and to possibly lead to significant empathy increases.



### **Limitations of the Review**

One limitation of this review is the relatively small number of studies included. Further, two studies did not report the magnitude of the effect size, making it difficult to ascertain the impact of the intervention and to integrate it into the analysis of study characteristics. While the review hoped to investigate several characteristics that interact with the efficacy of raising empathy, few of the studies reviewed consistently reported on (similar) interacting variables. Another limitation is that the vast majority of studies used self-report. This may have led to inaccuracies in the results, as students may have over-reported prosocial and empathic tendencies to conform with social expectations. Thus, observer results may differ from self-report results. Only two studies employed forms of observer reports (one having no significant effect), and these assessed a wide range of socio-emotional skills instead of specifically focusing on empathy.

While a high number of studies used scales with proven validity, such as the Basic empathy scale and the Individual Reactivity Index, some studies used self-constructed scales. As such, comparisons in the review may be limited by the differing of scales in content and by uncertainties about their validity and reliability. Nonetheless, studies included in the literature review generally included high numbers of participants in control and experimental groups, accounted for attrition rates in their statistical analysis, and used a methodologically rigorous randomized control trial design. However, studies included were predominantly from the Western hemisphere.

### **Directions and Implications for Future Research**

There is a clear need for further research focusing on interactions between individual and contextual characteristics and empathy intervention effects in school aged children, as well as on the differential effects of intervention components on intervention efficacy. This appears especially important given Garandean et al. (2021) noting that “whether the effects of

school-based interventions on empathy varied depending on individual characteristics of the participants or features of the classroom had never been investigated” (p. 523) prior to their intervention. Underscoring the importance of further investigating characteristics related to empathy gains towards specific out-groups, research indicates that “between-individual variations in empathy may be due to target differences: bullies and victims show less empathy for each other than for non-involved peers, and girls reported more empathy for girls than for boys” (Garandau et al., 2021, p. 525). For instance, analysis could possibly identify what individual and contextual characteristics (e.g. what in-group one belongs to) affect empathy towards different out-groups.

Future research may be needed to further elucidate interactions between study characteristics and empathy gains. Most of the studies reviewed here did not consistently investigate such characteristics, and rather conflicting results were revealed by studies which did investigate this question: On the one hand, Garandau et al. (2021) found that empathy gains in the intervention condition did not significantly interact with participant empathy, popularity of the child, bullying perpetrator status, classroom bullying norms and gender. Though the study only found a negligible effect size on affective empathy, this warrants future attempts at replication. On the other hand, Troxler et al. (2022) found that modifications to the intervention implementation (supplementing the main intervention, a theater visit with prior reading in preparation and / or a facilitated discussion), affected the magnitude of empathy gains and respective effect sizes.

In the same vein, intervention empathy gains may have been affected by being a member of a specific population subgroup, and such differences should be investigated further. For instance, studies found ringleader bullies to outscore follower bullies in emotion understanding (e.g., Sutton et al., 1999). Thus, interventions may have been more effective for follower bullies. Given that some studies found relations between cognitive empathy and

bullying and others did not, the degree to which the study population consists of ringleader or follower bullies may account for these discrepancies. This could be explored in future studies by designing questionnaires administered to specific subgroups of the student population, combined with measures of empathy towards specific peer subpopulations rather than general empathy measures. Likewise, future studies may wish to explore whether combinations of several variables (e.g. a combination of raising empathy, classrooms norms, bystander behavior training) had a larger effect on reductions in bullying together than any of these variables alone. In support of this, Garandeau et al. 2016 found that “bullies’ intention to change is the lowest when both empathy arousal and condemning of the behavior are at their lowest level and the highest when both empathy arousal and condemning of the behavior are at their highest level” (p. 1040).

Future studies could focus on identifying which intervention curriculum best activates both affective and cognitive empathy, the differences between affective and cognitive empathy in terms of their ability to be promoted in interventions, and what mechanisms underlie the magnitude of effect sizes for each construct. As recent as 2019, Ryzin and Roseth noted that their study “represents the first time... that an anti-bullying program has been found to have significant effects on both cognitive and affective empathy” (p. 649). Within this literature review, only Troxler et al. (2022) and Alan et al. (2021) additionally reported significant findings for both cognitive and affective empathy. Thus, more research on the potential for the promotion of affective and cognitive empathy through interventions is needed.

Another avenue for future research is to continue exploring whether affective and cognitive empathy (causally) raise prosocial outcomes, such as reductions in bullying. Only Ryzin and Roseth (2019) reported on relations between cognitive and affective empathy and bullying, and only Garandeau et al. (2021) additionally included measures of affective and

cognitive empathy and pointed to reduced bullying levels in the study their data was drawn from. As previously discussed, empathy did not always relate to bullying behaviors. While Garandeau et al. (2021) did not find an interaction of empathy gains with bullying characteristics, the KiVa program itself had less effect on the subgroup of highly popular bullies (Garandeau et al., 2021). Domitrovich et al. (2022) did not find an intervention effect for student willingness to intervene in hypothetical bullying vignettes. In the study by Ryzin and Roseth (2019), improvement in cognitive empathy had no effect on bullying. The intervention condition also had considerably little impact on bystander behavior in the study by Edwards et al. (2019). Thus, more research is needed to solidify knowledge on the role of affective and cognitive empathy in reducing bullying.

Empathy levels thus appear to depend on other factors than simply attempting to evoke affective and cognitive empathy in students, and these are in need of being explored. Garandeau et al. (2021) noted that affective and cognitive empathy for victimized classmates were lower in classrooms where popularity was a reward for bullying, potentially due to collective moral disengagement mechanisms. They recommend future interventions to take into account the effect of classroom bullying norms on empathy. Additionally, future research should explore other limiting effects on empathy gains, such as whether having few victimized peers in the classroom could potentially make them more deserving of their victim status to their peers and thus undermine empathy (Garandeau et al., 2021).

Changes in moral thinking may not produce consequent changes to empathy, and future interventions should further investigate empathy's relation to similar prosocial constructs. As suggested by the study of Barkoukis et al. (2015), despite empathy and its subscales being moderately correlated with moral disengagement and attitudes, no effects on empathy were identified. This indicates the need for future research to be wary of theoretical

and practical distinctions in these concepts, and that they may not necessarily overlap and mutually reinforce one another.

Finally, future research could also further explore the effect of context specific instruments for measuring empathy on empathy gains. It would be interesting for future research to measure context specific empathy levels post intervention, given that context specific empathy was targeted in the intervention, and then to compare these to global levels of empathy to ascertain if these had also changed. A question for future research is whether the predominantly negligible and small effect sizes found in the literature review would change if context specific instruments for empathy were specifically designed and applied. Importantly, research would also have to investigate whether context specific empathy and changes in context specific empathy relate to other prosocial behaviors, such as reductions in bullying.

### **Concluding Remarks**

In conclusion, the predominantly small effect sizes found by the studies included in the literature review suggest that school based interventions for children aimed at raising empathy did not have a substantial effect. Few studies explored interactions between study characteristics and empathy gains, highlighting the need for more comprehensive and rigorous future research that could potentially lead to more effective interventions.

### References

- Alan, S., Baysan, C., Gumren, M., & Kubilay, E. (2021). Building social cohesion in ethnically mixed schools: An intervention on perspective taking. *The Quarterly Journal of Economics*, *136*(4), 2147–2194. <https://doi.org/10.1093/qje/qjab009>
- Barkoukis, V., Lazuras, L., Ourda, D., & Tsorbatzoudis, H. (2015). Tackling Psychosocial Risk Factors for adolescent cyberbullying: Evidence from a school-based intervention. *Aggressive Behavior*, *42*(2), 114–122. <https://doi.org/10.1002/ab.21625>
- Bernstein, M. J., Chen, Z., Poon, K.-T., Benfield, J. A., & Ng, H. K. (2018). Ostracized but why? Effects of attributions and empathy on connecting with the socially excluded. *PLOS ONE*, *13*(8). <https://doi.org/10.1371/journal.pone.0201183>
- Boduszek, D., Debowska, A., Jones, A., Ma, M., Smith, D., Willmott, D., Jemmott, E. T., Da Breo, H., & Kirkman, G. (2019). Prosocial video game as an intimate partner violence prevention tool among youth: A randomised controlled trial. *Computers in Human Behavior*, *93*, 260–266. <https://doi.org/10.1016/j.chb.2018.12.028>
- Boncu, A., Costea, I., & Minulescu, M. (2017). A meta-analytic study investigating the efficiency of socio-emotional learning programs on the development of children and adolescents. *Romanian Journal of Applied Psychology*, 35–41. <https://doi.org/10.24913/rjap.19.2.02>
- Chan, Y. (2003). Randomised controlled trials (RCTs) – sample size: the magic number? *Singapore Medical Journal*, *44*(4), 172-174. <http://www.smj.org.sg/article/randomised-controlled-trials-rects-sample-size-magic-number>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>

- Cross, D., Barnes, A., Papageorgiou, A., Hadwen, K., Hearn, L., & Lester, L. (2015). A social-ecological framework for understanding and reducing cyberbullying behaviours. *Aggression and Violent Behavior, 23*, 109–117.  
<https://doi.org/10.1016/j.avb.2015.05.016>
- Davis, M. H., Luce, C., & Kraus, S. J. (1994). The heritability of characteristics associated with dispositional empathy. *Journal of Personality, 62*(3), 369–391.  
<https://doi.org/10.1111/j.1467-6494.1994.tb00302.x>
- Decety, J., Bartal, I. B., Uzefovsky, F., & Knafo, A. (2016). Empathy as a driver of prosocial behaviour: highly conserved neurobehavioural mechanisms across species. *Philosophical Transactions of the Royal Society B, 371*(1686), 20150077.  
<https://doi.org/10.1098/rstb.2015.0077>
- Ding, W. W., Shao, Y., Sun, B., Xie, R., Li, W., & Wang, X. (2018b). How Can Prosocial Behavior Be Motivated? The Different Roles of Moral Judgment, Moral Elevation, and Moral Identity Among the Young Chinese. *Frontiers in Psychology, 9*.  
<https://doi.org/10.3389/fpsyg.2018.00814>
- DiPerna, J. C., Lei, P., Cheng, W., Hart, S. C., & Bellinger, J. (2018). A cluster randomized trial of the Social Skills Improvement System-Classwide Intervention Program (SSIS-CIP) in first grade. *Journal of Educational Psychology, 110*(1), 1–16.  
<https://doi.org/10.1037/edu0000191>
- Domitrovich, C. E., Harris, A. R., Syvertsen, A. K., Morgan, N., Jacobson, L., Cleveland, M., Moore, J. E., & Greenberg, M. T. (2022). Promoting social and emotional learning in middle school: Intervention effects of facing history and ourselves. *Journal of Youth and Adolescence, 51*(7), 1426–1441. <https://doi.org/10.1007/s10964-022-01596-3>
- Duong, J. (2014). *Ecological predictors of children's social, emotional, and behavioral outcomes*. Doctoral dissertation, Johns Hopkins University.

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*(1), 405–432.

<https://doi.org/10.1111/j.1467-8624.2010.01564.x>

Edwards, K. M., Banyard, V. L., Sessarego, S. N., Waterman, E. A., Mitchell, K. J., &

Chang, H. (2019). Evaluation of a bystander-focused interpersonal violence prevention program with high school students. *Prevention Science, 20*(4), 488–498.

<https://doi.org/10.1007/s11121-019-01000-w>

Eisenberg, N., Fabes, R. A., & Spinrad, T. L. (2006). Prosocial Development. In N.

Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 646–718). John Wiley & Sons, Inc.

Espelage, D.L., King, M.T., Colbert, C.L. (2018). Emotional Intelligence and School-Based

Bullying Prevention and Intervention. In: Keefer, K., Parker, J., Saklofske, D. (eds) *Emotional Intelligence in Education. The Springer Series on Human Exceptionality.*

Springer. [https://doi.org/10.1007/978-3-319-90633-1\\_9](https://doi.org/10.1007/978-3-319-90633-1_9)

Fernandez, Y. M., & Marshall, W. L. (2003). Victim Empathy, Social Self-Esteem, and

Psychopathy in Rapists. *Sexual Abuse: A Journal of Research and Treatment, 15*(1),

11–26. <https://doi.org/10.1177/107906320301500102>

Fitzgibbons, R. P. (1986). The cognitive and emotive uses of forgiveness in the treatment of

anger. *Psychotherapy: Theory, Research, Practice, Training, 23*(4), 629–633.

<https://doi.org/10.1037/h0085667>



- Garaigordobil, M., & Martínez-Valderrey, V. (2014). Efecto del Cyberprogram 2.0 sobre la reducción de la victimización y la Mejora de la Competencia social en la adolescencia [Effect of cyberprogram 2.0 on reducing victimization and improving social competence in adolescence]. *Revista De Psicodidactica / Journal of Psychodidactics*, *19*(2), 289–305. <https://doi.org/10.1387/revpsicodidact.10239>
- Garandeau, C. F., Laninga-Wijnen, L., & Salmivalli, C. (2021). Effects of the kiva anti-bullying program on Affective and Cognitive Empathy in children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, *51*(4), 515–529. <https://doi.org/10.1080/15374416.2020.1846541>
- Garandeau, C. F., & Salmivalli, C. (2018). Le programme anti-harcèlement Kiva [The anti-bullying program Kiva]. *Enfance*, *N° 3*(3), 491–501. <https://doi.org/10.3917/enf2.183.0491>
- Greenberg, M. T., & Abenavoli, R. (2016). Universal interventions: Fully exploring their impacts and potential to produce population-level impacts. *Journal of Research on Educational Effectiveness*, *10*(1), 40–67. <https://doi.org/10.1080/19345747.2016.1246632>
- Garandeau, C. F., Vartio, A., Poskiparta, E., & Salmivalli, C. (2016). School bullies' intention to change behavior following teacher interventions: Effects of empathy arousal, condemning of bullying, and blaming of the perpetrator. *Prevention Science*, *17*(8), 1034–1043. <https://doi.org/10.1007/s11121-016-0712-x>
- Henry, J., & Ruffman, T. (2022). Supplemental material for a meta-analytic review on the social–emotional intelligence correlates of the six bullying roles: Bullies, followers, victims, bully-victims, defenders, and outsiders. *Psychological Bulletin*, *148*(3-4), 199–226. <https://doi.org/10.1037/bul0000364.supp>

- Hilliard, L. J., Buckingham, M. H., Geldhof, G. J., Gansert, P., Stack, C., Gelgoot, E. S., Bers, M. U., & Lerner, R. M. (2016). Perspective taking and decision-making in educational game play: A mixed-methods study. *Applied Developmental Science, 22*(1), 1–13. <https://doi.org/10.1080/10888691.2016.1204918>
- Holbrook, C., Hahn-Holbrook, J. (2022). Evolved to Learn: Emotions as Calibrational Adaptations. In Dukes, D., Samson, A. C., Walle, E. A. Walle (eds), *The Oxford Handbook of Emotional Development* (pp. 3-13). Oxford university press. <https://doi.org/10.1093/oxfordhb/9780198855903.013.14>
- Holtzworth-Munroe, A., & Meehan, J. C. (2004). Typologies of men who are maritally violent: Scientific and clinical implications. *Journal of Interpersonal Violence, 19*(12), 1369–1389. <https://doi.org/10.1177/0886260504269693>
- Influs, M., Pratt, M., Masalha, S., Zagoory-Sharon, O., & Feldman, R. (2018). A social neuroscience approach to conflict resolution: Dialogue intervention to Israeli and Palestinian youth impacts oxytocin and empathy. *Social Neuroscience, 14*(4), 378–389. <https://doi.org/10.1080/17470919.2018.1479983>
- Israelashvili, J., Sauter, D., & Fischer, A. H. (2020). Two facets of affective empathy: concern and distress have opposite relationships to emotion recognition. *Cognition & Emotion, 34*(6), 1112–1122. <https://doi.org/10.1080/02699931.2020.1724893>
- Jolliffe, D., & Farrington, D. P. (2010). Is low empathy related to bullying after controlling for individual and social background variables? *Journal of Adolescence, 34*(1), 59–71. <https://doi.org/10.1016/j.adolescence.2010.02.001>
- Kanske, P., Böckler, A., Trautwein, F., Lesemann, F. H. P., & Singer, T. (2016). Are strong empathizers better mentalizers? Evidence for independence and interaction between the routes of social cognition. *Social Cognitive and Affective Neuroscience, 11*(9), 1383–1392. <https://doi.org/10.1093/scan/nsw052>

- Knafo, A., Zahn-Waxler, C., Van Hulle, C., Robinson, J. L., & Rhee, S. H. (2008). The developmental origins of a disposition toward empathy: Genetic and environmental contributions. *Emotion, 8*(6), 737–752. <https://doi.org/10.1037/a0014179>
- Kärnä, A., Voeten, M., Little, T. D., Poskiparta, E., Kaljonen, A., & Salmivalli, C. (2011). A Large-Scale Evaluation of the KiVa Antibullying Program: Grades 4-6. *Child Development, 82*(1), 311–330. <https://doi.org/10.1111/j.1467-8624.2010.01557.x>
- Low, S., Cook, C. R., Smolkowski, K., & Buntain-Ricklefs, J. (2015). Promoting social-emotional competence: An evaluation of the Elementary version of Second Step®. *Journal of School Psychology, 53*(6), 463–477. <https://doi.org/10.1016/j.jsp.2015.09.002>
- Mahoney, J. L., Durlak, J. A., & Weissberg, R. P. (2018). An update on social and emotional learning outcome research. *Phi Delta Kappan, 100*(4), 18–23. <https://doi.org/10.1177/0031721718815668>
- Malinauskas, R., & Malinauskiene, V. (2021). Training the social-emotional skills of youth school students in Physical Education Classes. *Frontiers in Psychology, 12*. <https://doi.org/10.3389/fpsyg.2021.741195>
- Malti, T., Chaparro, M. P., Zuffianò, A., & Colasante, T. (2016). School-based interventions to promote empathy-related responding in children and adolescents: A developmental analysis. *Journal of Clinical Child & Adolescent Psychology, 45*(6), 718–731. <https://doi.org/10.1080/15374416.2015.1121822>
- Martingano, A. J., Herrera, F., & Konrath, S. (2021). Virtual reality improves emotional but not cognitive empathy: A meta-analysis. *Technology, Mind, and Behavior, 2*(1). <https://doi.org/10.1037/tmb0000034>

- Morawetz, C., Berboth, S., Kohn, N., Jackson, P. L., & Jauniaux, J. (2022). Reappraisal and empathic perspective-taking – more alike than meets the eyes. *NeuroImage*, 255, 119194. <https://doi.org/10.1016/j.neuroimage.2022.119194>
- Moore, R. C., Dev, S. I., Jeste, D. V., Dziobek, I., & Eyler, L. T. (2015). Distinct neural correlates of emotional and cognitive empathy in older adults. *Psychiatry Research: Neuroimaging*, 232(1), 42–50. <https://doi.org/10.1016/j.psychres.2014.10.016>
- Nieh, H.-P., & Wu, W.-C. (2018). Effects of a collaborative board game on Bullying intervention: A group-randomized controlled trial. *Journal of School Health*, 88(10), 725–733. <https://doi.org/10.1111/josh.12675>
- Niezink, L. W., Siero, F. W., Dijkstra, P., Buunk, A. P., & Barelds, D. P. H. (2012). Empathic concern: Distinguishing between tenderness and sympathy. *Motivation and Emotion*, 36(4), 544–549. <https://doi.org/10.1007/s11031-011-9276-z>
- NiMARE. (n.d.) *Meta-analytic functional decoding*. <https://nimare.readthedocs.io/en/latest/decoding.html>
- Nocentini, A., & Menesini, E. (2016). KiVa anti-bullying program in Italy: Evidence of effectiveness in a randomized control trial. *Prevention Science*, 17(8), 1012–1023. <https://doi.org/10.1007/s11121-016-0690-z>
- Raskauskas, J., & Stoltz, A. D. (2007). Involvement in Traditional and Electronic Bullying among Adolescents. *Developmental Psychology*, 43, 564-575. <https://doi.org/10.1037/0012-1649.43.3.564>
- 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(6), 922–934. <https://doi.org/10.1002/ejsp.504>
- Pinker, S. (2012). *The better angels of our nature: Why violence has declined*. Penguin Books.

- Roubin, A.M. (2014). *A resource manual of bullying intervention prevention programs for parents, educators, and community officials in the Los Angeles area*. Doctoral dissertation, Pepperdine University.
- Ryzin, M. J., & Roseth, C. J. (2019). Effects of cooperative learning on peer relations, empathy, and bullying in Middle School. *Aggressive Behavior, 45*(6), 643–651. <https://doi.org/10.1002/ab.21858>
- Rushton, J. P., Fulker, D. W., Neale, M. C., Nias, D. K. B., & Eysenck, H. J. (1986). Altruism and aggression: The heritability of individual differences. *Journal of Personality and Social Psychology, 50*(6), 1192–1198. <https://doi.org/10.1037/0022-3514.50.6.1192>
- Salay, D. M. (2018). *Walk in their shoes: How picture books and critical literacy instruction can foster empathy in first grade students*. Doctoral dissertation, Drexel University.
- Salavera, C., Bordás, C. S., Teruel, P., Urbón, E. F., & Murillo, V. H. V. (2021). School Bullying: Empathy among Perpetrators and Victims. *Sustainability, 13*(3), 1548. <https://doi.org/10.3390/su13031548>
- Smith, A. (2006). Cognitive empathy and emotional empathy in human behavior and evolution. *The Psychological Record, 56*(1), 3–21. <https://doi.org/10.1007/BF03395534>
- Spinrad, T. L., & Gal, D. E. (2018). Fostering Prosocial behavior and empathy in young children. *Current Opinion in Psychology, 20*, 40–44. <https://doi.org/10.1016/j.copsyc.2017.08.004>
- Teding van Berkhout, E., & Malouff, J. M. (2016). The efficacy of empathy training: A meta-analysis of randomized controlled trials. *Journal of Counseling Psychology, 63*(1), 32–41. <https://doi.org/10.1037/cou0000093>

- Sutton, J., Smith, P. K., & Swettenham, J. (1999). Social Cognition and bullying: Social inadequacy or skilled manipulation? *British Journal of Developmental Psychology*, *17*(3), 435–450. <https://doi.org/10.1348/026151099165384>
- Thompson, E., & Melchior, S. (2019). Improving empathy in children: Interactive read-aloud as a counseling intervention. *Journal of Creativity in Mental Health*, *15*(2), 199–211. <https://doi.org/10.1080/15401383.2019.1676857>
- van Baardewijk, Y., Stegge, H., Bushman, B. J., & Vermeiren, R. (2009). Psychopathic traits, victim distress and aggression in children. *Journal of Child Psychology and Psychiatry*, *50*(6), 718–725. <https://doi.org/10.1111/j.1469-7610.2008.02023.x>
- Troxler, R., Goldstein, T., Holochwost, S., Beekman, C., McKeel, S., & Shami, M. (2022). Deeper engagement with live theater increases middle school students' empathy and social perspective taking. *Applied Developmental Science*, 1–21. <https://doi.org/10.1080/10888691.2022.2096610>
- University of Cambridge. (n.d.). *Rules of thumb on magnitudes of effect sizes*. <https://imaging.mrc-cbu.cam.ac.uk/statswiki/FAQ/effectSize>
- van Noorden, T. H., Haselager, G. J., Cillessen, A. H., & Bukowski, W. M. (2014). Empathy and involvement in bullying in children and adolescents: A systematic review. *Journal of Youth and Adolescence*, *44*(3), 637–657. <https://doi.org/10.1007/s10964-014-0135-6>
- Vezzali, L., Stathi, S., Giovannini, D., Capozza, D., & Trifiletti, E. (2014). The greatest magic of Harry Potter: Reducing prejudice. *Journal of Applied Social Psychology*, *45*(2), 105–121. <https://doi.org/10.1111/jasp.12279>

- Wang, Y., Zhu, W., Xiao, M., Zhang, Q., Zhao, Y., Zhang, H., Chen, X., Zheng, Y., & Xia, L. (2018). Hostile Attribution Bias Mediates the Relationship Between Structural Variations in the Left Middle Frontal Gyrus and Trait Angry Rumination. *Frontiers in Psychology, 9*, Article 526. <https://doi.org/10.3389/fpsyg.2018.00526>
- Warrier, V., Grasby, K. L., Uzefovsky, F., Toro, R., Smith, P., Chakrabarti, B., & Lubke, G. (2018). Genome-wide meta-analysis of cognitive empathy: Heritability, and correlates with sex, neuropsychiatric conditions and cognition. *Molecular Psychiatry, 23*, 1402–1409. <https://doi.org/10.1038/mp.2017.122>
- Yeager, D. S. (2017). Social and emotional learning programs for adolescents. *The Future of Children, 27*(1), 73–94. <https://doi.org/10.1353/foc.2017.0004>
- Yeager, D. S., Fong, C. J., Lee, H. Y., & Espelage, D. L. (2015). Declines in efficacy of anti-bullying programs among older adolescents: Theory and a three-level meta-analysis. *Journal of Applied Developmental Psychology, 37*, 36–51. <https://doi.org/10.1016/j.appdev.2014.11.005>
- Zych, I., Ttofi, M. M., & Farrington, D. P. (2016). Empathy and callous–unemotional traits in different bullying roles: A systematic review and meta-analysis. *Trauma, Violence, & Abuse, 20*(1), 3–21. <https://doi.org/10.1177/1524838016683456>

## Appendix

### Supplementary Table 1

*Overview over the studies that were included into the full-text-reading: The first column shows study title and authors; the second column (S) its source: S - Database search or F - added during full-text reading (e.g., because the article was mentioned in a meta-analysis); the third column (E) the exclusion criteria: 0 – included, 1 – excluded for methodological reasons (e.g., a non-experimental design), 2 – excluded for only being a pilot study, 3 – excluded because of age or experimental group characteristics not covered by the inclusion criteria, 4 – excluded because of being non-empirical (i.e., reviews or meta-analyses), 5 – excluded because of irrelevant outcome measures (i.e., not including or related to empathy), and 6 – other (e.g., outside the time range to be included); the last column gives a more comprehensive, verbal explanation of the reasons for exclusion.*

Study title and authors	S	E	Comments
A cluster randomized trial of the Social Skills Improvement System-Classwide Intervention Program (SSIS-CIP) in first grade (DiPerna et al., 2018)	S	0	included
Building Social Cohesion in ethnically Mixed Schools: An intervention on Perspective Taking (Alan et al., 2021)	S	0	included
Deeper engagement with live theater increases middle school students' empathy and social perspective taking (Troxler, 2022)	S	0	included
Effects of a Collaborative Board Game on Bullying Intervention: A Group-Randomized Controlled Trial (Nieh & Wu, 2018)	S	0	included
Effects of cooperative learning on peer relations, empathy, and bullying in middle school (Ryzen & Roseth, 2019)	S	0	included



Study title and authors	S	E	Comments
Effects of Cyberprogram 2.0 on “face-to-face” bullying, cyberbullying, and empathy (Garaigordobil & Martinez-Valderrey, 2015)	S	0	included
Evaluation of a bystander-focused interpersonal violence prevention program with high school students (Edwards et al., 2019)	S	0	included
KiVa anti-bullying program in Italy: Evidence of effectiveness in a randomized control trial (Nocentini & Menesini, 2016)	S	0	included
Promoting social and emotional learning in middle school: Intervention effects of Facing History and Ourselves (Domitrovich et al., 2022)	S	0	included
Promoting social-emotional competence: An evaluation of the elementary version of Second Step® (Low et al., 2015)	S	0	included
Prosocial video game as an intimate partner violence prevention tool among youth: A randomised controlled trial (Boduszek et al., 2019)	S	0	included
School bullies’ intention to change behavior following teacher interventions: Effects of empathy arousal, condemning of bullying, and blaming of the perpetrator (Garandean et al., 2016)	S	0	included
Tackling psychosocial risk factors for adolescent cyberbullying: Evidence from a school-based intervention (Barkoukis et al., 2015)	S	0	included
Training the social-emotional skills of youth school students in physical education classes (Malinaskaus & Malinauskiene, 2021)	S	0	included
Effects of the KiVa Anti-Bullying Program on Affective and Cognitive Empathy in Children and Adolescents (Garandean et al., 2021)	F	0	included
Perspective taking and decision-making in educational game play: A mixed-methods study (Hilliard et al., 2016)	F	0	included

Study title and authors	S E Comments
Development of emotional skills in adolescents to prevent cyberbullying and improve subjective well-being (Schoeps et al., 2018)	S 1 quasi-experimental
Effectiveness of an empathy intervention for youths at-risk (Fullchange, 2017)	S 1 selection of the intervention site was not random, waitlist control.
Effects of an emotional intelligence intervention on aggression and empathy among adolescents (Castillo et al., 2013)	S 1 quasi-experimental
Effects of cultural awareness training in conjunction with an established bullying prevention program (Polanin, 2014)	S 1 non-randomized selection of only one school
Effects of the preventive intervention program "Media Heroes" on the association between traditional and cyberbullying and internalizing problems (Hess et al., 2020)	S 1 text not available; likely non-random as connected with previously excluded study
Emotional intelligence and mindfulness: Relation and enhancement in the classroom with adolescents (Rodriguez-Ledo et al., 2018)	S 1 only one school was chosen; lacks information on selection process
Enhancing cognitive and social-emotional development through a simple-to-administer mindfulness-based school program for elementary school children: A randomized controlled trial (Schonert-Reichl et al., 2015)	S 1 schools were selected based on them promoting students' social responsibility'; baseline differences between intervention and control condition on one self-report measures (empathy) and most peer behavioral assessment indices
Feeling cybervictims' pain-The effect of empathy training on cyberbullying (Schulze-Krumbholz et al., 2016)	S 1 self-selective sample of schools, schools randomly assigned classrooms themselves
Play to reduce bullying! Role-playing games are a useful tool for therapists and teachers (Bages et al., 2020)	S 1 quasi-experimental
Promoting emotional competence: Short-term effects of the emotion training with students (Diener et al., 2021)	S 1 non-random assignment to control and treatment group

Study title and authors	S E Comments
Putting yourself in someone else's shoes: The impact of a location-based, collaborative role-playing game on behaviour (Schmitz et al., 2015)	S 1 insufficient information on selection of schools
The effects of digital storytelling with group discussion on social and emotional intelligence among female elementary school students (Zarifsanaiey et al., 2021)	S 1 only partially randomized, and quasi-experimental; includes data from one female-only school; no direct empathy measure
The Efficacy of Emotional Intelligence Training for the Emotion Regulation of Bullying Students: A Randomized Controlled Trial (Lang, 2018)	S 1 control group problems, treatment group reported being more likely to participate
The Kids4Dementia education program is effective in improving children's attitudes towards dementia (Baker et al., 2018)	S 1 non-randomized
To understand is to forgive: Learning a simple model of appraisal leads to emotion knowledge transfer and enhances emotional acceptance and empathy (Lyashevsky et al., 2019)	S 1 high school graduates aged 18 to 25; insufficient randomization, conducted via online recruitment.
Walk in their shoes: How picture books and critical literacy instruction can foster empathy in first grade students (Salay, 2018)	S 1 site of research wasn't randomly chosen: "the researcher acknowledges the dual role as principal and researcher at the research site during the 2016-2017 school year".
Effectiveness of a Prevention Program for Gender-Based Intimate Partner Violence at a Colombian Primary School (Segura & Gonzalez, 2020)	F 1 quasi-experimental
Evaluation of the Olweus Bullying Prevention Program: A large scale study of U.S. students in grades 3–11 (Limber et al., 2018)	F 1 quasi-experimental
Evaluation of a virtual reality enhanced bullying prevention curriculum pilot trial (Ingram et al., 2018)	S 2 pilot study; pseudo-randomized
Promotion of social and emotional learning in a Chinese elementary school (An et al., 2021)	S 2 pilot study

Study title and authors	S E Comments
Forgiveness intervention for female South Korean adolescent aggressive victims (Park et al., 2013)	S 3 implemented in a juvenile correction facility only
A systematic review of the second step program (Moy & Hazen, 2018)	S 4 meta-analysis
The anti-bullying program KiVa (Garandeau & Salmivalli, 2018)	S 4 review paper of the KiVa Finland study
Assessing the Impact of Holocaust Education on Adolescents' Civic Values: Experimental Evidence from Arkansas (Lee & Beck, 2021)	S 5 no direct measure of empathy
Board Games on Emotional Competences for School-Age Children (Dell'Angela et al., 2020)	S 5 no measure of empathy
Clinical trial of Second Step© middle-school program: Impact on aggression and victimization (Espelage et al., 2014)	S 5 no measure of empathy or closely related construct
Comparing indirect and combined effects of mindfulness and compassion practice among schoolchildren on inter- and intra-personal abilities (Tarrasch & Berger, 2022)	S 5 schools included showed an interest in the intervention; no direct measure of empathy
Effectiveness of an individual school-based intervention for children with aggressive behaviour: A randomized controlled trial (Stoltz et al., 2013)	S 5 no direct empathy measure
Evaluating the implementation of an emotional wellbeing programme for primary school children using participatory approaches (Clarke et al., 2014)	S 5 no direct measure of empathy
Improving upper elementary students' humane attitudes and prosocial behaviors through an in-class humane education program (Samuels et al., 2016)	S 5 no direct empathy measure, focuses mainly on attitudes (towards animals and the environment).

Study title and authors	S E Comments
Tackling acute cases of school bullying in the KiVa anti-bullying program: a comparison of two approaches (Garandean et al., 2014)	S 5 not direct measure of empathy; content similar to: School bullies' intention to change behavior following teacher interventions: Effects of empathy arousal, condemning of bullying, and blaming of the perpetrator
The Art of Living Together: Reducing Stereotyping and Prejudicial Attitudes Through the Arab-Jewish Class Exchange Program (CEP; Berger et al., 2015)	S 5 schools were selected by researchers; no empathy measure
The impact of a middle school program to reduce aggression, victimization, and sexual violence (Espelage et al., 2013)	S 5 no empathy measure
The implementation and effectiveness of the KiVa antibullying program in Finland (Salmivalli et al., 2013)	S 5 other KiVa studies included in the review (that derive from the same data) focussed stronger on empathy
Emotion training with students: An effectiveness study concerning the relation between subjective well-being, emotional awareness, and emotion expression (Nitkowski et al., 2017)	S 6 text unavailable
Evaluation of a social-emotional and character development program: Methods and outcomes (Lewis, 2012)	S 6 published in 2012
Social-emotional competence: Evaluation of a short-term training program for elementary school children (Hess, 2020)	S 6 text unavailable
A cluster randomized controlled trial of child-focused psychiatric consultation and a school systems-focused intervention to reduce aggression (Fonagy et al., 2009)	F 6 the study was conducted in 2009