

WHY DO YOU LOOK SO ANGRY?

Understanding and changing social cognition
in boys with disruptive behavior problems



WIETEKE HIEMSTRA

Why do you look so angry?

Understanding and changing social cognition in boys with disruptive behavior problems.

Wieteke Hiemstra

ISBN 978-94-6375-458-3

Printed by Ridderprint BV | www.ridderprint.nl

Cover by Sophie Harbers

Copyright © W. Hiemstra, 2019.

For all published articles, the copyright has been transferred to the respective publisher. No part of this dissertation may be reproduced, stored in a retrieval system, or transmitted – in any form or by any means – without written permission of the author or, when appropriate, from the copyright-owning publisher.

Why do you look so angry?

Understanding and changing social cognition in boys with disruptive behavior problems.

Waarom kijk je zo boos?

Het begrijpen en veranderen van sociale cognitie in jongens met disruptief probleemgedrag.

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor

aan de Universiteit Utrecht

op gezag van de rector magnificus,

prof. dr. H.R.B.M. Kummeling,

ingevolge het besluit van het college voor promoties

in het openbaar te verdedigen op

vrijdag 30 augustus 2019

des middags te 12.45 uur

door

Wieteke Hiemstra

geboren op 4 december 1989 te Bemmelen

Promotoren : Prof. dr. B. Orobio de Castro

Prof. dr. S. Thomaes

Beoordelingscommissie:

Prof. dr. A.H.N. Cillessen, Radboud Universiteit Nijmegen

Prof. dr. J.J.S. Dubas, Universiteit Utrecht

Prof. dr. C. Finkenauer, Universiteit Utrecht

Prof. dr. M.A. van den Hout, Universiteit Utrecht

Prof. dr. R.W.H.J. Wiers, Universiteit van Amsterdam

Contents

Chapter 1	General introduction	9
Chapter 2	Self-Views and Aggression in Children Referred for Disruptive Behavior Problems: A Multi-Informant Approach	19
Chapter 3	Day-to-day Covariation of State Anger and Hostile Attributions in Children with Disruptive Behavior Problems	35
Chapter 4	Reducing Aggressive Children's Hostile Attributions: A Cognitive Bias Modification Procedure	49
Chapter 5	Improving Emotion Attribution in Children with Disruptive Behavior Problems Using Cognitive Bias Modification	73
Chapter 6	General discussion	91
	References	103
	Samenvatting/Summary in Dutch	121
	Dankwoord/Acknowledgements	129
	About the author	135



CHAPTER 1

GENERAL INTRODUCTION

General Introduction

“One of the most common referrals in child and family psychology is a boy in middle childhood who presents with conduct problems, specific learning difficulties and related family and school problems” (Carr, 2006, p. 361).

Childhood aggression is suggested to be the best-known predictor of (future) social adjustment difficulties (Card, Stucky, Sawalani, & Little, 2008). For example, Roff and Wirt (1984) showed that childhood aggression in boys emerged as the most prominent antecedent factor of later delinquency. Furthermore, research has shown that childhood aggression remains relatively stable over time (Burks, Laird, Dodge, Pettit, & Bates, 1999). This means that children with aggression problems often have a poor prognosis. Besides that, these children are remarkably unresponsive to treatment (Carr, 2006). From different models and existing literature, there are some important factors that come to front in development and maintenance of aggression. These factors provide entry points of intervention, and are therefore important to examine. In this dissertation, I focused on the social cognitive factors self-views, state-anger and hostile interpretation in boys with disruptive behavior problems. I aimed to get more insight of how these different factors interplay to produce aggressive behavior, but I also tested a new implicit intervention to see whether hostile interpretation is malleable and whether this affects aggressive behavior. To provide a context for the research, the following topics are theoretically outlined in this introduction: Aetiological theories of aggression problems; important personal factors in development and maintenance of aggression problems; differences between a clinical population and the general population; and existing interventions. The introduction concludes with the main aims of this dissertation and an outline of the content.

Aetiological theories of Aggression Problems

There are different influential theories about the aetiology of aggression problems. Amongst them are biological theories, psychodynamic theories, cognitive theories, social learning theories, and systems theories (Carr, 2006). Research has shown that contextual (parenting and peer) factors are very important in the development and maintenance of aggression problems (e.g. Hoeve, Dubas, Eichelsheim, van der Laan, Smeenk, & Gerris, 2009; Reijntjes, Kamphuis, Prinzie, Boelen, van der Schoot, & Telch, 2010). Besides that, there are some

specific personal pre-disposing and maintaining factors that put children at risk to develop and maintain aggressive behaviors. In this dissertation, I will mainly use two important models from existing literature as a framework: The General Aggression Model (GAM; Allen, Anderson, & Bushman, 2018), and the Social information Processing Model (SIP; Crick & Dodge, 1994).

General Aggression Model

The General Aggression Model is an integrative model which considers an interplay of biological, social, cognitive and personality factors in aggression. GAM separates two aspects: Proximate and Distal processes. Proximate processes explain individual episodes of aggression through person and situation factors, which influence the present internal state (where cognition, affect and arousal interplay). This internal state interacts with appraisal and decision processes. The distal processes include biological-, environmental-, and personality factors. The proximal processes are influenced by the distal processes, but in turn also add to the distal processes by adding experience and learning to the existing database of knowledge (personality).

Social Information Processing Theory

Crick and Dodge (1994) have proposed a SIP-model, which offers a more detailed theory of how children process social cues and how individual differences in social behaviour emerge. The SIP-model proposes different “processing” steps. The first step is the encoding of external and internal cues. In the second step, the cues are interpreted. The interpretation process is guided by existing social knowledge and schemata. As can be seen in the model, the interpretation process can include inferences about other’s intentions, causal analyses, and evaluations of self and others. The third step is the clarification or selection of a goal. At step 4 it is hypothesized that children access from memory different responses to the situation, or they may construct new responses. In step 5 the child decides on a response. Step 6 involves the behavioural enactment of the chosen response. It is assumed that the processing steps are occurring relatively rapid and in parallel, with numerous feedback loops (Lemerise & Arsenio, 2000). Within this model, an emotional component of SIP was added later. For example, mood, emotions and/or arousal can affect what is noticed about a social

encounter, and make the recollection of mood-congruent information more likely, thus influencing the interpretation of social cues.

The General Aggression Model (Allen, Anderson, & Bushman, 2018) and the Social Information Processing model (Crick & Dodge, 1994; Lemerise & Arsenio, 2000) both posit that aggressive behavior is a result of multiple factors influencing the interpretation, appraisal and decision making in social situations. Although GAM's framework is divided into proximal and distal processes in different stages, and the SIP model describes different steps that occur in parallel, both models focus on how an aggressive outcome is heavily influenced by a certain database or knowledge structures. This database is built through learning and experience, and therefore can become automatized with repeated practice.

Important personal factors in development and maintenance of aggression

From above models and existing literature, multiple factors are involved and important in the development and maintenance of aggression problems. In this dissertation, I will specifically focus on self-views and hostile intent attribution (hostile interpretation) of children referred for disruptive behavior problems.

Self-views: Narcissism and Self-esteem

It has long been assumed that children with aggression problems are prone to suffer from low self-esteem. Self-esteem refers to children's representations of their general, global feelings of worth (Rosenberg, 1986). Various classic authors have proposed that aggressive youth are likely to have low self-esteem, perhaps because low self-esteem weakens ties to society, leading to nonconformity to social norms (Rosenberg, 1965), or because children may protect themselves against negative self-feelings by externalizing blame for their own failures (e.g., Thomaes, Stegge, Olthof, Bushman, & Nezelek, 2011; Tracy & Robins, 2003). Accordingly, attempts to reduce children's aggression problems often seek to enhance children's level of self-esteem (e.g. Carr, 1999; Kusché & Greenberg, 1994; Mann, Hosman, Schaalma, & de Vries, 2004; Schonfeld et al., 2014). The literature on typically developing children, however, suggests that aggression may also stem from "threatened egotism" – inflated, narcissistic self-views that are jeopardized, rather than from low self-esteem per se (Barry, Thompson, Barry, Lochman, Adler, & Hill, 2007;

Baumeister, Smart, & Boden, 1996; David & Kistner, 2000; Thomaes, Bushman, Stegge, & Olthof, 2008). Narcissism refers to a sense of grandiosity and entitlement, and a strong need to be seen and admired by others (Thomaes, Brummelman, Reijntjes, & Bushman, 2013). Although extreme levels of narcissism may reflect pathology, narcissism is more typically conceived of as a trait on which children in the general population vary. One way in which trait narcissism differs from self-esteem, is that it involves unrealistically positive self-views and a strong sense of superiority, with a need for validation of these feelings by others (Brummelman, Thomaes, & Sedikides, 2016; Bushman & Baumeister, 1998; Raskin, Novacek, & Hogan, 1991).

Hostile Intent Attribution

Children displaying aggression differ in their social information processing compared to non-aggressive children, and these differences mediate between known risk factors for aggressive behavior and the development of aggressive behavior. One of the apparent characteristics of processing by aggressive children is called hostile intent attribution. It consists of an automatic tendency to interpret a variety of ambiguous social cues (including facial expressions) as hostile (Nasby, Hayden and DePaulo, 1980). This attribution is especially apparent in the schemata (database of experiences) and interpretation of social cues (De Castro, Veerman, Koops, Bosch, & Monshouwer, 2002), and leads to development and maintenance of aggressive behavior (Crick & Dodge, 1994; Dodge, 2006). It seems that aggressive children do not attend more to hostile cues compared to non-aggressive children, but nonetheless attribute more hostile intent than their non-aggressive peers (Horsley, Orobio de Castro, & Van der Schoot, 2010). As explained earlier, the interpretation of social cues is partly based on existing schemata. In this case, if the initial social cue is unfamiliar or ambiguous, the knowledge structure in children's database will facilitate the representational process by adding existing information to the cue (Burks et al., 1999). If this knowledge structure consists of events in which hostility was apparent, the child will be more likely to process information in a hostile manner. This processing is facilitated by heuristics (i.e. representativeness and availability) used to interpret and make judgements about information, and therefore becomes more rigid over time. For example, if children have had hostile experiences in the past, this is readily available information in the knowledge base, and will lead to a tendency to make an interpretation consistent with this

hostile category (Dodge, 2006). This knowledge structure is, in part, accountable for the stability of hostile attribution and aggressive behaviour over time (Burks et al., 1999). Not only does hostile intent attribution lead to the growth of aggressive behavior over time (Dodge, Bates, & Pettit, 1990), reciprocal effects are also likely. Aggressive behavior problems can, in their turn, lead to hostile reactions from the social environment that further promote the development of hostile attribution tendencies (Dodge, 2006). In this way, the different SIP may become a self-fulfilling prophecy (i.e. with time, the social environment may become more hostile because of the aggression the child shows; Crick & Dodge, 1996).

Emotionality is one of the biological predispositions that can influence social information processing (Lemerise & Arsenio, 2000). Children differ in the intensity of the emotions they experience, as well as in their emotion expression and regulation predispositions, which influences their social behavior. The “mood congruency effect” describes the tendency for people to render judgements in the direction of their own mood state (e.g. Bodenhausen & Sheppard, 1994). According to this effect, negative mood can activate knowledge structures that are associated with that affective state. Increased familiarity with experiencing and exposure to a particular emotion, may strengthen the available knowledge structures (Izard, 1977, in Schultz, Izard, & Bear, 2004). When individuals experience negative mood, like anger, and have a set of readily available negative knowledge structures, this can breed a negative interpretation of social stimuli. In children with aggressive behavior problems, an angry emotional state could thus trigger hostile schemata and hostile attributions of intent, which may instigate aggression.

Differences between clinical populations and the general population

From a clinical perspective, research in population samples is informative, but it is more important to know how children with the most serious aggression problems – i.e., those who receive treatment – typically view themselves, others and social situations. These children are not only more likely to exhibit personal vulnerabilities (e.g., low IQ, difficult temperament, attention problems, or impulsivity; e.g., Dodge, 2006), they also tend to have more prevalent histories of aversive social experience (e.g. coercive family interactions, peer rejection, neglect; e.g., Ettekal & Ladd, 2015; Olson, 1992; Smith et al., 2014; Patterson,

1976). These factors may well impact the link between different underlying factors and aggressive behavior. For example, parental neglect or lack of warmth is known to both undermine healthy self-esteem development and cultivate children's aggressive behavior (Brummelman, Thomaes, Nelemans, Orobio de Castro, Overbeek, & Bushman, 2015; Cicchetti, 2016) suggesting that low self-esteem may be a more potent risk factor for aggressive behavior in clinical samples of youth.

Existing interventions

Different interventions targeting childhood disruptive behavior problems have already been studied. The most effective interventions for boys with disruptive behavior problems in middle childhood to date are parent-training programs based on operant principles (McCart, Priester, Davies, & Azen, 2006), such as Incredible Years and Triple P (De Graaf, Speetjens, Smit, de Wolff, & Tavecchio, 2008). For children themselves, there are individual treatment programs that mostly focus on social skills and conscious re-training of intent attributions (e.g. Guerra & Slaby, 1990; Hudley & Graham, 1993; Lochman & Wells, 2002; Sukhodolsky, Smith, McCauley, Ibrahim, & Piasecka, 2016). However, these interventions have modest and heterogeneous effects. Therefore, it is important to seek for additional ways to intervene in childhood aggression. By doing this, we get better sight on which intervention techniques work and for which children. One promising intervention technique may be direct intervention on social cognitions that are known to evoke aggression.

Recent approaches have sought to modify this implicit process by directly influencing the target cognitive bias (i.e., cognitive bias modification (CBM); e.g. Hertel & Mathews, 2011; MacLeod, Koster, & Fox, 2009; MacLeod & Mathews, 2012). CBM constitutes of repeated practice on a specific task that is designed to slightly change attitudes and bias in information processing (MacLeod, Koster, & Fox, 2009). CBM does not require conscious reports of personal beliefs or attitudes and therefore does not depend on self-insight. Some studies using CBM to target hostile attributions, conducted in population-based and at-risk samples of children and adults, have shown promising results (Penton-Voak, Munafò, & Yen Looi, 2017; Penton-Voak, Thomas, Gage, McMurrin, McDonald, & Munafò, 2013; Stoddard et al., 2016; Vassilopoulos, Brouzos, & Andreou, 2015). However, CBM has not yet been tested yet in those who are in particular need of intervention to prevent adverse outcomes:

clinically referred children with aggressive behavior problems (e.g. De Castro et al., 2002; Fergusson, Horwood, & Ridder, 2005).

Aims and outline of this dissertation

The first aim of this dissertation was to get more insight in the social cognitions, mood, and self-views in boys who have been referred for disruptive behavior problems, in order to be better able to intervene. This aim is addressed in Chapters 2 and 3. More specifically, in Chapter 2 the role of different self-views (narcissism and self-esteem) is examined in relation to aggression problems. These relations are examined by carrying out two separate studies, using a multi-informant approach. In Chapter 3, a daily diary method is used to get more insight in the covariation of state-anger and hostile intent attributions in relation to aggression problems.

The second aim of this dissertation is to test the effects of an innovative cognitive bias modification procedure to (a) change social cognitions involved in aggressive behavior and (b) reduce aggressive behavior in boys with disruptive behavior problems. This aim is addressed in Chapters 4 and 5. In Chapter 4, an innovative implicit intervention is tested in two separate studies for effects on hostile intent attributions and aggressive behavior. Chapter 5 is a replication study of the findings in Chapter 4, and extends the use of the intervention to different emotion processing biases.

The main findings of the studies reported in Chapters 2 to 5 are discussed in Chapter 6. Theoretical and practical implications are discussed, as well as strengths and limitations and future directions. The dissertation closes with a summary in English and Dutch.

CHAPTER 2

SELF-VIEWS AND AGGRESSION IN CHILDREN REFERRED FOR DISRUPTIVE BEHAVIOR PROBLEMS: A MULTI-INFORMANT APPROACH

An adapted version of this Chapter was published as:

Hiemstra, W., Verhulp, E.E., Thomaes, S., & Orobio de Castro, B. (2019). Self-Views and Aggression in Children Referred for Disruptive Behavior Problems: Self-esteem, narcissism and their interaction, *European Child and Adolescent Psychiatry*, (), 1-9.

Orobio de Castro developed the study concept as part of a compounding study on NIHC/NWO study "Commanding Respect: How parents and peers contribute to the development of narcissistic aggression by youth (431-09-022). W. Hiemstra and E.E. Verhulp developed the design and performed data-collection. W.Hiemstra did the main literature search. S. Thomaes and B. Orobio de Castro provided advice and feedback, as well as critical revisions. W. Hiemstra drafted the manuscript, and E. E. Verhulp, B. Orobio de Castro and S. Thomaes provided critical revisions.

Abstract

How do children with aggression problems view themselves? The present research seeks to answer this question by examining the self-views (i.e., self-esteem and narcissistic traits) of boys referred for disruptive behavior problems. We used a multi-method assessment approach: In Study 1 ($N=85$, $Mage = 10.8$ years), we examined self-reported and parent-reported aggression; in Study 2 ($N=73$, $Mage=11.8$ years), we examined teacher-reported aggression. We found considerable support for a link between narcissism and aggression, which confirms previous findings in population-based samples of youth. Children with narcissistic traits were more aggressive according to both children themselves and their parents, and these links were independent of children's level of self-esteem. Moreover, as predicted, self-esteem was not associated with aggression. We encourage scholars to explore the possibility that interventions that target narcissistic self-views or its associated characteristics (e.g., perceived superiority, sensitivity to negative feedback) can effectively reduce aggression in subsets of referred boys.

Key words: Aggression; Self-Views; Narcissism; Self-Esteem; Conduct problems; Psychopathology

General Introduction

It has long been assumed that children with disruptive behavior problems like aggressive behaviors are prone to suffer from negative self-views. Accordingly, attempts to reduce children's aggressive behavior problems often seek to enhance children's level of self-esteem (e.g. Carr, 2006; Kusché & Greenberg, 1994; Mann, Hosman, Schaalma, & de Vries, 2004; Schonfeld et al., 2014). The literature on typically developing children, however, suggests that aggression may also stem from "threatened egotism" – inflated, narcissistic self-views that are jeopardized, rather than from low self-esteem per se (Barry, Thompson, Barry, Lochman, Adler, & Hill, 2007; Baumeister, Smart, & Boden, 1996; David & Kistner, 2000; Thomaes, Bushman, Stegge, & Olthof, 2008). One might be tempted to infer, then, that intervention-techniques that seek to boost children's self-esteem in order to reduce aggression are ill-advised. Unfortunately, however, research on the self-views of children referred for disruptive behavior problems – those for whom intervention is urgent – is scarce. Findings obtained in population samples of youth do not necessarily generalize to at risk or clinical samples: There are important differences in both individual (i.e., temperament, social cognition) and contextual (i.e., family and peer interaction) determinants and correlates of aggressive behavior in these groups (e.g. Dodge, 2006; Olson, 1992; Patterson, 1976). The present research seeks to contribute to filling this lacuna by examining the self-views (i.e., self-esteem and narcissistic traits) of boys referred to special education for their disruptive behavior problems.

Self-esteem and aggression

Self-esteem refers to children's global feelings of self-worth (Harter, 2012; Rosenberg, 1965). There are good reasons why it was long assumed that aggressive youth have low self-esteem. For example, low self-esteem may weaken ties to society, leading to nonconformity to social norms (Rosenberg, 1965). Also, children may be likely to protect themselves against negative self-feelings by externalizing blame for their own perceived shortcomings (Thomaes, Stegge, Olthof, Bushman, & Nezlek, 2011; Tracy & Robins, 2003). However, the link between self-esteem and aggression is empirically controversial. In population samples of youth, some studies found evidence for a small link between low self-esteem and aggression (e.g., Donnellan, Trzesniewski, Robins, Moffitt, & Caspi, 2005;

Fergusson & Horwood, 2002; Rosenberg, Schooler, & Schoenbach, 1989), but other research has suggested that such links may be at least partially accounted for by third variables (e.g., family characteristics such as parental education level) associated with both self-esteem and aggression, and thus, that it is not low self-esteem per se driving the effects on aggression (Boden, Fergusson, & Horwood, 2007). Still other research found no links between low self-esteem and aggression at all (e.g., Barry et al., 2007; Barry, Frick, & Killian, 2003; Thomaes, Bushman, Stegge, & Olthof, 2008), or found that inflated, unrealistically positive (rather than deflated) self-views are linked to aggressive behavior (David & Kistner, 2000; Orobio de Castro, Brendgen, van Boxtel, Vitaro, & Schaeppers, 2007). Thus, in population-based samples, there is no evidence for a substantial or robust link between low self-esteem and aggressive behavior. Research into the self-esteem-aggression link in clinical samples of children, however, is lacking.

Narcissism and aggression

Narcissism refers to a sense of grandiosity and entitlement, and a strong need to be seen and admired by others (Thomaes & Brummelman, 2016; Thomaes, Brummelman, Reijntjes, & Bushman, 2013). Although extreme levels of narcissism may reflect pathology, narcissism is more typically conceived of as a trait on which children in the general population vary. One way in which trait narcissism differs from self-esteem, is that it involves unrealistically positive self-views and a strong sense of superiority, with a need for validation of these feelings by others (Brummelman, Thomaes, & Sedikides, 2016; Bushman & Baumeister, 1998; Raskin et al., 1991). Narcissism has long been thought to predispose individuals to aggress (Kernberg, 1975). Indeed, when narcissists' fragile positive self-views are threatened (e.g., when they are criticized, rejected, or ridiculed) they tend to lash out aggressively, perhaps in an attempt to protect their ego (Baumeister et al., 1996; Kernis, 2003; Morf & Rhodewalt, 2001; Thomaes, Bushman, Stegge, & Olthof, 2008). Studies have shown consistent links between narcissism and various forms of aggressive behavior in population samples of children and adolescents (e.g., Barry et al., 2003; Bukowski, Schwartzman, Santo, Bagwell, & Adams, 2009; Golmaryami & Barry, 2009; Smith, Lynch, Stephens, & Kistner, 2015; Thomaes et al., 2008).

Importantly, self-esteem and narcissism are not two sides of the same coin, and show only weak to moderate overlap (Brummelman et al., 2015; Thomaes & Brummelman, 2016).

For example, narcissism differs from self-esteem in its characteristic cognitions (i.e., perceived superiority versus satisfaction with oneself) and origins (i.e., parental overvaluation versus parental warmth; Brummelman et al., 2016). Indeed, clinical theory and research suggest that there might be different subtypes of narcissists, who differ in their level of self-esteem (e.g., Kernberg, 1975; Kohut, 1977), suggesting that self-esteem might interact with narcissism in predicting aggression.

Self-views and aggression in clinical samples

From a clinical perspective, research in population samples is informative, but it is more important to know how children with the most serious aggression problems – i.e., those who receive treatment – typically view themselves. These children are not only more likely to exhibit personal vulnerabilities (e.g., low IQ, difficult temperament, attention problems, or impulsivity; e.g., Dodge, 2006), they also tend to have more prevalent histories of aversive social experience (e.g. coercive family interactions, peer rejection, neglect; e.g., Ettekal & Ladd, 2015; Olson, 1992; Smith et al., 2014; Patterson, 1976). These factors may well impact the link between children’s self-views and their aggressive behaviors. For example, parental neglect or lack of warmth is known to both undermine healthy self-esteem development and cultivate children’s aggressive behavior (Brummelman, et al., 2015; Cicchetti, 2016) suggesting that low self-esteem may be a more potent risk factor for aggressive behavior in clinical samples of youth.

A small body of evidence suggests that boys with co-occurring ADHD (attention deficit hyperactivity disorder) and aggression problems tend to display positive illusory self-perceptions (Gresham, MacMillan, Bocian, Ward, & Forness, 1998; Hoza, Pelham, Milich, Pillow, & McBride, 1993; Hoza et al., 2004). Furthermore, research has found that subsets of children with ASD (autism spectrum disorders), who often show deviant self-view development due to the difficulties they experience incorporating social feedback in their self-views (Lombardo & Baron-Cohen, 2010), display aggression (Farrar-Schneider, 1992). It is interesting to explore, then, if and how self-views are associated with aggression in this group. Our study will be the first to examine these links in a group of clinically referred children with disruptive behavior problems.

This research and hypotheses

The present research investigates the links between self-esteem, narcissism, and aggression in clinical samples of children that are referred to special education for disruptive behavior problems. In doing so, we used a multi-method assessment approach: In Study 1, we examined self-reported and parent-reported aggression as outcome measures; in Study 2, we examined teacher-reported aggression as outcome measure. We hypothesized that narcissism, but not self-esteem, would be positively linked to aggression. We also explored the potential moderating role of self-esteem. Finally, in both studies we explored whether the links between self-views and aggression differ for two prevalent subgroups of clinically referred children: Those who show co-occurring autism spectrum disorder (ASD), or another behavioral disorder (e.g., ODD, ADHD).

Study 1

Methods

Participants

Participants were 85 Dutch boys aged 8 to 13 (mean age = 10 years, 8 months). They were recruited from seven schools providing special elementary education for children with behavior problems. In the Netherlands, children are exclusively referred to this form of special education if the severity of their behavior problems significantly impairs social functioning and prohibits participation in regular education, as judged by parents, teachers, and diagnosticians. About 2,5% of the school-aged population is referred to this type of education (Statistics Netherlands, 2014). Participating schools were situated in rural and suburban areas in the centre and south of the Netherlands, but boys came from across the country. Participants were diagnosed with a DSM-IV/V disorder by multidisciplinary teams of professionals (obtained from school care records; 30.6% ADHD only, 8.2% ODD or CD only, 7.1% both ADHD and ODD or CD, 25.9% ASD, 20% ADHD and ASD, 8.2% other).

Measures

Instrument for Reactive and Proactive Aggression. We measured aggressive behavior using the Instrument for Reactive and Proactive Aggression (IRPA; Polman, Orobio de Castro, Thomaes & van Aken, 2009). Both children and parents completed this measure. The IRPA consists of seven “form” items (e.g. “How often did the child/you hit someone in

the past 6 months?”), and six “function” items (e.g. “How often did he do this because he was angry?”). For this study, only form-items were used. Items are rated along a 5-point Likert scale ranging from 0 (never) to 4 (multiple times). Items were averaged to obtain a total aggression score. Good discriminant, convergent, and construct validity have been demonstrated for the IRPA (Polman et al., 2009). Alpha in the current sample for child-report and parent-report aggression was good (Cronbach’s $\alpha = .79$ and $.82$, respectively).

Childhood Narcissism Scale. We measured narcissism using the Childhood Narcissism Scale (CNS), a one-dimensional self-report measure of relatively stable individual differences in childhood narcissism (Thomaes et al., 2008). Ten items (e.g., “I am a great example for other kids to follow”) are rated along a 4-point Likert scale ranging from 0 (not at all true) to 3 (completely true). Responses were averaged across items, reliability in the current sample was good (Cronbach’s $\alpha = .81$).

Self-Perception Profile for Children. We measured self-esteem using the Global Self-Worth subscale of the Self-Perception Profile for Children (Harter, 1985). Children rated six items (e.g., “Some kids are happy with themselves as a person”) along a 4-point Likert scale ranging from 0 (I am not like these kids at all) to 3 (I am exactly like these kids). Responses were averaged across items, with acceptable reliability (Cronbach’s $\alpha = .67$).

Procedure

Schools distributed parental consent forms. Only boys who received active consent from their parents/caregivers were included in the study. We asked schools to distribute the consent forms among 250 children. 96 Children got active consent to take part in the study (participation rate = 38%), 85 of them were boys. Diagnostic information was retrieved from school care records. Children completed questionnaires individually under the supervision of a trained research assistant (i.e., a female psychology student). After completing the questionnaire, we gave children stickers to thank them for their voluntary participation.

Missing values

Eight parents returned incomplete or no IRPA questionnaires. Data from these parents were imputed using multiple imputation in SPSS.

Results

Table 1 shows the correlations between the main variables in Study 1. Aggression reported by children and their parents was positively correlated. We calculated a composite

aggression score to be used in the main analyses. Participants with an ASD diagnosis did not differ significantly from participants without ASD diagnosis in self-esteem and narcissism ($p = .68$ and $p = .58$, respectively).

Table 1 | *Descriptive Statistics and Correlations.*

	Mean	SD	Aggression Parent	Aggression Child	Aggression Composite	Narcissism
Aggression Parent	0.99	.66	-	-	-	-
Aggression Child	1.00	.63	.39**	-	-	-
Aggression Composite	0.00	.85	.83**	.83**	-	-
Narcissism	1.40	.62	.22	.29**	.29**	-
Self-Esteem level	2.22	.49	-.11	-.18	-.18	.01

Notes. * $p < .05$; ** $p < .01$; Aggression Parent is child aggression rated by parents; Aggression Child is self-reported child aggression; Aggression Composite is the composite of parent and child rated child aggression.

We used regression analyses to examine the association between narcissism and aggression, and the hypothesized moderating role of self-esteem. The Aggression aggregate served as the dependent variable. Narcissism and Self-Esteem were centered and entered in Step 1, and their interaction in Step 2. A positive main effect of Narcissism on Aggression was found; as predicted, more narcissistic boys were more aggressive (see Table 2). No main effect of self-esteem was found. The interaction between Narcissism and Self-Esteem was also not significant.

Table 2 | *Summary of Regression of Aggression on Narcissism and Self-Esteem (N=85).*

Model / Variables	R^2	B	SE B	β
Model 1	.12**			
Narcissism		.39	.14	.29**
Self-Esteem		-.31	.18	-.18†
Model 2	.12*			
Narcissism		.37	.14	.28**
Self-Esteem		-.32	.18	-.19†
NarcxSE		.22	.36	.08

* $p < .05$; ** $p < .01$; † $p < .10$

We further explored whether the link between Narcissism and Aggression would be moderated by ASD diagnosis. We performed an additional regression analysis which included an interaction term with ASD. No main or interaction effects involving ASD were found ($\beta = -.13$, $t(81) = -1.27$, $p = .20$, and $\beta = -.06$, $t(79) = -.54$, $p = .59$, respectively).

Discussion

In Study 1, we examined the associations between self-views and aggression for a clinically referred group of boys. More aggressive children showed higher levels of narcissism, but not lower levels of self-esteem. We found no evidence that the link between narcissism and aggression was contingent upon children's level of self-esteem, or diagnosis with ASD.

Our findings are largely consistent with previous research in general population samples of youth, showing that higher levels of narcissism are associated with aggressive behavior (e.g., Barry et al., 2003; Bukowski et al., 2009; Golmaryami & Barry, 2009; Smith et al., 2015; Thomaes et al., 2008). The absence of a moderating role of self-esteem, however, is somewhat surprising, given clinical theory and research that do suggest that different subtypes of narcissists (characterized by differential levels of self-esteem) may show different levels of aggression (e.g., Kernberg, 1975; Kohut, 1977; Bushman et al., 2009). One possibility is that narcissistic aggression is mainly rooted in fluctuating self-esteem (e.g., triggered by feelings of threat or shame), rather than low self-esteem per se. In Study 2, we therefore included an index of self-esteem stability as a potential correlate of aggression and moderator of the narcissism-aggression link.

Study 2

The goal of Study 2 was to obtain a more comprehensive understanding of the links between self-views and aggression among clinically referred children, by (1) examining associations between children's self-views and teacher reported aggression, and (2) examining stability of self-esteem as an additional potential correlate and moderator of aggression.

Self-esteem stability refers to the magnitude of short-term fluctuations in immediate feelings of self-worth (Kernis, 2005). Self-esteem instability is a trait-like characteristic that is (negatively) associated with, but distinct from, level of self-esteem (Webster, Kirkpatrick,

Nezlek, Smith, & Paddock, 2007). Self-esteem stability has been considered a proxy for the vulnerability of children's self-views. Studies in adults have shown that individuals with high yet unstable self-esteem tend to be more hostile compared to those with (stable or unstable) low self-esteem (Kernis, 2005; Kernis, Grannemann, & Barclay, 1989).

Furthermore, research in adults has found some evidence that self-esteem stability is a unique predictor of aggression (Falkenbach, Howe, & Falki, 2013). We hypothesized that self-esteem stability would be a unique predictor of aggression, and would also act as a moderator of the link between narcissism and aggression.

Methods

Participants

The sample consisted of 73 Dutch boys aged 8 to 13, with a mean age of 11 years and 8 months. These boys were recruited from four schools providing special elementary education for children with disruptive behavior problems (i.e., the same school type as in Study 1), located in rural and suburban regions in the Netherlands. The schools were different from those taking part in Study 1. Participants were diagnosed with a DSM-IV/V disorder by multidisciplinary teams of professionals (obtained from their school care records; 17.8% ADHD only, 8.2% ODD or CD only, 8.2% both ADHD and ODD or CD, 27.4% ASD, 17.8% ADHD and ASD, 20.6% other).

Measures

As in Study 1, participants completed the Childhood Narcissism Scale (Thomaes et al., 2008; Cronbach's $\alpha = .71$), and the Self-Perception Scale for Children (Harter, 1985; Cronbach's $\alpha = .82$). For self-esteem stability, we used a slightly adapted version of this latter measure, by asking children to report on a day-to-day basis, for 5 consecutive days, how they feel "right now". We found acceptable to good reliability for the daily self-esteem assessments (α s ranging between .67 and .82). Following standard procedures (e.g., Kernis, Granneman, & Barclay, 1989) self-esteem stability was calculated as the standard deviation of state self-esteem scores across days. Teachers completed the Instrument for Reactive and Proactive Aggression (IRPA; Polman et al., 2009; Cronbach's $\alpha = .92$).

Procedure

We sent out parental consent forms via the schools. Only boys who received active consent from their parents/caregivers were included in the study. We asked schools to

distribute the consent forms among 144 boys. 87 boys took part in the study (participation rate = 60 %). Diagnostic information was retrieved from school care records. This study was part of a larger intervention study. Participants completed daily assessments of self-esteem in individual testing sessions at the school grounds. The other questionnaires were completed one week earlier in class. Participants received a small present (i.e., a pen or eraser) at the third and the fifth day to thank them for voluntary participation. We obtained ethics approval from our local ethics review board at Utrecht University, Faculty of Social and Behavioural Sciences.

Missing values

For 14 boys complete questionnaires were missing, and were excluded from the study. The questionnaires were missing because of illness or therapy-related school absence.

Results

Table 3 shows the correlations between the main Study 2 variables. No significant associations were found between any of the self-view measures and teacher-reported aggression. Self-Esteem Level and Self-Esteem Stability were significantly negatively correlated; participants with lower self-esteem showed more self-esteem instability. Again, no significant differences were found on any of the self-view variables for participants with and without ASD diagnosis.

Table 3 | *Descriptive Statistics and Correlations.*

	<i>Mean</i>	<i>SD</i>	<i>Aggression Teacher</i>	<i>Narcissism</i>	<i>S-E Level</i>
Aggression Teacher	.91	.88	-	-	-
Narcissism	1.37	.53	-.14	-	-
S-E Level	2.24	.71	-.18	-.03	-
S-E Stability	.30	.28	.04	.08	-.33**

Notes. * $p < .05$; ** $p < .01$; Aggression Teacher is child aggression rated by teachers; S-E is Self-Esteem.

To examine the links between narcissism and teacher-reported aggression and the moderating role of self-esteem level and stability, we performed two separate regression analyses. As in Study 1, we regressed Aggression on the centered variables Narcissism and Self-Esteem to analyze any main effects in Step 1, and added the Self-Esteem Level and

Stability interaction-terms with Narcissism to these models in Step 2.

We found no main effects for Narcissism or Self-Esteem Level, nor did we find an interaction between these variables (see Table 4). Again, we explored whether having an ASD diagnosis would influence the findings by adding ASD and its interaction with Narcissism to the regression equation (45% of the participants were diagnosed with ASD). A main effect was found ($\beta = -.32$, $t(69) = -2.84$, $p < .01$)—teachers reported higher levels of aggression for children without an ASD diagnosis—but the interaction was non-significant ($\beta = .07$, $t(67) = .58$, $p = .57$).

Table 4| *Summary of Regression of Aggression Reported by Teachers on Narcissism and Self-Esteem Level (N=73).*

Model / Variables	R^2	B	$SE B$	β
Model 1	.06			
Narcissism		-.25	.19	-.15
Self-Esteem		-.24	.15	-.19
Model 2	.06			
Narcissism		-.15	.19	-.15
Self-Esteem		-.24	.15	-.19
NarcxSE		.02	.27	.01

* $p < .05$; ** $p < .01$; † $p < .10$

In the second analysis, the main and interaction effects of Self-Esteem Stability were not significant (Table 5). In the additional analysis the same main effect was shown for children with ASD ($\beta = -.32$, $t(69) = -2.74$, $p < .01$) and no interaction between ASD and Narcissism ($\beta = .01$, $t(67) = .08$, $p = .94$).

Table 5| *Summary of Regression of Aggression Reported by Teachers on Narcissism and Self-Esteem Variability (N=73).*

Model / Variables	R^2	B	$SE B$	β
Model 1	.02			
Narcissism		-.24	.19	-.15
Self-Esteem Stability		.15	.37	.05
Model 2	.02			
Narcissism		-.25	.20	-.15
Self-Esteem Stability		.18	.40	.06
NarcxSEstab		-.15	.83	-.02

* $p < .05$; ** $p < .01$; † $p < .10$

General Discussion

This research is the first to examine the links between multiple aspects of children's self-views and aggression in boys referred for disruptive behavior problems. We used a multi-method approach to study the links between narcissism, self-esteem and aggression in two independent samples. We found considerable support for a link between narcissism and aggression, which confirms previous findings in population-based samples of youth (e.g., Barry et al., 2003, Bukowski et al., 2009; Golmaryami & Barry, 2009; Smith et al., 2015; Thomaes et al., 2008). Children with narcissistic traits were more aggressive according to both children themselves and their parents, and these links were independent of children's level of self-esteem. Moreover, as predicted, self-esteem was not associated with aggression (e.g., Barry et al., 2007; Barry, Frick, & Killian, 2003; Bushman et al., 2009; Thomaes, Bushman, Stegge, & Olthof, 2008).

Our findings for teacher reported aggression were somewhat inconsistent with the other findings. In particular, we found no evidence that children with narcissistic traits were considered more aggressive by their teachers. One explanation may be found in the situation specificity of children's aggression. Narcissistic aggression is typically triggered in situations in which children experience ego-threat (i.e., interpersonal contexts; e.g., Baumeister, Smart, & Boden, 1996). Such aggression may be less likely to occur in the settings in which special education teachers typically observe their students: Highly structured and regulated, small group settings.

We also took the opportunity to explore whether relations between self-views and aggression differed for those with and without ASD diagnosis. No differences were found, except that teachers reported higher levels of aggression for children without ASD. Thus, our research suggests that the link between self-views and aggression generalizes across children with and without ASD diagnosis.

This research has limitations. Although we used a multi-method approach, we relied on reported rather than behavioral measures of aggression. Observing the link between self-views and aggression in an *in vivo* setting (e.g., Thomaes, Bushman, Stegge, & Olthof, 2008) would be a valuable complement to this work, tapping the interpersonal context in which narcissistic aggression may be most apparent. Besides that, we used cross-sectional data to look at factors related to aggressive behavior. To conclude anything about causality,

experimental or longitudinal research should confirm our findings.

This research also has several strengths. First, it is the first to comprehensively test the links between self-views and aggression in samples of referred youth. Second, we used multiple informants to test our hypotheses. Aggression is highly context-dependent, and so the inclusion of multiple informants helps to validly assess individual differences in aggression and inform intervention possibilities.

Future research should further explore the links between self-views and aggression in clinically referred samples, preferably using longitudinal designs. Because clinical populations tend to be heterogeneous, differential developmental trajectories should be explored. For example, whereas deviant self-views may be an antecedent of aggression for some children, they may be a consequence of children's aggressive predispositions as well. Furthermore, there are different forms and functions in aggression (Polman et al., 2007). Some narcissists may use more proactive, instrumental aggression (e.g., because they feel they can gain status from behaving aggressively; Seah & Ang, 2008), others might reactively aggress because they see others as being hostile and unjust (e.g., Fossati, Borroni, Eisenberg, & Maffei, 2010).

Research has shown that children with positive views of themselves but negative views of their peers ("I'm OK but you're not") are especially aggressive (Salmivalli, Ojanen, Haanpää, & Peets, 2005). It may be that narcissistic children, who typically hold such discrepant self- and peer-views, consider their aggression as a justified response to perceived hostility. This would be consistent with experimental research on the mediating role of hostile intent attribution in the link between narcissism and aggression (Reijntjes et al., 2011). This would suggest that rather than boosting children's level of self-esteem (perhaps unwittingly contributing to the "I'm OK" part of the aggression equation), interventions would better focus on helping children adopt a more realistic and accepting view of one's strengths and weaknesses (Thomaes & Brummelman, 2016; Thomaes et al., 2009).

Our research challenges the notion that boosting aggressive children's level of self-esteem should be a common aim for intervention. Indeed, even among referred children, there is little evidence that aggressive children typically suffer from low self-esteem. We encourage scholars, instead, to explore the possibility that interventions that target

narcissistic self-views or its associated characteristics (e.g., perceived superiority, sensitivity to negative feedback) are effective to reduce aggression in subsets of referred boys.

CHAPTER 3

DAY-TO-DAY COVARIATION OF STATE ANGER AND HOSTILE ATTRIBUTIONS IN CHILDREN WITH DISRUPTIVE BEHAVIOR PROBLEMS

W. Hiemstra developed the study concept, developed the design and performed data-collection. B. Orobio de Castro and S. Thomaes gave advice and feedback during this process. W. Hiemstra did the main literature search and performed the data-analysis. W. Hiemstra, B. Orobio de Castro and S. Thomaes performed interpretation. W. Hiemstra drafted the manuscript, and B. Orobio de Castro and S. Thomaes provided critical revisions.

Abstract

Children with aggressive behavior problems experience anger more often than other children. The frequent experience of anger itself, can in turn impair children's social information processing and social behaviors. To better understand this interplay, the current research used a naturalistic approach by testing day-to-day associations between emotion and cognition in boys with severe aggression problems. The aim of the present study was to assess whether boys with aggressive behavior problems attribute more hostility to others' facial expressions on days they experience increased anger themselves. Our secondary goal was to analyze possible moderation of function of aggression at the between-subjects level. Participants were 76 boys ages 8 to 14 ($M_{age} = 11.6$ years), referred to special education for their disruptive behavior problems. We used multilevel analysis in Mplus to examine covariation of affective and cognitive processes at the within-subjects level. Contrary to our hypothesis, we found that day-to-day fluctuations in children's hostile interpretations were independent of fluctuations in state-anger. Furthermore, we found no evidence that the covariation between state-anger and hostile interpretation depends on children's function of aggression predispositions. Our results do show day-to-day variation in hostile interpretation of facial expressions. Future research should try and catch these processes in a shorter time period, and look at specific subgroups of boys for whom this covariation might exist.

Key words: Anger; Hostile Interpretation; Aggression; Social information processing; Covariation

Introduction

Children with disruptive behavior problems experience anger more often than other children (e.g. Schultz, Izard, & Bear, 2004). Although the exact reasons may be diverse, such anger proneness is typically due to sustained transactions between child temperament and negative life-events, such as peer rejection, negative parent-child interaction, and interparental conflict (e.g. Crockenberg & Langrock, 2001; Dodge, et al., 2003; Patterson, 1976). The frequent experience of anger itself, can in turn impair children's social information processing and social behaviors (Lemerise & Arsenio, 2000). To better understand this interplay, the current research uses a naturalistic approach and tests day-to-day associations between emotion and cognition in boys with disruptive behavior problems.

Both the General Aggression Model (GAM; Allen, Anderson, & Bushman, 2018) and the Social Information Processing model (SIP; Crick & Dodge, 1994) suggest how aggressive behavior results from an interplay of biological predispositions, knowledge structures, emotional state, and social-cognitive processes. Knowledge structures are built over the course of development, and influence the way children perceive and interpret social stimuli. One of the most important social-cognitive deviances that is linked to aggression, is hostile attribution of intent or hostile attribution bias (Crick & Dodge, 1994; De Castro et al., 2002; Dodge et al., 2015): An automatic tendency to interpret ambiguous social cues as hostile (Crick & Dodge, 1996; Dishion, French, & Patterson, 1995; Dodge, 2006).

Emotionality is one of the biological predispositions that can influence social information processing (Lemerise & Arsenio, 2000). Children differ in the intensity of the emotions they experience, as well as in their emotion expression and regulation predispositions, which influences their social behavior. The "mood congruency effect" describes the tendency for people to render judgements in the direction of their own mood state (e.g. Bodenhausen & Sheppard, 1994). According to this effect, negative mood can activate knowledge structures that are associated with that affective state. Increased familiarity with experiencing and exposure to a particular emotion, may strengthen the available knowledge structures (Izard, 1977, in Schultz, Izard, & Bear, 2004). When individuals experience negative mood, like anger, and have a set of readily available negative knowledge structures, this can breed a negative interpretation of social stimuli. In children with aggressive behavior problems, an angry emotional state could thus trigger hostile schemata and hostile attributions of intent, which may instigate aggression. Such aggression

in turn may provoke hostile reactions from others which will further increase one's anger, thereby potentially triggering a negative recursive cycle (Dodge, 2006).

Although many studies have reported on associations between trait-anger (i.e. anger-proneness) and hostile attribution (e.g. Anderson & Bushman, 2002; Schultz, Izard, & Bear, 2004), few studies have directly studied the influence of state-anger, or angry mood, on hostile attribution of intent. Dodge and Somberg (1987) studied groups of aggressive-rejected and nonaggressive-popular boys, who rated intentions of peers in hypothetical situations before and after a manipulation. Results showed that the manipulation triggered more hostile intent attributions in the aggressive-rejected group. However, the manipulation could have triggered both negative affect and representations of hostile intent, because it was threat-related. This finding led Orobio de Castro and colleagues (Orobio de Castro, Slot, Bosch, Koops, & Veerman, 2003) to do an experiment in which the emotion induction was unrelated to threat. They found similar results: Negative feelings exacerbated attribution of hostile intent in highly aggressive boys. In the present study, we extend this work by testing the links between aggressive children's negative feelings (i.e., anger) and hostile intent attributions in a naturalistic setting, also allowing for a test of recursive effects over time.

Daily diary methods are the method of choice to examine dynamic emotional and cognitive processes as they occur in daily life (Bolger, Davis, & Rafaeli, 2003; Hamaker, Ceulemans, Grasman, & Tuerlinckx, 2015). The aim of the present study was to assess whether boys with disruptive behavior problems attribute more hostility to others' facial expressions on days they experience increased anger themselves. Earlier studies have shown that experimentally induced negative emotion increases hostile attribution of intent (e.g. Orobio de Castro et al., 2003). Also, research has shown that the mood-congruent effects can have particularly strong effects on constructive social judgements, such as perception of facial expressions (Forgas & East, 2008). We obtained daily reports of state-anger and emotion attribution. Boys with disruptive behavior problems were included to increase clinical relevance. There are important differences with population samples in both individual (i.e. temperament, emotion regulation) and contextual (i.e., family and peer interaction) determinants and correlates of aggressive behavior in this group (e.g. Dodge, 2006; Olson, 1992; Patterson, 1976). Therefore, findings obtained in population samples do not necessarily generalize to clinical samples. Our study exclusively included boys. It would have been valuable to sample girls as well, but earlier studies show that there are important

differences between boys and girls in expression and consequences of anger (Buntaine & Costenbader, 1997; Card, Stucky, Sawalani, & Little, 2008). We sampled children ages 8 to 12, since behavioral and emotional self-regulation are important aspects of competence during these years (Shields, Cicchetti, & Ryan, 1994).

The secondary aim of the study was to explore whether a key individual difference characteristic—the function of aggressive behaviors that children typically engage in—would influence the day-to-day link between anger and attribution. Theoretically, children’s reactive aggression (hostile), but not proactive aggression (instrumental; Crick & Dodge, 1996), should be anger- and hostile attribution-driven. This hypothesis has been supported in a number of studies (e.g. Hubbard, Dodge, Cillessen, Coie, & Schwartz, 2001; Orobio de Castro, Merk, Koops, & Veerman, 2005). Children showing higher levels of reactive function of aggression therefore are quicker influenced by their own anger and thus show more hostile intent. This suggests that the relation between state-anger and hostile attribution of intent is stronger in these children. Therefore, we hypothesized that the covariation between anger and hostile attribution will depend on the child’s level of reactive aggression (and not proactive aggression).

Method

Design

A repeated measures design was used, with daily assessments of state-anger and hostile attributions over days within the same week (3-5 measurements per participant).

Participants

Participants were 76 boys ages 8 to 14 ($M_{age} = 11$ years and 6 months). They were recruited from four schools providing special elementary education for children with behavior problems. In the Netherlands, about 2.5% of the school-aged population is referred to this type of school (Statistics Netherlands, 2014). Criteria for referral include: (1) Children meet criteria for at least one DSM-IV/V disorder, (2) the severity of children’s behavior problems falls within a clinical range and prohibits participation in regular education, and (3) children are treated for their problems in a clinical setting. A multidisciplinary examination committee decides on placement in this type of school. Education is provided in small classes (of about 10 children), by specialized teachers trained to provide individualized supervision and guidance. The participating schools are situated in rural and suburban areas

in the middle and south of the Netherlands, although boys came from across the country.

Our sample of participants was taken from a larger study that tested the effectiveness of a cognitive bias modification training (see Hiemstra, Orobio de Castro, & Thomaes, 2018). Only boys in the active control condition (i.e., those who did not receive the training) were included in the present analyses. The initial sample of boys in this control condition consisted of 78 participants, but 2 of them dropped out of the study (they were absent from school on most days due to illness).

Measures

Hostile attribution. We assessed hostile attribution following a recently developed computer-based approach (Penton-Voak et al., 2013). Participants were presented a set of pictures of facial expression of emotion. Specifically, from two pictures of non-ambiguous happiness and anger expressions, we created a morphed continuum consisting of 15 pictures which gradually differ in the extent to which they are ambiguously happy or angry (see Figure 1).

Pictures were presented in random order, and displayed for 500 ms each. Participants responded to a two-alternative forced-choice judgment whether the facial expression was happy or angry. These responses enabled us to obtain an estimation of the “balance point,” i.e., the point along the continuum of facial expressions from which a participant’s happy attributions shift to anger attributions. This balance point indexes hostile attribution. It was calculated as the number of ‘angry’ responses as a proportion of the total number of trials. For a detailed description of the training procedure, see Hiemstra, et al. (2018).

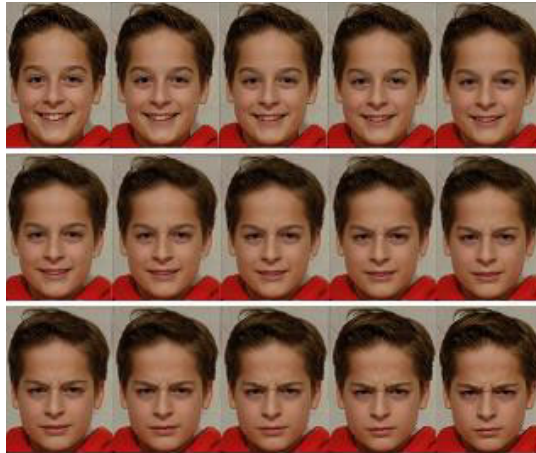


Figure 1. Example of One of the Morphing Continua: Unambiguously Happy, Ambiguous, and Unambiguously Angry (From Left to Right, Top to Bottom, Respectively).

Emotional state. State anger and state positive mood were measured every day immediately before the training by asking participants to indicate “how you feel right now” on an “anger thermometer” and a “happy thermometer” (e.g. Elias, 2004). Responses were rated along a 4-point scale, ranging from 0 (“*not at all angry*”) to 3 (“*very angry*”).

Function of aggression. Reactive and proactive functions of aggression were measured with the Instrument for Proactive and Reactive Aggression (IRPA; Polman, Orobio de Castro, Thomaes & van Aken, 2009). This is a teacher-report instrument that assesses forms and functions of aggression separately. The IRPA contains 8 form items (e.g. “How often did the child hit someone in the past week?”), and 7 function items (e.g. “How often did he do this because he was angry?”). Responses are rated along a 5-point scale ranging from 0 (*never*) to 4 (*multiple times a day*) for form; and from 0 (*never*) to 4 (*always*) for function. For the purposes of this study, we chose not to distinguish functions for each aggression form separately, yielding total scores for function exclusively (and thus shortening the measure). Alpha’s in the current sample were adequate ($\alpha_{\text{proactive aggression}} = 0.76$, and $\alpha_{\text{reactive aggression}} = 0.71$; Cohen, 1988). There was a significant correlation ($r = .52$, $p < .01$) between proactive and reactive aggression in our sample, indicating 25% shared variance between the concepts.

Procedure

We obtained ethics approval from the Utrecht University Institutional Review Board. Schools for special education for children with disruptive behavior problems across the Netherlands were requested to take part in the study. The schools distributed consent letters to parents/caregivers -- only boys who received active parental/caregiver consent were allowed to take part in the study (consent rate was approximately 68%).

On each of the days of data collection, boys were taken from their class at a random moment of the day and completed the mood thermometers prior to the computer training. Prior to the week of data collection, teachers completed the IRPA questionnaire. Boys received a small present (a pen or eraser) to thank them for their participation.

Data analysis

We used multilevel analysis in Mplus (Muthén & Muthén, 2015) to analyze our data. Our primary goal was to examine co-variation of affective and cognitive processes at the within-subjects level (Level 1); our secondary goal was to analyze possible moderation of function of aggression at the between-subjects level (Level 2). We used Maximum Likelihood Estimation to analyze our models.

From six participants, complete aggression measures were missing due to teachers not completing the questionnaire. Therefore, analysis on Level 2 involved $N=70$ participants. The six participants with missing aggression measures did not differ on their variation in hostile attributions compared to the group with complete aggression measures ($F(1,74) = .120, p = .73$).

Results

Descriptive statistics

Descriptive statistics and correlations for the main study variables are presented in Table 1. Surprisingly, hostile interpretation and state-anger at measurement 1 (T1) were not correlated. Moreover, mean variations over time (as reflected in individuals' SDs) in state-anger and hostile interpretation were not correlated.

Table 1 | *Descriptive Statistics and Correlations.*

	<i>Mean</i>	<i>SD</i>	<i>HI T1</i>	<i>S-A T1</i>	<i>SD Hostile Int.</i>	<i>SD S-A.</i>	<i>Reactive Aggr.</i>
Hostile Int. T1	.49	.10	-	-	-	-	-
State-Anger T1	1.32	.66	.03	-	-	-	-
SD Hostile Int.	.09	.04	-.02	.09	-	-	-
SD State-Anger	.33	.41	-.07	.48**	.06	-	-
Reactive Aggression	1.35	1.05	.21	-.10	-.12	-.002	-
Proactive Aggression	.96	.93	.08	-.06	-.17	.01	.52**

* $p < .05$; ** $p < .01$; Hostile Int. T1 is the first measure of hostile interpretation (balance point), State-Anger T1 is the first measure of state anger.

Main Analyses

We first tested an unconditional model (random-intercept) to check individual differences in the variance of hostile interpretation over time. This is necessary to see whether there is variance that can be explained with other variables. If there is no variance to begin with, no further steps in the model are taken. With this unconditional model, it is possible to derive basic descriptive statistics as well as the proportion of variance that can be explained at each level. The average level of hostile interpretation in our sample is .50 on a scale from 0 to 1. The variance at the first (within-persons) level is .009 ($p < .001$), and variance at the second (between-persons) level is .003 ($p < .001$). So, total variance in our data is .012 (1.2%). Since the variance at both levels is a significant amount, further steps in the model are executed.

In Figure 2 we present an example of variation across assessments in hostile interpretation of 10 random participants. The figure shows how such variation is substantial for some, less so for others.

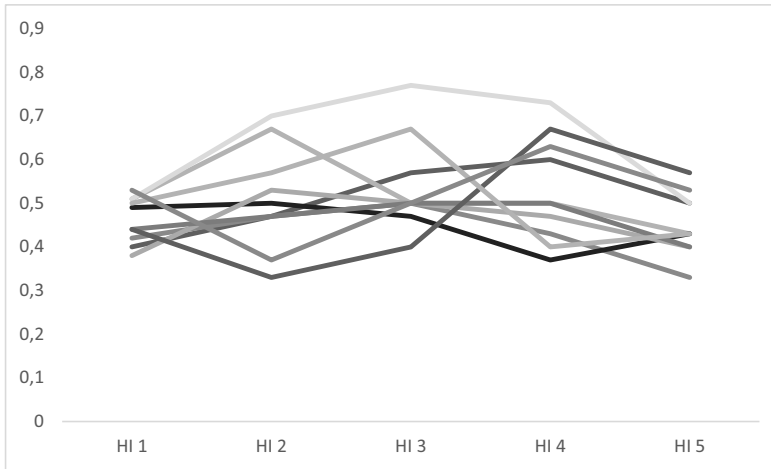


Figure 2. *Example of Variation in Hostile Interpretation over Days from Ten Randomly Chosen Participants.*

The second step in our multi-level model was to add our repeated measure of state-anger as a predictor to Level-1. This step was taken to see whether state-anger and hostile interpretation do co-vary across days. The coefficient of the slope (Level-1) is 0.007, $SE = 0.010$, $t = 0.731$, $p = 0.465$. Thus, at the within-person level, we do not find significant co-variation of hostile interpretation and state-anger.

Lastly, we added the moderator, reactive aggression, at Level-2 to see whether the individuals' slope would depend on the height of their reactive function score. The coefficient of reactive aggression is .002, $SE = 0.010$, $t = 0.208$, $p = 0.835$. Participants with higher levels of reactive aggression do not differ in their co-variation between state-anger and hostile interpretation. Also, adding proactive aggression instead of reactive aggression at Level-2 showed no significant effects. The coefficient of proactive aggression is .014, $SE = 0.020$, $t = 0.722$, $p = 0.470$.

Discussion

The aim of the present study was to assess whether state-anger in boys with aggressive behavior problems covaries with their interpretation of negative emotions in others. Contrary to our hypothesis, we found that day-to-day fluctuations in children's hostile interpretations were independent of fluctuations in state-anger. This finding is at odds with theory and previous experimental research (e.g. De Castro et al., 2003). Furthermore, we

found no evidence that the co-variation between state-anger and hostile interpretation depends on children's function of aggression predispositions.

One explanation for not finding co-variation between state-anger and hostile interpretation over time may be that there actually is a relation and co-variation over time between hostile attribution and state anger, but only for a small subgroup of children. We tested for a role of reactive aggression in identifying such a potential sub-group, but there are other potentially relevant personal characteristics or cognitive processes that that we did not test. For example, children scoring high on narcissistic traits tend to be reactive to perceived threat or rejection (e.g. Baumeister et al., 1996; Thomaes, Bushman, Stegge, & Olthof, 2008). When children high on narcissism experience such an event, this might thus trigger higher peaks of anger and hostile attribution in the moment. However, we did not measure daily stressors, and thus do not know whether such events took place in the week we measured our concepts. Another example comes from Wilkowski and Robinson (2008), who describe how hostile attribution, trait-anger and aggression are related and influenced by different cognitive processes. The authors describe how cognitive processes like effortful control can 'correct' automatic hostile attributions. This would mean that for children low on effortful control, the covariation between anger and interpretation might be stronger. Future research should include whether and how narcissism and effortful control influence covariation between anger and hostile interpretation.

An alternative explanation is that the time-frame of our within-person measurement was too broad. Although diary methods often use measurement across days (which we chose to adhere to), co-variation of the concepts we measured may actually occur in seconds or minutes rather than days. Affective states tend to fluctuate throughout the day, and might fluctuate even more quickly in children with decreased emotion regulation, such as in our sample (e.g. Chang, Schwartz, Dodge, & McBride-Chang, 2003). Thus, we may have simply missed fast cycles of co-variation between state-anger and hostile interpretation in our participants. A result from this explanation would be that children with aggression problems do not merely experience a bad day or a good day, but more so bad moments and good moments. Intervention should then focus on those peaks of negative states to influence cognition and behavior.

This study has some limitations. As mentioned above, we might not have caught the right time frame and could have measured the concepts in an even more real-time manner.

We tested the children in a highly structured setting which might have offered them a distraction. Valid real-time anger scores are hard to obtain from these children, but for future research one might consider measuring anger without taking participants out of their direct context (i.e. classroom). Furthermore, we did not use an exhausting measure of hostile attributions. We focused on hostile interpretation of facial expressions, since facial expressions have a key function in regulating social interaction and can convey signs of intent (Walker-Andrews, 1997). Future research should consider to measure hostile attributions in different ways to make sure to grasp the whole concept. Lastly, since we only focused on a clinical population with a small sample size, we might have had some reduction of range in our measurements. With this sample we might have had too little power to detect existing relations between the concepts. Future studies might include non-clinical groups to make a comparison in results.

This study also has different strengths. It is the first to focus on the within-person covariance between affect and cognition in children using a daily measures approach. It provides methodological information in how these concepts can be measured and analyzed on daily basis. Second, we worked with a clinical sample, for whom hostile schemata may be relatively rigid and more difficult to change (Dodge, 2006). Therefore, it is important to get insight in how these cognitive processes evolve over small periods of time, to find ways to improve existing interventions or form new ones. Lastly, we used ecologically valid pictures of boys in our measure of hostile interpretation. Although this might not capture the whole concept of hostile intent attribution in these children, social interaction with peers plays an important role.

In conclusion, to our knowing we are the first to study co-variation between affective and cognitive processes as within-person factors in a clinically relevant sample. Our results show day-to-day variation in hostile interpretation of facial expressions, which are not related to state-anger. Future research should try and catch these processes in a shorter time period, and look at specific subgroups of boys for whom this covariation might exist. This way, intervention can be informed on what specific process to tap on to reduce hostile attributions and in the end possibly reduce aggressive behavior problems.

CHAPTER 4

REDUCING AGGRESSIVE CHILDREN'S HOSTILE ATTRIBUTIONS: A COGNITIVE BIAS MODIFICATION PROCEDURE

An adapted version of this Chapter was published as:

Hiemstra, W., Orobio de Castro, B., & Thomaes, S. (2018). Reducing Aggressive Children's Hostile Attributions: A Cognitive Bias Modification Procedure. *Cognitive Therapy and Research*, 43(2), 387-398.

W. Hiemstra developed the study concept, developed the design and performed data-collection. B. Orobio de Castro and S. Thomaes gave advice and feedback during this process. W. Hiemstra did the main literature search and performed the data-analysis. W. Hiemstra, B. Orobio de Castro and S. Thomaes performed interpretation. W. Hiemstra drafted the manuscript, and B. Orobio de Castro and S. Thomaes provided critical revisions.

Abstract

Aggressive children tend to over-attribute hostile intentions to others, often in ambiguous social situations. Among clinically referred children with disruptive behavior problems, hostile schemata may be relatively rigid and more difficult to change due to prevalent histories of aversive social experience and/or personal vulnerability. The present study examined the effectiveness of a cognitive bias modification (CBM) training to reduce hostile attributions in clinically referred aggressive boys. We conducted two experiments, using a procedure that targets interpretation of ambiguous facial expressions. Both Study 1 ($N=59$, $Mage=11.8$) and Study 2 ($N=75$, $Mage=11.5$) showed that the training effectively reduced hostile attributions. However, we found no generalization to relevant other outcomes, including state anger and aggression (assessed using a multi-method approach), and hostile attribution assessed in a game context. Taken together, this research illustrates the potential of employing the CBM procedure to reduce aggressive boys' hostile attributions: The procedure is time and cost-efficient, and relatively easily implemented. Despite this potential, the context-dependency and generalization of effects should be better understood before the procedure can be taken to scale in at risk populations.

Key words: Aggression; hostile attribution; social information processing; cognitive bias modification; psychopathology; intervention

Introduction

Children with aggressive behavior problems tend to over-attribute hostile intentions to others in ambiguous social situations (De Castro, Veerman, Koops, Bosch, & Monshouwer, 2002; Dodge et al., 2015; Nasby, Hayden, & DePaulo, 1980;). Interventions that focus on changing these hostile attributions exist (e.g. Guerra & Slaby, 1990; Hudley & Graham, 1993; Lochman & Wells, 2002; Sukhodolsky, Smith, McCauley, Ibrahim, & Piasecka, 2016), but only show modest effects. One reason for the modest effects may be that the interventions mainly target conscious social processing skills for change, such as by using role play and re-training of intentionality. However, making hostile attributions is a relatively automatic and implicit process (Wilkowski & Robinson, 2008). Moreover, many children with aggression problems may not be particularly motivated to change attributions that provide them a defense-mechanism that they have often relied on in the past. Accordingly, recent approaches have sought to modify this implicit process by directly influencing the target cognitive bias using cognitive bias modification (CBM) (e.g. Hertel & Mathews, 2011; MacLeod, Koster, & Fox, 2009; MacLeod & Mathews, 2012).

CBM constitutes of repeated practice on a specific task that is designed to slightly change attitudes and bias in information processing (MacLeod, Koster, & Fox, 2009). CBM does not require conscious reports of personal beliefs or attitudes and therefore does not depend on self-insight. Some studies using CBM to target hostile attributions, conducted in population-based and at-risk samples of children and adults, have shown promising results (Penton-Voak, Munafò, & Yen Looi, 2017; Penton-Voak, Thomas, Gage, McMurrin, McDonald, & Munafò, 2013; Stoddard et al., 2016; Vassilopoulos, Brouzos, & Andreou, 2015). However, CBM has not yet been tested yet in those who are in particular need of intervention to prevent adverse outcomes: clinically referred children (e.g. De Castro et al., 2002; Fergusson, Horwood, & Ridder, 2005). Although it is possible that results obtained in population-based or at-risk samples will generalize to these children, this cannot be simply assumed. Among clinically referred children with disruptive behavior problems, hostile schemata may be relatively rigid and more difficult to change due to prevalent histories of aversive social experiences or personal vulnerabilities (e.g., low IQ, difficult temperament, or impulsivity; Dodge, 2006). The present study therefore tested the effectiveness of a CBM

training to target hostile attributions in clinically-referred children with disruptive behavior problems.

Hostile attributions

Children with aggressive behavior problems differ in their social information processing (SIP) from less aggressive peers: They have an automatic tendency to interpret ambiguous social cues as hostile, known as “hostile attribution bias” (Crick & Dodge, 1994; De Castro et al., 2002; Dodge et al., 2015; Nasby, Hayden and DePaulo, 1980), which is especially apparent in their interpretation of social cues (Crick & Dodge, 1996; Dishion, French, & Patterson, 1995; Dodge, 2006). Hostile attribution bias seems to be a trans-diagnostic driver of aggression (e.g., Henry, Moses, Castellini, & Scott, 2015).

People use heuristics (i.e., representativeness and availability heuristics) to interpret information, which may lead SIP to become more rigid over time. For example, if children have experienced hostility in the past, this often becomes readily available information in their schemata, and will influence them to make hostile interpretations again in the future (Dodge, 2006). These hostile interpretations instigate more aggression, which might in turn provoke others to respond with hostility as well, thereby triggering a negative recursive cycle.

One factor that is particularly important in the interpretation of social cues, is the perception of others’ facial expressions of emotion (Ekman, 1993; Etcoff & Magee, 1992; Hendriks & Vingerhoets, 2006). Facial expressions of emotion have a key function in regulating social interaction between humans and can convey signs of intent (Walker-Andrews, 1997). Therefore, facial expressions are an important basis of hostile attributions. Ambiguous facial expressions are interpreted more negatively by aggressive individuals than by others (Fairchild et al., 2009; Mellentin, Dervisevic, Stenager, Pilegaard, & Kirk, 2015; Schönberg & Jusyte, 2014), which may contribute to the development and maintenance of aggression (Crick & Dodge, 1994; Dodge, 2006). This phenomenon provides an entry point for intervention: To the extent that we may be able to help aggressive children perceive ambiguous facial expressions in less hostile ways, we may reduce their hostile attribution bias and associated behavior problems. There is now growing evidence that supports a causal role of biased facial expression recognition in conduct problems (Penton-Voak, Munafò, & Yen Looi, 2017).

Intervening with hostile attributions

Most present-day interventions that target hostile attributions in children and adolescents use explicit techniques such as re-training intention attribution or behavioral responses to cues (i.e. through role play; Guerra & Slaby, 1990; Hudley & Graham, 1993; Lochman & Wells, 2002). These interventions often lead to modest behavior changes, in both clinical and non-clinical samples, which illustrates that hostile attributions can be altered, and may function as a mediator for behavioral change.

One limitation of these intervention procedures, however, is that they typically rely on explicit cognitive behavioral training that requires reflection, while hostile attributions are implicit processes (Wilkowski & Robinson, 2008). Such explicit training may attempt to help children suppress automatic social information processing, and replace it with a non-stereotype response. However, conscious change of an automatic process is hard to achieve and potentially inefficient (Monteith, Sherman, & Devine, 1998). For children with aggression problems, who tend to be impulsive, high in emotionality, somewhat less (verbally) intelligent, and low on effortful control, this might be especially problematic (De Castro et al., 2002; Eisenberg et al., 2005). Furthermore, these children may often lack motivation to change patterns of cognition that offer them a defense-mechanism.

To circumvent these problems, it may be promising to *directly* influence hostile attributions by targeting the automatic process that gives rise to them. By targeting such an automatic process, and not letting children be conscious of the cognitions they are expected to change, we are less dependent on children's ability and motivation to change. Implicit interventions in modifying attributional styles have already proven to be effective in other domains of psychopathology, especially anxiety. A meta-analysis showed that CBM experimental paradigms can reduce maladaptive biases and associated anxiety symptoms in healthy undergraduates and adults (Hallion & Ruscio, 2011).

Recently, studies have adapted CBM to target hostile attributions and associated aggressive behavior. One study in 10-to-12 year olds that used vignettes to assess hostile versus benign responding to provocation found reductions in hostile attributions and self-reported aggression following CBM (Vassilopoulos et al., 2015). However, this study did not use a control-group, and so causality could not be established. Penton-Voak and colleagues (2013) conducted CBM experiments with healthy adults and high-risk adolescents in residential care. They successfully reduced the extent to which participants perceived anger

in ambiguous facial expressions, which led to reductions in hostile attributions and self- and staff-reported aggressive behavior in both groups. However, none of this work included clinically referred children. In the present study we sought to begin to fill this gap, and examined effects of a CBM procedure on hostile attributions in clinically referred children with disruptive behavior problems.

A recent meta-analysis and review, which both focused on mental health outcomes other than hostile attributions, suggested that CBM interventions typically are (moderately) effective to reduce cognitive bias, but less so to change behavioral outcomes (Cristea, Mogoase, David, & Cuijpers, 2015; Krebs, Pile, Grant, Degli Espositi, Montgomery, & Lau, 2017). Although this work primarily focused on anxiety and depression outcomes, it begs the question of how readily potential effects on hostile attribution bias will generalize to reducing aggression, especially in clinically referred children. The aggressive predispositions of these children tend to be robust, particularly when they grow up in social environments that are harsh, reinforcing their hostile attributions (e.g. Nix et al., 1999).

Aim of the present research

The aim of the present research was to systematically test whether an implicit bias modification training in the interpretation of facial expressions (Penton-Voak et al., 2013) reduces hostile attributions in clinically aggressive boys. To this end, we conducted two experiments. We only studied boys, because boys show more direct aggression, which is key to disruptive externalizing problems (Card, Stucky, Sawalani, & Little, 2008). We focused on children in late childhood and early adolescence, an age period when hostile attributions tend to have a pronounced impact on child problem behavior (De Castro et al., 2002).

Aggression is a heterogeneous phenomenon, with various manifestations (i.e., forms) and various underlying intentions (i.e., functions; Polman, Orobio de Castro, Thomaes & van Aken, 2009). Concerning functions, a distinction can be made between “reactive” and “proactive” aggression. Reactive aggression is a reaction to a perceived threat, whereas proactive aggression is planned, relatively unemotional behavior (Polman et al., 2007). Hostile attribution is a key determinant of, specifically, reactive aggression (e.g. Crick & Dodge, 1996; De Castro et al., 2002; Dodge, Price, Bachorowski, & Newman, 1990). We hypothesized that boys prone to aggression with reactive function would benefit most from our training (i.e., we hypothesized that reactive function of aggression would moderate the

effects of the training on interpretation in Studies 1 and 2).

We additionally explored whether the putative effects of the training on hostile attribution would generalize to aggressive behavior (Studies 1 and 2). Previous work found that modification of bias in emotional processing also has effects on mood (Penton-Voak et al., 2013), and so we investigated whether our training would reduce state anger (Study 2). As such, we explored the therapeutic potential of direct bias modification (MacLeod & Mathews, 2012). Given that severe aggression problems are multi-determined, we did not expect our focused training to exert effects of similar magnitude as more standard, comprehensive interventions that target multiple determinants. Rather, we sought to explore whether targeting a single maladaptive process (i.e., hostile attribution bias) would lead to observable reductions in aggression.

Study 1

We conducted an experiment, following procedures developed by Penton-Voak et al. (2013), in a sample of clinically-referred boys with aggressive behavior problems. The primary goal was to investigate whether the training would reduce hostile attribution and whether this effect would be moderated by reactive function of aggression; a secondary goal was to explore whether aggressive behavior would decrease accordingly.

Methods

Participants

The sample consisted of 59 boys aged 9 to 14, with a mean age of 11 years and 8 months. They were recruited from four schools providing special elementary education for children with behavior problems. In the Netherlands, about 2.5% of the school-aged population is referred to this type of school (Statistics Netherlands, 2014). Criteria for referral include: (1) Children have at least one DSM-diagnosis, (2) the severity of children's behavior problems falls within a clinical range and prohibits participation in regular education, and (3) children are treated for their problems in a clinical setting. A multidisciplinary examination committee decides on placement in this type of school. Education is provided in small classes (of about 10 children), by specialized teachers trained to provide individualized supervision and guidance.

The participating schools are situated in rural and suburban areas in the middle and south of the Netherlands, although boys came from across the country. The initial sample

consisted of 78 participants, but 19 of them dropped out of the study, either because they were absent from school on training days (e.g., due to illness, or therapy-related absence) or because teachers failed to complete both aggression measures. Children who dropped out of the study did not differ from those in the final sample in terms of their baseline hostile attribution (if available; $F(1,74)=.33, p= .568$).

Measures

Instrument for Reactive and Proactive Aggression. Boys' aggressive behavior was measured using the Instrument for Reactive and Proactive Aggression (IRPA; Polman, Orobio de Castro, Thomaes & van Aken, 2009). This is a teacher-report instrument that assesses forms and functions of aggression separately. The IRPA contains 8 form items (e.g. "How often did the child hit someone in the past week?"), and 7 function items (e.g. "How often did he do this because he was angry?"). Ratings are given on a Likert scale ranging from 0 (*never*) to 4 (*multiple times a day*) for form, and from 0 (*never*) to 4 (*always*) for function. We created a total sum score for aggression, and a separate sum score for reactive function. We slightly deviated from the original IRPA format and chose not to distinguish functions for each form separately, yielding a briefer measure. Good discriminant, convergent, and construct validity have been reported for the IRPA (Polman et al., 2009). Alpha's in the current sample were all acceptable to good (α total = .92, α proactive aggression = .83, and α reactive aggression = .73; Cohen, 1988).

Tally form. To obtain an objective aggression measure, teachers also completed an aggression tally form. Teachers tallied each occurrence of three form scales of aggression during a school day at pre-test and at post-test, including 8 items for (a) physical aggression (i.e., hitting, kicking, pushing), (b) verbal aggression (i.e., name calling, arguing, threatening), and (c) covert aggression (i.e., doing sneaky things, refusing to do something). For each of these behaviors, teachers also tallied which function the aggressive act served. A sum score was made to see how many times a child behaved reactively aggressive. Since we designed this tally list for this study, no information on validity or reliability is available.

Emotional state. State anger and state positive mood were measured every day right before the training by asking participants to indicate "how you feel right now" on an "anger thermometer" and a "happiness thermometer" (e.g. Elias, 2004). Responses ranged from 0 ("*not at all angry*") to 3 ("*very angry*"). We used the anger thermometer in the preliminary

analyses to see whether boys' state anger influenced their interpretation on the task.

Computer training. Following Penton-Voak et al. (2013), we created a computer-based training for this study using the software package OpenSesame (Mathôt, Schreij, & Theeuwes, 2012). We made photos of 9 boys ages 10 to 15, recruited in an acting class at a secondary school in the Netherlands. Informed consent to make and use the photographs was obtained from the boys themselves and their parents. The boys were photographed while asked to look either happy or angry. The happy and angry images were used as endpoints to generate 9 linear morph sequences consisting of 15 images that gradually transition from happy to angry, with emotionally ambiguous images in the middle (see Figure 1 for an example). These morph sequences were used as experimental stimuli in the training. Three versions of the computer task were created with a different order of the 9 morphing sequences. This was done to be sure that any effects of the training did not depend on a specific morphing sequence. Participants receiving the three different versions did not differ significantly on their first measurement of interpretation (balance point; $F(2, 56) = 0.15, p = .860$).

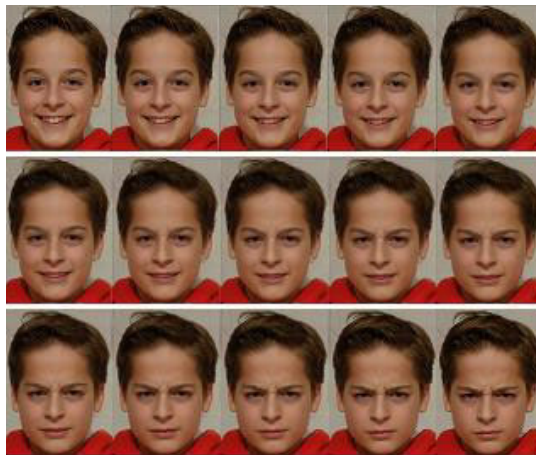


Figure 1. Example of One of the Morphing Continua: Unambiguously Happy, Ambiguous, and Unambiguously Angry (From Left to Right, Top to Bottom, Respectively).

Images were presented for 500 ms (Vlamings, Jonkman, & Kemner, 2010), preceded by a fixation cross. The pictures within the morphing sequences were presented in random order for every participant. Stimulus presentation was followed by a prompt asking participants to respond to a two-alternative forced-choice judgment whether the morphed

image they saw was happy or angry. Responses to such morphed images tend to shift categorically (either happy or angry) from one response to the other across the sequence (Etcoff & Magee, 1992; Young et al., 1997). This shift enabled us to obtain an estimation of the balance point at which the child was equally likely to decide on happy or angry in the presented morph. This was done by calculating the number of 'angry' responses as a proportion of the total number of trials. The balance point was used as a measure of interpretation.

The training took place on 5 consecutive (school) days. Each day, we first established a balance point in a baseline phase (using 45 trials on day 1 [= 3 complete morph continua]), and 30 trials on the subsequent days [= 2 complete morph continua]), after which we offered training (using 60 trials each day). On day 5 we also offered a new baseline phase, to establish the "final" balance point for each participant. The first balance point and the last balance point were calculated based on the same morph continua.

During the training phase, participants were provided computer-generated feedback on the responses they gave, based on their personal balance points that we established that day. The feedback that participants received on each response they gave was either "correct" (i.e., if they identified images two morph steps above their balance point image as "happy" or faces above that image as "angry") or "incorrect (in all other cases). The feedback was displayed visually on screen for 1 second. The exact feedback message read: "Correct/incorrect! That face was happy/angry". To illustrate, suppose that a child's balance point was at the seventh morph along the sequence. In the training condition, the child then received "correct" feedback only if he gave a "happy" response for all morphs up until the ninth morph along the sequence. In the control condition, no feedback was given. Overall, the boys were encouraged to finish the task and pay attention during the task (i.e., which took about 5-10 minutes each day). When children's attention appeared to drift, the experimenter helped them focus on the task again, or offered to have a short break.

Procedure

We obtained ethics approval from our local ethics review board. Next, twelve schools for special education for children with aggressive behavior problems across the Netherlands were requested to take part in the study, and five chose to do so. The schools distributed consent letters to parents/caregivers -- only boys who received active parental/caregiver

consent were allowed to take part in the study (consent rate was approximately 75%). Participants were randomly assigned to the experimental or control conditions. Teachers and schools were blind to conditions.

On each of the five days of training, boys completed the mood thermometers prior to the computer task. If a boy appeared agitated, we let him return to the classroom and try again later that day (this occurred 11 times). After the training, boys could play a fun computer game for a maximum of 5 minutes if they wanted to, to heighten their motivation to take part in the study. Boys also received a small present (a pen or eraser) at the third and the fifth day to thank them for their participation. To establish treatment fidelity, we kept a log and recorded any incidents where boys were 'not focusing on the task' / 'just pushing the same button repeatedly' (this occurred about 5 times).

Teachers completed the tally list and the IRPA the week before and the week after the boys received the training. The tally list was completed every day during these weeks; the IRPA was completed once at the end of the week.

Missing values and data reduction

Prior to data analysis, we checked treatment fidelity. We considered boys' data to be missing if they had two or more 'unrealistic' values. 'Unrealistic' values were opposite scores on the extreme morphs in the task (i.e. boys choosing "angry" for the two most happy pictures, and vice versa; this occurred for 1.7% of all balance point measures). Missing values were handled as follows: If the first and/or the last balance point missed, this value was replaced with the balance point from the previous day (and with the day 2 balance point for day 1; this was the case for 3.4%). For the tally list, any missing values were replaced with the mean of aggressive behaviors from the other days (this was the case for 3%). Some teachers did not return their reports for the second assessment, and so in the analyses of effects on aggression we left out the data from 11 boys who were in the initial sample. These boys did not significantly differ from the boys included in the aggression analyses on pre-test hostile interpretation, aggression, or reactive function of aggression. Because the tally list and the IRPA were highly correlated ($r = .87, p < .001$), we chose to aggregate these measures. Because reactive function of aggressive behavior was based on the aggression measures (IRPA and tally list), the two sum scores of reactive aggression were also aggregated ($r = .51, p < .001$).

Results

Preliminary analyses

We present descriptive statistics in Table 1. The distribution of the aggregated aggression score (z-scores) was slightly skewed; teachers reported low levels of aggression scores for about 25% of the participants. However, mean levels of aggression were still significantly higher than mean levels reported in studies using the IRPA in a general population sample ($t(57) = 5.20, p < .001; M = 0.83$ in this study, compared to $M = 0.25$ in Polman et al., 2007). State-anger, as measured by the anger thermometer, was not significantly associated with interpretation of the facial expressions at baseline ($r = .14, p = .317$).

Table 1 | *Descriptive Values of All Main Variables Used in Analyses Split by Condition (Study 1; N=59, HI = Hostile Interpretation).*

Variables	Experimental Group		Control Group	
	M	Range	M	Range
Age	11y8m	9y5m – 13y10m	11y10m	9y2m – 14y1m
Pre-test Hostile Interpretation	.50	.29 - .69	.48	.24 - .71
Pre-test State Anger	1.42	1.00 - 4.00	1.36	1.00 - 3.00
Pre-test Aggression (z-score)	.08	-.98 – 3.04	.01	-.98 – 3.02
Pre-test Reactive Aggression (z-score)	-.14	-1.09 – 1.93	.12	-1.09 – 2.42
Post-test Hostile Interpretation	.29	.13 - .47	.52	.27 - .73
Post-test Aggression (z-score)	-.31	-.78 - .73	-.33	-.78 – 1.67

Pre-test aggression was indexed by teacher' reports on aggressive behaviors in *one week*.

Main analyses

We used multiple regression analysis to test the effect of the training on hostile attribution. Hostile interpretation, defined as the “final” balance point, served as the dependent variable (see Table 2 for the complete model). The training significantly decreased hostile interpretation ($F(4,54) = 23.12, p < .001$). Figure 2 presents differences in balance points (i.e., hostile interpretation) between conditions. Participants in the experimental condition showed a shift in balance point of approximately .20 (on a 0 – 1 scale) relative to participants in the control condition. On average, boys in the experimental condition rated 8 morphs along the sequence as angry at baseline, while they did so for only 5 morphs after the training. Thus, boys in the experimental condition shifted from a hostile bias to an unbiased interpretation.

Table 2 | Summary of Regression Analysis for Variables Predicting Post-test Hostile Interpretation (Study 1; N=59).

	R^2	B	$SE B$	β
Final Model	.63**			
Pre-test Hostile Interpretation		.46	.14	.28**
Condition		.11	.01	.70**
Reactive Aggression		.03	.02	.16
Condition*Reactive Agg.		.03	.02	.14

* $p < .05$; ** $p < .01$

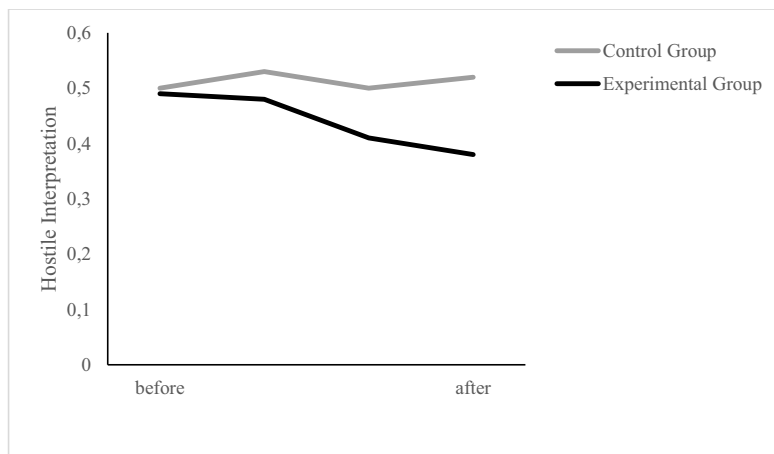


Figure 2. Effect of the on the Boys' Hostile Interpretation for the Five Consecutive Days (Last Day Includes Two Measurements).

Next, we analyzed whether the boys' reactive function of aggression moderated the condition effect. For this question, reactive aggression and its interaction with condition (based on centered variables) were added to the model. Neither reactive aggression, nor the interaction significantly predicted change in hostile interpretation.

Next, we conducted another regression analysis to test the effects of the manipulation on aggressive behavior. Aggressive behavior at post-test served as the dependent variable. No significant effects emerged ($\beta = -.01$, $SE = .09$, $p = .923$): both the experimental and the control group decreased significantly in aggression at post-test ($F(1,18) = 6.37$, $p = .021$, and $F(1,26) = 10.38$, $p = .003$, respectively).

Discussion

In Study 1, we investigated whether a CBM training could reduce referred boys' hostile interpretation. Boys in the experimental condition showed a shift in interpretation compared to boys in the control condition. This is a promising finding, especially because altering hostile attributions in children with disruptive behavior problems can be notoriously difficult (Dodge, 2006). However, as reported by teachers, the training did not impact aggressive behavior, nor did we find any moderating effect of boys' reactive aggressive predispositions.

We considered the possibility that a lack of effect on aggression was due, in part, to the aggression measure we used. Teacher reports of aggression, although valuable, can only provide information on aggressive behavior in the school-context. Therefore, we chose to conduct another experiment to investigate possible effects on an alternative aggression measure. We also investigated the emotional impetus to (reactive) aggression: feelings of anger. Anger can drive the link between situational variables and aggression (Anderson & Bushman, 2002). Since we target a situational variable (interpretation of faces), and explore effectiveness on aggression, it is relevant to explore effects for anger. Accordingly, Penton-Voak et al. (2013) tested an effect of their training on anger (along with aggression) as well.

Study 2

We conducted Study 2 to replicate our finding that the training reduces aggressive boys' hostile interpretation. We used a shorter training procedure, given that in Study 1 we found the biggest shift in interpretation after day three. Furthermore, to clarify whether generalization to multiple outcomes occurs, Study 2 investigated the effect of the training on both anger and an alternative measure of aggression.

Methods

Participants

The sample consisted of 75 boys ages 8 to 13, with a mean age of 11 years and 5 months. They were recruited from five schools providing special elementary education for children with behavior problems. Participating schools are situated in rural and suburban areas in the centre and south of the Netherlands, but boys came from across the country. The initial sample consisted of 87 participants; 12 boys dropped out of the study, because

they failed to complete the training (e.g., due to illness, or therapy-related absence from school). Participants were diagnosed with a DSM-IV disorder by multidisciplinary teams of professionals (obtained from school care records; 34.6% Autism Spectrum Disorder (ASD), 12.8% Attention-Deficit Hyperactivity Disorder (ADHD), 21.8% both ADHD and ASD, 11.5% Oppositional Defiant Disorder (ODD)/Conduct Disorder (CD), 6.4% both ADHD and ODD, 12.9% other/unknown diagnosis).

Measures

Computerized training. We used the same stimuli as in Study 1, but implemented the training as an app installed on an I-pad. Again, three versions of the training were created with a different order of the 9 morphing sequences. This was done to validate the pictures and be sure that any effects of the training did not depend on a specific morphing sequence. Participants receiving the three different versions did not differ significantly on their first measurement of interpretation (balance point; $F(2, 72) = 0.58, p = .562$).

Testing took place during 3 consecutive days. The session at the first two days consisted of two phases: a baseline phase and a training phase. The last day consisted of three phases: a baseline, training, and new baseline phase.

State anger questionnaire. To measure state anger we used the State Anger subscale of the State-Trait Anger Expression Inventory for Children and Adolescents (STAXI-CA; del Barrio, Aluja, & Spielberger, 2004). This measure was derived from the original adult version (STAXI; Spielberger, 1988), and shows acceptable psychometric properties for use in children. The State Anger subscale consists of 12 statements regarding state anger, such as “At this moment, I feel like hitting someone”. The items can be rated along a scale from 0 (“not at all true”) to 3 (“completely true”). Cronbach’s alpha was good in our sample ($\alpha = .95$; Cohen, 1988).

Aggression measure. To assess aggression we used another, self-created app based on the ‘Survivor’ procedure that has been used in previous research (e.g., Reijntjes et al., 2010). The procedure lets participants create a ‘profile page’ on which they get feedback from (imaginary) other boys. First, a photo was taken that was shown on participants’ profile page. Next, participants completed some information on their personal page, like their hobbies, occupational goals, and things they like and dislike about themselves.

Subsequently, participants received alerts that they other boys gave them feedback. They

could see the feedback at the bottom of their profile page, along with a picture of the boy who ‘reacted’. By clicking on the feedback, participants could visit the profile pages of these boys, where they could post feedback themselves. Participants could then ‘comment’ on the reactions they received. All participants received four ambiguous reactions¹ to their profile, which were “Whuut? He is kind of funny”, “Wow..!”, “Seems like a nice boy, but I don’t know if I would want to be friends with him” and “..Sick!”. These ambiguous reactions were chosen after consulting existing profile pages and reactions of children about the same age. The order in which participants received these reactions, and the way photos and profile pages were coupled to these reactions, were presented in random order. Participants’ responses (‘comments’) on these reactions served as the first measure of aggression (frequency of aggressive responses), coded by two independent coders. They coded each reaction is one of four categories (999= missing/unclear, 0= nice, 1= neutral, 2= clearly mean/aggressive). Inter-rater reliability for these four coding categories was $\kappa = .825$, $\kappa = .827$, $\kappa = .887$, $\kappa = .888$, respectively; almost perfect agreement (Landis & Koch, 1977).

In addition, participants were given the opportunity to aggress against the peers who responded to their profile by influencing the amount of money these peers received for participation in the game. The default fee was €2. Participants could leave this amount unchanged, subtract €1 or €2, or add €1 or €2.

Next, we asked participants to rate four statements, two of which were designed to measure hostile attributions. The statements were “They wanted to be mean to me”, and “They wanted to say bad things about me”. The participants rated these statements along a scale ranging from 0 (completely not agree) to 4 (completely agree). The two statements ($r = .63$, $p < .001$) were averaged to a total mean score of hostile intent attribution.

Procedure

We again obtained ethics approval from our local ethics review board. Nine schools for special education for children with disruptive behavior problems across the Netherlands were requested to take part in the study, and five chose to do so. The schools distributed consent letters to parents/caregivers -- only boys who received active parental/caregiver were allowed to take part in the study (consent rate was approximately 70%). Participants

¹ In the original Survivor game, feedback on participant profiles was manipulated to be positive, neutral or negative. Consistent with the present study purposes—to investigate interpretations of and reactions to ambiguous situations—we designed the feedback in this study to be ambiguous.

were randomly assigned to the experimental or control conditions. Teachers and schools were blind to conditions.

Each participant engaged in three sessions, once a day for three consecutive days in the same school week. On the last day, participants played the ‘Survivor game’ directly after the training. They completed the state anger questionnaire on the first day (before training) and the last day (after training and ‘Survivor’). Again, to establish treatment fidelity, we kept a log and recorded any incidents where boys were ‘not focusing on the task’/ ‘just pushing the same button repeatedly’ (this occurred 3 times).

Participants were thoroughly debriefed. They were told that the peers and the feedback they received were fictitious, and they were informed about the purpose of the study. Research has shown that these debriefing procedures are effective for young adolescents (Hurley & Underwood, 2002). Previous versions of the Survivor procedure have been used by us and others and are generally positively evaluated and enjoyed by participants. Participants received a small present (a pen or eraser) at the last day to thank them for their participation.

Missing values and data reduction

Prior to data analysis, we checked treatment fidelity. As in Study 1, we considered participants’ data to be missing if they had two or more ‘unrealistic’ values (this was the case for 1.6% of the balance points measured). Missing values were handled as in Study 1. Single missing values on the questionnaire were replaced by the participant mean (0.2% single missing values). For the ‘Survivor’ game, about 33% of participants responded aggressively at least once. Accordingly, the distribution of this variable was not normal. We thus chose to make this a categorical variable (aggression: yes, at least once/no).

Results

Main analyses

Descriptive statistics for the main variables are presented in Table 3. We used multiple regression analysis to test the effect of the training on hostile attribution. Hostile interpretation, the “final” balance point, served as the dependent variable. Figure 3 presents differences in balance points between conditions. Again, the training decreased hostile attribution significantly ($F(2, 72) = 37.77, p < .001$). Participants in the experimental

condition showed a shift in balance point of .11 (on a 0 – 1 scale) relative to participants in the control condition ($\beta = .37, SE = .03, p < .001$). On average, boys in the experimental condition rated 7.5 morphs along the sequence as angry at baseline, while they did so for only 6 morphs after the training. Thus, replicating Study 1, the CBM procedure was effective at reducing participants' hostile attribution bias.

Table 3 | *Descriptive Values of All Main Variables Used in Analyses Split by Condition (Study 2; N=75).*

Variables	Experimental Group		Control Group	
	M	Range	M	Range
Age	11y8m	8y9m - 13y7m	11y1m	8y2m - 13y0m
Pre-test Hostile Interpretation	.49	.23 - .83	.50	.33 - .70
Pre-test State Anger	.36	.00 - 1.92	.68	.00 - 3.00
Post-test Hostile Interpretation	.41	.10 - .83	.52	.20 - .83
Post-test State Anger	.21	.00 - 1.92	.24	.00 - 1.50
Post-test Aggression (coins)	2.21	.75 - 3.50	2.16	.50 - 4.00
Post-test Aggression (reaction)	.30	.00 - 1.00	.16	.00 - 1.00
Post-test Hostile Intent	1.30	.00 - 4.00	1.55	.00 - 4.00

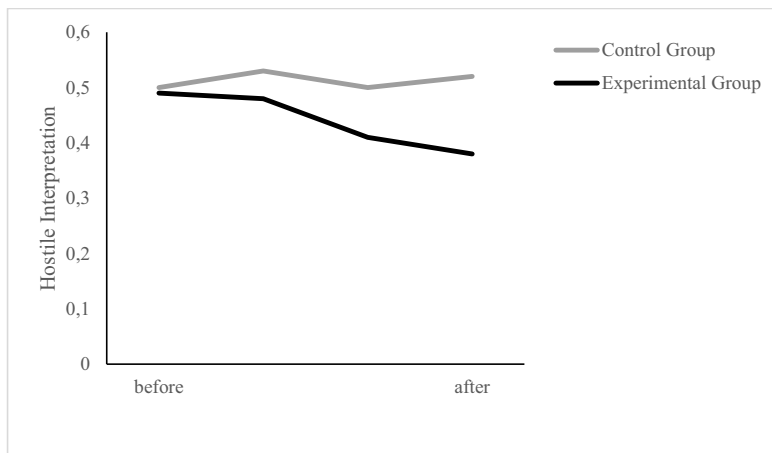


Figure 3. Effect of the Training on the Boys' Hostile Interpretation for the Three Consecutive Days (Last Day Includes Two Measurements).

Next, we conducted another regression analysis to test the effects of the manipulation on state anger. We first examined links between hostile attribution and reported state anger at baseline. The first balance point and pre-training state anger were

significantly negatively correlated, $r = -.34, p = .003$. Thus, higher levels of anger in the boys were associated with more hostile interpretation. In the regression analysis, state anger post-test served as the dependent variable. No significant effects of the training were found for state anger ($\beta = .09, SE = .09, p = .372$), only pre-test anger was a significant predictor in the model ($F(2,72) = 14.17, p < .001$).

Lastly, we conducted three additional regression analyses to test the effects of the manipulation on the intent attribution- and aggression measures within the ‘Survivor’ game. The training was not effective in reducing aggression, regardless of whether aggression was indexed by how much money participants took away from their peers ($F(1,73) = 0.12, p = .746, \beta = .04, SE = .16$), or how aggressively they reacted on others’ profile pages ($\chi^2(1, N = 75) = 2.22, p = .136$). Also, the training was not effective in reducing hostile intent attribution as indexed by the two relevant statements in ‘Survivor,’ ($F(1,73) = 0.85, p = .358, \beta = -.11, SE = .27$).

Discussion

In Study 2, we again investigated whether a CBM training could reduce referred boys’ hostile interpretation. Boys in the experimental condition showed a shift in interpretation compared to boys in the control condition. Compared to Study 1, this effect was reached in a smaller time-frame. However, comparable to Study 1, the training did not impact aggressive behavior and hostile intent attribution as measured in a game-context, nor did we find an effect on state-anger.

General Discussion

We conducted two experiments to investigate whether a CBM training targeting interpretation of facial expressions reduces hostile attributions in aggressive boys. It does. In both studies, boys in the experimental condition showed a benign attributional shift compared to boys in the control condition. However, we found no generalization to state-anger and aggression, as measured using a multi-method approach, nor to hostile intent attribution in a game context.

Our finding that the training reduces hostile interpretation of facial expression builds on and extends similar findings obtained in other populations (i.e., typically developing children and adults, and high-risk adolescents; Penton-Voak et al., 2013; Penton-Voak, Munafò, & Yen Looi, 2017; Stoddard et al., 2016; Vassilopoulos, Brouzos, & Andreou, 2015;).

The implicit process of hostile attribution is strongly influenced by existing schemata, and may therefore become increasingly rigid over time (Dodge, 2006). However, our two experiments show that aggressive children's predisposition to make hostile attributions can be changed using a time-efficient procedure, with little burden for children taking part. As such, CBM may be a promising means to help break the cycle of hostility that aggressive children often face.

We found no generalization from curbed hostile attributions to self-reported state anger (Study 2) and aggressive behavior (regardless of whether it was assessed by teacher-reports or behaviorally). This is at odds with previous work, which did find such generalization (Penton-Voak et al., 2013, Vassilopoulos et al., 2015). That said, our findings do dovetail with the outcomes of a recent meta-analysis, which found that in other mental health domains (especially anxiety; Cristea et al., 2015), generalization of CBM effects to behavioral outcomes is rare.

Several factors may explain why we did not find an effect on aggressive behavior. One possibility is that CBM has a genuine effect on the specific bias that we targeted (i.e., children's hostile attributions of ambiguous facial expressions of emotion), but somehow this effect fails to transpire to other contexts (e.g., hostile attributions in the game task), and to related emotional- and behavioral responses. Perhaps CBM does not sufficiently target the constellation of cognitive factors underlying severe aggression problems, to be able to prevent aggressive behaviors. Indeed, other social information processes, such as cognitive and regulatory control, may be important to reduce aggression as well (e.g., Saleminck & Wiers, 2011; Wilkowski, Crowe, & Ferguson, 2015). Furthermore, aggressive children may often suffer from recursive, relatively ingrained interpersonal problems; to really break the vicious cycles that maintain their aggression problems, additional intervention targeting children's immediate social environment may be necessary (e.g. classroom intervention, parent training). Another possibility is that children did not truly change their interpretational style, even if they learned how to interpret facial expressions as less angry. We intended to teach participants that ambiguous expressions are not as hostile as they may think, but we cannot rule out the possibility that participants learned from the procedure that people may not show angry facial expressions even if they actually are angry and hostile inside. We do note, however, that this possibility would not be able to account for earlier findings (e.g., Penton-Voak et al., 2013).

Another limitation concerns the nature of our samples. All participants were clinically referred for aggressive behavior problems, but we did not include structured diagnostic interviews, nor do we have information on possible medication use. Although we assume that hostile intent attribution is a trans-diagnostic driver of aggression, future research is needed to establish whether CBM procedures exert similar effects for children with different diagnoses (e.g., children with an autism spectrum disorder vs. those with an oppositional defiant disorder).

This research has several methodological strengths. We conducted two experiments, which allowed us to replicate and extend our key findings. We also used a multi-method assessment approach, thus circumventing subjectivity and mono-method biases. The CBM procedure we used was unobtrusive and relatively time- and cost-efficient. And lastly, we sampled, for the first time in this literature, boys referred for aggressive behavior problems – those who are in particular need of intervention (De Castro et al., 2002).

Future work may consider the school context when examining the effectiveness of the procedure. Children's aggression levels differ between schools—as such, it will be valuable to consider classroom and school system characteristics (Alexander, 2001). For example, future work may assess classroom sociometrics, or real-time peer and teacher interactions (e.g. Hudley & Graham, 2007). Similarly, future work may assess aggression as it occurs outside of the school context (e.g., at home), and identify contextual factors influencing the effectiveness of the training. For example, strict classroom management and teacher-control might overrule any effects of the training, simply because there are less situations in class in which a child can behave aggressively. In less structured contexts outside the school, however, training effects may be more likely to be found. Second, future work may seek to more comprehensively target other implicit processes that are known to contribute to children's aggressive behaviour. For example, aggressive children high on psychopathic traits show deficits in the processing of fearful facial expressions (Blair, 1999; Frick et al., 2003). These children might benefit from a CBM procedure to target this deficit (Dadds et al., 2006). A procedure targeting both hostile attributions and deficits in fear recognition in this clinical group might be especially promising.

Third, longitudinal research should examine the long-term effects of this procedure. It is possible that it produces sleeper effects; that its long-term effects are larger than its immediate effects. Indeed, social information processing theory (Crick & Dodge, 1994)

predicts that interventions are effective to the extent that they modify recursive processes that unfold over time (Yeager & Walton, 2011). To strengthen such recursive processes, it may be interesting to test additional effects of this training as an add-on to evidence-based interventions, such as parent training or classroom management. It is possible that CBM procedures, such as the one we used, may help amplify the putative effects of such interventions that target the family or school context.

Overall, our research shows promising results; the CBM procedure that we used effectively reduced aggressive boys' hostile attributions. The training is time- and cost-efficient, and relatively easily implemented. However, we found no generalization in terms of children's emotion (i.e., anger) and behavior (i.e., aggression). We conclude that the training is not yet ready to be implemented as an intervention on a larger scale, but its potential deserves further empirical scrutiny.

CHAPTER 5

IMPROVING EMOTION ATTRIBUTION IN CHILDREN WITH DISRUPTIVE BEHAVIOR PROBLEMS USING COGNITIVE BIAS MODIFICATION

W. Hiemstra developed the study concept, developed the design and performed data-collection. B. Orobio de Castro and S. Thomaes gave advice and feedback during this process. W. Hiemstra did the main literature search and performed the data-analysis. W. Hiemstra, B. Orobio de Castro and S. Thomaes performed interpretation. W. Hiemstra drafted the manuscript, and B. Orobio de Castro and S. Thomaes provided critical revisions.

Abstract

Deviations in attributing emotions may contribute to children's aggressive behavior problems. Specifically, anger attribution may contribute to reactive aggression and reduced fear attribution may contribute to aggression by children high on psychopathy. These deviations may provide entry points for intervention. Aim of the present study was to test whether a cognitive bias modification training may reduce anger attributions, and increase fear attribution in children referred for disruptive behavior problems. To this end, a randomized controlled trial was conducted with 70 boys referred to special education for their disruptive behavior problems, aged 8 to 13. The training significantly reduced anger attribution in the experimental group, specifically for children whose reactive aggression scores were at the 37th percentile or higher. The training also marginally significantly increased fear attribution. The training did not reduce aggressive behavior. Overall, the potential of using CBM to target emotion processing deficits seems substantial, and warrants further empirical scrutiny.

Key words: Aggression; intervention; cognitive bias modification; emotion attribution; psychopathy

Introduction

The perception of others' facial expressions of emotion is crucial in the interpretation of social cues, and has a key function in regulating social interaction (Ekman, 1993; Etcoff & Magee, 1992; Hendriks & Vingerhoets, 2006; Walker-Andrews, 1997). From infancy onwards, the ability to attribute emotions to facial expressions plays an important role in the development and maintenance of relationships (Mc Clure, 2000). More specifically, facial expressions of emotions have a key function in regulating social interactions as they can convey signs of interaction partners' intentions (Walker-Andrews, 1997).

Not surprisingly, then, difficulty in attributing emotions is theorized to contribute social behaviour problems, including youth's aggressive behaviour problems. According to the Social Information Processing model (SIP; Crick & Dodge, 1994) and the General Aggression Model (GAM; Allen, Anderson, & Bushman, 2018), problematic encoding and interpretation of social cues such as facial expression of emotion may have repercussions for how children react and behave in social situations (Dodge, 2006).

Research has identified two deviancies in emotion attribution in aggressive behaviour problems in youth. First, ambiguous facial expressions are more often interpreted as "angry" by aggressive youth than by others (Fairchild, Van Goozen, Calder, Stollery, & Goodyer, 2009; Mellentin, Dervisevic, Stenager, Pilegaard, & Kirk, 2015; Schöenberg & Jusyte, 2014;). Such mis-attributions of anger are part of the more general tendency among aggressive children, and especially reactively aggressive children (i.e. those who tend to aggress in response to perceived threat; Crick & Dodge, 1996; De Castro, Veerman, Koops, Bosch, & Monshouwer, 2002; Dodge, Price, Bachorowski, & Newman, 1990) to interpret ambiguous social cues as more hostile (i.e., hostile attribution bias; Nasby, Hayden, & DePaulo, 1980). This bias is known to maintain or foster children's aggressive behavior (Cadesky, Mota, Schachar, 2000; Fairchild et al., 2009; Mellentin et al., 2015), and contribute to higher levels of aggression over time (Schultz, Izard, & Ackerman, 2000).

Second, aggressive children high on psychopathic traits (i.e., children who lack empathy, show shallow emotions, and are inclined to instrumentally use others for their own gain; Frick, Ray, Thornton, & Kahn, 2014) are known to have difficulties processing signs of emotional distress in others, such as recognizing expressions of fear (Blair, 1999; Frick, Cornell, Barry, Bodin, & Dane, 2003; Marsh & Blair, 2008; Wilson, Juodis, & Porter, 2011).

These two emotion processing deviancies, anger attribution (i.e., in children prone to

reactive aggression) and reduced fear attribution (i.e., in children high on psychopathy), may provide entry points for intervention for children with aggression problems. If we are able to help these children process others' emotional expressions more accurately, we may possibly reduce their aggressive behavior problems.

One promising procedure to do so is Cognitive Bias Modification (CBM; e.g. Koster, Fox & MacLeod, 2009; MacLeod & Mathews, 2012). CBM involves the repeated practice on a specific task designed to unobtrusively change attitudes and bias in information processing (Koster et al., 2009). CBM does not require conscious, introspective reflection on personal beliefs or attitudes; rather, it is an implicit procedure that seeks to influence automatic processes. Pioneering work has already shown that CBM can be beneficial in reducing problem behaviors, including aggression, for some children (e.g. Dadds et al., 2006; Penton-Voak, Thomas, Gage, McMurran, McDonald, & Munafò, 2013; Stoddard et al., 2016). However, this work focused narrowly on changing either anger attribution or fear attribution—it did not target these processes simultaneously.

A recent meta-analysis of the effects of CBM on youth mental health problems (primarily anxiety and depression) found that CBM typically is moderately effective to reduce emotion attribution biases, but less so to change behavioral outcomes (Cristea, Mogoase, David, & Cuijpers, 2015; Krebs et al., 2017). This issue may be especially pertinent in clinically referred children, for whom aggressive behavior problems are relatively stable (e.g. Willoughby, Mills-Koonce, Gottfredson, & Wagner, 2014). Specifically, the aggression-reinforcing cognitive schemata of these children may be relatively rigid, due to prevalent histories of aversive social experience (e.g. coercive family interaction, rejection by peers; Olson, 1992; Patterson, 1976) and personal vulnerabilities that sustain those schemata (e.g., low IQ, difficult temperament, or impulsivity; Dodge, 2006). Indeed, our previous research has shown that, although changes in anger attributions may be induced by CBM in these children, generalization to aggressive behavior is more difficult to establish (Hiemstra, Orobio de Castro, & Thomaes, 2018).

However, given that clinically referred children's aggression problems are rooted in diverse emotion attribution deviances, the use of single emotion-focused CBM procedures may not be ideal. More comprehensive CBM procedures that simultaneously target multiple emotion attributions might be more effective to influence actual behavior. The present study explored the effectiveness of a cognitive bias modification training that simultaneously

targets two such deviances: excessive anger attribution and impaired fear attribution in others' facial emotional expressions.

Aim of the present study

Aim of the present study was to test whether an implicit bias modification training in the attribution of facial expressions (Penton-Voak et al., 2013; Hiemstra et al., 2018) may reduce anger attributions, and increase fear attribution in children referred for disruptive behavior problems. Furthermore, we explored whether the putative effects of the training on children's emotion attribution would generalize to their interpretation of vignettes describing hypothetical peer provocations, and their actual aggressive behavior.

Participants were boys referred to special education for their disruptive behavior problems, aged 8 to 13. No previous research has used CBM procedures to try to modify multiple emotion processing biases in clinical samples of children. We only included boys because they tend to show more direct aggression than girls (Card, Stucky, Sawalani, & Little, 2008), and therefore their behavior patterns are more often experienced as disruptive.

Since anger attribution is a determinant of, specifically, reactive aggression (e.g. Crick & Dodge, 1996; De Castro et al., 2002; Dodge et al., 1990), we hypothesized that boys showing higher levels of reactive aggression would benefit most from the anger-focused part of the training. Moreover, because fear attribution deficits are especially pronounced among children with elevated psychopathic traits, we predicted that the fear-focused part of the training would be especially effective for these children.

Methods

Participants

Participants were 70 Dutch boys with severe disruptive behaviour problems, aged 8 to 13 (mean age = 11 years, 8 months). They were recruited from four schools providing special elementary education for children with disruptive behavior problems. In the Netherlands, children are referred to this type of education only if the severity of their behavior problems significantly impairs their social functioning and prohibits participation in regular education, as agreed on by parents, teachers, and psychologists. This type of education is offered in small classes (typically about 10 children) by specialized teachers who provide individualized teaching and guidance. About 2,5% of the Dutch school-aged

population is referred to this type of education (Statistics Netherlands, 2014). Almost all children in special education also receive treatment in mental health care. Participating schools were situated in rural and suburban areas in the middle and south of the Netherlands, but boys came from all over the country. The initial sample consisted of 87 participants; 17 of them dropped out of the study either because they were absent from school on training days (9 boys; e.g., due to illness, or therapy-related absence), couldn't finish the training properly (6 boys; see missing values), or because teachers failed to complete the aggression measure (2 boys). The boys who dropped out did not differ from the final sample of participants in pre-training aggression ($F(1,82) = .181, p = .67$). Mean levels of aggression were significantly higher than mean levels reported in earlier studies in a general population sample (see below for a detailed description of the instrument used; $t(69) = 6.19, p < .01; M = .91$ in this study, compared to $M = .25$ in the study done by Polman, Orobio de Castro, Koops, van Boxtel, & Merk, 2007).

Measures

Instrument for Reactive and Proactive Aggression. We measured aggressive behavior using the Instrument for Reactive and Proactive Aggression (IRPA; Polman, Orobio de Castro, Thomaes & van Aken, 2009). This teacher-report measure assesses forms and functions of aggression separately. The IRPA contains 8 aggression form items (e.g. "How often did the child fight with others *this week?*"), and 6 aggression function items (e.g. "How often did he do this because he felt threatened?"). Items are rated along 5-point Likert scales (0 = *never*, 4 = *multiple times a day* for form; 0 = *never*, 4 = *always* for function). We slightly deviated from the original IRPA format and chose not to distinguish functions for each form separately, in order to obtain a shorter measure. Items were averaged to obtain frequency of aggression (form), reactive function and proactive function scores. Good discriminant, convergent, and construct validity have been reported for the IRPA (Polman et al., 2009). Alpha's in the current sample were all acceptable to good ($\alpha_{aggression} = .93, \alpha_{proactive aggression} = .70$, and $\alpha_{reactive aggression} = .77$).

Antisocial Process Screening Device. We measured trait psychopathy using the Antisocial Process Screening Device (APSD; Frick & Hare, 2001). This is a 20-item teacher-report instrument measuring callous/unemotional, narcissistic, and impulsive facets of psychopathy in youth (e.g., "*Is concerned about the feelings of others*", "*Uses or cons other people to get what he wants*", and "*Acts without thinking of consequences*"). Items are rated

along a 3-point Likert scale (0 = *not at all true*; 2 = *definitely true*). Acceptable to good validity and reliability have been reported for the APSD (e.g., Poythress et al., 2006). Items were averaged to obtain a total psychopathy score, of which the alpha in the current sample was good ($\alpha = .86$).

Emotional state. We assessed state anger and positive mood on each day just before the training was offered. We asked participants to indicate “*how you feel right now*” on an “anger thermometer” and a “happy thermometer” (Elias, 2004). Responses ranged from 0 (“*not at all angry/happy*”) to 3 (“*very angry/happy*”). We used this variable to check the possibility that the training scores would be influenced by mood.

Computer training. Conform procedures used by Penton-Voak and colleagues (2013), we developed a CBM training for this study, using the software package OpenSesame (Mathôt, Schreij, & Theeuwes, 2012). The training was offered for 5 to 10 minutes during 5 consecutive days. In the training, participants were presented pictures of facial expressions of emotion, which differed in level of ambiguity. To create these pictures 9 boys around the same age as the participants were photographed, expressing non-ambiguous happiness, anger and fear. These non-ambiguous pictures were used at both ends of 15-picture morph sequences. The sequences thus consisted of pictures gradually displaying different levels of mixed emotion.

The first three days of training focused on changing perceptions of anger; the last two days focused on changing perceptions of fear. Sequences used on days 1-3 therefore contained expressions ranging from non-ambiguously happy to non-ambiguously angry; sequences used on day 4-5 contained expressions ranging from non-ambiguously fearful to non-ambiguously angry. Pictures from these sequences were presented in random order, and displayed for 500 ms each. After presentation of each picture, participants responded to a two-alternative forced-choice judgment whether the face was happy or angry (days 1-3), or fearful or angry (days 4-5). Children tend to categorize such emotional expressions as either happy, fearful or angry (Young et al., 1997; Etcoff & Magee, 1992). As such, participants’ responses enabled us to obtain an estimation of the balance point for the emotion display at which an individual participant was equally likely to decide on happy (or fearful) vs. angry in the presented face. We estimated balance points by calculating the number of ‘happy’ or ‘fearful’ responses as a proportion of the total number of trials. These balance points were used as indications of current anger and fear attribution in the training.

Each day, we first established a balance point (on day 1 and 4 using 45 trials; on day 2, 3, and 5 using 30 trials), after which we provided the training itself (using 60 trials). On days 3 and 5 (i.e., the days that concluded training for each emotion sequence), we also established a new balance point after the training itself (again using 30 trials). For a more detailed description of the training procedure, see Hiemstra, Orobio de Castro, and Thomaes (2018).

Vignettes. Eight vignettes describing hypothetical, ambiguous peer provocations were used to assess children's hostile attribution and aggressive response access (Orobio de Castro, Merk, Koops, Veerman, & Bosch, 2005). The vignettes were read to participants. We administered four vignettes before the training, and four after (presented in random, counterbalanced order across participants). Hostile attribution was assessed by asking two questions following each vignette. The first was an open question: "*Why did he do that?*". Answers were coded in one of four categories: benevolent (1), on accident (2), ambiguous (3), or hostile (4). The question was coded by two independent coders, with an intra-class correlation of .94. The second question was closed: "*How did he intend it?*". Participants rated their response as nice (1), accidental/nice (2), accidental (3), mean/accidental (4) or mean (5). For both questions, the average of the scores were calculated. Because the two questions were significantly correlated ($r = .64$), we calculated a composite score after standardizing the variables. To assess aggressive response access, a third question was asked: "*What would you do?*". We coded answers in one of three categories: no aggression, mild aggression (coercion or verbal), severe aggression (physical or material damage). Answers were again coded by two independent coders, with an intra-class correlation of .96.

Procedure

Four schools for special education were willing to take part in our study. The schools distributed consent letters to parents/caregivers; only boys who received active consent were allowed to take part in the study. We asked schools to distribute the consent forms among 144 boys. 87 boys took part in the study (participation rate = 60 %). Participants were randomly assigned to the experimental or control condition. Teachers were blind to conditions. Ethics approval was given by the ethics review board of Utrecht University, Faculty of Social and Behavioural Sciences.

On each of the five days of training, participants completed the mood thermometers

prior to the computer task. If a participant impressed as excessively angry or restless, the trainer decided whether or not to proceed with the training. In case he could not proceed, we asked him to return to the classroom and try again later that day (this occurred in 1,5% of the testing occasions). Participants received a small present (a pen or eraser) at the third and the fifth day to thank them for their voluntary participation. To establish treatment fidelity, we kept a log and recorded any incidents where boys were 'not focusing on the task' or 'pushing the same button repeatedly' (this occurred 8 times). On the first day before training, and on the last day after training, participants completed the vignettes.

Teachers completed the APSD and the IRPA in the week before training. They completed the IRPA again in the week after training.

Missing values and data reduction

After data collection, but before analyses, we checked the log and considered the responses of boys who did not appear to finish the training properly. If their responses showed two or more 'unrealistic' values, these were handled as missing data. We operationalized 'unrealistic' values as opposite scores on the unambiguous pictures in the task (i.e., "angry" responses at the two unambiguously happy pictures, and vice versa; this happened multiple times for 6 boys and once for 7 boys). If the first and/or the last balance point was missing, the value was replaced by the balance point from the day after or before, respectively. Single missing values on the vignettes were replaced by the mean. Three boys were excluded from analyses because they left more than one question unanswered or answered "I don't know". Some teachers did not return the second aggression measure; the concerning participants ($N = 11$) were not included in the analyses of aggressive behavior.

Results

Preliminary analyses

At baseline, there was a slight (i.e., 6-month) but significant age difference between the experimental and the control group ($F(1, 68) = 5.170, p = .03$). Accordingly, we controlled for age in the main analyses. There were no other significant differences between the experimental and the control group on any of the study variables at baseline (see Table 1). Table 2 presents correlations between the main variables.

Table 1 | *Descriptive Values at Pretests by Group (n=70).*

Variables	Experimental Group		Control Group	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Age	11y11m	1y1m	11y5m	1y0m
Anger attribution	.48	.09	.51	.09
Fear attribution	.59	.10	.62	.12
Hostile attribution	.03	.99	-.04	.84
Aggressive response	.52	.55	.52	.56
Total Aggression	.80	.76	1.02	1.00
Reactive function	1.58	1.22	1.43	1.09
Psychopathy	.79	.40	.79	.38

Total aggression was indexed by teacher' reports on aggressive behaviors in *one week*, with a mean score of 1.00 meaning that the boys showed one to two aggressive act(s) per day, on average.

Table 2 | *Correlations Between the Independent and Dependent Variables at Pretest (n=70).*

	1	2	3	4	5	6	7
1. Anger attribution	-	-	-	-	-	-	-
2. Fear attribution	.06	-	-	-	-	-	-
3. Hostile attribution	.23	.02	-	-	-	-	-
4. Aggressive response	-.15	.24*	.23	-	-	-	-
5. Total aggression	.11	.21	.06	.09	-	-	-
6. Reactive function	.18	.19	-.01	.09	.59**	-	-
7. Proactive function	.09	.18	.03	.15	.62**	.64**	-
8. Psychopathy	.12	.27*	.04	.23	.72**	.56**	.71**

* Correlation is significant at the .05 level (2-tailed)

** Correlation is significant at the .01 level (2-tailed)

Main analyses

We used multiple regression analysis to test the effect of the intervention on anger attribution. Anger attribution, defined as the balance point along the happy-angry continuum established on day three (i.e., after the final training session), served as the dependent variable. Pre-test anger attribution, as well as Age, were entered in Step 1. Condition was entered in Step 2. Adding Condition to the model produced a significant R^2 change of .13 ($F(3, 66) = 10.62, p < .001$). The intervention thus significantly reduced anger attribution: Participants in the experimental condition showed a shift in balance point (indexing anger attribution) of .10 on a 0 – 1 scale ($\beta = -.37, t(67) = 3.54, p < .01$), whereas participants in the control condition showed no shift at all. After the training, participants in the training condition considered an average of only 6 morphs as angry, while they considered 8 morphs as angry at baseline. Thus, during the training, participants shifted from a hostile bias to an unbiased interpretation (Figure 1).

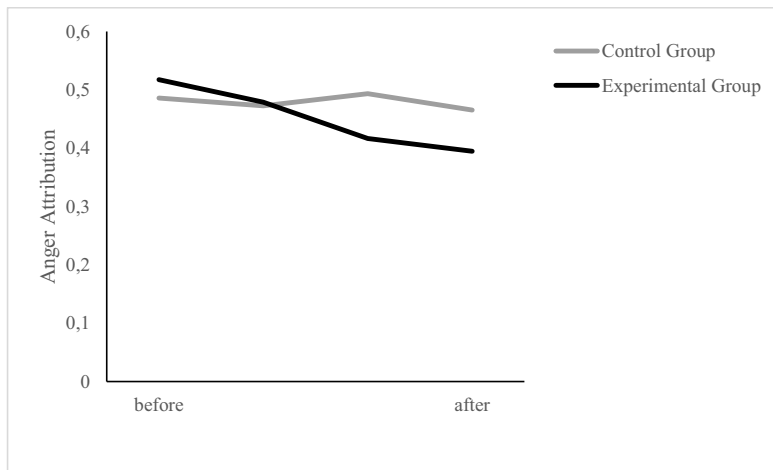


Figure 1. Effect of the Intervention on Participants' Anger Attribution for the Three Consecutive Days (Last Day Includes Two Measurements).

Next, we tested whether Reactive Aggression at Pre-test moderated the training effect. To address this question, Reactive Aggression and its interaction with Condition (continuous variables were centered) were added to the model in Steps 2 and 3, respectively. There was no main effect of Reactive Aggression on the effectiveness of the training, but the interaction was marginally significant (R^2 change = .03, $\beta = .18, t(66) =$

1.77, $p = .08$). We probed the interaction using the Johnson-Neyman (1936) technique (Bauer & Curran, 2005). Specifically, the effect of the intervention was significant for children whose reactive aggression scores were are the 37th percentile or higher (but not lower; see Figure 2).

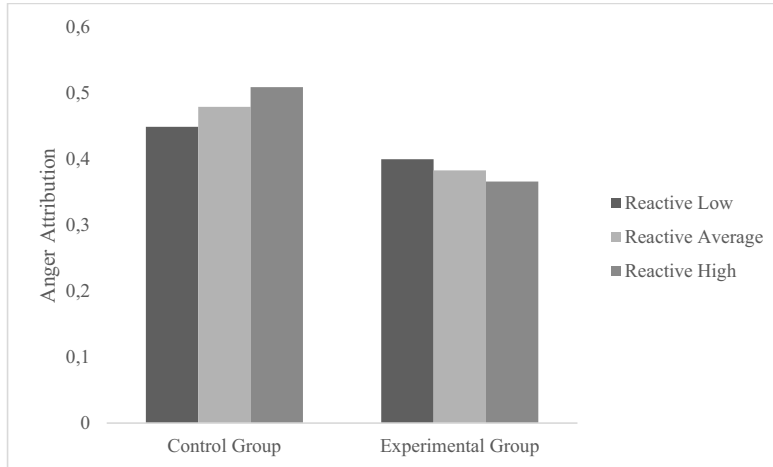


Figure 2. Effect of the Intervention on the Boys' Final Anger Attribution Scores for Boys Showing Low, Average and High Levels of Reactive Function for aggression.

To examine the effect of the intervention at increasing fear attribution, we used a similar multiple regression analysis. Fear attribution, defined as the final balance point along the fearful-angry continuum, served as the dependent variable. Fear attribution at baseline and Age were entered in Step 1. Condition was entered in Step 2. The intervention marginally significantly increased fear attribution ($\beta = .23$, $t(67) = 1.92$, $p = .06$). Participants in the experimental condition showed a shift in balance point of .06 (on a 0 – 1 scale), whereas participants in the control condition showed no shift. After the training, participants considered an average of 5.5 morphs as angry, while they considered 6.2 morphs as angry at baseline (Figure 3).

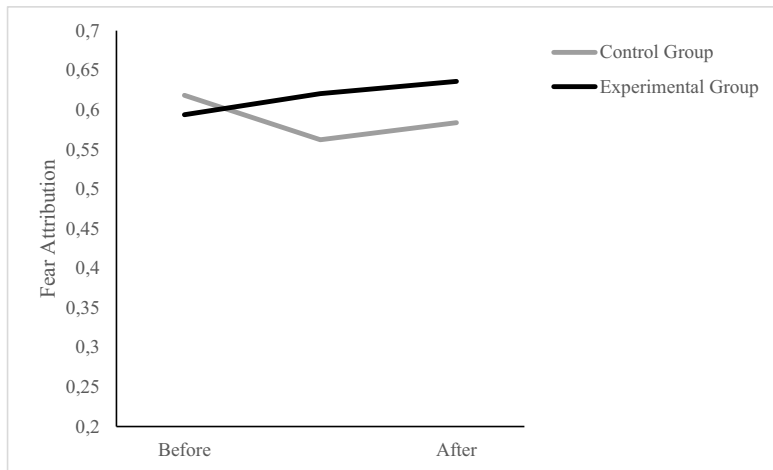


Figure 3. Effect of the Intervention on Participants' Fear Attribution for the Two Consecutive Days (Last Day Includes Two Measurements).

Next, we tested whether trait Psychopathy moderated the training effect. For this question, Psychopathy and its interaction with Condition (based on centered variables) were added to the model in Step 2 and Step 3, respectively. There was no main effect of Psychopathy on the effectiveness of the training ($\beta = .02$, $t(66) = 0.17$, $p = .87$). The interaction was also non significant (R^2 change = .004, $\beta = .07$, $t(65) = 0.55$, $p = .58$).

To explore whether effects of the training on emotion recognition generalized to hostile attribution and aggressive response access assessed with vignettes, we conducted two additional multiple regression analyses. For the first analysis, Hostile attribution post-training served as the dependent variable, Hostile attribution at baseline and Age were entered in Step 1. Condition was entered in Step 2. There was no effect for Condition ($\beta = .03$, $t(64) = 0.21$, $p = .84$). Both Reactive function and Psychopathy were also entered in Step 2, and the interaction-terms in Step 3. No main- or moderator effects were found.

In the second analysis, aggressive response access post-training served as the dependent variable, aggressive response access at baseline and age were entered in Step 1. Condition was entered in Step 2. Again, there was no main effect of condition ($\beta = -.08$, $t(64) = -0.82$, $p = .41$). Both reactive function and psychopathy were also entered in Step 2, and their interaction-terms with Condition in Step 3. No main effects of both variables were found. However, the interaction-term of Reactive function again was (marginally) significant. Again, we probed the interaction using the Johnson-Neyman (1936; in Bauer & Curran, 2005)

technique. Specifically, the effect of the intervention was significant for children whose reactive aggression scores were at the 95th percentile or higher (but not lower; see Figure 4).

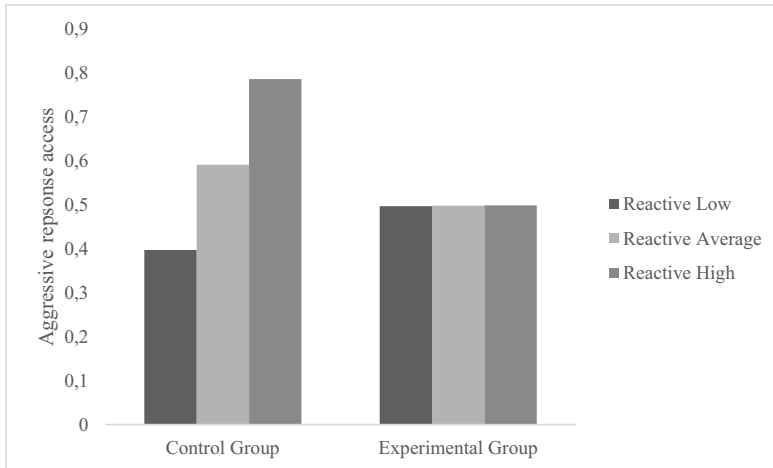


Figure 4. Effect of the Intervention on the Boys' Aggressive Response Access Final Scores for Boys Showing Low, Average and High Levels of Reactive Function

Finally, we examined whether the combination of both training components was effective at reducing Aggressive behavior. We used multiple regression analysis to test the effect of the training on teacher-reported aggression. Aggression post-training served as the dependent variable, aggression at baseline and age were entered in Step 1. Condition was entered in Step 2. The training did not have a significant effect on aggressive behavior ($\beta = .01$, $t(57) = 0.56$, $p = .96$). Both reactive function and psychopathy were entered as moderators in Step 2, and the interaction-terms in Step 3. No main or moderation effects were found.

Discussion

The aim of the present study was to test whether a cognitive bias modification training in the attribution of facial expressions would reduce anger attribution, and increase fear attribution, in clinically referred boys with aggressive behavior problems. We found that the CBM procedure is effective in reducing anger attribution and, tentatively, in increasing fear attribution in these children. These findings are in line with previous research on changing attribution through CBM procedures (e.g. Dadds et al., 2006; Penton-Voak et al.,

2013; Stoddard et al., 2016). For anger attribution, as expected, the training appeared especially effective for children who show average and high levels of reactive aggression (e.g. Crick & Dodge, 1996; De Castro et al., 2002; Dodge et al., 1990).

Furthermore, we explored whether the putative effects of the training on children's emotion attribution would generalize to their interpretation and response access in hypothetical peer provocations, and to their actual aggressive behavior. For response access we did find an additional moderation effect: The training was, again, most effective in changing response access for boys high in reactivity. Thus, for some children, changes in emotion attribution influence the way they think they would respond in ambiguous certain situations. However, we did not find effects on hostile attribution and actual aggressive behavior. This is at odds with previous work, which did find such generalization (Penton-Voak et al., 2013, Vassilopoulos et al., 2015). Our findings do dovetail, however, with meta-analytic findings indicating that generalization of CBM effects to behavioral outcomes in other mental health domains (especially anxiety; Cristea et al., 2015), is rare.

Several factors might explain why we did not find effects on hostile attribution and aggressive behavior. Our participants might attribute less anger to ambiguous expressions, but not incorporate these emotion perceptions in their consequent information processing when interpreting intent or retrieving an actual aggressive behavioural response. Although facial expressions are a key feature in conveying signs of intent (Walker-Andrews, 1997), research also shows that emotion attribution and intent attribution are both unique correlates in aggression (Orobio de Castro et al., 2005). Another explanation may be found in the specificity of the task: Although we already looked at the putative effects of two processing deficits, other social information processes may be important to reduce aggression as well, like cognitive and regulatory control (e.g. Wilkowski, Crowe, & Ferguson, 2015; Saleminck & Wiers, 2011).

This study has limitations. We looked at the added value of targeting multiple emotional processes, by focusing on anger and fear. However, we are unlikely to have caught the complete spectrum of deficits in these children. For example, research shows that the recognition deficits in children with psychopathy might be across modalities (facial, vocal, and postural) and emotions (Dawel, O'Kearney, McKone, & Palermo, 2012). Besides that, we did not focus on training the children on where to look to recognize emotions. Earlier studies have shown that fear recognition in children high on psychopathy may be

impaired because of reduced eye gaze (Dadds, El Masry, Wimalaweera, & Guastella, 2008). Lastly, our aggression measure might not have been sensitive enough to catch slight changes in aggressive behavior in real life. We only looked at effects within one week – future research could focus on a broader time range.

The study also has several strengths. First, we examined an intervention that is cost- and time efficient and does not depend on trainer abilities. Second, CBM procedures circumvent problems that children referred for disruptive behavior can have at this age with explicit forms of cognitive behavioral therapy (McCart, Priester, Davies, & Azen, 2006). Third, we used a multi-method measurement approach, and so our findings do not hinge on subjectivity or mono-method biases. And lastly, we included boys referred for aggressive behavior problems – those who may benefit from intervention the most (De Castro et al., 2002). We examined important subgroups by testing different moderators, capitalizing on the heterogeneity of our sample.

Future work may consider targeting an even broader range of emotion recognition deficits. Moreover, future work could include measures of multiple modalities. For example, research has shown that reduced fear attribution might be related to a failure to attend to informative face regions, such as the eyes (Martin-Key, Graf, Adams, & Fairchild, 2017). Lastly, social information processing theory (Crick & Dodge, 1994) predicts that interventions are effective to the extent that they may impact on recursive processes that unfold over time (Yeager & Walton, 2011). To strengthen such recursive processes, potential additional effects of this training as an add-on to evidence-based interventions (e.g., parent training or classroom management) may be tested both in the short and longer term.

Overall, our study shows promising results regarding the effects of the CBM procedure on aggressive boys' emotion attributions, but not on their aggressive behavior. The intervention is time- and cost-efficient, and relatively easily implemented. CBM procedures may be especially effective in children who show reactive aggression problems. The potential of using CBM to target emotion processing deficits seems substantial, and warrants further empirical scrutiny.



CHAPTER 6

GENERAL DISCUSSION

General Discussion

Aims of this dissertation

One of the most common referrals in child and family psychology are boys in middle childhood who present with disruptive behavior and related family- and school problems (Carr, 2006). Different interventions targeting childhood disruptive behavior problems have already been studied. The most effective interventions for boys with disruptive behavior problems in middle childhood to date are parent-training programs based on operant principles (McCart, Priester, Davies, & Azen, 2006) and conscious re-training of intent attributions (e.g. Guerra & Slaby, 1990; Hudley & Graham, 1993; Lochman & Wells, 2002; Sukhodolsky, Smith, McCauley, Ibrahim, & Piasecka, 2016). However, these interventions have modest and heterogeneous effects. Therefore, it is important to seek for additional ways to intervene in childhood aggression. By doing this, we get better sight on which intervention techniques work and for which children. One promising intervention technique may be direct intervention on social cognitions that are known to evoke aggression.

The first aim of this dissertation was to get more insight in the social cognitions, mood, and self-views in boys who have been referred for disruptive behavior problems, in order to be better able to intervene. The second aim of this dissertation was to test the effects of an innovative cognitive bias modification procedure to (a) change social cognitions involved in aggressive behavior and (b) reduce aggressive behavior in boys with disruptive behavior problems. A mixed methodological and multi-informant approach was used, consisting of cross-sectional and experimental methods. Also, we examined between- and within subjects variance to see individual differences in cognitive, emotional and behavior change over short periods of time.

In the current chapter, first, an integrative summary of the findings is presented, in which the findings are related to theory, and compared to the findings of previous studies. Second, the theoretical and practical implications of these findings are discussed. This is followed by the strengths, limitations and suggestions for future directions. Finally, the conclusion summarizes the take home message of the current dissertation.

Integrative summary of the findings

The General Aggression Model (Allen, Anderson, & Bushman, 2018) and the Social Information Processing model (Crick & Dodge, 1994; Lemerise & Arsenio, 2000) both posit that aggressive behavior is a result of multiple factors influencing social cognitive processes that in turn trigger aggressive behavior. Both models focus on how aggressive behavior results from information processing that is guided by a social database or knowledge structure. This database is built through learning and experience, and becomes automatized with repeated experience. Children who have been clinically referred for disruptive behavior problems typically have developed databases built on aversive social experiences, that result in over-attribution of hostility towards others (i.e. hostile schemata; De Castro, 2002). One would expect such databases to be relatively rigid and difficult to change, because they result from prevalent histories of aversive social experience and/or personal vulnerability (i.e. impulsivity, lower IQ, difficult temperament; Dodge, 2006), and because they protect the individual from a potentially threatening environment. These hostile schemata, in combination with person factors (e.g., self-views) and situational factors, influence further processing of social information, in which affect, cognition and arousal interplay to produce behavior. In our studies, we have focused on different processes within this framework. First, I will describe our findings concerning social cognitions, mood and self-views underlying disruptive behavior. Second, I will focus on the outcomes of our intervention studies concerning these factors.

Self-views, affect and cognition

In this dissertation, we focused on the social cognitive factors self-views, state-anger and hostile interpretation in boys with disruptive behavior problems. We found considerable support for a link between narcissism and aggression (Chapter 2), which extends previous findings in population-based samples of youth to boys with clinically significant behavioral problems. Children with narcissistic traits were more aggressive according to both children themselves and their parents, and these links were independent of children's level or stability of self-esteem. Moreover, in contrast to common clinical intuition, self-esteem was not associated with aggression, suggesting that the current focus in many treatment programs on increasing self-esteem may miss the mark.

However, children with narcissistic traits were not considered more aggressive by

their teachers. Situational specificity of narcissistic aggression might provide an explanation for this latter finding: Narcissistic aggression is typically triggered in situations in which children experience ego-threat (i.e. interpersonal contexts, e.g. Baumeister, Smart, & Boden, 1996). The highly structured school setting within special education might lead to less opportunities in which this type of aggression can be evoked. In terms of the above described models, narcissism seems to be one important personal factor that, given the presence of situational cues, triggers social information processing patterns that spark aggressive behaviors.

Next, we focused on the interplay between affect and cognition in producing disruptive behaviors in their everyday context (Chapter 3). By doing this, we wanted to improve validity of lab findings and inform intervention that specifically targets processes apparent in daily life. To better understand this interplay, we used a naturalistic approach to focus on day-to-day associations between state-anger (affect) and hostile interpretation (cognition). We found that our sample showed day-to-day fluctuations in hostile interpretation of facial expression, but these fluctuations did not co-occur with fluctuations in state-anger, which was in turn independent of children's level of reactive aggression.

Cognitive Bias Modification procedure as an intervention to change social cognition

Since social information processing and related behavioral outcomes are strongly influenced by existing databases (Allen, Anderson, & Bushman, 2018; Crick & Dodge, 1994; Lemerise & Arsenio, 2000), it may be promising to look at ways to change these structures underlying behavior problems. The Cognitive Bias Modification procedure has been designed to do so by changing the contents of the database to repeated practice on a specific task that is designed to slightly change attitudes and bias in information processing (MacLeod, Koster, & Fox, 2009). We examined the effects of CBM to reduce hostile interpretations, and thereby change hostile attributions in general, state-anger, and aggressive outcomes (Chapters 4 and 5). We conducted three studies to test whether this procedure reduces hostile interpretations of facial expressions in boys with disruptive behavior problems. All studies showed that it did. Specifically, one study (Chapter 5) showed that the procedure is especially effective for the subgroup of children who show aggression primarily with a reactive function. In addition, we showed that the procedure improves fear recognition.

However, we found little evidence for generalization of the effects to relevant outcomes – hostile attributions in a different context, state-anger, or aggressive behavior. Only a small generalization effect was found on response access for the subgroup of highly reactively aggressive boys.

Theoretical and practical implications

Theoretical Implications

Self-views, affect and cognition

Whether a link exists between self-views and aggression is somewhat controversial. In clinical practice, much energy is devoted to increase presumed low self-esteem in clients with disruptive behavior problems (e.g. Carr, 2006). Yet scholars have argued that low self-esteem may be neither prevalent nor explanatory for disruptive behavior. Empirically, in population samples, there is no evidence for a robust link between low self-esteem and aggressive behavior, while a link between narcissism and aggressive behavior has been found consistently (e.g. Barry et al., 2007; Boden et al., 2007; Bukowski et al., 2009; Golmaryami & Barry, 2009; Smith et al., 2015; Thomaes et al., 2008). However, despite the primary relevance of this debate to clinical practice, research in clinical samples is limited, and some have argued that there is ample reason not to generalize findings from population samples to often disadvantaged children who present with behavior problems in clinical practice. I hypothesized that personal vulnerabilities and histories of aversive social experience, which are often apparent in clinically referred children, would impact the link between self-views and aggression. For example, parental lack of warmth is known to both undermine healthy self-esteem development and cultivate aggressive behavior (Brummelman et al., 2015; Cicchetti, 2016). My research shows that narcissism, but not self-esteem, is related to clinically severe disruptive behaviors. Thus, my findings challenge the notion that among clinically referred children, aggression may stem from low self-esteem.

Moreover, I found no evidence for day-to-day co-variation of state-anger and hostile interpretations. This seems at odds with both the GAM and the SIP model, which describe how affect and cognition interplay to produce behavior. I think this finding might be a result of other variables influencing the process and acting as a moderator. For example, children scoring high on narcissistic traits might be more reactive to perceived threat or rejection caused by daily stressors (e.g. Baumeister et al., 1996; Thomaes, Bushman, Stegge, & Olthof,

2008). When children high on narcissism experience such an event, this might thus trigger higher peaks of anger and hostile attribution in that moment. However, we did not measure daily stressors, and thus do not know whether and for whom such events took place in the week we measured our concepts. The modest variance over time in our study suggests that there may have been too few daily stressors to set off parallel changes in mood and cognition.

Cognitive Bias Modification procedure as an intervention to change social cognition

Previous literature found a robust causal relation between deviant social cognition and aggressive behavior. Accordingly, I anticipated to be able to influence aggressive behavior via children's social cognition, especially using implicit procedures. However, although we showed that social cognition is malleable, and can be changed using an implicit intervention, generalization to interpretation and behavior in general was not achieved.

In terms of the GAM and SIP models, it might be that the CBM procedure we used fails to target the stable (rigid) generalized database of experience and knowledge that was shaped by numerous social experiences over the life course. This database, in combination with situational factors, might overrule the context-specific interpretations induced by CBM and lead to other responses. Thus perhaps, this procedure changes hostile interpretation of specific practiced stimuli, but this change does not generalize to changes in existing schemata. This implies that a procedure that only targets context-specific interpretations might not be enough to change behaviors in this clinical sample. Moreover, for these children, this pattern of cognition from their database offers them a defense-mechanism in aversive social situations. Therefore, it seems logical that if the environment does not change, the database is resistant also to change.

Another view on how our findings fit within these models, is that there are factors involved that are not explicitly captured by the models and existing literature. It is not completely clear what targeting one single aspect (e.g. self-view, affect or cognition) of the whole social information processing results in in the long term; one would expect, from the GAM and SIP models, that other factors change with it. However, it is not yet well understood how these recursive processes take place and over what time. Besides, in theory all different factors involved in social information processing are given the same weight in

their contribution in producing a behavioral outcome, but what if some factors overrule others? From our research, it seems as if personal and contextual factors and the database of knowledge overrule temporal changes in interpretations and affect.

Since previous studies using CBM have shown larger effects on actual behavior (e.g. Penton-Voak, et al., 2017; Penton-Voak, et al., 2013; Stoddard et al., 2016; Vassilopoulos, et al., 2015), it seems a promising addition to other (already existing) programs. We think that the reason why we did not find effects on behavior, and others did, suggests heterogeneity: The CBM procedure might be especially effective for some children, but not for others. Yet until now, researchers have not often focused on possible moderators of CBM effects on behavioral change (Cristea et al., 2015; Krebs et al., 2018). We made a start by looking at possible moderating effects of reactive function of aggression, and found small effects. Perhaps the effects of CBM are moderated by other factors as well, such as task characteristics (e.g., number of trials, way of presenting the pictures, what kind of pictures are used). For example, research on CBM for anxiety problems shows that effects of CBM on cognition depend on whether words or pictures are used as stimuli, with larger effects for words (Kindt & Brosschot, 1999). Although the Cognitive Bias Modification procedure we examined does not seem sufficient to achieve long term change in behavior in a clinical sample, it does show promising effects on changes in social cognition.

Practical Implications

The first practical implication of my research concerns self-views in relation to disruptive behaviors. Treatments in clinical settings are still often focused on enhancing self-esteem in children. But my research, together with previous literature in population samples, suggests that self-esteem is not a correlate of aggressive behavior. Focusing on narcissism might be more effective to reduce disruptive behavior. Narcissism could be assessed as an explanatory factor in psychological assessment procedures. It could be included in treatment for children themselves, for example by practicing responding to criticism or failure. It may also be addressed by instructing parents and teachers, since we know that specific parenting practices foster narcissism (Brummelman, 2016). For example, at present, treatment advice to parents often includes giving compliments to heighten self-esteem. However, if not executed correctly, such procedures might backfire to the extent that parental “overvaluation” can lead to heightened levels of narcissism (Brummelman et

al., 2016). Children with ADHD and other problem behaviors often already show unrealistically high positive self-views; it is not inconceivable that giving excessive compliments might reinforce this view (Hoza et al., 2004). A plausible alternative might be to help children handle criticism or failure without resorting to self-defensive strategies like denial and aggression, and adopt a more realistic self-views with acceptance of one's strengths and weaknesses.

A second practical implication concerns the complex interplay between different factors within social information processing (e.g. self-views, affect and cognition) and in the environment (e.g. social cues, peer relations) in producing disruptive behaviors. Our results suggest that targeting one of these factors in this cycle might not be enough to produce change in behavior. In the assessment cycle, explanatory factors of behavior are tested and treatment is indicated based on these multiple factors. However, in evaluation research goals of treatment are often limited to changes in actual behavior. This is reinforced by national requirements, such as Routine Outcome Monitoring, which focuses on reductions of problem behaviors. What exactly happens when targeting the explanatory factors in a specific individual, and how these in turn influence actual behavior, is often not specifically monitored. These processes are assumed based on literature, and not often assessed or reflected on during the treatment process. We would suggest to set treatment goals for the specific explanatory factors, and monitor how changes in these factors result in changes in actual behavior. That is, to not only focus on effect sizes based on changes in problem behaviors, but also measure effects on the specific factors that are targeted. To get insight in the specific factors resulting in aggressive behavior for each individual, therapists should execute personalized function analysis for each client. By using this method, one can see which factors might overrule others and what methods might be appropriate to include in treatment. Also, this is a way to see how adaptive behavior really is. In some contexts, for some children, aggression might be adaptive behavior. Focusing on the different contexts for every child in treatment, should therefore add to effectiveness of the treatment.

A third practical implication stems from our intervention studies. Our studies show the potential of employing the CBM procedure to change interpretation of emotion. It is cost- and time efficient, and relatively easily implemented. Despite this potential, the context-dependency and lack of generalization of effects should be better understood before the procedure can be taken to scale in clinical populations.

Strengths, limitations and future directions

One strength of our research is our sample. We only included boys who were clinically referred for their disruptive behaviors. Few previous studies on self-views, SIP, and CBM included this clients in this (age)group, while findings with general population samples are frequently overgeneralized to clinical groups, and in practice this is by far the biggest group that is referred for disruptive behaviors. Targeting behaviors and underlying factors in this group might result in better life-long outcomes, and can save costs on the long term. Second, we used Randomized Controlled Trials with active control conditions to rule out that we would have found any effects as a by-product of attention. We also used a multi-method assessment approach, thus circumventing subjectivity and mono-method biases. Lastly, we used multiple experiments and studies to replicate and extend our key findings.

This research has some methodological limitations. Our samples of clinically referred children were relatively small, and so statistical power to detect significant effects was somewhat lower than is typical in population-based studies. Also, we did not know whether all children were sufficiently biased to profit from CBM, because we did not have norm-groups for the CBM procedure we tested. Lastly, although our studies took place in the school-context, children were still taken out of their daily context (the classroom). Since most participants liked being taken from the classroom to do something else, this might have led to underreporting of, for example, state-anger.

Future research should be more comprehensive in examining the interplay between self-views, affect, and cognition to produce disruptive behaviors. More specifically, moderating factors should be studied more extensively (what works for whom). Children can be more receptive for change when multiple factors are targeted at once, because there might be perpetuating factors at work. Different implicit personal factors might be of influence as well, like low glucose levels (Bushman, DeWall, Pond, & Hanus, 2014) or amount of physical activity (Hoza et al., 2015). Second, future research should examine the effects of CBM procedures as an addition to existing (parenting) programs, to see any combined effects. We tested the effects of a very specific element, while literature already showed how effective interventions usually require a combination of multiple elements. Concerning CBM, it seems theoretically plausible that effects on behavior would require a concerted effort to change hostile attributions with CBM together with a concerted change in hostility form the environment (for example by starting in a new classroom).

This begs the question whether one should test complete, comprehensive interventions on effectiveness, or focus on effects of specific intervention elements like CBM. I think there are pro's and con's for both. If we only test parts of interventions, we examine an isolated effect. In practice, often a combinations of parts cause effects, and by doing this one might not find an effect that would have actually been present when combined with other elements. However, it is not cost- and time efficient to always complete a whole comprehensive program if only some elements are actually effective. Moreover, if we would only test complete multicomponent intervention packages, we would never be able to detect which elements cause its effect. I don't think there is one answer to this question, but I would recommend to test separate elements for indications of potential effectiveness by first measuring effects on underlying mechanisms that are targeted as proximal outcomes. So, for example, to examine whether CBM is a promising candidate for inclusion in intervention packages, we first tested its effects on the proximal outcome hostile intent attribution. To see whether these elements affect behavior in itself or in combination with other elements, one could change the design to include within-person sequences of elements, to check for combination and order-effects. Clearly, progress in this area will require innovative designs to test the effectiveness of (combinations of) specific elements.

In sum, the studies in this dissertation show that aggression is related to narcissism in children who have been referred for disruptive behavior problems, and not with self-esteem. The implicit CBM procedure reduces hostile interpretation, but this effect does not generalize to aggressive behavior. This research stresses the complexity of factors involved in producing disruptive behaviors in clinically referred boys. Future research may continue to look for implicit processes and the usability of CBM procedures for this clinical sample. To get more insight in how different factors are at interplay in a specific individual, practitioners should carry out function analyses to see which factors should be targeted in treatment. During treatment, one should monitor the changes in these factors, together with changes in actual behavior. In the absence of solid evidence for substantial effects of CBM on aggressive behavior, I would advise practitioners to use existing programs that have already proven to be effective in changing disruptive behaviors.

The background of the page is a watercolor wash. It features a mix of colors: light blue, pale yellow, and soft grey. The colors are blended together, creating a textured, organic feel. There are some darker, more saturated areas, particularly in the blue and grey tones, which contrast with the lighter, more washed-out areas. The overall effect is artistic and somewhat abstract.

REFERENCES

- Alexander, R.J. (2001). Border crossings: towards a comparative pedagogy. *Comparative Education, 37*(4), 507-523.
- Allen, J. J., Anderson, C. A., & Bushman, B. J. (2018). The general aggression model. *Current Opinion in Psychology, 19*, 75-80.
- Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology, 53*, 27–51.
- Barry, C.T., Frick, P.J., & Killian, A.L. (2003). The Relation of Narcissism and Self-Esteem to Conduct Problems in Children: A Preliminary Investigation, *Journal of Clinical Child & Adolescent Psychology, 32*(1), 139-152.
- Barry, T. D., Thompson, A., Barry, C. T., Lochman, J. E., Adler, K., & Hill, K. (2007). The importance of narcissism in predicting proactive and reactive aggression in moderately to highly aggressive children. *Aggressive Behavior, 33*(3), 185-197.
- Baumeister, R.F., Smart, L., & Boden, J.M. (1996). Relation of threatened egotism to violence and aggression: The dark side of high self-esteem. *Psychological Review, 103*(1), 5-33.
- Blair, R.J.R. (1999). Responsiveness to distress cues in the child with psychopathic tendencies. *Personality and Individual Differences, 27*(1), 135-145.
- Blair, R.J.R. (2001). Neurobiological basis of psychopathy. *The British Journal of Psychiatry, 182*(1), 5-7.
- Boden, J. M., Fergusson, D. M., & Horwood, L. J. (2007). Self-esteem and violence: Testing links between adolescent self-esteem and later hostility and violent behavior. *Social Psychiatry and Psychiatric Epidemiology, 42*(11), 881-891.
- Bodenhausen, G. V., Sheppard, L. A., & Kramer, G. P. (1994). Negative affect and social judgment: The differential impact of anger and sadness. *European Journal of Social Psychology, 24*(1), 45-62.
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*(1), 579-616.

- Brummelman, E., & Thomaes, S., Nelemans, S.A., Orobio de Castro, B., Overbeek, G., & Bushman, B.J. (2015). Origins of narcissism in children. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(12), 3659-3662.
- Brummelman, E., Thomaes, S., & Sedikides, C. (2016). Separating narcissism from self-esteem. *Current Directions in Psychological Science*, *25*(1), 8-13.
- Bukowski, W.M., Schwartzman, A., Santo, J., Bagwell, C., & Adams, R. (2009). Reactivity and distortions in the self: Narcissism, types of aggression, and the functioning of the hypothalamic-pituitary-adrenal axis during early adolescence. *Development and Psychopathology*, *21*, 1249-1262.
- Buntaine, R. L., & Costenbader, V. K. (1997). Self-reported differences in the experience and expression of anger between girls and boys. *Sex Roles*, *36*(9-10), 625-637.
- Bushman, B.J., & Baumeister, R.F. (1998). Threatened egotism, narcissism, self-esteem and direct and displaced aggression: Does self-love lead to violence? *Journal of Personality and Social Psychology*, *75*(1), 219-229.
- Bushman, B. J., DeWall, C. N., Pond, R. S., & Hanus, M. D. (2014). Low glucose relates to greater aggression in married couples. *Proceedings of the National Academy of Sciences*, 201400619.
- Cadesky, E. B., Mota, V. L., & Schachar, R. J. (2000). Beyond words: how do children with ADHD and/or conduct problems process nonverbal information about affect?. *Journal of the American Academy of Child & Adolescent Psychiatry*, *39*(9), 1160-1167.
- Card, N.A., Stucky, B.D., Sawalani, G.M., & Little, T.D. (2008). Direct and indirect aggression during childhood and adolescence: A meta-analytic review of gender differences, intercorrelations, and relations to maladjustment. *Child Development*, *79*(5), 1185-1229.
- Carr, A. (2006). *The handbook of child and adolescent clinical psychology: A contextual approach*. Hove, East Sussex: Routledge.
- Chang, L., Schwartz, D., Dodge, K. A., & McBride-Chang, C. (2003). Harsh parenting in relation to child emotion regulation and aggression. *Journal of Family Psychology*, *17*(4), 598.

- Cicchetti, D. (2016). Socioemotional, personality, and biological development: Illustrations from a multilevel developmental psychopathology perspective on child maltreatment. *Annual Review of Psychology, 67*, 187-211.
- Crick, N.R., & Dodge, K.A. (1994). A review and reformulation of social information processing mechanisms in children's social adjustment. *Psychological Bulletin, 115*(1), 74-101.
- Crick, N.R., & Dodge, K.A. (1996). Social information-processing mechanisms in reactive and proactive aggression. *Child Development, 67*(3), 993-1002.
- Cristea, I.A., Mogoase, C., David, D., & Cuijpers, P. (2015). Practitioner review: Cognitive bias modification for mental health problems in children and adolescents: a meta-analysis. *Journal of Child Psychology & Psychiatry, 56*(7), 723-734.
- Crockenberg, S., & Langrock, A. (2001). The role of specific emotions in children's responses to interparental conflict: A test of the model. *Journal of Family Psychology, 15*(2), 163.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Second edition). NY: Routledge Academic.
- Dadds, M.R., El Masry, Y., Wimalaweera, S., & Guastella, A.J. (2008). Reduced eye gaze explains "fear blindness" in childhood psychopathic traits. *Journal of the American Academy of Child & Adolescent Psychiatry, 47*(4), 455-463.
- Dadds, M.R., Perry, Y., Hawes, D.K., Merz, S., Riddell, A.C., Haines, D.J., ... Abeygunawardane, A.I. (2006). Attention to the eyes and fear-recognition deficits in child psychopathy. *The British Journal of Psychiatry, 189*(3), 280-281.
- David, C.F., & Kistner, J.A. (2000). Do positive self-perceptions have a "dark side"? Examination of the link between perceptual bias and aggression. *Journal of Abnormal Child Psychology, 28*(4), 327-337.
- Dawel, A., O'Kearney, R., McKone, E., & Palermo, R. (2012). Not just fear and sadness: meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neuroscience & Biobehavioral Reviews, 36*(10), 2288-2304.

- De Castro, B., Veerman, J.W., Koops, W., Bosch, J.P., & Monshouwer, H.J. (2002). Attribution of intent and aggressive behavior: A meta-analysis. *Child Development, 73*(3), 916-934.
- De Graaf, I., Speetjens, P., Smit, F., de Wolff, M., & Tavecchio, L. (2008). Effectiveness of the Triple P Positive Parenting Program on behavioral problems in children: A meta-analysis. *Behavior Modification, 32*(5), 714-735.
- Del Barrio, V., Aluja, A., & Spielberger, C. (2004). Anger assessment with the STAXI-CA: Psychometric properties of a new instrument for children and adolescents. *Personality and Individual Differences, 37*(2), 227-244.
- Dishion, T.J., French, D.C., & Patterson, G.R. (1995). The development and ecology of antisocial behavior. In D. Cicchetti & D.J. Cohen (Eds.), *Developmental Psychopathology. Vol.2: Risk, Disorder, and Adaptation* (pp. 421-471). New York: John Wiley & Sons.
- Dodge, K.A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. *Development and Psychopathology, 18*, 791-814.
- Dodge, K.A., Bates, J.E., & Pettit, G.S. (1990). Mechanisms in the cycle of violence. *Science, 250*(4988), 1678-1683.
- Dodge, K. A., Lansford, J. E., Burks, V. S., Bates, J. E., Pettit, G. S., Fontaine, R., & Price, J. M. (2003). Peer rejection and social information-processing factors in the development of aggressive behavior problems in children. *Child Development, 74*(2), 374-393.
- Dodge, K. A., Malone, P. S., Lansford, J. E., Sorbring, E., Skinner, A. T., Tapanya, S., ... Bacchini, D. (2015). Hostile attributional bias and aggressive behavior in global context. *Proceedings of the National Academy of Sciences, 112*(30), 9310-9315.
- Dodge, K.A., Price, J.M., Bachorowski, J.-A., & Newman, J.P. (1990). Hostile attributional biases in severely aggressive adolescents. *Journal of Abnormal Psychology, 99*(4), 385-392.
- Dodge, K. A., & Somberg, D. R. (1987). Hostile attributional biases among aggressive boys are exacerbated under conditions of threats to the self. *Child Development, 58*, 213-224.
- Donnellan, M.B., Trzesniewski, K.H., Robins, R.W., Moffitt, T.E., & Caspi, A. (2005). Low self-esteem is

- relation to aggression, antisocial behavior, and delinquency. *Psychological Science*, *16*(4), 328-335.
- Eisenberg, N., Sadovsky, A., Spinrad, T.L., Fabes, R.A., Losoya, S.H., Valiente, C., ... Shepard, S.A. (2005). The relations of problem behavior status to children's negative emotionality, effortful control, and impulsivity: Concurrent relations and prediction of change. *Developmental Psychology*, *41*(1), 193-211.
- Ekman P. (1993). Facial expression and emotion. *American Psychologist*, *48*(4), 384-392.
- Elias, M.J. (2004). The connection between social-emotional learning and learning disabilities: Implications for intervention. *Learning Disability Quarterly*, *27*(1), 53-63.
- Etcoff, N.L., & Magee, J.J. (1992). Categorical perception of facial expressions. *Cognition*, *44*(3), 227-240.
- Ettekal, I., & Ladd, G. W. (2015). Developmental pathways from childhood aggression–disruptiveness, chronic peer rejection, and deviant friendships to early-adolescent rule breaking. *Child Development*, *86*(2), 614-631.
- Fairchild, G., Van Goozen, S.H.M., Calder, A.J., Stollery, S.J., & Goodyer, I.A. (2009). Deficits in facial expression recognition in male adolescents with early-onset of adolescence-onset conduct disorder. *The Journal of Child Psychology and Psychiatry*, *50*(5), 627-636.
- Falkenbach, D.M., Howe, J.R., & Falki, M. (2013). Using self-esteem to disaggregate psychopathy, narcissism, and aggression. *Personality and Individual Differences*, *54*, 815-820.
- Farrar-Schneider, D. (1992). Aggression and noncompliance: Behavior modification. *Autism in Children and Adults: Etiology, Assessment, and Intervention*, 181-191.
- Fergusson, D. M., Horwood, L.J., & Ridder, E. M. (2005). Show me the child at seven: the consequences of conduct problems in childhood for psychosocial functioning in adulthood. *Journal of Child Psychology and Psychiatry*, *46*(8), 837-849.
- Forgas, J. P., & East, R. (2008). How real is that smile? Mood effects on accepting or rejecting the veracity of emotional facial expressions. *Journal of Nonverbal Behavior*, *32*(3), 157-170.

- Fossati, A., Borroni, S., Eisenberg, N., & Maffei, C. (2010). Relations of proactive and reactive dimensions of aggression to overt and covert narcissism in nonclinical adolescents. *Aggressive Behavior, 36*(1), 21-27.
- Frick, P.J., Cornell, A.H., Barry, C.T., Bodin, S.D., Dane, H.E. (2003). Callous-Unemotional traits and conduct problems in the prediction of conduct problem severity, aggression, and self-report of delinquency. *Journal of Abnormal Child Psychology, 31*(4), 457-470.
- Frick, P.J., & Hare, R.D. (2001). *Antisocial process screening device: APSD*. Toronto: Multi-Health Systems.
- Frick, P.J., Ray, J.V., Thornton, L.C., & Kahn, R.E. (2014). Can callous-unemotional traits enhance the understanding, diagnosis, and treatment of serious conduct problems in children and adolescents? A comprehensive review. *Psychological Bulletin, 140*(1), 1-57.
- Golmaryami, F.N., & Barry, C.T. (2009). The associations of self-reported and peer-reported relation aggression with narcissism and self-esteem among adolescents in a residential setting. *Journal of Clinical Child & Adolescent Psychology, 39*(1), 128-133.
- Gresham, F. M., MacMillan, D. L., Bocian, K. M., Ward, S. L., & Forness, S. R. (1998). Comorbidity of hyperactivity-impulsivity-inattention and conduct problems: Risk factors in social, affective, and academic domains. *Journal of Abnormal Child Psychology, 26*(5), 393-406.
- Guerra, N.G., & Slaby, R.G. (1990). Cognitive mediators of aggression in adolescent offenders: 2. Intervention. *Developmental Psychology, 26*(2), 269-277.
- Hallion, L.S., & Ruscio, A.M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin, 137*(6), 940-958.
- Hamaker, E. L., Ceulemans, E., Grasman, R. P. P., & Tuerlinckx, F. (2015). Modeling affect dynamics: State of the art and future challenges. *Emotion Review, 7*(4), 316-322.
- Harter, S. (1985). *The Self-Perception Profile for Children: Revision of the Perceived Competence Scale for Children*. Denver, CO: University of Denver.
- Harter, S. (2012). *The construction of the self*. New York: The Guilford Press.

- Hendriks, M.C.P., & Vingerhoets, J.J.M. (2006). Social messages of crying faces: Their influence on anticipated person perception, emotions and behavioural responses. *Cognition and Emotion, 20*(6), 878-886.
- Henry, J. D., Moses, E., Castellini, J., & Scott, J. (2015). Mental health problems in adolescence and the interpretation of unambiguous threat. *PloS One, 10*(6), e0127167.
- Hertel, P. T., & Mathews, A. (2011). Cognitive bias modification past perspectives, current findings, and future applications. *Perspectives on Psychological Science, 6*(6), 521–536.
- Hoeve, M., Dubas, J. S., Eichelsheim, V. I., Van Der Laan, P. H., Smeenk, W., & Gerris, J. R. (2009). The relationship between parenting and delinquency: A meta-analysis. *Journal of Abnormal Child Psychology, 37*(6), 749-775.
- Horsley, T. A., de Castro, B. O., & Van der Schoot, M. (2010). In the eye of the beholder: Eye-tracking assessment of social information processing in aggressive behavior. *Journal of Abnormal Child Psychology, 38*(5), 587-599.
- Hoza, B., Gerdes, A. C., Hinshaw, S. P., Arnold, L. E., Pelham Jr, W. E., Molina, B. S., ... & Wigal, T. (2004). Self-perceptions of competence in children with ADHD and comparison children. *Journal of Consulting and Clinical Psychology, 72*(3), 382-391.
- Hoza, B., Pelham, W. E., Milich, R., Pillow, D., & McBride, K. (1993). The self-perceptions and attributions of attention deficit hyperactivity disorder and nonreferred boys. *Journal of Abnormal Child Psychology, 21*(3), 271-286.
- Hoza, B., Smith, A. L., Shoulberg, E. K., Linnea, K. S., Dorsch, T. E., Blazo, J. A., ... & McCabe, G. P. (2015). A randomized trial examining the effects of aerobic physical activity on attention-deficit/hyperactivity disorder symptoms in young children. *Journal of Abnormal Child Psychology, 43*(4), 655-667.
- Hubbard, J. A., Dodge, K. A., Cillessen, A. H., Coie, J. D., & Schwartz, D. (2001). The dyadic nature of social information processing in boys' reactive and proactive aggression. *Journal of Personality and Social Psychology, 80*(2), 268.

- Hudley, C., & Graham, S. (1993). An attributional intervention to reduce peer-directed aggression among African-American boys. *Child Development, 64*(1), 124-138.
- Hurley, J.C., & Underwood, M.K. (2002). Children's understanding of their research rights before and after debriefing: Informed assent, confidentiality, and stopped participation. *Child Development, 73*(1), 132-144.
- Kernberg, O. F. (1975). A systems approach to priority setting of interventions in groups. *International Journal of Group Psychotherapy, 25*(3), 251-275.
- Kernis, M. H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry, 14*(1), 1-26.
- Kernis, M. H. (2005). Measuring self-esteem in context: The importance of stability of self-esteem in psychological functioning. *Journal of Personality, 73*(6), 1569-1605.
- Kernis, M. H., Grannemann, B. D., & Barclay, L. C. (1989). Stability and level of self-esteem as predictors of anger arousal and hostility. *Journal of Personality and Social Psychology, 56*(6), 1013.
- Kindt, M., & Brosschot, J. F. (1999). Cognitive bias in spider-phobic children: Comparison of a pictorial and a linguistic spider Stroop. *Journal of Psychopathology and Behavioral Assessment, 21*(3), 207-220.
- Kohut, H. (1977). *The restoration of the self*. Madison, CT: International Universities Press Inc.
- Koster, E. H., Fox, E., & MacLeod, C. (2009). Introduction to the special section on cognitive bias modification in emotional disorders. *Journal of Abnormal Psychology, 118*(1), 1.
- Krebs, G., Pile, V., Grant, S., Degli Esposti, M., Montgomery, P., & Lau, J. Y. (2017). Research Review: Cognitive bias modification of interpretations in youth and its effect on anxiety: a meta-analysis. *Journal of Child Psychology and Psychiatry, 59*(8), 831-844.
- Kusché, C. A., & Greenberg, M. T. (1994). *The PATHS curriculum*. Seattle, WA: Developmental Research and Programs.
- Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data.

- Biometrics*, 33(1), 159-174.
- Lemerise, E. A., & Arsenio, W. F. (2000). An integrated model of emotion processes and cognition in social information processing. *Child Development*, 71(1), 107-118.
- Lochman, J.E., & Wells, K.C. (2002). Contextual social-cognitive mediators and child outcome: A test of the theoretical model in the Coping Power program. *Development and Psychopathology*, 14(4), 945-967.
- Lombardo, M. V., & Baron-Cohen, S. (2010). Unraveling the paradox of the autistic self. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(3), 393-403.
- MacLeod, C., Koster, E. H.W., & Fox, E. (2009). Whither cognitive bias modification research? Commentary on the special section articles. *Journal of Abnormal Psychology*, 118(1), 89-99.
- MacLeod, C., & Mathews, A. (2012). Cognitive bias modification approaches to anxiety. *Annual Review of Clinical Psychology*, 8, 189-217.
- Mann, M. M., Hosman, C. M., Schaalma, H. P., & De Vries, N. K. (2004). Self-esteem in a broad-spectrum approach for mental health promotion. *Health Education Research*, 19(4), 357-372.
- Marsh, A. A., & Blair, R. J. R. (2008). Deficits in facial affect recognition among antisocial populations: a meta-analysis. *Neuroscience & Biobehavioral Reviews*, 32(3), 454-465.
- Martin-Key, N. A., Graf, E. W., Adams, W. J., & Fairchild, G. (2017). Facial emotion recognition and eye movement behaviour in conduct disorder. *Journal of Child Psychology and Psychiatry*, 59(3), 247-257.
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44(2), 314- 324.
- McCart, M. R., Priester, P. E., Davies, W. H., & Azen, R. (2006). Differential effectiveness of behavioral parent-training and cognitive-behavioral therapy for antisocial youth: A meta-analysis. *Journal of Abnormal Child Psychology*, 34(4), 525-541.
- McClure, E.B. (2000). A meta-analytic review of sex differences in facial expression processing and

- their development in infants, children, and adolescents. *Psychological Bulletin*, 126(3), 424-453.
- Mellentini, A.I., Dervisevic, A., Stenager, E., Pilegaard, M., & Kirk, U. (2015). Seeing enemies? A systematic review of anger bias in the perception of facial expressions among anger-prone and aggressive populations. *Aggression and Violent Behavior*, 25, 373-383.
- Monteith, M.J., Sherman, J.W. & Devine, P.G. (1998). Suppression as a stereotype control strategy. *Personality and Social Psychology*, 2(1), 63-82.
- Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry*, 12(4), 177-196.
- Muthén, L. K., & Muthén, B. (2015). Mplus. *The comprehensive modelling program for applied researchers: user's guide*, 5.
- Nasby, W., Hayden, B., & DePaulo, B.M. (1980). Attributional bias among aggressive boys to interpret unambiguous social stimuli as displays of hostility. *Journal of Abnormal Psychology*, 89(3), 459-468.
- Nix, R.L., Pinderhughes, E.E., Dodge, K.A., Bates, J.E., Pettit, G.S., & McFadyen-Ketchum, S.A. (1999). The relation between mothers' hostile attribution tendencies and children's externalizing behavior problems: The mediating role of mothers' harsh discipline practices. *Child Development*, 70(4), 896-909.
- Olson, S. L. (1992). Development of conduct problems and peer rejection in preschool children: A social systems analysis. *Journal of Abnormal Child Psychology*, 20(3), 327-350.
- Orobio de Castro, B., Brendgen, M., van Boxtel, H., Vitaro, F., & Schaepers, L. (2007). "Accept me, or else...": Disputed overestimation of social competence predicts increases in proactive aggression. *Journal of Abnormal Child Psychology*, 35(2), 165-178.
- Orobio de Castro, B., Merk, W., Koops, W., Veerman, J. W., & Bosch, J. D. (2005). Emotions in social information processing and their relations with reactive and proactive aggression in referred aggressive boys. *Journal of Clinical Child and Adolescent Psychology*, 34(1), 105-116.

- Orobio de Castro, B., Slot, N. W., Bosch, J. D., Koops, W., & Veerman, J. W. (2003). Negative feelings exacerbate hostile attributions of intent in highly aggressive boys. *Journal of Clinical Child and Adolescent Psychology, 32*(1), 56-65.
- Patterson, G. R. (1976). The aggressive child: Victim and architect of a coercive system. *Behavior Modification and Families, 1*, 267-316.
- Penton-Voak, I.S., Munafò, I.S., & Yen Looi, C. (2017). Biased facial-expression perception in mental health disorders: A possible target for psychological intervention? *Current Directions in Psychological Science, 26*(3), 294-301.
- Penton-Voak, I.S., Thomas, J., Gage, S.H., McMurrin, M., McDonald, S., & Munafò, M.R. (2013). Increasing recognition of happiness in ambiguous facial expressions reduces anger and aggressive behavior. *Psychological Science, 24*(5), 688-697.
- Polman, H., Orobio de Castro, B., Koops, W., van Bortel, H.W., & Merk, W.W. (2007). A meta-analysis of the distinction between reactive and proactive aggression in children and adolescents. *Journal of Abnormal Child Psychology, 35*, 522-535.
- Polman, H., Orobio de Castro, B., Thomaes, S., & van Aken, M. (2009). New directions in measuring reactive and proactive aggression: Validation of a teacher questionnaire. *Journal of Abnormal Child Psychology, 37*, 183-193.
- Raskin, R., Novacek, J., & Hogan, R. (1991). Narcissism, Self-Esteem, and Defensive Self-Enhancement. *Journal of Personality, 59*(1), 19-38.
- Reijntjes, A., Thomaes, S., Kamphuis, J. H., Bushman, B. J., De Castro, B. O., & Telch, M. J. (2011). Explaining the paradoxical rejection-aggression link: The mediating effects of hostile intent attributions, anger, and decreases in state self-esteem on peer rejection-induced aggression in youth. *Personality and Social Psychology Bulletin, 37*(7), 955-963.
- Roff, J. D., & Wirt, R. D. (1984). Childhood aggression and social adjustment as antecedents of delinquency. *Journal of Abnormal Child Psychology, 12*(1), 111-126.
- Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University

Press.

- Rosenberg, M., Schooler, C., & Schoenbach, C. (1989). Self-esteem and adolescent problems: Modeling reciprocal effects. *American Sociological Review*, 1004-1018.
- Salemink, E., & Wiers, R.W. (2011). Modifying threat-related interpretive bias in adolescents. *Journal of Abnormal Child Psychology*, 39(7), 967-976.
- Salzer Burks, V., Laird, R. D., Dodge, K. A., Pettit, G. S., & Bates, J. E. (1999). Knowledge structures, social information processing, and children's aggressive behavior. *Social Development*, 8(2), 220-236.
- Salmivalli, C., Ojanen, T., Haanpää, J., & Peets, K. (2005). "I'm OK but you're not" and other peer-relational schemas: explaining individual differences in children's social goals. *Developmental Psychology*, 41(2), 363.
- Schonfeld, D. J., Adams, R. E., Fredstrom, B. K., Weissberg, R. P., Gilman, R., Voyce, C., ... & Speese-Linehan, D. (2015). Cluster-randomized trial demonstrating impact on academic achievement of elementary social-emotional learning. *School Psychology Quarterly*, 30(3), 406-420.
- Schönenberg, M., & Justyte, A. (2014). Investigation of the hostile attribution bias toward ambiguous facial cues in antisocial violent offenders. *European Archives of Psychiatry & Clinical Neuroscience*, 264(1), 61-69.
- Schultz, D., Izard, C. E., & Ackerman, B. P. (2000). Children's anger attribution bias: Relations to family environment and social adjustment. *Social Development*, 9(3), 284-301.
- Schultz, D., Izard, C. E., & Bear, G. (2004). Children's emotion processing: Relations to emotionality and aggression. *Development and Psychopathology*, 16(2), 371-387.
- Seah, S. L., & Ang, R. P. (2008). Differential correlates of reactive and proactive aggression in Asian adolescents: Relations to narcissism, anxiety, schizotypal traits, and peer relations. *Aggressive Behavior*, 34(5), 553-562.
- Shields, A. M., Cicchetti, D., & Ryan, R. M. (1994). The development of emotional and behavioral self-regulation and social competence among maltreated school-age children. *Development and*

- Psychopathology*, 6(1), 57-75.
- Smith, J. D., Dishion, T. J., Shaw, D. S., Wilson, M. N., Winter, C. C., & Patterson, G. R. (2014). Coercive family process and early-onset conduct problems from age 2 to school entry. *Development and Psychopathology*, 26(4), 917-932.
- Smith, S. D., Lynch, R. J., Stephens, H. F., & Kistner, J. A. (2015). Self-perceptions and their prediction of aggression in male juvenile offenders. *Child Psychiatry & Human Development*, 46(4), 609-621.
- Spielberger, C. D. (1988). *Manual for the State-Trait Anger Expression Inventory*. Odessa, FL: Psychological Assessment Resources.
- Statistics Netherlands (2014, July 21). 70 Thousand pupils in special education ("70 Duizend leerlingen in het special basisonderwijs"). Retrieved from <https://www.cbs.nl/nl-nl/nieuws/2014/30/70-duizend-basisschoolleerlingen-in-speciaal-onderwijs>.
- Stoddard, J., Sharif-Askary, B., Harkins, E.A, Frank, H.R., Brotman, M.A., Penton-Voak, I.S., ... Leibenluft, E. (2016). An open pilot study of training hostile interpretation bias to treat disruptive mood dysregulation disorder. *Journal of Child and Adolescent Psychopharmacology*, 26(1), 49-57.
- Sukhodolsky, D. G., Smith, S. D., McCauley, S. A., Ibrahim, K., & Piasecka, J. B. (2016). Behavioral interventions for anger, irritability, and aggression in children and adolescents. *Journal of Child and Adolescent Psychopharmacology*, 26(1), 58-64.
- Thomaes, S., & Brummelman, E. (2016). Narcissism. *Developmental Psychopathology*, 3(16), 1-47.
- Thomaes, S., Brummelman, E., Reijntjes, A., & Bushman, B.J. (2013). When narcissus was a boy: Origins, nature and consequences of childhood narcissism. *Child Development Perspectives*, 7(1), 22-26.
- Thomaes, S., Bushman, B.J, Orobio de Castro, B., Cohen, G.L., & Denissen, J. (2009). Reducing narcissistic aggression by buttressing self-esteem: An experimental field study. *Psychological Science*, 20(12), 1536-1542.

- Thomaes, S., Bushman, B.J., Stegge, H. & Olthof, T. (2008). Trumping shame by blasts of noise: Narcissism, self-esteem, shame, and aggression in young adolescents. *Child Development*, 79(6), 1792-1801.
- Thomaes, S., Stegge, H., Bushman, B.J., Olthof, T., & Denissen, J. (2008). Development and validation of the childhood narcissism scale. *Journal of Personality Assessment*, 90(4), 382-391.
- Thomaes, S., Stegge, H., Olthof, T., Bushman, B. J., & Nezlek, J. B. (2011). Turning shame inside-out: "humiliated fury" in young adolescents. *Emotion*, 11(4), 786-793.
- Tracy, J. L., & Robins, R. W. (2003). "Death of a (Narcissistic) Salesman:" An Integrative Model of Fragile Self-Esteem. *Psychological Inquiry*, 14(1), 57-62.
- Vassilopoulos, S.P., Brouzos, A., & Andreou, E. (2015). A multi-session attribution modification program for children with aggressive behaviour: Changes in attributions, emotional reaction estimates, and self-reported aggression. *Behavioural and Cognitive Psychotherapy*, 43(5), 538-548.
- Vlamings, P.H.J.M., Jonkman, L.M., & Kemner, C. (2010). An eye for detail: An event-related potential study of the rapid processing of fearful facial expressions in children. *Child Development*, 81(4), 1304-1319.
- Walker-Andrews, A.S. (1997). Infants' perception of expressive behaviors: Differentiation of multimodal information. *Psychological Bulletin*, 121(3), 437-456.
- Webster, G. D., Kirkpatrick, L. A., Nezlek, J. B., Smith, C. V., & Paddock, E. L. (2007). Different slopes for different folks: Self-esteem instability and gender as moderators of the relationship between self-esteem and attitudinal aggression. *Self and Identity*, 6(1), 74-94.
- Wilkowski, B.M., Crowe, S.E., & Ferguson, E.L. (2015). Learning to keep your cool: Reducing aggression through the experimental modification of cognitive control. *Cognition and Emotion*, 29(2), 251-265.
- Wilkowski, B.M., & Robinson, M.D. (2008). The cognitive basis of trait anger and reactive aggression:

- An integrative analysis. *Personality and Social Psychology Review*, 12(1), 3-21.
- Willoughby, M. T., Mills-Koonce, W. R., Gottfredson, N. C., & Wagner, N. J. (2014). Measuring callous unemotional behaviors in early childhood: Factor structure and the prediction of stable aggression in middle childhood. *Journal of Psychopathology and Behavioral Assessment*, 36(1), 30-42.
- Wilson, K., Juodis, M., & Porter, S. (2011). Fear and loathing in psychopaths: A meta-analytic investigation of the facial affect recognition deficit. *Criminal Justice and Behavior*, 38(7), 659-668.
- Yeager, D.S., & Walton, G.M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research*, 81(2), 267-301.
- Young, A.W., Rowland, D., Calder, A.J., Etcoff, N.L., & Perrett, D.I. (1997). Facial expression megamix: Tests of dimensional and category accounts of emotion recognition. *Cognition*, 63(3), 271-313.



SAMENVATTING / SUMMARY IN DUTCH
WAAROM KIJK JE ZO BOOS?
HET BEGRIJPEN EN VERANDEREN VAN SOCIALE COGNITIE IN
JONGENS MET DISRUPTIEF PROBLEEMGEDRAG.

Inleiding

Eén van de meest voorkomende verwijzingen in de geestelijke gezondheidszorg is van jongens tussen de 8 en 11 jaar met disruptieve gedragsproblemen en daaraan gerelateerde gezins- of schoolse problemen. Op dit moment bestaan er verschillende interventies die gericht zijn op deze problemen, waaronder oudertrainingen en cognitieve gedragstherapie. Deze interventies laten echter bescheiden resultaten zien, die wisselend van aard zijn. Hierdoor is het van belang om verder onderzoek te doen naar interventies die hierbij van toegevoegde waarde kunnen zijn, om de effectiviteit van de programma's te vergroten. Een veelbelovende interventie techniek zou het direct en onbewust trainen van sociale cognities kunnen zijn.

Het eerste doel van dit proefschrift was het verschaffen van meer inzicht in de rol van sociale cognities, stemming en zelfwaardering bij jongens met disruptieve gedragsproblemen, om hiermee (nieuwe) interventies te informeren. Het tweede doel van dit proefschrift was het onderzoeken van een innovatieve cognitieve bias modificatie procedure om (a) sociale cognities te veranderen die leiden tot agressief gedrag en (b) agressief gedrag te verminderen bij jongens met disruptief probleemgedrag. Hiervoor zijn verschillende methoden en meerdere informanten gebruikt, waarbij zowel cross-sectionele als experimentele studies zijn gedaan.

Achtergrond van het onderzoek

Twee verschillende modellen die beschrijven hoe agressief gedrag tot stand komt (General Aggression Model; GAM, Allen, Anderson, & Bushman, 2018, Social Information Processing model; SIP, Crick & Dodge, 1994; Lemerise & Arsenio, 2000) laten zien dat agressief gedrag een resultaat is van sociaal-cognitieve processen, die op hun beurt door meerdere factoren beïnvloed worden. Beide modellen suggereren dat iedereen een sociale database aan kennis heeft, die is gevormd door leren en ervaringen en die automatisch wordt naarmate bepaalde ervaringen vaker voorkomen. Kinderen die aangemeld worden bij de geestelijke gezondheidszorg voor disruptief probleemgedrag hebben vaker dan anderen een database die is gevormd door negatieve sociale ervaringen. Deze negatieve ervaringen kunnen er toe leiden dat een kind, vaker dan anderen, sociale situaties op een negatieve (vijandige) manier interpreteert (vijandige schemata; De Castro, 2002). Men zou verwachten dat deze

schemata moeilijk te veranderen en vrij rigide zijn, omdat ze gebaseerd zijn op meerdere eerdere negatieve ervaringen en/of persoonlijke kwetsbaarheid (bijvoorbeeld impulsiviteit, lager IQ, moeilijk temperament; Dodge, 2006). Deze negatieve schemata, gecombineerd met verschillende contextuele en persoonlijke factoren (bijvoorbeeld kwetsbare zelfwaardering), beïnvloeden verdere sociale informatieverwerking. Hierbij bestaat een wisselwerking tussen hoe een kind zich voelt en hoe hij zichzelf en zijn omgeving waarneemt en interpreteert, dat resulteert in bepaald gedrag, zoals agressie.

Samenvatting van de resultaten

Zelfwaardering, stemming en cognities

In dit proefschrift hebben we de rol van zelfwaardering, boosheid en vijandige interpretatie onderzocht bij jongens met disruptieve gedragsproblemen. We hebben gevonden dat narcisme samenhangt met agressie (Hoofdstuk 2). Dit komt overeen met eerdere bevindingen uit onderzoek in de algemene populatie; deze kennis is nu uitgebreid naar klinisch verwezen jongens met gedragsproblemen. Volgens henzelf en hun ouders, waren kinderen met meer narcistische zelfbeelden agressiever, onafhankelijk van de hoogte of stabiliteit van hun zelfbeeld. Daarnaast vonden wij geen link tussen een laag zelfbeeld en agressie, wat suggereert dat interventies gericht op het verhogen van het zelfbeeld om disruptief probleemgedrag te verminderen wellicht hun doel voorbij schieten.

Onze studies lieten echter ook zien dat kinderen met meer narcistische persoonlijkheidstrekken niet als agressiever werden beoordeeld door hun leerkrachten. Wij denken dat de context hiervoor een belangrijke verklaring biedt: Narcistische agressie wordt met name opgeroepen in situaties waarin kinderen bedreiging van hun ego ervaren (met andere woorden, een interpersoonlijke context; bijv. Baumeister, Smart, & Boden, 1996). De hoog gestructureerde schoolsetting waarin wij ons onderzoek hebben gedaan (het speciaal onderwijs) geeft wellicht minder ruimte voor dit soort situaties met leerkrachten. In het kader van de hierboven beschreven modellen lijkt narcisme één van de belangrijke persoonlijke factoren te zijn, die in een specifieke context van invloed is op het uitlokken van agressief gedrag.

Daarnaast hebben we gekeken naar de wisselwerking tussen stemming en cognitie in het produceren van agressief gedrag in de dagelijkse context (Hoofdstuk 3). We beoogden

eerdere bevindingen vanuit het laboratorium te valideren en hiermee interventie te informeren. Om deze wisselwerking goed te kunnen begrijpen, hebben we een naturalistische benadering gebruikt om dagelijkse associaties tussen boosheid (stemming) en vijandige interpretatie (cognitie) te bekijken. We vonden dat de jongens dagelijkse fluctuaties laten zien in hun vijandige interpretatie van gezichtsexpressies, maar dat deze fluctuaties niet samenhangen met fluctuaties in boosheid. Ook hingen zij niet samen met fluctuaties in hoeveelheid agressie.

Cognitieve bias modificatie procedure als interventie om sociale cognitie te veranderen

Omdat sociale informatieverwerking en resulterende gedragingen sterk beïnvloed worden door bestaande schemata vanuit de kennisdatabase, kan het veelbelovend zijn om manieren te onderzoeken die deze structuren kunnen veranderen. Cognitieve bias modificatie (CBM) focust hierop door herhaalde oefening op een specifieke taak, die ontworpen is om afwijkingen in de informatieverwerking te veranderen (MacLeod, Koster, & Fox, 2009). We hebben de effecten van CBM onderzocht om vijandige interpretatie van gezichtsexpressies te verminderen, en daarbij ook de effecten hiervan onderzocht op vijandige interpretatie in het algemeen, boosheid en agressie (Hoofdstukken 4 en 5).

Onze drie experimentele studies laten allemaal zien dat CBM vijandige interpretatie van gezichtsexpressies bij jongens met disruptief probleemgedrag kan verminderen. Eén van de studies laat daarbij zien dat deze procedure met name effectief is voor een bepaalde groep kinderen die vooral reactieve agressie laat zien (Hoofdstuk 5). Daarnaast hebben we ook laten zien dat CBM effectief is in het versterken van het herkennen van angstige gezichtsexpressies. Echter laten alle drie de studies *geen* generalisatie van de effecten zien naar de andere relevante uitkomstmaten – vijandige interpretatie in het algemeen, boosheid en/of agressief gedrag. We vonden alleen een zeer klein effect op het bedenken van gedragsalternatieven voor een subgroep van hoog reactief agressieve jongens.

Praktische implicaties

De eerste implicatie van mijn onderzoek betreft de link tussen zelfwaardering en disruptief probleemgedrag. Behandeling van disruptief probleemgedrag maakt vaak gebruik van

technieken om de zelfwaardering van kinderen te verhogen. Echter laat mijn onderzoek zien, tezamen met eerdere studies in algemene populaties, dat laag zelfbeeld geen of slechts een zeer beperkte link heeft met agressief gedrag. Focussen op narcisme zou wellicht meer effectief kunnen zijn om agressie te verminderen. Dat zou kunnen door narcisme te overwegen als potentieel verklarende factor in de diagnostiek en daaropvolgend ook mee te nemen in de behandeling, bijvoorbeeld door te oefenen hoe kinderen kunnen reageren op kritiek of falen. Daarnaast kan het ook gebruikt worden in ouder- of leerkrachtbegeleiding, met name omdat we al weten dat sommige opvoedstrategieën narcisme in kinderen kunnen vergroten (Brummelman, 2016). Nu wordt in ouderbegeleiding vaak aangeraden om bijvoorbeeld complimenten te gebruiken om succeservaringen (en daarmee zelfbeeld) te verhogen. Maar als dit niet goed wordt uitgevoerd, bijvoorbeeld als complimenten heel frequent worden gegeven, gebruik maken van interpersoonlijke vergelijking, of opgeblazen zijn, is er risico op een negatieve 'bijwerking': ouderlijke overwaardering kan narcisme van kinderen voeden (Brummelman et al., 2016). Uiteraard staat het geven van een compliment niet gelijk aan overwaardering, maar we dienen alert te zijn op mogelijk ongewenste effecten van positieve bekrachtiging in de behandeling van kinderen met disruptieve gedragsproblematiek. Kinderen met ADHD en andere probleemgedragingen laten vaak al onrealistisch hoge zelfwaardering zien; het is niet uitgesloten dat excessieve complimenten deze blik bekrachtigen (Hoza et al., 2004). Een mooi alternatief zou zijn om kinderen te helpen meer accepterend om te gaan met kritiek of falen, zonder terug te hoeven vallen op ontkenning of agressie. Het uiteindelijke doel zou dan zijn om een meer realistisch zelfbeeld te ontwikkelen, waarbij sterke en zwakke kanten beiden worden geaccepteerd.

Een tweede praktische implicatie betreft de complexe wisselwerking tussen verschillende factoren binnen de sociale informatieverwerking (zelfwaardering, stemming en cognitie) en context in het produceren van agressie. Onze resultaten suggereren dat het interveniëren op één van deze factoren in de cyclus niet voldoende is om gedragsverandering teweeg te brengen. In de diagnostiekfase moeten daarom zowel persoonlijke- als omgevingsfactoren meegenomen worden als verklarende en versterkende factoren, op basis waarvan behandeling moet worden ingezet. Op dit moment wordt in de evaluatie van de behandeling vaak gefocust op veranderingen in gedrag, in nationale richtlijnen en *Routine Outcome Monitoring*. Wat er precies gebeurt met sociale cognities, emoties en zelfbeeld voor een individuele client wanneer behandeling wordt ingezet op

verklarende factoren, wordt vaak niet gemonitord. Deze processen worden vaak aangenomen op basis van bestaande literatuur. Uit mijn en ander onderzoek blijkt echter dat bij verschillende cliënten verschillende processen een rol kunnen spelen (equifinaliteit) en dat die ook nog eens tussen contexten kunnen verschillen. Ik zou daarom behandelaren willen aanmoedigen om specifieke doelen op te stellen als het gaat om verklarende factoren en om deze te monitoren en bij evaluatie te bekijken in hoeverre veranderingen hierin hebben geleid tot vermindering van het probleemgedrag. Om inzicht te verkrijgen in de onderliggende factoren bij agressie voor elk individu, zou de diagnosticus gepersonaliseerde functieanalyse moeten uitvoeren. Hiermee kan ook bekeken worden in welke context agressief gedrag wellicht adaptief kan zijn en in welke niet, waarmee verschillende contexten een verschillende benadering vragen.

Een derde praktische implicatie komt voort uit onze interventiestudies. Onze studies hebben de potentie van het gebruik van CBM bij het verminderen van vijandige interpretatie aangetoond. CBM is kost- en tijdefficiënt en makkelijk te implementeren. Ondanks deze potentie, maakt het gebrek aan generalisatie en de context-afhankelijkheid dat wij adviseren deze procedure eerst verder te bestuderen alvorens het te implementeren als behandeltechniek.

Conclusie

Samenvattend laat ons onderzoek zien dat narcisme, en niet lage zelfwaardering, gerelateerd is aan agressie bij jongens die naar de geestelijke gezondheidszorg zijn doorverwezen in verband met disruptief probleemgedrag. De impliciete CBM procedure kan vijandige interpretatie van gezichtsexpressies verminderen, maar dit effect generaliseert niet naar agressief gedrag. Ons onderzoek illustreert de complexiteit van de wisselwerking tussen verschillende factoren in het produceren van agressief gedrag bij kinderen in deze klinische populatie. Om per kind meer zicht te krijgen op hoe deze factoren een rol spelen in agressie, zouden diagnostici en behandelaren individuele functieanalyses moeten uitvoeren. In de afwezigheid van voldoende bewijs voor aanzienlijke effecten van CBM bij agressie, zouden we adviseren op dit moment bestaande programma's te gebruiken die al bewezen effectief gebleken zijn.

The background is an abstract watercolor wash. It features large, irregular shapes in shades of blue, grey, and yellow. The colors are layered and blended, creating a soft, textured effect. The blue tones range from light sky blue to deep, dark navy. The grey is a muted, dusty shade. The yellow is a pale, buttery hue. The overall composition is organic and fluid, with no discernible figures or objects.

DANKWOORD

Dit onderzoek had niet gedaan kunnen worden zonder alle jongens, hun ouders/verzorgers, leerkrachten, intern begeleiders, schooldirecteuren, zorgcoördinatoren en gedragswetenschappers die aan het onderzoek hebben meegewerkt. Het was super leuk en motiverend om samen te werken. Heel erg bedankt voor jullie tijd en moeite, wie weet komen we elkaar in de toekomst (of op korte termijn) weer tegen. Ook wil ik de studenten die zich vol enthousiasme en flexibiliteit hebben ingezet om de data te verzamelen en in te voeren bedanken. Bedankt Govert dat jij de computertaak in een app hebt omgezet! Zonder jullie had ik dit niet kunnen uitvoeren!

Natuurlijk ben ik ook enorm dankbaar voor mijn twee promotoren. Bram, jij hebt mij de kans gegeven van mijn scriptie een promotietraject te maken. Van jouw energie, positieve en praktische insteek heb ik enorm veel geleerd. Jouw inzet om naast de wetenschap ook de praktijk te verbeteren, vind ik heel inspirerend. Na elke meeting had ik, ondanks alle actiepunten die daar uit voortkwamen, altijd het idee dat het allemaal wel goed zou komen. Ik heb met enorm veel plezier met je samengewerkt en ik hoop dat we elkaar in de toekomst nog regelmatig tegen gaan komen.

Sander, ondanks dat jij later in het project kwam, voelt het alsof je er vanaf het begin bij bent geweest. Jouw kritische en vooruitkijkende blik, daarnaast je prachtig woordgebruik, heeft mijn stukken enorm verbeterd. Bedankt dat ik dit van jou mocht leren.

Aan de leden van de beoordelingscommissie: Bedankt dat jullie de tijd en energie hebben genomen om mijn proefschrift te lezen en bij mijn verdediging te opponeren.

Aan mijn (oud) collega's van OWP en CAS in het algemeen: Bedankt voor de gezellige en fijne samenwerking, koffiemomenten, lunches, afdelingsuitjes, etc. Ik heb me altijd heel prettig gevoeld op de afdeling, ondanks dat ik er niet dagelijks te vinden was. Iedereen heeft zich open opgesteld en is bereid anderen te helpen, voor mij een ontzettend fijne werksfeer! Een speciaal bedankje voor de Jonkie's Anouk, Meike, Ivy, Inge, Ziyen, Yue en Lianne: Ik heb veel van jullie geleerd, mooi om zo onze ervaringen en vragen te kunnen delen.

A special thanks to prof. dr. Graham, for giving me the opportunity to visit their lab at the University of California Los Angeles. Thank you for your time and effort to show me around, for our discussions on different topics, for introducing me to other great researchers and by doing all this giving me a valuable learning opportunity.

Ook wil ik mijn collega's en het MT van RIOzorg bedanken. Super leuk om te merken dat onderzoek zo belangrijk wordt gevonden en initiatieven hierin gewaardeerd worden.

Bedankt dat ik hier verder mag leren en ontwikkelen!

Last but not least: mijn familie en vrienden. Pap, mam, bedankt voor jullie steun en liefde. Jullie hebben altijd geïnteresseerd en enthousiast gereageerd. Bedankt ook voor de relaxte en rustige thuisbasis, waar ik na een week hard werken weekendjes kon bijkomen met een gezellige borrel en lekker eten. Daarnaast ook enorm bedankt voor de mooie animatie, pap! Yas, fijn om in hetzelfde proces te hebben gezeten, zo konden we extra meeleven met elkaars frustraties en successen! Daarnaast natuurlijk ook de afleiding met alle andere dingen waar we uren over kunnen kletsen. Super leuk dat wij elkaars paranimf kunnen zijn! Lieve oma, bedankt voor jouw interesse en altijd lovende woorden.

Matthijs, bedankt dat je al bijna 10 jaar mijn vriend bent. Met jou kan ik mijn gezeur delen, maar vooral genieten van het leven. Je staat altijd voor me klaar en weet me altijd weer aan het lachen te krijgen. Laten we nog vele jaren samen op adventure!

Lieve vriendinnen en vrienden, NYMA en NEC-fanclub, wat een heerlijke mensen heb ik om mij heen. Jullie weten me altijd intellectueel uit te dagen, maar vooral te laten zien hoe je naast ambities ook veel plezier met elkaar kan hebben. Bedankt voor alle mooie momenten, dat er nog vele mogen komen! Ren, bedankt dat jij mijn paranimf bent, maar vooral bedankt voor onze vriendschap. Jij weet zaken te relativieren en tegelijkertijd op waarde te schatten. Super blij dat ik jou niet alleen als vriendin heb gehad deze periode, daarnaast ook met jouw bijdrage als coördinator dataverzameling! Soph, mijn lieve vriendin waarmee ik uren kan kletsen en die mij aan het denken zet. Bedankt dat jij jouw creatieve kracht wilde inzetten om een mooie omslag te ontwerpen, ik ben je heel erg dankbaar voor het eindresultaat!

The background is an abstract watercolor wash. It features large, irregular shapes in shades of blue, ranging from light sky blue to deep navy. Interspersed are areas of pale yellow and soft grey, creating a layered, ethereal effect. The colors bleed into each other, with some darker blue areas showing hints of green or purple. The overall texture is soft and painterly.

ABOUT THE AUTHOR

Wieteke Hiemstra (1989) obtained a Bachelor's degree in Child and Youth Psychology at Utrecht University in 2012. Within her bachelors, she studied abroad at the University of California Berkeley for a semester and lectured in the course Methods and Statistics at Utrecht University. She volunteered as a research assistant at the Baby Research Lab at Radboud University during summer break. In 2014, she obtained a Master's degree from the research master's program Development and Socialisation in Childhood and Adolescence at Utrecht University. At the end of this program, she started her PhD project on social cognition in boys with disruptive behavior problems, after obtaining the Personal Support Grant from Education and Learning Sciences at Utrecht University. In 2015, she obtained a second Master's degree in Clinical Child & Adolescent Psychology at Utrecht University. During her PhD project, she worked at Altrecht and RIOzorg as a psychologist.

During her PhD project, she spent one month at the laboratory Education and Diversity at the University of California Los Angeles, where she was welcomed by prof. dr. S. Graham. Wieteke presented her work at international conferences, including meetings of the European Society of Child and Adolescent Psychiatry (ESCAP; Madrid, 2015), Society of Research on Adolescence (SRA; Baltimore, 2016), and European Conference on Developmental Psychology (ECDP; Utrecht, 2017).

During her PhD project, Wieteke supervised theses of master's students, internships and certificates of psychological assessment (BAPD). Currently, Wieteke works as a psychologist in the Post-Master's healthcare training program (i.o.t. GZ-psycholoog) at RIOzorg.

