HHS Public Access

Author manuscript

Dev Rev. Author manuscript; available in PMC 2016 September 01.

Published in final edited form as:

Dev Rev. 2015 September 1; 37: 41–65. doi:10.1016/j.dr.2015.05.001.

Children's emotion understanding: A meta-analysis of training studies

Manuel Sprung,
International Psychology & Psychotherapy Center, Vienna (Austria)

Hannah M. Münch,
Department of General Internal Medicine and Psychosomatics, University Hospital Heidelberg (Germany)

Paul L. Harris,
Graduate School of Education, Harvard University (USA)

Chad Ebesutani, and
Department of Psychology, Duksung Women's University (South Korea)

Stefan G. Hofmann
Department of Psychological and Brain Sciences, Boston University (USA)

Abstract

BACKGROUND—In the course of development, children show increased insight and understanding of emotions—both of their own emotions and those of others. However, little is known about the efficacy of training programs aimed at improving children's understanding of emotion.

OBJECTIVES—To conduct an effect size analysis of trainings aimed at three aspects of emotion understanding: external aspects (i.e., the recognition of emotional expressions, understanding external causes of emotion, understanding the influence of reminders on present emotions); mental aspects (i.e., understanding desire-based emotions, understanding belief-based emotions, understanding hidden emotions); and reflective aspects (i.e., understanding the regulation of an emotion, understanding mixed emotions, understanding moral emotions).

DATA SOURCES—A literature search was conducted using PubMed, PsycInfo, the Cochrane Library, and manual searches.

REVIEW METHODS—The search identified 19 studies or experiments including a total of 749 children with an average age of 86 months (*S.D.*=30.71) from seven different countries.

RESULTS—Emotion understanding training procedures are effective for improving external (Hedge's g = 0.62), mental (Hedge's g = 0.31), and reflective (Hedge's g = 0.64) aspects of emotion understanding. These effect sizes were robust and generally unrelated to the number and lengths of training sessions, length of the training period, year of publication, and sample type.

Correspondence concerning this article should be addressed to Stefan G. Hofmann, Ph.D., Department of Psychological and Brain Sciences, Boston University, 648 Beacon Street, 6th Fl, Boston, MA 02215, Tel: 617 353 9233, Fax: 617 353 0609, shofmann@bu.edu.

However, training setting and social setting moderated the effect of emotion understanding training on the understanding of external aspects of emotion. For the length of training session and social setting, we observed significant moderator effects of training on reflective aspects of emotion.

CONCLUSION—Emotion understanding training may be a promising tool for both preventive intervention and the psychotherapeutic process. However, more well-controlled studies are needed.

Keywords

Emotion Understanding; Training; Meta-analysis; Emotion comprehension; Emotional competency; Intervention; Training studies; Test of Emotion Comprehension; Theory of Mind

Emotional competence can be analyzed in terms of various domains (Saarni, 1999), including the ability to experience basic emotions such as joy, sadness, fear, and anger, the ability to express emotions, to recognize emotions (both those of the self and those of others), to control their expression and to regulate the subjective experience of emotion. Finally, competence also includes a cognitive understanding of emotions, including their nature, causes, consequences, and strategies for regulating them. These different domains of emotional competence are interconnected. In particular, improved emotion understanding has been linked to greater competence in various other domains, such as emotion recognition, control and regulation (Harris, 2008). Emotion understanding (EU) is also connected to the broader construct of Theory-of-Mind (i.e., knowledge and awareness of mental states—including desires, beliefs and emotions). With respect to several different aspects of emotion understanding, marked developmental differences (especially in preschool and school-age children) as well as individual differences have been reported (Pons, de Rosnay, Anderson & Cuisinier, 2010).

An empirically-derived model of emotion understanding by Pons, Harris and de Rosnay (2004; Pons & Harris, 2005) identified three hierarchically organized levels of emotion understanding (with three components per level). The first level – *external* – consists of the following components: recognizing emotional expressions (e.g., of sadness, anger, happiness, etc.); understanding external or situational causes of emotions (e.g., the death of a pet causes sadness), and understanding the impact of external reminders (e.g., understanding that re-encountering a situation with emotional significance in the past can re-activate the emotion). The second level – *mental* – consists of the following components: understanding the role of desires (e.g., understanding that different people like/hate/fear/… different entities); understanding the role of beliefs (e.g., recognizing that a person's belief about a situation determines his or her emotional reaction to it); and understanding hidden emotions (e.g., realizing that expressed and felt emotions can differ). The third level – *reflective* – includes the following components: understanding emotion regulation (e.g., understanding that different coping strategies have different effects); understanding mixed emotions (e.g., understanding that the same situation can make someone feel both excited and anxious); and understanding moral emotions (e.g., understanding that transgression elicits guilty feelings). Based on this model, Pons and Harris (2005; Pons, Harris & deRosnay, 2004) have

developed a comprehensive test of emotion understanding, the Test of Emotion Comprehension (TEC).

Children's understanding of emotion is an important aspect of cognitive development that has been linked with a variety of outcomes. Children and adolescents with good emotion understanding show better academic performance (e.g., Doudin, Martin & Albanese, 2001; Jones, Brown & Aber, 2011; Jones, Brown, Hoglund & Aber, 2010; Lecce, Caputi & Hughes, 2011; Pons, Harris & Doudin, 2002), and are more successful in their social interactions with peers and teachers (e.g., Bosacki & Astington, 1999; Cassidy, Parke, Butkovsky, & Braungart, 1992; McDowell, O'Neil, & Parke, 2000). Conversely, children with poor emotion understanding show poorer academic performance, are more likely to be rejected by their peers and teachers, and are at an increased risk of being expelled from regular classrooms (e.g., Doudin & Erkohen, 2000; Lafortune & Mongeau, 2002). In several studies, Rieffe and colleagues have linked children's improved emotional awareness to fewer symptoms of anxiety and depression (Rieffe & de Rooij, 2012; Rieffe, Oosterveld, Miers, Terwogt & Ly, 2008; Rieffe, Terwogt, Petrides, Cowan, Miers & Tolland, 2007). Impairments in emotion understanding have been reported in hearing-impaired children (e.g., Kouwenberg, Rieffe, Theunissen & Oosterveld, 2012; Rieffe, 2012), children with autism (e.g., Rieffe, Ketelaar, & Wiefferink, 2010; Rieffe, Terwogt & Kotronopoulou, 2007) and children with frequent somatic complaints (e.g., Jellesma, Rieffe, Terwogt, & Kneepkens, 2006; Rieffe, Terwogt, & Bosch, 2004; Rieffe et al., 2007). Maltreated children have also been found to be impaired in emotion understanding (e.g., Pears & Fisher, 2005) although this impairment may be moderated by intellectual functioning and language skills.

Studies to foster children's emotion understanding have mostly been conducted to reduce impairment or delay among children with various clinical conditions, such as autism (e.g., Begeer, Gevers, Clifford, Verhoeve, Kat, Hoddenbach & Boer, 2011; Ozonoff & Miller, 1995), hearing impairment (Dyck & Denver, 2003), disability (cognitive, speech and language or motor delay) (DeLuca, 2004), or anxiety (Fox et al., 2012) as well as socio-economic disadvantage (Smith, 2011). Several other training studies (mostly with typically developing children) have sought to investigate the potentially beneficial effects of improved emotion understanding on academic performance (e.g., Pons, Harris & Doudin, 2002). Some training studies have been conducted to examine the relationship between mental state language and emotion understanding (e.g., Grazzani & Ornaghi, 2011; Ornaghi, Brockmeir & Grazzani, 2011).

The methods and designs employed in these studies have varied widely, with respect to components of emotion that have been targeted as well as the nature of the outcome measures. In terms of training components, the majority of studies have focused on improving children's understanding of the *external* aspects of emotion (i.e., the recognition and differentiation of emotions; Bauminger, 2002, 2007a,b; Calabro, 2003; DeLuca, 2004; Doyle, 2001; Dyck & Denver, 2003; Hadwin, Baron-Cohen, Howlin, & Hill, 1996; Pons, Harris, & Doudain, 2002; Solomon, Goodlin-Jones, & Anders, 2004; Steerneman & Huskens, 1996), and/or the understanding of which emotions are triggered by which situations (Bauminger, 2002, 2007a, b; Bennett & Hiscock, 1993; DeLuca, 2004; Doyle, 2001; Dyck et al., 2003; Hadwin et al., 1996; Pons et al., 2002; Steerneman & Huskens,

1996; Tenenbaum, Alfieri, Brooks, & Dunne, 2008). Some studies have also focused on improving children's understanding of the *mental* aspects of emotion (e.g., understanding hidden emotions; Pons et al., 2002; Tenenbaum et al., 2008). Finally, several studies have focused on improving children's understanding of the *reflective* aspects of emotion (e.g., understanding mixed feelings; Bauminger, 2007ab; Bennet et al., 1993; Peng, Johnson, Pollock, Glasspool, & Harris, 1992; Tenenbaum et al., 2008), or understanding emotion regulation strategies (Bauminger, 2007a; Calabro, 2003; DeLuca, 2004; Dyck et al., 2003; Pons et al., 2002; Solomon et al., 2004; Steerneman et al., 1996).

Training procedures have included a variety of materials (e.g., stories, picture books, videos, games.) akin to the materials used in EU assessment tools, such as the TEC (Pons & Harris, 2005; Pons, Harris & deRosnay, 2004). Common interventions have included discussion of the different components of emotion understanding, with or without corrective feedback (e.g., Bauminger, 2002, 2007ab; DeLuca, 2004; Grazzani & Ornaghi, 2011; Peng et al., 1992; Tenenbaum et al,. 2008), prompted use of the imagination (e.g., Fox et al., 2012; Schonert-Reichl, Smith, Zaidman-Zait & Herztman, 2012; Steerneman & Huskens, 1996), modeling, and role-play (e.g., Calabro, 2003; Fox et al., 2012; Schonert-Reichl et al., 2012).

Primary outcome measures have also been quite diverse, although the TEC was used in several studies (Grazzani & Ornaghi, 2011; Grazzani, Brockmeier & Ornaghi, 2011; Pons et al., 2002; Tenenbaum et al., 2008). In line with the components targeted during training, outcome measures in most studies have assessed children's understanding of external aspects of emotions. However, some studies have also assessed the mental and reflective components of emotion understanding.

Despite the significant number of EU training studies, no quantitative review of their impact has yet been conducted. Given the potential social, clinical and academic benefits to be had with improved emotion understanding skills, it is important to establish the effectiveness of interventions aimed at improving children's emotion understanding, and to assess whether there are any specific training characteristics that are associated with more favorable outcomes. The aim of the present meta-analysis was to evaluate not only the overall efficacy of EU training but also the efficacy of trainings that target external, mental, or reflective aspects of emotion, as well as combinations of those aspects. The analysis also examines the influence of various different training characteristics: the quality of the study, the study year, country, the number of training sessions, the length of each training session, the length of the training period, the length of the delay between pre- and posttest, training setting, sample type and specific training composition.

Methods

We followed the PRISMA guidelines (standing for the Preferred Reporting Items for Systematic Reviews and Meta-Analyses) to aid in transparent reporting of systematic reviews and meta-analyses (Liberati et al, 2009). Following PRISMA guidelines, Table 1 presents a checklist of items to include when reporting a systematic review or meta-analysis.

Search

Studies were identified by searching PsycINFO, PubMed, and the Cochrane Library. Methods of analysis and inclusion criteria were specified in advance. We conducted searches for studies published between the first available year for any given database and April 30, 2012, using the following search term combinations: ("*theory of mind*" OR "*emotion understanding*" AND *children* AND (*training* OR *teaching* OR *intervention* OR *program* OR *therapy*). Additionally, an extensive manual review of reference lists of relevant studies and review articles extracted from the database searches was conducted. Articles determined to be related to the topic of emotion understanding or theory-of mind training were selected for further examination. To identify unpublished studies, the authors of studies selected for inclusion in the meta-analysis were contacted for unpublished studies.

Selection

Studies were included if the training program: a) aimed to improve the understanding of emotion ; b) included one or more pre- and post-test measure of emotion understanding; c) included youth samples (up through an average age of 18 years); d) provided sufficient data to perform effect size analyses (i.e., means and standard deviations, *F*-values, or change scores). If all criteria except for criterion d) were met, then the authors were contacted for additional data.

Data Abstraction

We extracted data from tasks designed to measure children's understanding of emotions. Each task was classified according to the three hierarchical levels of emotion understanding described earlier (i.e., external, mental, reflective) by two independent raters (Sprung and Münch). We performed an inter-rater agreement analysis using kappa, and any disagreements were resolved by discussion. Additionally, numerical and categorical data were extracted for the purpose of conducting moderator analyses. If a study did not report one or more variables targeted in the moderator analyses, these data were requested from the authors. Moreover, the training components and procedures in each study were also classified according to the three hierarchical levels of emotion understanding ability (external, mental, reflective) by two independent raters (first and second author). Again an inter-rater agreement analysis using kappa statistics was calculated, and any disagreements were resolved by discussion. All effect size analyses were conducted based on the three levels of emotion understanding.

Study Characteristics

To explore variability in study results, we examined whether effect size estimates varied depending on methodological quality using the Effective Public Health Practice Project (EPHPP) criteria Quality Assessment Tool for Quantitative Studies (Thomas, Ciliska, Dobbins, & Micucci, 2004). The Global Rating of the EPHPP contains the following Rating Components: (A) Selection Bias, (B) Study Design, (C) Confounders, (D) Blinding, (E) Data Collection Methods (only the outcome specific measurements were considered), and (F) Withdrawals and dropouts. Possible quality ratings for each study were strong, moderate, or weak.

Quantitative Data Synthesis

We calculated effect sizes for continuous measures of emotion understanding using Hedge's *g* (a variation of Cohen's *d* that corrects for biases due to small sample size [Hedges & Olkin, 1985]) and its 95% confidence interval. The within-group effect sizes were based on the differences in gains from pre- to post within the training group. In addition, a between group effect size analysis was conducted using all studies that included a control group.

The magnitude of Hedge's *g* corresponds to Cohen's recommendations for interpreting effect sizes as small (0.2), medium (0.5), and large (0.8). We calculated effect size estimates using a random-effects model, which assumes that the studies in the meta-analysis do not share a common true effect and also assumes the existence of significant heterogeneity between the studies (Hedges & Vevea, 1998; Moses, Mosteller, & Buehler, 2002). Consequently, the weighting of a single study does not depend on the study size but rather on the effect size distribution within the study (i.e., exceptionally large studies do not have a dominant impact and studies with a small sample size do not get ignored in the overall effect sizes estimation). When several measures of external, mental or reflective emotion understanding were used within the same study, average Hedge's *g* effect sizes were calculated for each of them. Because the correlation between pre- and post-treatment measures could not be determined from the study reports, we followed the recommendation of Rosenthal (1993) and assumed a conservative estimate of $r = 0.7$.

We conducted separate analyses of overall Hedge's g effect sizes for studies that included measures of the understanding of external, mental, and reflective levels of understanding. We pooled effect sizes across studies to obtain a summary statistic for each level of understanding.

Risk of bias

We assessed the risk of publication bias by calculating the *fail-safe N* to estimate the number of unpublished studies with an effect size of zero needed to nullify the significant effect (Rosenthal, 1991; Rosenthal & Rubin, 1988). The effect size can be considered robust if the required number of studies to reduce the overall effect size to a nonsignificant level exceeds $5K + 10$, where K is the number of studies included in the analysis (Rosenthal, 1991). Additionally, we constructed a funnel plot to give a visual idea of potential publication bias. Following convention, the precision of the studies are plotted along the y-axis with the more precise studies (e.g., larger N) at the top and the less precise studies (e.g., smaller N) at the bottom of the graph. To determine the precision, we divided 1 by standard error for each study. Studies were plotted along the x-axis depending on their effect sizes estimates. A symmetrical distribution of the studies effect size estimates around the mean effect size is assumed with more variability at the bottom and the more precise studies at the top, closer to the estimated mean effect size. It is assumed that if there are more studies at the right (high effect sizes) than at the left (small or nul effects) side of the graph, the presence of publication bias is suggested. Additionally, we used Duval and Tweedie's Trim and Fill method (Duval & Tweedie, 2000ab), which uses an iterative process in which the most extreme small studies are removed and a new effect size is computed until the funnel plot is symmetrical. Subsequently, we examined the number of studies that might be missing on the

left side of the funnel plot and obtained the effect size including the estimated missing studies. We completed all analyses manually and also using the software program Comprehensive Meta-Analysis, Version 2 (Borenstein, Hedges, Higgins, & Rosenstein, 2005).

Moderator Analyses

We conducted meta-regression analyses to examine whether the effect sizes varied as a function of the following continuous variables: number of training sessions, length of each training session (in hours), total length of the training period (in months), and post-test delay (i.e., number of weeks between the last day of training and the post-test). Meta-regression analyses, however, were not calculated for age and gender distribution[1]. We also calculated *Q* statistics to determine whether the effect sizes varied as a function of the following categorical moderators: environmental setting (classroom, area in school, lab), social setting (children trained in a group vs. children trained individually), sample type (typically developing children and children with a clinical condition), country (America, Europe, others), and the Global Rating based on the EPHPP assessment tool for quantitative studies (1= Strong, 2= Moderate, 3= Weak).

First, we grouped the studies according to the categorical variable being assessed. Then, we calculated $Q_{between}$ using random effect weights with pooled estimates of τ^2 to determine whether heterogeneity existed between the groups. If the $Q_{between}$ value was significant (i.e., if heterogeneity existed between the groups), the grouping variable was considered a moderator.

Finally, the different training programs were also grouped based on the particular level of training (external, mental, reflective) that was targeted or on the specific combination of levels that were targeted (external+mental, external+reflective, mental+reflective, external +mental+reflective).

Results

Study Selection

Figure 1 illustrates our study selection process. Of the 485 articles initially identified, 17 were included in the meta-analysis. The majority of studies either did not contain any type of training (n = 193) or were review articles (n = 141). Training studies that measured or trained only theory of mind (n = 29) were excluded because the focus of this study was EU and because the outcome measures and training procedures for these constructs were not comparable. The remaining 17 articles included 19 studies or experiments. In studies with multiple training or control conditions, we compared the most complete training condition (i.e., the training condition that the authors hypothesized to be the most effective) to the most active, but neutral control condition (i.e., the condition that controlled for nonspecific

[1]The interpretation of meta-regression effects based on study or group averages (e.g. sample mean age or gender distribution) is problematic, because the relationship between participants' outcome—in this case the ability to understand emotion—and age or gender distribution is not necessarily the same across different studies as compared to within an individual study (Thompson & Higgins, 2002).

aspects of training, such as the number and length of session or the type of training material but did not include any emotional content). In cases where the author(s) included multiple groups in order to analyze the contribution of different training components (e.g., self-explanation vs. experimenter-explanation) or different age groups, we treated them as subgroups within the study and conducted separate comparisons.

The inter-rater agreement of the task classification based on the three hierarchical levels of emotion understanding (external, mental, reflective) was high with a kappa of 0.86 ($p < 0.001$), CI 95% (0.70; 1.01). Any disagreements were resolved by discussion. There were no inter-rater disagreements between the two independent raters concerning training classification.

Study Characteristics

Table 2 provides a description of the characteristics of the studies included in the present analysis. Studies were conducted in seven different countries (Australia, Canada, Israel, Italy, Netherlands, UK, and USA) and included a total of 1,308 children. The age of children ranged from 35 to 207 months, with an average mean age of 86 months (*SD*= 30.71; median=103). The number of training sessions ranged from 1–60 sessions (mean=16.58, *SD*=16.23; median=14.50), with individual session length between 15 and 180 minutes (mean=59.25, *SD*=47.88, median=180), and overall training periods spanning from 1 to 280 days (mean=81.52, *SD*=86.35, median=98.5). The Global Rating of the EPHPP Quality Assessment Tool for Quantitative Studies indicated a full range of studies rated as weak, moderate or strong in quality. Details of the Global Rating and the single Rating Components (A) to (F) are outlined in Table 3.

Quantitative Data Synthesis

In the following effect size estimates for external, mental and reflective emotion understanding are presented. First, the within group effect size estimates—based on the gain from pre-to post-test—for all studies and subgroups in studies are presented below. Figure 2 presents the overall effect sizes for external, mental and reflective understanding. Additional analyses examining the risk of publication bias and moderators—based on these within group effect sizes—are presented further below. Further, the average effect size estimates from the additional controlled (between-group) effect size analysis—only with studies that included a control group—are presented below. Finally, we present the average effect size estimates from an additional within group effect size analyses—including only studies that were used for calculating between group effects sizes—so that direct comparisons can be made.

Understanding external aspects of emotion—We included a total of 16 studies to analyze the within group effect size of training to improve the understanding of external aspects of emotion. Some studies included different subgroups and so we conducted a total of 25 effect size calculations. This led us to an average medium within-study effect size estimate (Hedge's g) of 0.62 (SE = 0.08; 95% CI [0.47; 0.78]; $z = 8.01^{***}$, $p < 0.001$). The controlled (between group) effect size analysis included 10 of the 25 studies and was based on a total of 15 comparisons with a pooled medium-to-large between-group effect size

estimate (Hedge's g = 0.60, SE = 0.09, 95% CI [0.41; 0.78], z = 6.35***, p < 0.001). The within group effect size analysis—only with studies that were used for the between group effect sizes analysis—revealed an average effect size of Hedge's g = 0.65 (SE = 0.10; 95% CI [0.45; 0.86]; z = 6.39***; *p* < 0.001).

Understanding mental aspects of emotion—We included a total of 5 studies (11 comparisons) that measured the understanding of mental aspects of emotion in the within group effect size analysis—resulting in a small average effect size estimate (Hedge's g = 0.31, SE = 0.10, 95% CI [0.11; 0.50], z = 3.05**, *p* < 0.01). Five of these studies (9 comparisons) also used control groups. The results revealed a small-to-medium *between-group* effect size of Hedge's g = 0.55 (SE = 0.12, 95% CI [0.32; 0.78], z = 4.73***, *p* < 0.001). The pooled within-group effect size of studies (comparisons)—that were used for the between group effect size analysis—was Hedge's g = 0.36 (SE = 0.12; 95% CI [0.13; 0.60]; z = 3.02**; *p* < 0.01).

Understanding reflective aspects of emotion—Our within group effect size analysis of training to improve the understanding of reflective aspects of emotion was based on 13 studies (21 comparisons). It revealed a medium effect with a tendency to a large effect (Hedge's g =0.64, SE = 0.07, 95% CI [0.50;0.79], z = 8.87***, *p* < 0.001). There were 7 studies (14 comparisons) that included a control group, used for the between group effect size, revealing an average-to-large controlled pooled effect size (Hedge's g = 0.68, SE = 0.12, 95% CI [0.45; 0.91], z = 5,86***, *p* < 0.001). The average within group effect sizes of studies (comparisons)—that were used for the between group effect size analysis—was Hedge's g = 0.63 (SE = 0.09; 95% CI [0.45; 0.82]; z = 6.78***; *p* < 0.001). In summary, for all three levels of emotion understanding (external, mental and reflective), the overall effect size estimate revealed a significant within training group gain from pre-to post. As suggested by the controlled effect sizes, emotion understanding training seem to improve external and reflective emotion understanding as compared to the control groups. For mental emotion understanding, surprisingly, the average controlled effect size was higher than the within effect size, possibly because control groups worsened from pre to post test (although it should be noted that this effect is based on only 5 studies).

Publication bias

For each of the three levels of emotion understanding, a fail-safe N was calculated to estimate the robustness of the effect size analyses by estimating the number of unpublished studies with zero effect that would be necessary to nullify the observed effect. For the understanding of external aspects of emotion, the fail-safe N was 1475, which is far greater than the critical value of 135, as calculated following Rosenthal (1991), and may therefore be considered robust. The fail-safe N for the understanding of mental aspects of emotion was 68, which also exceeds the critical value of 65 and can also be considered as robust. The fail-safe N for understanding the reflective aspects of emotion was 1052, which exceeds the critical value of 115, indicating a robust result.

An additional analysis was conducted using Duval and Tweedie's Trim and Fill method to address the question of the impact of publication bias on the observed effect size. Figure 3

presents the funnel plots for the understanding of external, mental and reflective aspects of emotion including the observed studies and the missing studies imputed via the Trim and Fill method for symmetrical distribution around the average effect size. Using the Trim and Fill method, the number of missing studies was n = 10 studies for the analysis of the ability to understand external aspects. Assuming a random-effects model, the new imputed mean effect size was Hedge's g = 0.34 (95% CI [0.17; 0.51]). For the analysis of the ability to understand mental aspects, n = 2 studies were missing and the new imputed effect size was Hedge's g = 0.22 (95% CI [0.02; 0.42]). For the effect size estimation of the understanding of reflective aspects n = 5 studies were assumed to be missing. The estimated effect size after imputing these studies was Hedge's g = 0.52 (95% CI [0.37; 0.68]).

The fail-safe N indicates that for the two levels of emotion understanding aspects external and reflective, the number of missing studies needed to nullify the results clearly exceeds the cutoff and can be considered robust. Similarly, the fail-safe N for mental aspects of emotion understanding exceeded the critical value, but by only 3 additional studies. Therefore, this effect size should be interpreted with caution. The results of the Trim and Fill method suggest that for external emotion understanding 10 unpublished studies are missing; imputing them decreased the observed medium to large effect to a small to medium effect. Imputing the 2 unpublished studies, which are assumed to be missing in the analyses for mental emotion understanding lead to no change in the initial observed small to medium effect of the training. For reflective emotion understanding, an imputation of the 5 missing studies led to a small to medium effect size estimate.

Moderator Analyses

Table 4 outlines the results of the meta-regression analyses conducted to assess the moderating effects of continuous variables. There were no significant moderator effects for any of the continuous variables on the effect of training on the understanding of external and mental aspects of emotion. However, the effect size for training on the understanding of reflective aspects of emotion was moderated by the length of session ($\beta = 0.19$, SE = 0.08, 95% CI [0.04; 0.33]; $p < 0.05$), year of publication ($\beta = -0.03$, SE = 0.01, 95% CI [-0.05; −0.01]; $p < 0.01$), and the delay of data assessment after the last session of training ($\beta = -0.19$, SE = 0.06, 95% CI [-0.32; −0.07]; $p < 0.01$). Accordingly, emotion understanding training improved the understanding of reflective aspects of emotion and emotion understanding training was more effective when sessions were longer in duration. Additionally, the effect sizes were larger in older studies and in studies with shorter post-test delay.

Environmental and Social setting—The influence of specific environmental settings on the effectiveness of training was examined. The groups of different environmental settings revealed a significant degree of heterogeneity for external aspects of emotion understanding ($Q_{between} = 9.84$; df = 2; $p < 0.01$) but not for mental ($p = 0.70$) or reflective ($p = 0.83$) aspects (see Table 5, "Environmental Setting" section, for summary of results regarding moderating effects of environmental setting). We further examined whether the effectiveness of training was moderated by individual versus group settings. The degree of heterogeneity between the groups was not significant for mental aspects of emotion

understanding ($p = 0.42$). The degree of heterogeneity for external and reflective aspects, however, was significant ($Q_{between} = 3.87$; df = 1; $p < 0.05$ and $Q_{between} = 15.68$; df = 1; $p < 0.001$). The effects of training on *external* aspects of emotion were larger when the training was conducted in group settings. The effects of training on *reflective* aspects of emotion, however, were larger when the training was conducted in individual one-to-one settings. Table 5 ("Social Setting" section) summarizes results regarding moderating effects of social setting. Both the environment where the training took place as well as the social setting had a moderating effect on the effectiveness of training to improve external aspects of emotion understanding. The effectiveness of training to improve reflective aspects of emotion understanding was moderated by social setting only.

Sample Type—To evaluate the effects of training on different sample types, the samples were first classified as either typically developing children or as children with some type of a special condition (i.e. clinical psychiatric condition). Subsequently, an additional comparison within the different sub-types of clinical condition was calculated. More specifically, we grouped studies that included children with autism and compared them to studies of children with a clinical condition other than autism. Because studies of the mental aspects of emotion included only one sub-group—namely children with autism—no further 'within clinical group' analyses were conducted.

The degree of heterogeneity between the groups was not significant for typically developing children versus children with a clinical condition for any of the three aspects of emotion understanding. When comparing children with autism with the sub-group of children with any other clinical condition, the degree of heterogeneity between the groups was also not significant; the same was true for both external and reflective aspects of emotion understanding. However, there was a trend towards significance for a moderating effect of the type of clinical condition in improving reflective emotion understanding abilities through training ($p = 0.07$). Thus, children with autism seemed to benefit more from training than children with other disabilities or psychiatric conditions. However, the effect size for children with other disabilities was based on only one training study (see Table 5, "Sample Type" section, for summary of results regarding moderating effects of sample type).

Country—To assess the impact of country on the effectiveness of training, studies were divided into the following three groups: North America, Europe, and others (including Australia and Israel). For external aspects of emotion understanding the degree of heterogeneity was significant ($Q_{between} = 12.68$; df = 2; $p < 0.01$), suggesting that children participating in European training programs benefitted more than children participating in North American and other programs. For mental and reflective aspects, no moderating effects were detected. However, for mental aspects of emotion understanding, there was a trend towards significance ($p = 0.06$). Table 5 ("Country" section) summarizes the analyses conducted to measure the results regarding possible moderating effects of country.

Risk of Bias—To further examine the influence of study quality, the global rating of the EPHPP Quality Assessment Tool for Quantitative Studies was applied as a categorical moderator containing the following three categories: "Strong", "Moderate" and "Weak." There was no significant degree of heterogeneity in the effectiveness of training for any of

the three aspects—external, mental and reflective emotion understanding (see Table 5, "EPHPP Global Rating" section, for a summary of results regarding moderating effects of study quality).

Analysis of training composition

The effect of training specific combination of levels was also assessed (see Table 6). The degree of heterogeneity between the groups was not significant for external, mental or reflective aspects of emotion. However, there was a trend towards significance for external aspects ($p = 0.07$), suggesting that there was a moderating effect of the specific combination of training components, with higher effect sizes for training programs focused only on external aspects of emotion and those focused on a combination of mental and reflective aspects of emotion.

Discussion

The results of the present meta-analysis suggest that EU training procedures may be effective for improving children's understanding of emotion. The trainings were associated with medium to large average effects size estimates (Hedge's *g*) for external (g = 0.62) and reflective (g=0.64) aspects of emotion, and a small to medium average effect size estimate for mental aspects (g=0.31) of emotion. For training of mental aspects of emotion, however, the fail-safe N exceeded the critical value by only 3 additional studies. When examining the publication bias on effect sizes—using Duval and Tweedie's Trim and Fill method—suggest that the effect of training on mental and reflective aspects of emotion may be robust. The effect of training on external aspects of emotion, however, may be less robust. Imputing 10 unpublished studies—as estimated by the Trim and Fill method—decreased the observed medium to large effect to a small to medium average effect size. Therefore, these effect sizes should be interpreted with caution.

The preliminary moderator analyses suggest that the overall effect size for the training of external aspects of emotion was moderated by training setting (environmental and social setting) and country of the study. None of the other training characteristics (number or length of training session, length of the training period), sample characteristics (typically developing or clinical), or study characteristics (year of publication, post-test delay, study quality) moderated the overall effect size for external aspects. There was, however, a trend towards a significant moderating effect of training composition, suggesting that training programs focused only on external aspects of emotion were most effective. The pooled effect size of training programs focused on a combination of mental and reflective aspects of emotion was also higher than training programs focused on other combinations, but this effect was only based on the results from two studies. One interpretation for the greater benefits of trainings focused on external aspects of emotion is that the integration of more complex aspects of emotion may be confusing for children who are still learning about the more basic aspects of emotion. These findings suggest that it is possible to improve children's recognition of emotions, their understanding of external causes of emotion, and their understanding of the impact of reminders even with a few short training sessions over a brief training period. Moreover, results suggested that training is effective regardless of

children's clinical status and its benefits were enduring. The benefits of training on external aspects of emotion, however, were greater if the training was conducted in a group setting with other children and in a European country—but not in the regular classroom setting. The effects of training on external aspects of emotion may also be greater if the training is focused solely on external aspects of emotion.

The overall effect size for the training of mental aspects of emotion (understanding desire-based emotions, understanding belief-based emotions, understanding hidden emotions) was not moderated by any of the training, sample or study characteristics. There was, however, a trend towards a significant moderator effect of the country in which the training was conducted—specifically, children participating in European training programs appeared to benefit more from training than children participating in North American or other training programs. By implication, training directed at the mental aspect of emotion is effective across a range of procedures, settings and groups, but its benefits tended to be the greatest in training programs conducted in European countries.

The overall effect size for the training of reflective aspects of emotion (understanding emotion regulation, mixed emotions, and moral emotions) was moderated by the length of the training session, study year (i.e., year of publication), the delay of data assessment after the last session of training (i.e., post-test delay), and the social training setting. None of the other training, sample or study characteristics moderated the overall effect size for reflective aspects of emotion. There was, however, a trend towards a significant moderator effect for sample characteristics—specifically, children with a clinical psychiatric condition or disability seemed to benefit more than typically developing children (see Table 5, "Sample Type" section). By implication, training directed at the reflective aspect of emotion appears to be effective across a range of procedures, settings and groups, but its effectiveness may be improved by increased length of training sessions and conducting training in individual (rather than in group) settings. One interpretation of this finding is that because an understanding of the reflective aspects of emotion (i.e., emotion regulation, mixed emotions, moral emotions) calls for more advanced thinking—both from a conceptual and an empirical perspective—training is more effective when conducted via one-to-one dialogue. Further, the effects of training on reflective aspects of emotion were less stable (i.e. the effects decreased with longer post-test delays). The overall effect size of training directed at reflective aspects of emotion was also moderated by the year in which the study was published. The effect sizes of older studies were generally higher than those of more recent studies. Although this could be due to the increasing rigor of more recent studies, the study quality—as judged by EPHPP Global Rating—did not moderate the effect sizes. Moreover, the effects of training of reflective aspects of emotion may be especially beneficial for children with autism or another psychiatric condition. However, this effect was not statistically significant. Future studies are needed to explore this issue further.

The specific training composition—combination of levels or aspects of emotion understanding being trained—did not have a moderating effect on any of the three aspects of emotion. Training that targeted fewer and less complex levels of emotion was not more or less effective in improving the understanding of external, mental or reflective aspects of emotion— suggesting that the addition of a greater number and more complex aspects of

emotion to trainings may not result in greater gains. There was, however, a trend towards a significant moderating effect of the specific training composition on external aspects of emotion understanding. By implication, training that primarily focuses on less complex aspects (i.e., external aspects) of emotion can also impact more advanced levels of understanding (i.e., reflective aspects) and vice versa—although for training of external aspect of emotion it might be more effective to just focus on external aspects of emotion.

Although it would be important to also investigate any moderating effects of age or gender distribution, interpretation of meta-regression effects based on study or group averages is problematic (see footnote 1; Thompson & Higgins, 2002). Therefore, we recommend that future EU training studies define specific age groups for their respective study sample (in addition to reporting the mean age, standard deviation and range of the overall sample). This would allow for a better assessment of the influence of age on training effectiveness.

This meta-analytic review has a number of limitations. First and foremost, the number of included studies is relatively small, especially for conducting moderator analyses. Moreover, very few studies included a control group. Because the estimates of between group effect sizes were based on a very small number of studies, the results need to be interpreted with caution. Another limitation is that recent studies of the training of emotion understanding show considerable diversity both in terms of the aspects of emotion understanding that are targeted during training, the procedures used, as well as the outcome measures employed. Although several studies used the same outcome measure (i.e., the TEC), the specific training protocols were largely idiosyncratic. In future studies, support for the effectiveness of training could be strengthened if both the outcome measures and the training protocol were aligned with an established model of emotion understanding (e.g., the levels and components of the TEC). Another limitation of the existing evidence is the paucity of studies aimed at improving the understanding of mental aspects of emotion. Future training studies should aim to include this aspect of emotion in their training protocols.

The present meta-analysis focused on the training of emotion understanding and it was, therefore, based on the relatively narrow search term “emotion understanding.” Alternatively, we could have used broader search terms, such as “emotion or affect and understanding or knowledge or recognition or identification.” We decided, however, to limit the present metaanalysis to the training of “emotion understanding” in order to identify a homogenous sample of studies. A meta-analysis using broader search terms will likely result in a much more heterogeneous sample of studies—with highly diverse training procedures and outcome measures—limiting comparability across studies.

Given that a good understanding of emotion is linked to various beneficial outcomes (e.g., better social skills, better academic performance, as well as fewer psychological problems), the results of the present meta-analysis underscore the possibility and feasibility of training in this domain. This suggests the potential use of training for emotion understanding in preventive interventions. Indeed, several existing preventive intervention programs (Bierman, Coie, Dodge, Greenberg, Lochman, McMahon & Pinderhughes, 2010; Jones, et al., 2011; 2010) already include elements that are akin to the training procedures in the training of emotion understanding. Several of the training studies included in the present

meta-analysis were part of a more extensive clinical intervention program (i.e., cognitive behavioral therapy or rational emotive therapy). The ability to report and reflect on one's thoughts and feelings is also fundamental for the identification and treatment of many psychological problems (e.g., internalizing symptoms). Thus, training directed at emotion understanding may be helpful in the psychotherapeutic process, especially with young children and impaired individuals.

In sum, the study shows that children may benefit from emotion training and the results point to length of the training session and the social setting as potential moderators of this effect. However, given the number of studies, it remains unknown whether the interventions will reliably influence group means. More well-controlled studies are clearly needed.

Acknowledgments

This research was supported in part by a grant from the Austrian Science Foundation (i.e., Erwin-Schrödinger Fellowship Abroad to Manuel Sprung, Project number: J2853-G16).

We would like to thank the following authors for providing additional data, which were not included in the original published articles: Sander Begeer, Heather Smith, Harriet Tenenbaum, Francisco Pons, and Ilaria Grazzani Gavazzi.

Dr. Hofmann receives support from NIH/NCCIH (R01AT007257), NIH/NIMH (R01MH099021, R34MH099311, R34MH086668, R21MH102646, R21MH101567, K23MH100259), and the Department of the Army for work unrelated to the studies reported in this article. He receives compensation for his work as an advisor from the Palo Alto Health Sciences and Otsuka America Pharmaceutical, Inc., and for his work as a Subject Matter Expert from John Wiley & Sons, Inc. and SilverCloud Health, Inc. He also receives royalties and payments for his editorial work from various publishers.

References

References marked with an asterisk indicate studies included in the meta-analysis.

*. Bauminger N. The Facilitation of Social-Emotional Understanding and Social Interaction in High-Functioning Children with Autism: Intervention Outcomes. Journal of Autism and Developmental Disorders. 2002; 32(4):283–298. [PubMed: 12199133]

*. Bauminger N. Brief Report: Group Social-Multimodal Intervention for HFASD. Journal of Autism and Developmental Disorders. 2007a; 37(8):1605–1615. [PubMed: 17072752]

*. Bauminger N. Brief Report: Individual Social-Multimodal Intervention for HFASD. Journal of Autism and Developmental Disorders. 2007b; 37(8):1593–1605. [PubMed: 17072753]

*. Begeer S, Gevers C, Clifford P, Verhoeve M, Kat K, Hoddenbach E, Boer F. Theory of mind training in children with autism: A randomized controlled trial. Journal of Autism and Developmental Disorders. 2011; 41(8):997–1006. [PubMed: 20976617]

Bennett M, Hiscock J. Children's understanding of conflicting emotions: A training study. The Journal of Genetic Psychology. 1993; 154(4):515–524. [PubMed: 8176393]

Borenstein, M.; Hedges, L.; Higgins, J.; Rothstein, H. Comprehensive meta-analysis, version 2. Englewood, NJ: Biostat Inc; 2005.

Bosacki S, Astington JW. Theory of mind in preadolescence: Relations between social understanding and social competence. Social Development. 1999; 8(2):237–255.

Calabro E. Rational emotive behavior play therapy vs. client-centered therapy. Dissertations Abstracts International: Section B: The Science and Engineering. 2003; 64(02):957.

Cassidy J, Parke RD, Butkovsky L, Braungart JM. Family-peer connections: The roles of emotional expressiveness within the family and children's understanding of emotions. Child Development. 1992; 63(3):603–618. [PubMed: 1600825]

*. DeLuca CM. Rational-Emotive Education (REE): Consultation vs. intervention with special education preschoolers. Dissertations Abstracts International: Section B: The Science and Engineering. 2003; 64(10):5212.

Doudin, P-A.; Erkohen, M. Violence à l'école: Fatalité ou défi?. Bruxelles: De Boeck; 2000.

Doudin, P-A.; Martin, D.; Albanese, O. Metacognition et education: theorie et practique. Berne: Peter Lang; 2001.

Doyle MM. An interplay based social skills group for children with Asperger's Syndrome. Dissertations Abstracts International: Section A: The Humanities and Social Sciences. 2001; 62(04):1583.

Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics. 2000a; 56(2):455–463. [PubMed: 10877304]

Duval S, Tweedie R. A nonparametric "Trim and Fill" method of accounting for publication bias in meta-analysis. Journal of the American Statistical Association. 2000b; 95(449):89–98.

*. Dyck DJ, Denver E. Can the emotion recognition ability of deaf children be enhanced? A pilots Study. Journal of Deaf Studies and Deaf Education. 2003; 8(3):348–356. [PubMed: 15448058]

*. Fox JK, Warner CM, Lerner AB, Ludwig K, Ryan JL, Colognori D, Lucas CP, Brotman LM. Preventive Intervention for Anxious Preschoolers and Their Parents: Strengthening Early Emotional Development. Child Psychiatry and Human Development. 2012; 43(4):544–559. [PubMed: 22331442]

*. Grazzani GI, Ornaghi V. Emotional state talk and emotion understanding: a training study with preschool children. Journal of Child Language. 2011; 38(5):1124–1139. [PubMed: 21426620]

*. Hadwin J, Baron-Cohen S, Howlin P, Hill K. Can we teach children with autism to understand emotions, belief, or pretence? Development and Psychopathology. 1996; 8(2):345–365.

Harris, PL. Children's understanding of emotions. In: Lewis, M.; Haviland-Jones, J.; Barrett, L. Feldman, editors. Handbook of emotions. 3rd Edition. New York: The Guilford Press; 2008. p. 320-331.

Hedges, LV.; Olkin, I. Statistical methods for meta-analysis. New York, NY: Academic Press; 1985.

Hedges LV, Vevea JL. Fixed- and random-effects models in meta-analysis. Psychological Methods. 1998; 3(4):486–504.

Jellesma FC, Rieffe C, Terwogt MM, Kneepkens CMF. Somatic complaints and health care use in children: Mood, emotion awareness and sense of coherence. Social Science and Medicine. 2006; 63(10):2640–2648. [PubMed: 16916567]

Jones SM, Brown JL, Aber J. Two-year impacts of a universal school-based social-emotional and literacy intervention: An experiment in translational developmental research. Child Development. 2011; 82(2):533–554. [PubMed: 21410922]

Jones SM, Brown JL, Hoglund WL, Aber J. A school-randomized clinical trial of an integrated social-emotional learning and literacy intervention: Impacts after 1 school year. Journal of Consulting and Clinical Psychology. 2010; 78(6):829–842. [PubMed: 21114343]

Kouwenberg M, Rieffe C, Theunissen SC, Oosterveld P. Pathways underlying somatic complaints in children and adolescents who are deaf or hard of hearing. Journal of Deaf Studies and Deaf Education. 2012; 17(3):319–332. [PubMed: 22193292]

Lafortune, L.; Mongeau, P. L'affectivité dans l'apprentissage. Sainte-Foy: Presses de l' Université du Québec; 2002.

Lecce S, Caputi M, Hughes C. Does sensitivity to criticism mediate the relationship between theory of mind and academic achievement? Journal of Experimental Child Psychology. 2011; 110(3):313–331. [PubMed: 21624614]

Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JPA, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. British Medical Journal. 2009; 339:b2700. [PubMed: 19622552]

McDowell DJ, O'Neil R, Parke RD. Display rule application in a disappointing situation and children's emotional reactivity: Relations with social competence. Merrill-Palmer Quarterly: Journal of Developmental Psychology. 2000; 46(2):306–324.

Moses LE, Mosteller F, Buehler JH. Comparing results of large clinical trials to those of meta-analyses. Statistics in Medicine. 2002; 21(6):793–800. [PubMed: 11870817]

*. Ornaghi V, Brockmeier J, Gavazzi IG. The role of language games in children's understanding of mental states: A training study. Journal of Cognition and Development. 2011; 12(2):239–259.

Ozonoff S, Miller JN. Teaching theory of mind: A new approach to social skills training for individuals with autism. Journal of Autism and Developmental Disorders. 1995; 25(4):415–433. [PubMed: 7592252]

*. Peng M, Johnson C, Pollock J, Glasspool R, Harris P. Training young children to acknolwedge mixed emotions. Cognition and Emotion. 1992; 6(5):387–401.

Peras KC, Fisher PA. Emotion understanding and theory of mind among maltreated children in foster care: Evidence of deficits. Development and Psychopathology. 2005; 17:47–65. [PubMed: 15971759]

Pons F, Harris PL. Longitudinal change and longitudinal stability of individual differences in children's emotion understanding. Cognition and Emotion. 2005; 19(8):1158–1174.

Pons, F.; de Rosnay, M.; Andersen, BG.; Cuisinier, F. Emotional competences: Development and intervention. In: Pons, F.; de Rosnay, M.; Doudin, PA., editors. Emotions in research and practice. Aalborg Universitetsforlag; 2010. p. 203-238.

Pons F, Harris PL, de Rosnay. Emotion comprehension between 3 and 11 years: Developmental periods and hierarchical organization. European Journal of Developmental Psychology. 2004; 1(2):127–152.

*. Pons F, Harris PL, Doudin P-A. Teaching emotion understanding. European Journal of Psychology of Education. 2002; 17(3):293–304.

Rieffe C. Awareness and regulation of emotions in deaf children. British Journal of Developmental Psychology. 2012; 30(4):477–603. [PubMed: 23039328]

Rieffe C, de Rooij M. The longitudinal relationship between emotion awareness and internalising symptoms during late childhood. European Child and Adolescent Psychiatry. 2012; 21(3):349–356. [PubMed: 22466448]

Rieffe C, Ketelaar L, Wiefferink CH. Assessing empathy in young children: Construction and validation of an empathy questionnaire (EmQue). Personality and Individual Differences. 2010; 49(5):362–367.

Rieffe C, Oosterveld P, Miers AC, Terwogt MM, Ly V. Emotion awareness and internalising symptoms in children and adolescents: The Emotion Awareness Questionnaire revised. Personality and Individual Differences. 2008; 45(8):756–761.

Rieffe C, Terwogt MM, Bosch JD. Emotion understanding in children with frequent somatic complaints. European Journal of Developmental Psychology. 2004; 1(1):31–47.

Rieffe C, Terwogt MM, Kotronopoulou K. Awareness of single and multiple emotions in high-functioning children with autism. Journal of Autism and Developmental Disorders. 2007; 37(3): 455–465. [PubMed: 16868846]

Rieffe C, Terwogt MM, Petrides KV, Cowan R, Miers AC, Tolland A. Psychometric properties of the emotion awareness questionnaire for children. Personality and Individual Differences. 2007; 43(1): 95–105.

Rosenthal, R. Meta-analytic procedures for social research. rev. ed.. Thousand Oaks, CA: Sage Publications Inc; 1991.

Rosenthal, R. Meta-analytic procedures for social research. Newbury Park, CA: Sage Publications; 1993.

Rosenthal R, Rubin DB. Comment: Assumptions and procedures in the file drawer problem. Statistical Science. 1988; 3(1):120–125.

Saarni, C. The development of emotional competence. New York: Guilford Press; 1999.

*. Schonert-Reichl KA, Smith V, Zaidman-Zait A, Hertzman C. Promoting Childrens's Prosocial Behaviors in School: Impact of the "Roots of Empathy" Program on the Social and Emotional Competence of School-Aged Children. School Mental Health. 2012; 4(1):1–21.

*. Smith H. The effects of a drama-based language intervention on the development of theory of mind and executive function in urban kindergarten children. Dissertation Abstract International: Section B: The Sciences and Engineering. 2010; 72(5-B):3125.

*. Solomon M, Goodlin-Jones BL, Anders TF. A Social Adjustment Enhancement Intervention for High Functioning Autism, Asperger`s Syndrome, and Pervasive Development Disorder NOS. Journal of Autism and Developmental Disorders. 2004; 34(6):649–668. [PubMed: 15679185]

*. Steerneman P, Huskens B. The development of a social cognition training for autistic children. Psychologie in Erziehung und Unterricht. 1996; 43(4):291–301. Retrieved from http://www.reinhardt-verlag.de/de/zeitschriften/peu/.

*. Tenenbaum HR, Alfieri L, Brooks PJ, Dunne G. The effects of explanatory conversations on children's emotion understanding. British Journal of Developmental Psychology. 2008; 26(2): 249–263.

Thomas BH, Ciliska D, Dobbins M, Micucci S. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. Worldviews on Evidence-Based Nursing. 2004; 1(3):176–184. [PubMed: 17163895]

Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted? Statistics in Medicine. 2002; 21(11):1559–1573. [PubMed: 12111920]

Highlights

- A quantitative review of Emotion Understanding (EU) training programs was conducted
- We identified 19 studies/experiments totaling 749 children (mean age=86 months (SD=30.71)
- EU trainings were effective for all three aspects of EU (Hedges' g = 0.62; 0.31; 0.64)
- Specific EU skill-related outcomes were moderated by training setting, social setting and the length of training session
- We conclude that EU training procedures can effectively enhance EU in children

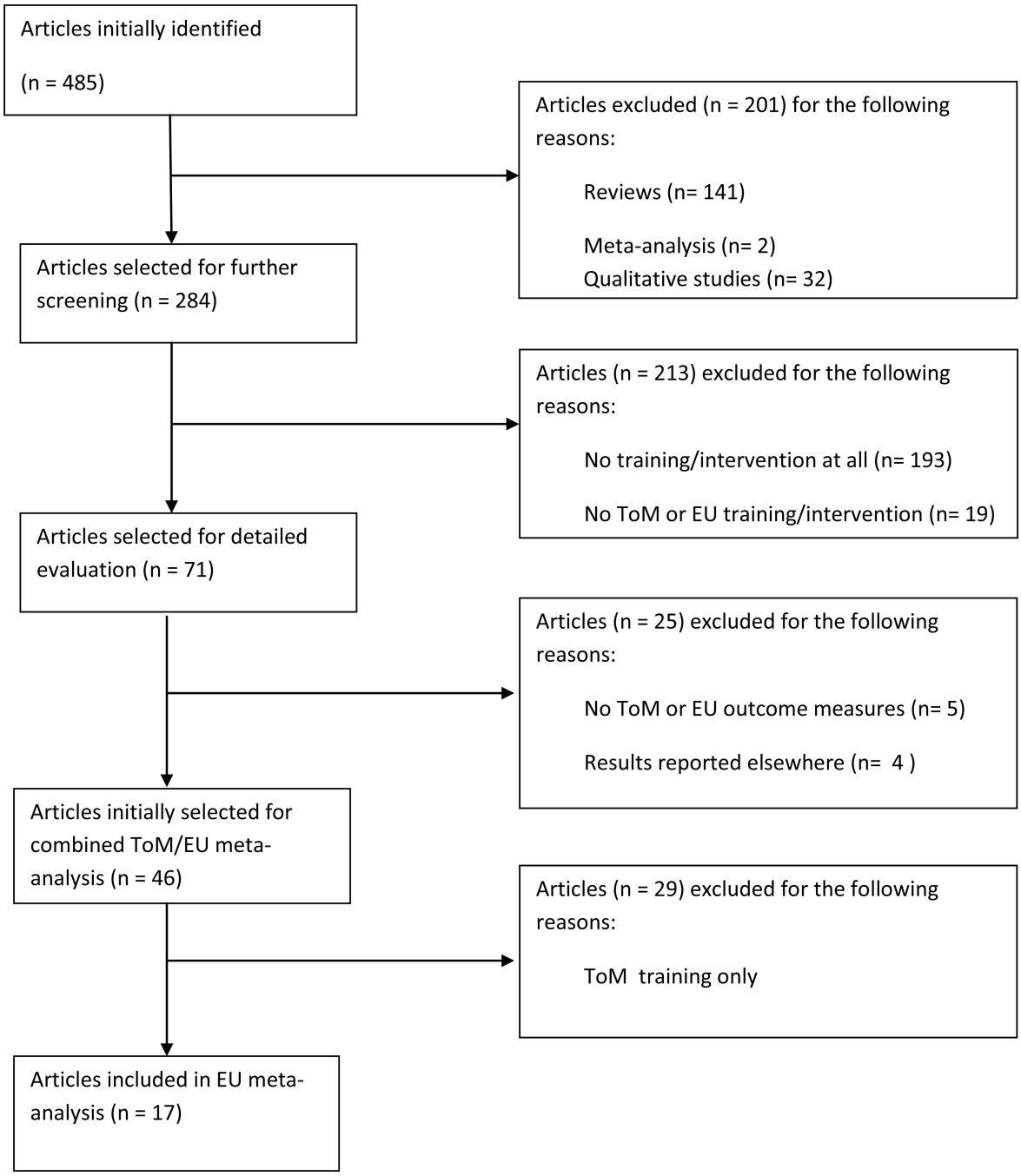
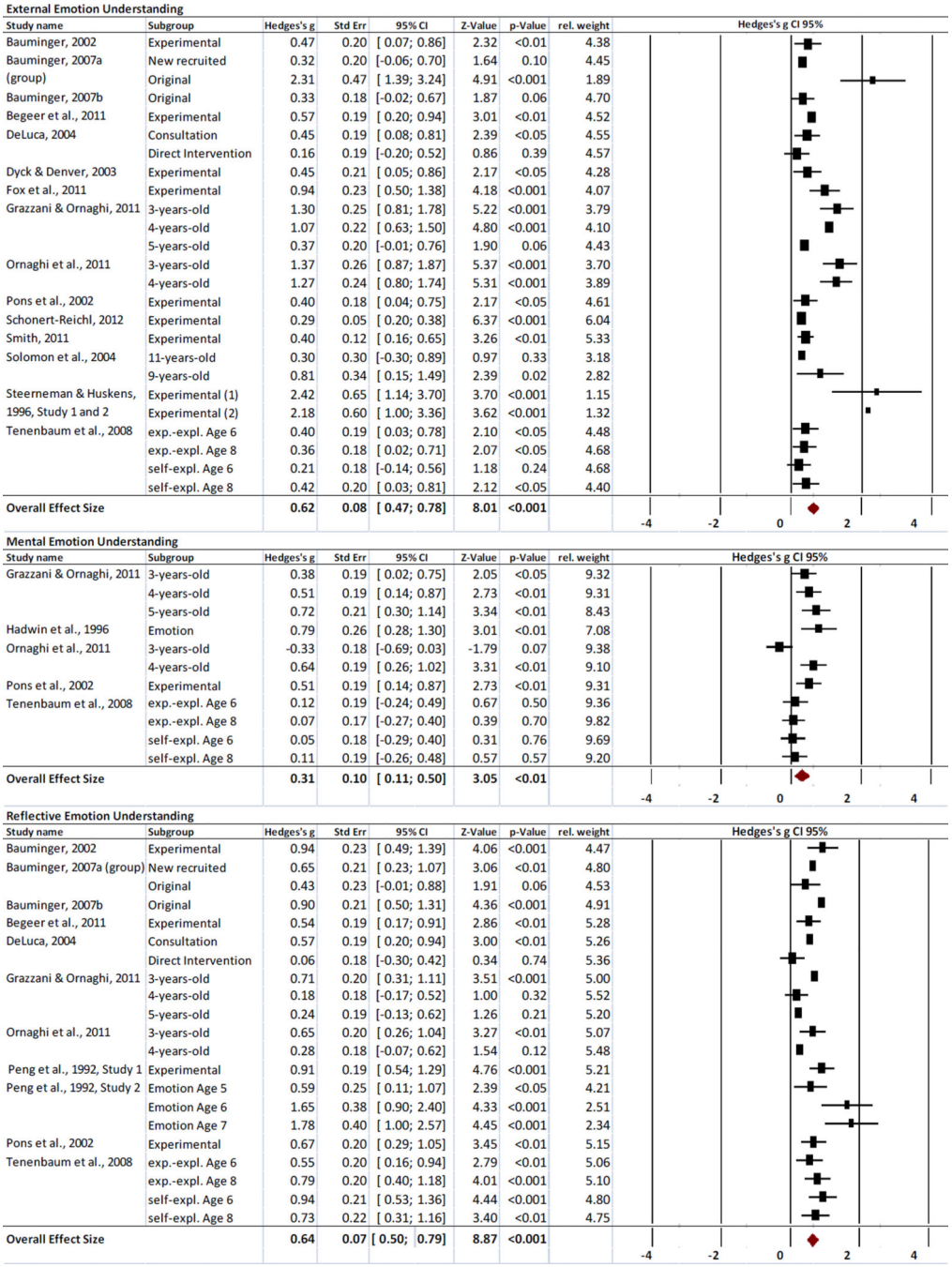


Figure 1.
Flow diagram of the study selection process



External Emotion Understanding

Study name	Subgroup	Hedges's g	Std Err	95% CI	Z-Value	p-Value	rel. weight
Bauminger, 2002	Experimental	0.47	0.20	[0.07; 0.86]	2.32	<0.01	4.38
Bauminger, 2007a	New recruited	0.32	0.20	[-0.06; 0.70]	1.64	0.10	4.45
(group)	Original	2.31	0.47	[1.39; 3.24]	4.91	<0.001	1.89
Bauminger, 2007b	Original	0.33	0.18	[-0.02; 0.67]	1.87	0.06	4.70
Begeer et al., 2011	Experimental	0.57	0.19	[0.20; 0.94]	3.01	<0.01	4.52
DeLuca, 2004	Consultation	0.45	0.19	[0.08; 0.81]	2.39	<0.05	4.55
	Direct Intervention	0.16	0.19	[-0.20; 0.52]	0.86	0.39	4.57
Dyck & Denver, 2003	Experimental	0.45	0.21	[0.05; 0.86]	2.17	<0.05	4.28
Fox et al., 2011	Experimental	0.94	0.23	[0.50; 1.38]	4.18	<0.001	4.07
Grazzani & Ornaghi, 2011	3-years-old	1.30	0.25	[0.81; 1.78]	5.22	<0.001	3.79
	4-years-old	1.07	0.22	[0.63; 1.50]	4.80	<0.001	4.10
	5-years-old	0.37	0.20	[-0.01; 0.76]	1.90	0.06	4.43
Ornaghi et al., 2011	3-years-old	1.37	0.26	[0.87; 1.87]	5.37	<0.001	3.70
	4-years-old	1.27	0.24	[0.80; 1.74]	5.31	<0.001	3.89
Pons et al., 2002	Experimental	0.40	0.18	[0.04; 0.75]	2.17	<0.05	4.61
Schonert-Reichl, 2012	Experimental	0.29	0.05	[0.20; 0.38]	6.37	<0.001	6.04
Smith, 2011	Experimental	0.40	0.12	[0.16; 0.65]	3.26	<0.01	5.33
Solomon et al., 2004	11-years-old	0.30	0.30	[-0.30; 0.89]	0.97	0.33	3.18
	9-years-old	0.81	0.34	[0.15; 1.49]	2.39	0.02	2.82
Steerneman & Huskens,	Experimental (1)	2.42	0.65	[1.14; 3.70]	3.70	<0.001	1.15
1996, Study 1 and 2	Experimental (2)	2.18	0.60	[1.00; 3.36]	3.62	<0.001	1.32
Tenenbaum et al., 2008	exp.-expl. Age 6	0.40	0.19	[0.03; 0.78]	2.10	<0.05	4.48
	exp.-expl. Age 8	0.36	0.18	[0.02; 0.71]	2.07	<0.05	4.68
	self-expl. Age 6	0.21	0.18	[-0.14; 0.56]	1.18	0.24	4.68
	self-expl. Age 8	0.42	0.20	[0.03; 0.81]	2.12	<0.05	4.40
Overall Effect Size		**0.62**	**0.08**	**[0.47; 0.78]**	**8.01**	**<0.001**	

Mental Emotion Understanding

Study name	Subgroup	Hedges's g	Std Err	95% CI	Z-Value	p-Value	rel. weight
Grazzani & Ornaghi, 2011	3-years-old	0.38	0.19	[0.02; 0.75]	2.05	<0.05	9.32
	4-years-old	0.51	0.19	[0.14; 0.87]	2.73	<0.01	9.31
	5-years-old	0.72	0.21	[0.30; 1.14]	3.34	<0.01	8.43
Hadwin et al., 1996	Emotion	0.79	0.26	[0.28; 1.30]	3.01	<0.01	7.08
Ornaghi et al., 2011	3-years-old	-0.33	0.18	[-0.69; 0.03]	-1.79	0.07	9.38
	4-years-old	0.64	0.19	[0.26; 1.02]	3.31	<0.01	9.10
Pons et al., 2002	Experimental	0.51	0.19	[0.14; 0.87]	2.73	<0.01	9.31
Tenenbaum et al., 2008	exp.-expl. Age 6	0.12	0.19	[-0.24; 0.49]	0.67	0.50	9.36
	exp.-expl. Age 8	0.07	0.17	[-0.27; 0.40]	0.39	0.70	9.82
	self-expl. Age 6	0.05	0.18	[-0.29; 0.40]	0.31	0.76	9.69
	self-expl. Age 8	0.11	0.19	[-0.26; 0.48]	0.57	0.57	9.20
Overall Effect Size		**0.31**	**0.10**	**[0.11; 0.50]**	**3.05**	**<0.01**	

Reflective Emotion Understanding

Study name	Subgroup	Hedges's g	Std Err	95% CI	Z-Value	p-Value	rel. weight
Bauminger, 2002	Experimental	0.94	0.23	[0.49; 1.39]	4.06	<0.001	4.47
Bauminger, 2007a (group)	New recruited	0.65	0.21	[0.23; 1.07]	3.06	<0.01	4.80
	Original	0.43	0.23	[-0.01; 0.88]	1.91	0.06	4.53
Bauminger, 2007b	Original	0.90	0.21	[0.50; 1.31]	4.36	<0.001	4.91
Begeer et al., 2011	Experimental	0.54	0.19	[0.17; 0.91]	2.86	<0.01	5.28
DeLuca, 2004	Consultation	0.57	0.19	[0.20; 0.94]	3.00	<0.01	5.26
	Direct Intervention	0.06	0.18	[-0.30; 0.42]	0.34	0.74	5.36
Grazzani & Ornaghi, 2011	3-years-old	0.71	0.20	[0.31; 1.11]	3.51	<0.001	5.00
	4-years-old	0.18	0.18	[-0.17; 0.52]	1.00	0.32	5.52
	5-years-old	0.24	0.19	[-0.13; 0.62]	1.26	0.21	5.20
Ornaghi et al., 2011	3-years-old	0.65	0.20	[0.26; 1.04]	3.27	<0.01	5.07
	4-years-old	0.28	0.18	[-0.07; 0.62]	1.54	0.12	5.48
Peng et al., 1992, Study 1	Experimental	0.91	0.19	[0.54; 1.29]	4.76	<0.001	5.21
Peng et al., 1992, Study 2	Emotion Age 5	0.59	0.25	[0.11; 1.07]	2.39	<0.05	4.21
	Emotion Age 6	1.65	0.38	[0.90; 2.40]	4.33	<0.001	2.51
	Emotion Age 7	1.78	0.40	[1.00; 2.57]	4.45	<0.001	2.34
Pons et al., 2002	Experimental	0.67	0.20	[0.29; 1.05]	3.45	<0.01	5.15
Tenenbaum et al., 2008	exp.-expl. Age 6	0.55	0.20	[0.16; 0.94]	2.79	<0.01	5.06
	exp.-expl. Age 8	0.79	0.20	[0.40; 1.18]	4.01	<0.001	5.10
	self-expl. Age 6	0.94	0.21	[0.53; 1.36]	4.44	<0.001	4.80
	self-expl. Age 8	0.73	0.22	[0.31; 1.16]	3.40	<0.01	4.75
Overall Effect Size		**0.64**	**0.07**	**[0.50; 0.79]**	**8.87**	**<0.001**	

Figure 2.

Overall effect sizes of External, Mental and Reflective Emotion Understanding abilities

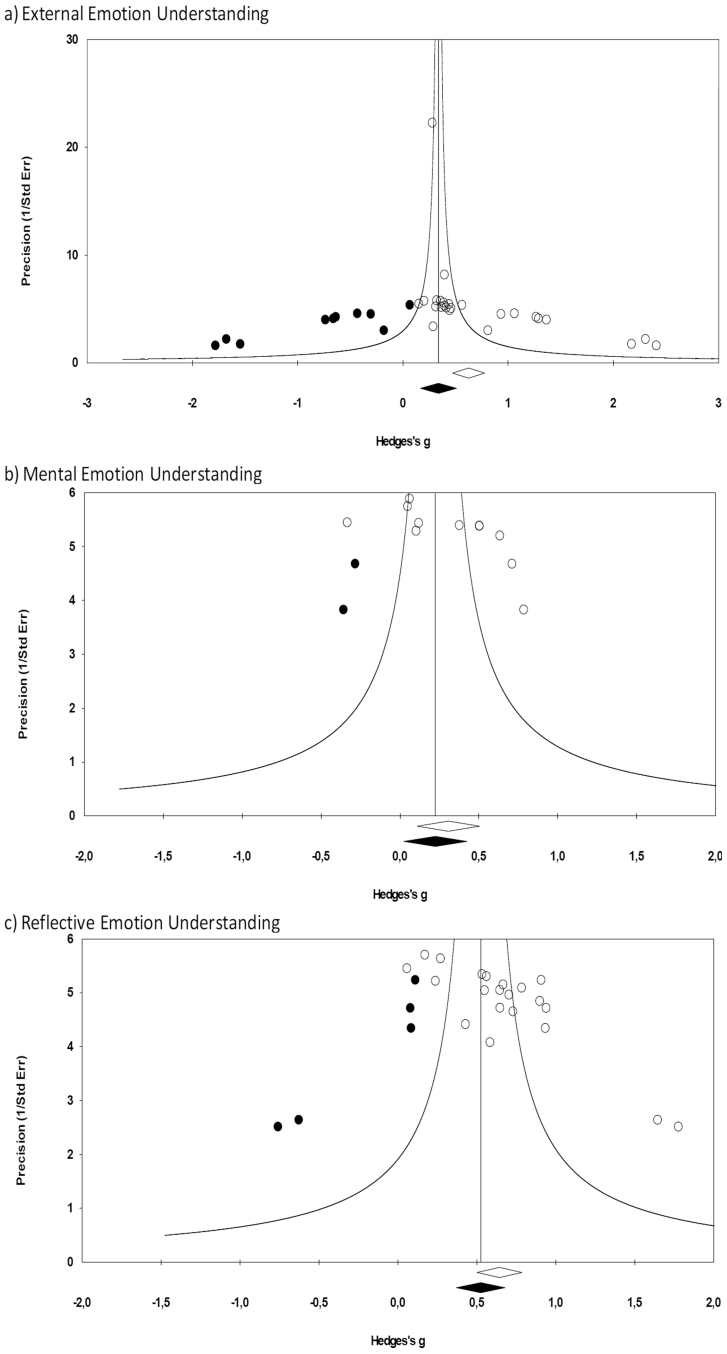


Figure 3.

Funnel Plot of precision by Hedge's g for external, mental and reflective emotion understanding abilities

Note: Includes recalculated effect size by the Trim and fill method and the imputed missing studies (filled diamond and dots). For the analysis of training impact of a) external emotion understanding eleven unpublished studies, for b) mental emotion understanding two

unpublished studies, and for c) reflective emotion understanding no unpublished study is assumed to be missing.

Table 1

PRISMA Checklist of items to include when reporting a systematic review or meta-analysis

Section/topic	Item No.	Checklist item	Reported on page No.
Title			
Title	1	Identify the report as a systematic review, meta-analysis, or both	1
Abstract			
Structured Summary	2	Provide a structured summary including, as applicable, background, objectives, data sources, study eligibility criteria, participants, interventions, study appraisal and synthesis methods, results, limitations, conclusions and implications of key findings, systematic review registration number	2
Introduction			
Rationale	3	Describe the rationale for the review in the context of what is already known	3–7
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS)	3–7
Methods			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (such as web address), and, if available, provide registration information including registration number	7
Eligibility Criteria	6	Specify study characteristics (such as PICOS, length of follow-up) and report characteristics (such as years considered, language, publication status) used as criteria for eligibility, giving rationale	8
Information sources	7	Describe all information sources (such as databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched	7–8
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated	7–8
Study selection	9	State the process for selecting studies (that is, screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis)	8
Data collection process	10	Describe method of data extraction from reports (such as piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators	8–9
Data items	11	List and define all variables for which data were sought (such as PICOS, funding sources) and any assumptions and simplifications made	8–9
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias on individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis	9
Summary measures	13	State the principal summary measure (such as risk ratio, difference in means)	9–10
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (such as I^2 statistic) for each meta-analysis	9–10
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (such as publication bias, selective reporting within studies)	10–11
Additional analyses	16	Describe methods of additional analyses (such as sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified	11–12
Results			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusion at each stage, ideally with a flow diagram	12–13, Figure 1
Study characteristics	18	For each study, present characteristics for which data were extracted (such as study size, PICOS, follow-up period) and provide the citations	13, Table 2
Risk of bias within studies	19	Present data on risk of bias of each study and, if applicable, any outcome-level assessment (see item 12)	18 Table 4

Section/topic	Item No.	Checklist item	Reported on page No.
Results of individual studies	20	For all outcomes considered (benefits or harms), present for each study (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot	Figure 2
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency	13–15, Figure 2
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see item 15)	15–16, Figure 3
Additional analyses	23	Give results of additional analyses, if done (such as sensitivity or subgroup analyses, meta-regression) (see item 16)	17–20, Tables 4–6
Discussion			
Summary of evidence	24	Summarize the main findings including strength or evidence for each main outcome; consider their relevance to key groups (such as health care providers, users, and policy makers)	20–23
Limitations	25	Discuss limitations at study and outcome level (such as risk of bias), and at review level (such as incomplete retrieval of identified research, reporting bias)	23–24
Conclusions	26	Provide general interpretation of the results in the context of other evidence, and implications for future research	24–25
Funding			
Funding	27	Describe sources of funding for the systematic review and other support (such as supply or data) and role of funders for the systematic review	1

Table 2

Description of studies

Study	Primary domain targeted in training	Emotion Understanding training (n)	Comparison condition (n)	Total sample size	Total sample mean age (months)	Sample type	External	Mental	Reflective	No. of training sessions	length of session (min)	environmental setting	social setting	country	post-test delay (days)
Bauminger, 2002	Recognition, External Causes	Social-Emotion Intervention (15)	None	15	135	HFASD	Em. Inv.: Simple Emotions	None	Em. Inv.: Complex Emotions	28	180	classroom	individual	Israel	0
Bauminger, 2007a (group)	Regulation, Mixed Emotions, Moral	Group Curriculum (25)	None	25	108	HFASD	Em. Inv.: Simple Emotions; Affective Matching Measure	None	Em. Inv.: Complex Emotions	56	n.p.	separate room in school	group	Israel	0
Bauminger, 2007b	Recognition, External Causes	Social-Emotion Intervention (19)	None	19	105	HFASD	Em. Inv.: Simple Emotions; Affective Matching Measure	None	Em. Inv.: Complex Emotions	28	180	separate room in school	individual	Israel	0
Begeer et al., 2011	Recognition, External Causes	Manualized ToM Training (19)	Waiting-list (17)	36	123	ASD	LEAS-C: Total Score	None	LEAS-C: Mixed and Complex Emotions	16	90	Outpatient Clinic	group	Nether lands	14
*DeLuca, 2004	Recognition, External Causes, Regulation	Rational Emotive Education Lessons (23)	Non-treatment (12)	46	49	Children with a Disability	Ability to Verbalize Emotions; Ability to Identify Emotion Causality	None	Measure of the Ability to Generate Appropriate Alternatives	12	45	class room	group	United States of America	n.p.
Dyck & Denver, 2003	External Causes, Regulation, Moral	Funny Face Program (14)	None	14	142	Deaf Children	FES; CT; EVT	None	None	11	45	class room	group	Australia	2,5
Fox et al., 2011	Recognition	Strengthening Early Emotional Development (16)	None	16	49	Anxious Children	Preschool Emotion Interview	None	None	10	60	Out patient Clinic	group	United States of America	14

Study	Primary domain targeted in training	Emotion Understanding training (n)	Comparison condition (n)	Total sample size	Total sample mean age (months)	Sample type	External	Mental	Reflective	No. of training sessions	length of session (min)	environmental setting	social setting	country	post-test delay (days)
Grazzani & Ornaghi, 2011	External Causes	Conversational Language Game (50)	Free Play (50)	100	52	TD	TEC Level 1	TEC Level 2	TEC Level 3	16	15	separate room in school	group	Italy	14
Hadwin et al., 1996	Recognition, External Causes, Desires, Beliefs	Teaching Emotional Understanding (10)	Teaching Belief Understanding (10)	20	108	ASD	None	Concept Understanding of 5 hierarchical Emotion-	None	8	30	Lab	individual	United Kingdom	1
Ornaghi, Brockmeier & Grazzani, 2011	External Causes	Language Games (35)	Free Play (35)	70	46	TD	TEC Level 1	TEC Level 2	TEC Level 3	16	20	separate room in school	group	Italy	14
Peng et al., 1992, Study 1	External Causes, Mixed Emotions	Mixed Emotion Training (22)	Active Training (19)	31	74	TD	None	None	Question about the feelings provoked through a given story	4	n.p.	quiet area in school	individual	United Kingdom	0
Peng et al., 1992, Study 2	External Causes, Mixed Emotions	Mixed Emotion Training (37)	Active (38)	75	72	TD	None	None	Question of the possibility to feel two ambivalent emotions at the same time	9	n.p.	quiet area in school	individual	United States of America	0
Pons et al., 2002	Recognition, External Causes, Regulation, Hiding, Moral	Schools Matters in Lifeskills Education (18)	Noneactive (18)	36	109	TD	TEC Level 1	TEC Level 2	TEC Level 3	1	30	classroom	n.p.	United Kingdom	10,5
Schonert-Reichl, 2012	Recognition	Roots of Empathy Program (306)	Treatment as usual (279)	585	127	TD	Understanding of an infants crying	None	None	27	45	classroom	group	Canada	21
Smith, 2011	Recognition, External Causes	The Georgia Wolftrap program (41)	Standard Curriculum (42)	83	68	low SES	NEPSY-II: Subscale ToM-Contextual Portion	None	None	13	50	classroom	group	United States of America	14

Study	Primary domain targeted in training	Emotion Under standing training (n)	Com-parison condition (n)	Total sample size	Total sample mean age (months)	Sample type	External	Mental	Reflective	No. of training sessions	length of session (min)	environmental setting	social setting	country	post-test delay (days)
Solomon et al., 2004	Recognition, External Causes, Regulation, Moral	The Social Adjustment Enhancement Curriculum (9)	Waiting-list (9)	18	117	HFASD	DANVA: 2-AF and 2-CF	None	None	20	90	lab	group	United States of America	0
Steerneman & Huskens, 1996, Study 1	External Causes	Innovative group training (5)	Waiting-list (5)	10	108	ASD	Test of Perception of Emotions from Facial Expression and Posture Cues	None	None	27	60	n.p.	both	Nether lands	n.p.
Steerneman & Huskens, 1996, Study 2	Recognition, External Causes	Innovative group training 2 (5)	None	5	118	ASD and PDDNOS	Test of Perception of Emotions from Facial Expression and Posture Cues	None	None	21	60	n.p.	group	Nether-lands	n.p.
Tenenbaum et al., 2008	External Causes, Hiding, Mixed Emotions	Nine Vignettes: Self- and Experimenter-explanation (69)	Nine Vignettes: Self-Summary (33)	102	83	TD	TEC Level 1	TEC Level 2	TEC Level 3	1	n.p.	multiple locations	individual	United States of America	0

*Dissertation

Em. Inv. = Emotion Inventory
HFASD = High-functioning Children with autism spectrum disorder
ASD = Children with autism spectrum disorder
PDD-NOS = pervasive developmental disorder not otherwise specified
TD = typically developed children
Low SES = low social economic status
FES = Fluid Emotions Scale
CT = Comprehension Task
EVT = Emotion Vocabulary Test
NEPSY II – NEuroPSYchological second edition, an instrument designed to assess neuropsychological development in preschool and school-age children
DANVA-2-AF = Diagnostic Analysis of Non-verbal Accuracy-2-Adult Facial Expression
DANVA-2-AF = Diagnostic Analysis of Non-verbal Accuracy-2-Child Facial Expression

Table 3

EPHPP - Quality Assessment Tool for Quantitative Studies

Study	(A) Selection Bias	(B) Study Design	(C) Confounder	(D) Blinding	(E) Data Collection Method	(F) With drawals and Drop-outs	GLOBAL RATING
Bauminger, 2002	2	2	3	2	2	1	**2**
Bauminger, 2007a (group)	2	2	3	2	2	1	**2**
Bauminger, 2007b	2	2	3	2	2	1	**2**
Begeer, et al, 2011	1	1	1	2	1	1	**1**
[X]DeLuca, 2004	2	1	3	2	3	3	**3**
Dyck & Denver, 2003	2	2	3	2	1	1	**2**
Fox et al, 2012	1	2	3	2	1	1	**2**
Grazzani & Ornaghi, 2011	2	1	1	2	1	1	**1**
Hadwin, Baron-Cohen, Holwin & Hill, 1996	2	1	3	2	3	1	**3**
Ornaghi, Brockmeier & Grazzani, 2011	2	1	1	2	1	1	**1**
Peng et al, 1992 Study 1	2	1	3	2	3	1	**3**
Peng et al, 1992 Study 2	2	1	3	2	2	1	**2**
Pons et al, 2002	2	1	3	2	1	1	**2**
Schonert-Reichl et al, 2012	1	1	1	2	2	1	**1**
[X]Smith, 2010	2	1	1	2	1	1	**1**
Solomon et al, 2004	2	1	1	2	2	1	**1**
Steerneman and Huskens, 1996 Study 1	2	1	3	2	3	1	**3**
Steerneman and Huskens, 1996 Study 2	2	2	3	2	3	1	**3**
Tenenbaum et al, 2008	2	1	1	2	1	1	**1**

Note.

[X]Dissertation

1 = Strong, 2 = Moderate, 3 = Weak

Table 4

Moderator analyses (continuous variables)

		B value	SE	95% CI	Z	*p*-value
Number of sessions						
	External	0.004	0.01	[−0.01; 0.01]	0.81	0.42
	Mental	0.01	0.01	[−0.01; 0.02]	1.13	0.26
	Reflective	−0.003	0.004	[−0.01; 0.004]	−0.84	0.40
Length of each Session (hours)						
	External	−0.17	0.12	[−0.39; 0.06]	−1.44	0.15
	Mental	0.55	1.54	[−2.46; 3.56]	0.36	0.72
	Reflective	0.19	0.08	[0.04; 0.33]	2.46*	<0.05
Length of Training period (months)[a]						
	External	0.03	0.03	[−0.03; 0.09]	0.98	0.33
	Mental	0.11	0.09	[−0.07; 0.28]	1.19	0. 24
	Reflective	−0.01	0.03	[−0.07; 0.05]	−0.43	0.67
Year of publication						
	External	<0.001	0.02	[−0.04; 0.04]	−0.01	0.99
	Mental	−0.02	0.02	[−0.07; 0.02]	−1.01	0.31
	Reflective	−0.03	0.01	[−0.05; −0.01]	−2.66**	<0.01
Posttest Delay (weeks)[aa]						
	External	0.10	0.08	[−0.06; 0.26]	1.25	0.21
	Mental	0.10	0.11	[−0.01; 0.31]	1.00	0.32
	Reflective	−0.19	0.06	[−0.32; −0.07]	−3.09**	<0.01

Note.

* p < 0.05,

** p < 0.01,

*** p < 0.001

[a] Excluded: Bauminger 2007a (group), Peng et al. 1992 study 1 and 2, and Tenenbaum et al. 2008; because there are no information provided

[aa] Excluded: DeLuca, 2004, and Steerneman and Huskens, 1996, Study 1 and 2, since no information provided.

Table 5

Between-group moderator analyses (categorical variable)

	N	Hedge's g	S.E.	95% CI	*z*-value	*p*-value
Environmental Setting						
External						
Classroom	7	0.37	0.11	[0.14; 0.59]	3.21**	< .01
Lab	4	0.66	0.18	[0.31; 1.02]	**3.65** ***	< .001
Area in school	8	0.89	0.12	[0.65; 1.13]	7.24***	< .001
Mental						
Classroom	1	0.51	0.42	[−0.32; 1.33]	1.21	0.23
Lab	1	0.79	0.46	[−0.11; 1.69]	1.72	0.09
Area in school	5	0.38	0.19	[0.01; 0.75]	1.99*	< .05
Reflective						
Classroom	4	0.55	0.19	[0.18; 0.91]	2.94*	< .01
Lab	1	0.54	0.37	[−0.18; 1.25]	1.47	0.14
Area in school	12	0.67	0.11	[0.45; 0.89]	5.94***	< .001
Social Setting						
External						
Group	17	0.72	0.10	[0.53; 0.91]	7.39 ***	< .001
Individual	6	0.36	0.15	[0.06; 0.66]	2.38*	< .05
Mental						
Group	5	0.38	0.16	[0.07; 0.68]	2.42*	< .05
Individual	5	0.20	0.16	[−0.11; 0.50]	1.28	0.20
Reflective						
Group	10	0.42	0.08	[0.27; 0.56]	5.45***	< .001
Individual	10	0.88	0.09	[0.71; 1.05]	10.12 ***	< .001
Sample Type						
External						

	N	Hedge's g	S.E.	95% CI	*z*-value	*p*-value
Typically developed	11	0.64	0.12	[0.41; 0.87]	5.40***	< .001
Clinical condition	14	0.63	0.12	[0.40; 0.85]	5.47***	< .001
- Children with Autism	9	0.76	0.16	[0.45; 1.06]	4.85***	< .001
- Other than Autism	5	0.47	0.18	[0.12; 0.82]	2.66**	< .01
Mental						
Typically Developed	10	0.27	0.10	[0.07; 0.47]	2.67**	< .01
Clinical Condition	1	0.79	0.37	[0.07; 1.51]	2.14*	< .05
Reflective						
Typically Developed	14	0.68	0.09	[0.50; 0.86]	7.45***	< .001
Clinical Condition	7	0.57	0.13	[0.33; 0.82]	4.57***	< .001
- Children with Autism	5	0.69	0.12	[0.46; 0.92]	5.85***	< .001
- Disability-Autism	2	0.31	0.17	[−0.03; 0.65]	1.81	0.07
Country						
External						
America	11	0.41	0.10	[0.22; 0.59]	4.20***	< .001
Europe	9	0.96	0.12	[0.72; 1.20]	7.80***	< .001
Other	5	0.55	0.15	[0.25; 0.85]	3.55 ***	< .001
Mental						
America	4	0.09	0.15	[−0.20; 0.38]	0.59	0.55
Europe	7	0.44	0.12	[0.21; 0.67]	3.78 ***	< .001
Reflective						
America	9	0.74	0.12	[051; 0.96]	6.33***	< .001
Europe	8	0.52	0.11	[0.30; 0.74]	4.75***	< .001
Other	4	0.73	0.17	[0.40; 1.06]	4.34***	< .001
EPHPP Global Rating						
External						

	N	Hedge's g	S.E.	95% CI	*z*-value	*p*-value
Strong	14	0.62	0.11	[0.41; 0.83]	5.83***	< .001
Moderate	7	0.61	0.16	[0.30; 0.91]	3.92***	< .001
Weak	4	0.76	0.24	[0.28; 1.23]	3.14*	< .05
Mental						
Strong	9	0.24	0.11	[0.03; 0.46]	2.24*	< .05
Moderate	1	0.51	0.33	[−0.13; 1.15]	1.57	0.12
Weak	1	0.79	0.37	[0.06; 1.52]	2.11*	< .05
Reflective						
Strong	10	0.55	0.10	[0.36; 0.74]	5.59***	< .001
Moderate	8	0.85	0.12	[0.61; 1.09]	6.86***	< .001
Weak	3	0.51	0.18	[0.16; 0.86]	2.88**	< .01

Note.

* $p < 0.05$,

** $p < 0.01$,

*** $p < 0.001$

Table 6

Effect of Training compositions

Outcome - Training composition	N	Hedge's g	SE	95% CI	z-value	p-value
External						
External aspects	12	0.83	0.13	[0.59; 1.07]	6.65***	< 0.001
External and Reflective aspects	6	0.44	0.17	[0.10; 0.78]	2.54*	<0.05
Mental and Reflective aspects	2	0.96	0.34	[0.30; 1.62]	2.86**	< 0.01
Combination of All three level	5	0.36	0.18	[0.01; 0.71]	1.99*	< 0.05
Mental						
External aspects	5	0.37	0.15	[0.09; 0.66]	2.58*	< 0.05
External and Mental aspects	1	0.79	0.37	[0.06; 1.52]	2.13*	< 0.05
Combination of All three level	5	0.17	0.14	[-0.11; 0.45]	1.20	0.23
Reflective						
External aspects	7	0.54	0.13	[0.29; 0.79]	4.26***	< 0.001
External and Reflective aspects	7	0.73	0.14	[0.46; 1.00]	5.29***	< 0.001
Mental and Reflective aspects	2	0.55	0.25	[0.06; 1,03]	2.21*	< 0.05
Combination of All three level	5	0.74	0.15	[0.44; 1.03]	4.85***	< 0.001