WHEN EMOTIONS RUN HIGH

Using Interactive Virtual Reality to Understand Social Information Processing in Boys with Aggressive Behavior Problems

ROGIER E.J. VERHOEF

WHEN EMOTIONS RUN HIGH

Using Interactive Virtual Reality to Understand Social Information Processing in Boys with Aggressive Behavior Problems

Rogier E.J. Verhoef

When Emotions Run High: Using Interactive Virtual Reality to Understand Social Information Processing in Boys with Aggressive Behavior Problems

Rogier E.J. Verhoef

ISBN: 978-94-6458-102-7

Provided by thesis specialist Ridderprint, ridderprint.nl
Printing: Ridderprint
Layout and design: Mila Slappendel, persoonlijkproefschrift.nl

When Emotions Run High:

Using Interactive Virtual Reality to Understand Social Information Processing

in Boys with Aggressive Behavior Problems

Wanneer de Emoties Hoog Oplopen:

Het Gebruik van Interactieve Virtual Reality om Sociale

Informatieverwerkingsprocessen

te Begrijpen bij Jongens met Agressieve Gedragsproblemen

(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 dinsdag 5 april 2022 des middags te 4.15 uur

door

Rogier Elbert Jan Verhoef

geboren op 19 maart 1985

te Naarden

Promotoren: Prof. dr. B. Orobio de Castro
Prof. dr. S. Thomaes

Copromotoren: Dr. E.E. Verhulp
Dr. A. van Dijk

Support for the research in this dissertation was provided by a grant from the Netherlands Organisation for Scientific Research (NWO grant number 453-15-004/511) to Prof. dr. B. O. de Castro.

BEOORDELINGSCOMMISSIE

Prof. dr. M. Deković	Utrecht University
Prof. dr. S.J.T. Branje	Utrecht University
Prof. dr. J.J.A. Denissen	Utrecht University
Prof. dr. W.A. Veling	University of Groningen
Dr. T.A.M. Lansu	Radboud University

TABLE OF CONTENTS

Chapter 1	General introduction	9
Chapter 2	Hostile Intent Attribution and Aggressive Behavior in Children Revisited: A Meta-Analysis	19
Chapter 3	Interactive Virtual Reality Assessment of Aggressive Social Information Processing in Boys with Behavior Problems: A Pilot Study	83
Chapter 4	Interactive Virtual Reality versus Vignette-Based Assessment of Children's Aggressive Social Information Processing	103
Chapter 5	Detecting Social Information Processing Profiles of Boys with Aggressive Behavior Problems: An Interactive Virtual Reality Approach	133
Chapter 6	A Dual-Mode Social-Information-Processing Model to Explain Individual Differences in Children's Aggressive Behavior	161
Chapter 7	Summary & General Discussion	183
Appendices	References	197
	Summary in Dutch (Samenvatting in het Nederlands)	220
	Acknowledgements (Dankwoord)	229
	About the Author	232
	Scientific Publications	233

CHAPTER 1

General Introduction

GENERAL INTRODUCTION

Aggressive behavior problems in childhood are among the most prevalent psychological problems in children and have a detrimental impact on children, their environment, and society at large (Dodge et al., 2006; Polanczyk et al., 2016; Romeo et al., 2006). The prognosis of children with aggressive behavior problems is generally poor, with an increased risk for various forms of psychopathology, delinquent and criminal behavior, dysfunctional family and peer relations, lower school performance, and unemployment (Cornacchio et al., 2018; Hubbard et al., 2010; Romeo et al., 2006; Vitaro et al., 2006). To better understand and treat children's aggressive behavior, we need to acquire more insight in how children's aggression may originate.

A useful approach to explain children's aggression is through their mental processing of social events. For instance, a boy who is bumped in the back by a peer during play and thinks the peer did this on purpose, will be more likely to respond aggressively than a boy who thinks it was an accident. The social information processing (SIP) model distinguishes several processing steps children engage in before responding to social events: (1) encoding and (2) interpreting social cues, (3) setting interaction goals, (4) generating and (5) evaluating responses, and (6) enacting a selected response (Crick & Dodge, 1994). Children's aggression has been shown to derive from deviations in each of these SIP steps, such as perceiving more hostile cues, attributing more hostile intentions to others, pursuing revenge or instrumental goals more often, generating more aggressive responses, and evaluating aggressive responses more positively (for reviews, see De Castro & Van Dijk, 2017; Dodge, 2011). Empirical work has demonstrated that children's deviant SIP patterns do not only precede, but may also maintain their aggressive behavior, and mediate relations between contextual risk factors (e.g., dysfunctional parenting and peer relations) and future aggressive behavior patterns (for reviews, see: De Castro et al., 2015; Dodge et al., 2006; Dodge, 2006). Moreover, children's SIP has shown to be an effective target for cognitive-behavioral interventions to reduce aggressive behavior problems (e.g., Lochman et al., 2017; Lochman et al., 2019).

Despite a wealth of empirical work on how children's SIP contributes to their aggressive behavior, SIP research thus far only explains relatively small proportions of variation in children's concurrent and future real-life aggression (e.g., De Castro et al., 2002; Dodge et al., 1986; Lansford et al., 2006). This may be due to three knowledge gaps in the current SIP literature: (1) the limited ecological validity of current SIP assessment methods; (2) the use of analytical methods that may not uncover the individual differences in children's aggressive SIP patterns because they study average SIP patterns over groups of children instead of clustering children based on their SIP patterns; (3) the current theoretical SIP models that only provide a clear explanation of how children's calm, deliberate SIP may contribute to aggression, but do not capture the emotional nature of children's SIP that may lead them to respond aggressively without any deliberate thought.

The current dissertation aims to fill these knowledge gaps in the current SIP literature by answering three research questions:

1. Can interactive Virtual Reality provide a more ecologically valid assessment of children's aggressive SIP?
2. Is it possible to distinguish SIP profiles of children with aggressive behavior problems?
3. Can we specify an improved theoretical model to explain aggressive SIP and behavior?

Limitations of Current SIP Assessment Methods

Current methods to assess aggressive SIP and behavior have three important shortcomings: 1) they are often not emotionally engaging; 2) when they are, they lack ecological validity or standardization; 3) they rarely assess actual aggressive behavior.

First, children's emotional engagement may have been limited in studies conducted with questionnaires and interviews about hypothetical events (i.e., vignettes). In a typical vignette-study, children are asked to imagine that a hypothetical social event is actually happening to them and to reflect on their SIP and behavior in response to this hypothetical event. Many children, however, only show aggressive SIP when they experience strong emotions such as anger, frustration, or jealousy (Anderson & Bushman, 2002; De Castro et al., 2003; De Castro, 2004; Dodge & Somberg, 1987; Lemerise & Arsenio, 2000; Reijntjes et al., 2011). For example, many children may only attribute hostile intent to others when they feel angry or frustrated, or only pursue instrumental goals when they strongly desire an object. Hence, although the majority of SIP research thus far has used vignettes, it seems unlikely that vignettes evoke sufficient emotional engagement to fully trigger the aggressive SIP patterns leading up to children's real-life aggressive behavior.

Second, measures may have lacked ecological validity and/or standardisation. Studies using video-game tasks to assess children's SIP (e.g., Yaros et al., 2014) may have evoked sufficient emotional engagement, but seem to lack ecological validity because it involves a computer game rather than a real-world-like social interaction. A few studies have used more ecologically valid methods to assess children's SIP, such as staged social interactions with alleged peers or child actors (Hubbard et al., 2001; Kempes et al., 2008; Steinberg & Dodge, 1983). Such methods may indeed evoke strong emotions in children and appear more ecologically valid, but they seem challenging in terms of adequate standardization and ethics (Underwood, 2005).

Third, a particularly sensitive aspect of ecological validity is the assessment of children's aggressive behavior in response to social events (Ritter & Eslea, 2005). Aggressive behavior is defined as 'any behavior directed towards another individual with the intent to cause harm' (Anderson & Bushman, 2002, p. 27). In response to vignettes, children reflect on their hypothetical behavior in response to hypothetical events, but they do not actually engage in aggressive behavior and thus do not actually inflict harm to another person. This limitation may provide another explanation for why children's

SIP assessed through vignettes only explains relatively small proportions of variance in their real-life aggression, as children may imagine (or report) quite different responses to vignettes than the actual responses they would give in real life. Thus, to better explain children's real-life aggressive behavior through their SIP patterns, it is important that their aggressive responses to events do actual harm to another person. Given the ethical repercussions of this requirement, it seems sensible that current assessment methods have rarely been able to achieve this.

To overcome the shortcomings of current assessments methods for both children's aggressive SIP and behavior, there is a need for innovative methods that use highly emotionally engaging, realistic, standardized social interactions that allow children to actually aggress against their interactional partner and inflict harm, whilst being ethically acceptable. To this end, we developed and tested an interactive Virtual Reality (VR) environment as a new method to assess children's aggressive SIP and behavior.

Can Interactive Virtual Reality Provide a More Ecologically Valid Assessment of Children's Aggressive SIP?

We designed an interactive VR classroom where children could play games with virtual peers and were confronted with distinct types of social events known to evoke aggression in children (see Figure 1a and 1b). Our interactive Virtual Reality (VR) may provide a viable solution for limitations encountered by previous research on children's aggressive SIP and behavior. In our interactive VR, participants are visually completely immersed in a simulated virtual classroom. As opposed to non-interactive VR, interactive VR engages participants actively in a realistic and truly interactive environment that, just like the real world, responds naturally to each single motion. As in real life, participants can freely walk around and interact with virtual characters using verbal and physical behaviors. Participants use controllers that mimic their hands in VR, allowing them to use objects and thus for example to play games. Interactive VR allows researchers to engage children in standardized social events within an emotionally engaging, realistic environment. Children may experience strong emotions such as anger, frustration, desire, or jealousy because they are actually provoked and tempted to use aggression by virtual peers. Moreover, interactive VR may allow children to actually aggress against virtual peers, instead of reporting on their hypothetical aggressive responses as with traditional vignettes. Consequently, interactive VR may evoke SIP and behavior patterns more similar to real life than traditional vignette-based methods. Despite its potential, to our knowledge, no studies to date have used interactive VR to assess children's aggressive SIP and behavior. Therefore, we directly tested whether interactive VR provides a more ecologically valid assessment of children's aggressive SIP and behavior than a traditional vignette-based assessment.

Figure 1a & 1b *Interactive Virtual Classroom where Children Played Games with Virtual Peers.*

Is it Possible to Distinguish SIP Profiles of Children with Aggressive Behavior Problems?

Interactive VR may also help to acquire more insight in the large individual differences in children's aggressive SIP patterns, allowing to better explain and thus treat children's aggressive behavior. For instance, for some children their aggressive behavior may predominantly derive from a tendency to attribute hostile intent to others and take revenge, whereas for other children their aggressive behavior may predominantly derive from a desire to pursue instrumental goals and their positive outcome expectancies for aggression (Arsenio et al., 2009; De Castro et al., 2005; Crick & Dodge, 1996). A fundamental issue that may hinder effective treatment of children with aggressive behavior problems is that these children are often classified based on labels that provide a description of their behavior patterns (e.g., through clinical syndromes such as oppositional deviant disorder or conduct disorder), while children with the same diagnostic classification may differ considerably in the underlying processes contributing to their aggressive behavior. The large individual differences in aggressive SIP between children suggest that distinct subgroups of children with aggressive behavior problems could be identified based on their aggressive SIP patterns. If this is true, current cognitive behavioral interventions can be more effectively tailored to the unique SIP style of individual children. Interactive VR seems particularly suited to identify subgroups of children based on their aggressive SIP patterns, because it taps into children's SIP in an emotionally engaging context—a context in which individual differences in SIP may become most evident (Anderson & Bushman, 2002; Lemerise & Arsenio, 2000). Therefore, we examined whether it is possible to distinguish SIP profiles among boys with aggressive behavior problems using interactive VR.

Can we Specify an Improved Theoretical Model to Explain Aggressive SIP and Behavior?

The current literature on children's SIP stresses the need for a more ecologically valid and fine-grained explanation of children's aggressive SIP and behavior (e.g., De Castro & Van Dijk, 2017). However, the appreciation of emotional engagement as pivotal to our understanding of SIP appears to be more fundamental than limitations of assessment: Current theoretical models of SIP do not specify precisely how and why emotional engagement is expected to have such a profound influence on children's SIP (Anderson & Bushman, 2002; De Castro, 2004). For instance, when children are actually provoked by a peer, many of them may experience high levels of emotional arousal (e.g., anger or anxiety), triggering a direct aggressive response without any reflection on intent attributions or decision processes (De Castro et al., 2005, 2012; Helmsen et al., 2012). Aggressive behavior of these children may therefore primarily derive from fast, automatic, emotional SIP. Although the current SIP models posit that aggressive behavior may be a consequence of skipping part of these deliberate SIP steps (Crick & Dodge, 1994; even due to emotions: Lemerise & Arsenio, 2002), they do not describe how children who skip parts of these processing steps would actually engage in automatic processing or what would make children skip parts of the deliberate SIP steps. In fact, SIP models do not provide any specific handle for emotions to influence this process. The shortcomings of current theoretical SIP models may limit further progress in our understanding of children's aggressive behavior. A new SIP model is therefore needed that extends current SIP models by explaining how children's emotional arousal may contribute to fast, emotion-driven aggression preceded by automatic SIP, as well as deliberate, controlled aggression preceded by reflective SIP. Therefore, this dissertation proposes a new theoretical SIP model that explains what determines whether children process social information automatically or reflectively, how these processes take place and how this may lead to aggression.

Aims and Outline of this Dissertation

The **overarching aim** of this dissertation was to improve our understanding of the SIP patterns underlying children's aggressive behavior. For this purpose, three research questions were answered, using four empirical studies and a theoretical article in which we propose a new theoretical SIP model (see Table 1 for an overview of this dissertation).

This dissertation starts with a meta-analysis **(Chapter 2)** that aims to provide more insight in child-specific and methodological determinants of the strength of the relation between children's SIP (specifically, the tendency to attribute hostile intent to others) and their real-life aggression (using parent-report, teacher-report, self-report, observation, and peer nomination measures). Based on the SIP literature, we hypothesized that the relation between children's tendency to attribute hostile intent and their real-life aggression is stronger when emotionally engaging SIP assessment methods are used.

1. Can Interactive Virtual Reality Provide a More Ecologically Valid Assessment of Children's Aggressive SIP?

For Chapters 3-4, we developed and investigated an interactive VR environment to assess children's aggressive SIP and behavior. **Chapter 3** includes a pilot study that tested whether our interactive VR measure provides a valid assessment of children's SIP. **Chapter 4** directly tested whether interactive VR provides a more ecologically valid assessment of children's SIP, explaining more variance in children's real-life aggressive behavior, than a traditional vignette-based measure.

2. Is it Possible to Distinguish SIP Profiles of Children with Aggressive Behavior Problems?

For **Chapter 5**, we used our interactive VR environment to examine whether SIP profiles can be distinguished among children with aggressive behavior problems. For this purpose, we conducted latent profile analyses (LPA) on children's aggressive SIP patterns assessed in interactive VR

3. Can we Specify an Improved Theoretical Model to Explain Aggressive SIP and Behavior?

Building on insights from Chapters 2-5, **Chapter 6** provides a new theoretical framework to explain individual differences in children's aggressive SIP and behavior. We propose a dual mode SIP model that extends current SIP models by distinguishing between an automatic versus reflective SIP mode. Both SIP modes may contribute to children's aggression but in different ways. The automatic mode is characterized by fast automatic processing and impulsive behavioral responses, whereas the reflective mode is characterized by deliberate processing, leading to controlled behavioral responses. Our dual-mode SIP model aims to provide a more ecologically valid explanation of children's unique SIP styles, allowing to more effectively tailor treatment to children's individual needs.

This dissertation ends with the general discussion **(Chapter 7)**, starting with a summary of the findings of this dissertation. Next, the strengths and limitations of this dissertation and the implications of using interactive VR to assess children's aggressive SIP and behavior are discussed. Last, several directions for future research are provided that may further our understanding of individual differences in the SIP patterns leading up to children's aggressive behavior.

Table 1. *Overview of Research aims and Characteristics of the Chapters in this Dissertation*

Chapter	Design	Sample	Age	Dissertation Aim	Study Aim
Two	Meta-analysis	K = 111	3-17 years		To provide insight in child-specific and methodological determinants of the strength of the relation between children's SIP and real-life aggression.
Three	Cross-sectional	N = 32	8-13 years	SIP assessment	To examine whether our interactive VR measure provides a valid assessment of children's SIP.
Four	Cross-sectional	N = 184[a]	7-13 years	SIP assessment	To examine whether our interactive VR measure provides a more ecologically valid assessment of children's SIP than a traditional vignette-based measure.
Five	Cross-sectional	N = 181[a]	7-13 years	SIP profiles	To distinguish SIP profiles of children with aggressive behavior problems.
Six	Theoretical paper	—	—	Theoretical SIP model	To propose a new theoretical SIP model to explain individual differences in children's aggressive SIP and behavior.

[a] These chapters include the same sample (i.e., 184 participants completed the vignette-based SIP assessment and 181 participants completed the VR-based SIP assessment).

CHAPTER 2

Hostile Intent Attribution and Aggressive Behavior in Children Revisited: A Meta-Analysis

This chapter is published as:
Verhoef, R. E. J., Alsem, S. C., Verhulp, E. E., & De Castro, B. O. (2019). Hostile intent attribution and aggressive behavior in children revisited: A meta-analysis. Child Development, 90(5), 525-547. http://doi.org/10.1111/cdev.13255

Author contributions:
RV conceptualized the study, SA, EV, and BC gave advice and feedback; RV, SA, and EV conducted the literature search and coding; RV analyzed the data and drafted the manuscript; SA, EV, and BC provided feedback on the analyses and manuscript.

ABSTRACT

To test specific hypotheses about the relation between hostile intent attribution (HIA) and children's aggressive behavior, a multi-level meta-analysis was conducted on 111 studies with 219 effect sizes and 29.272 participants. A positive association between HIA and aggression was found, but effect sizes varied widely between studies. Results suggested that HIA is a general disposition guiding behavior across a broad variety of contexts, while the strength of the relation between HIA and aggression depends on the level of emotional engagement. The relation is stronger for more reliable HIA measures, but is not stronger for reactive aggression or co-morbid ADHD than for aggression in general. The importance of understanding specific moderators of effect size for theory development is discussed.

Keywords: Hostile intent attribution, aggression, children, meta-analysis

HOSTILE INTENT ATTRIBUTION AND AGGRESSIVE BEHAVIOR IN CHILDREN REVISITED: A META-ANALYSIS

Hostile intent attribution (HIA) is defined as the tendency to attribute hostile intent to others in social situations with a negative outcome for the individual, where the intention of the other person is ambiguous. In a typical study, HIA is measured by presenting children social situations with a negative outcome caused by a peer, who's intentions are ambiguous, and subsequently asking about the intentions of the peer in the presented social situation. Social-cognitive models propose that children who frequently interpret the intentions of others as hostile in ambiguous situations will be more prone to respond aggressively, as a way to retaliate or defend themselves, than children who attribute non-hostile intent after being hindered (Crick & Dodge, 1994; Dodge, 1980). Moreover, social-cognitive theory states that HIA not only causes aggressive behaviors but also maintains aggressive behavior patterns. The latter follows from the assumption that aggressive children, as a result of their aggressive behavior, will more frequently be confronted with problematic social interactions. These problematic social interactions prohibit aggressive children to challenge their hostile beliefs about the intentions of others and limit the opportunity to acquire prosocial behavioral strategies. The crucial role of HIA in the development and maintenance of aggression has been supported in experimental (e.g., Dodge, 1980; Lochman & Dodge, 1998), longitudinal (e.g. Dodge et al., 2003; Lansford et al., 2010), and longitudinal-experimental studies (e.g. Lochman & Wells, 2002), making HIA a plausible target for effective cognitive-behavioral interventions (CBT) to reduce aggressive behavior in children (e.g., Hudley & Graham, 1993; Lochman & Wells, 2002).

The construct HIA has much potential to further our understanding of the development of aggressive behavior problems and to improve clinical practice. HIA may mediate links between aggression, distal risk factors in children (such as executive functioning deficits or difficult temperament), and their environments (such as early harsh life experiences, rejection by peers, and coercive family interactions (e.g., Dodge, 2006)). More specifically, social-cognitive theory states that the tendency to attribute hostile intentions to others derives from transactions between early aversive child experiences such as harsh parenting and peer rejection on the one hand and child susceptibility to such experiences on the other hand (Dodge, 1980; Dodge et al., 2003; Dodge et al., 1995; Lansford et al., 2010; Weiss et al., 1992). Thus, children who experienced harsh parenting and peer rejection, and exhibit underlying vulnerabilities, such as executive functioning deficits or difficult temperament, could be particularly prone to develop hostile attribution styles and subsequent aggressive behavior patterns.

However, further progress in our understanding of HIA in aggressive behavior seems to be thwarted by unexplained variation in the strength of the relation between HIA and aggression. The last meta-analysis on the relation between aggressive behaviors in children and HIA demonstrated a modest robust relation (d = .35, fail-safe number of studies: 3.411) that did, however, vary widely between studies (De Castro, et al., 2002).

This meta-analysis was conducted in 2002 and showed that the relation between HIA and aggression was stronger for children exhibiting more severe aggressive behavior (clinically-referred aggressive children vs non-referred children), children between 8-12 years, children low on sociometric status and in studies that did not control for children's intelligence. Moreover, the use of staged interactions (standardized real-time interactions with a peer) and hypothetical stories read to or by children yielded higher effect sizes than the use of hypothetical stories presented through video-clips and pictures. Effect sizes were not related to aggression function (e.g., reactive aggression, general aggression), type of social context (e.g., provocation, non-provocation), setting (e.g., individual, group), response format (e.g., open responses, rating scales or multiple choice), and type of HIA scoring (e.g., hostile responses, hostile minus benign attributions).

Despite identifying several moderators of effect, this meta-analysis could not explain the significant variation in effect sizes between studies properly. In addition, this meta-analysis did not formulate specific hypotheses about moderators of the relation between aggressive behavior in children and HIA. Fortunately, since 2002 a number of important reviews and theoretical papers have suggested adaptations to social information processing (SIP) theory that may help to explain the divergent findings between studies. For example, De Castro (2004) suggested how HIA may be most evident in emotionally engaging situations, Peets and colleagues (2007) suggested that HIA may be unique to interactions with specific familiar peers (i.e., disliked peers), whereas both Dodge (2006) and Schultz and colleagues (2010) suggested that HIA may be specific to particular developmental stages. As far as we know, it has not yet been tested whether these hypotheses are supported by actually explaining variance in findings between studies. To test specific hypotheses about moderators of the relation between HIA and aggression in children, we conducted a new meta-analysis. Advances in theory suggest five specific hypotheses about moderators of the relation between HIA and aggression in children:

First, the relation between HIA and aggression may be stronger in emotionally engaging situations. Social-cognitive theories postulate that for many children the actual processes leading up to aggression only occur when they are emotionally and personally involved (Anderson & Bushman, 2002; Lemerise & Arsenio, 2000). Moreover, empirical research suggests that aggression is often associated with excessive anger or anxiety (Granic, 2014; Hubbard et al., 2002) and that the induction of negative emotions results in more severe HIA and aggression (De Castro et al., 2003; Dodge & Somberg, 1987; Reijntjes et al., 2011). An explanation might be that strong emotions (e.g., excessive anger) derail cognitive resources and thereby inhibit deliberate reflective processing. Strong emotions may force individuals to mainly rely on automatic SIP driven by hostile beliefs about the intentions of others established through early aversive child experiences. Since HIA and aggression are associated with aversive social experiences such as peer rejection (e.g., Dodge et al., 2003; Lansford et al., 2011), it seems that strong emotions in aggressive children steer the automatic interpretation of the intention of others in future social situations congruent with hostile memories of previous social interactions. This line of reasoning suggests that particularly social situations that are emotionally involving elicit the automatic and emotional processes that activate HIA. Thus, based on social-cognitive theories we hypothesized that the strength of the relation between HIA and aggression increases with the level of emotional involvement the social situation elicits. This would have direct implications for clinical practice since it implies that CBT should target HIA using emotionally engaging- and personally involving situations.

Second, the relation between HIA and aggression may be stronger in social situations with familiar others, encountered in previous problematic social situations (i.e., disliked others), than towards unfamiliar others. Social-cognitive theory proposes that the tendency to attribute hostile intent to others is a general cognitive disposition towards both familiar and unfamiliar others. This is based on the assumption that HIA steers SIP across a broad variety of contexts. However, several empirical studies suggest that HIA may only be present in social situations with others who were encountered in previous problematic encounters (Hubbard et al., 2001; Peets et al., 2007; Peets et al., 2008). If HIA would be limited to interactions with specific familiar peers, this would have serious implications for social-cognitive theory and clinical practice. It would suggest that HIA is context-specific and only guides SIP in social situations with disliked others known from previous problematic encounters. Importantly, all current evidence-based CBTs are based on the assumption that a general cognitive disposition needs to be targeted to establish significant and prolonged changes in SIP and subsequent behaviors across a wide range of contexts. If HIA were person-specific, such broad generalization would not take place, which would question our expectations of CBT treatment potential. In line with the SIP model, we hypothesized that the relation between HIA and aggression is present in social situations with both unfamiliar and familiar others (e.g., Dodge, 2006). In addition, we expected this relation to be stronger in situations with familiar others encountered in previous problematic social encounters (i.e., disliked others) than with unfamiliar others.

Third, the relation between HIA and aggression is expected to be present irrespective of the sociometric status of participants. Social-cognitive theory postulates that HIA is a general cognitive disposition that guides SIP across contexts. Thus, social cognitive models propose that the tendency to attribute hostile intent to peers is not uniquely related to specific past experiences of peer rejection but could also be a result of other aversive social experiences (e.g., harsh parenting). Therefore it could be expected that both aggressive non-rejected children and aggressive-rejected children make hostile attributions about peers' intentions. Nonetheless, the previous meta-analysis suggested that the relation between HIA and aggression was stronger for aggressive-rejected samples than for generally aggressive samples. This finding suggests that the relation between aggression and HIA might be stronger when the social situation matches specific memories of being rejected by peers. We therefore hypothesized that the relation between HIA and aggression would be present in both aggressive- and aggressive-rejected samples, yet would be particularly pronounced in aggressive-rejected samples.

Fourth, the relation between HIA and aggression may be stronger when aggression is operationalized as reactive aggression. Reactive and proactive aggression are proposed to have distinct etiologies (Dodge & Coie, 1987; Frick et al., 2003; Polman et al., 2007; Poulin & Boivin, 2000; Raine et al., 2006, but see Bushman & Anderson, 2001, for a critique). Reactive aggression is defined as an emotional, impulsive aggressive response to a perceived threat, provocation, or frustration aimed at defending oneself or retaliation. In contrast, proactive aggression is defined as coldblooded, planned aggressive behavior aimed at instrumental, material, or social personal gain (Dodge, 1991). It can be assumed that children who frequently attribute hostile intent to others will be more likely to perceive threats or provocations in other's behaviors and thereby engage in reactive aggressive behaviors. In addition, based on the same theory no relation between HIA and proactive aggression would be expected. The previous meta-analysis (De Castro et al, 2002) did not find an effect of function of aggression. However, this finding was based on only four studies. As suggested by the authors, a lack of power may explain this null-finding. Based on theory, we therefore hypothesized that the relation between HIA and aggression is stronger for reactive aggression and weaker for aggression measured as a general construct (with no differentiation between reactive- and proactive aggression).

Fifth, the strength of the relation between HIA and aggression may be positively associated with the proportion of children meeting criteria for attention-deficit hyperactivity disorder (ADHD). Social-cognitive theories state that aggression driven by HIA is partly due to limited cognitive capacities (e.g., Dodge & Pettit, 2003) and this seems to be supported by empirical research (e.g., Ellis et al., 2009). Moreover, research demonstrated that ADHD is positively associated with both aggression and executive functioning deficits (Doyle, 2006; Hummer et al., 2010; King & Waschbusch, 2010; Waschbusch, 2002). Given the important role of executive functioning deficits in SIP (e.g. Van Nieuwenhuijzen, et al., 2006; Ellis et al., 2009), it is expected that particularly aggressive children with executive functioning deficits may find it difficult to accurately process information from the social environment, making them more susceptible to attribute hostile intent to others in social situations. Therefore we hypothesized that the strength of the relation between HIA and aggression increases with the proportion of ADHD diagnoses in the aggressive sample.

Methodologically, the previous meta-analysis included too few studies to analyze important combinations of moderators, such as studies combining a clinical sample with in vivo provocation. Fortunately, while the 2002 meta-analysis only contained studies up to January 1998, many excellent studies into the relation between childhood HIA and aggression have been carried out since. The present extension of this meta-analysis allowed to include all eligible studies within a timeframe over 40 years (instead of 25 years in the previous meta-analysis). Moreover, due to statistical limitations at the time (e.g., inability to model dependency in effect sizes), the previous

meta-analysis was only able to derive a single effect size from each study. As a result of statistical developments (e.g., multilevel meta-analysis), our extension of this meta-analysis could accommodate dependency in effect sizes and therefore allowed to derive multiple effect sizes from each study.

To test specific hypotheses about moderators of the relation between HIA and aggression in children, we conducted a new meta-analysis to test specific hypotheses, including more than double the number of studies, more variance and more precise assessment of moderators than the 2002 meta-analysis, and using statistical innovations to model effects. As explained above, methodological characteristics that were hypothesized to influence effect sizes included the type of stimulus presentation and provocateur's status in the presented social situation. Child characteristics that were hypothesized to influence effect sizes included sociometric status, function of aggression and proportion of ADHD diagnoses in the sample. In addition, we coded all variables included in the previous meta-analysis (De Castro et al., 2002) and exploratively tested whether the moderator effects were replicated.

METHODS

Study Selection

Child aggression was operationalized as all behaviors leading to psychological, physical or material harm of others. Thus, this operationalization covered a broad range of behaviors including categorizations on a syndrome-level (e.g., diagnoses of disruptive behavioral disorders), categorizations on a symptom-level (e.g., starting fights), and behavioral outcomes measured on a continuum (e.g., externalizing behaviors). HIA was operationalized as the attribution of hostile intent to peer's behaviors in social situations where the peer's intentions are ambiguous or differ systematically across situations (e.g., partly ambiguous, partly hostile, and partly benign).

All empirical studies into the relation between childhood aggression and the attribution of hostile intent to peer's behavior conducted between January 1998 and October 2017 were searched in the following databases: PsycINFO, Web of Science, PubMed and Google Scholar. Within all search databases the following strings were searched: "aggress*" OR "violence" OR "violent behavior*" OR "behavior problem*" OR "conduct disorder*" OR "conduct problem*" OR "antisocial behavior*" OR "behavior disorder*" OR "oppositional defiant disorder*" OR "disruptive behavior*" in combination with "attribution*" OR "hostil*" OR "social cognit*" OR "social perception" OR "interpretation bias" OR "social information-processing" OR "cognitive style" OR "cognitive bias" OR "Kenneth. A. Dodge". The search was limited to human participants, childhood (0-12 years) or adolescence (13-17 years), and English language. It is important to note that the literature search of this extension started where the literature search from the previous meta-analysis ended (De Castro et al., 2002). This search resulted in 6834 studies. In addition, all studies that cited the original meta-analysis were also searched in the Web of Science database. This search retrieved 329 additional studies resulting in 7163 studies total. After removal of duplicates, 4973 potential studies remained for further evaluation of eligibility. The authors acknowledge that although the search process was extensive and thorough, the possibility that specific studies were not identified cannot be ruled out.

The strategy to evaluate study eligibility consisted of two steps. First, all retrieved studies were scanned on title and abstract for exclusion. Second, for all remaining articles full-texts were evaluated for eligibility. A flow diagram for the search and identification of studies is depicted in Figure 1. Thus, 4973 studies were scanned on title and abstract, which resulted in the exclusion of 4653 studies. Subsequently, the 320 remaining articles full-texts were evaluated for eligibility. The current meta-analysis applied identical inclusion- and exclusion criteria as the 2002 meta-analysis. The inclusion- and exclusion criteria were the following:

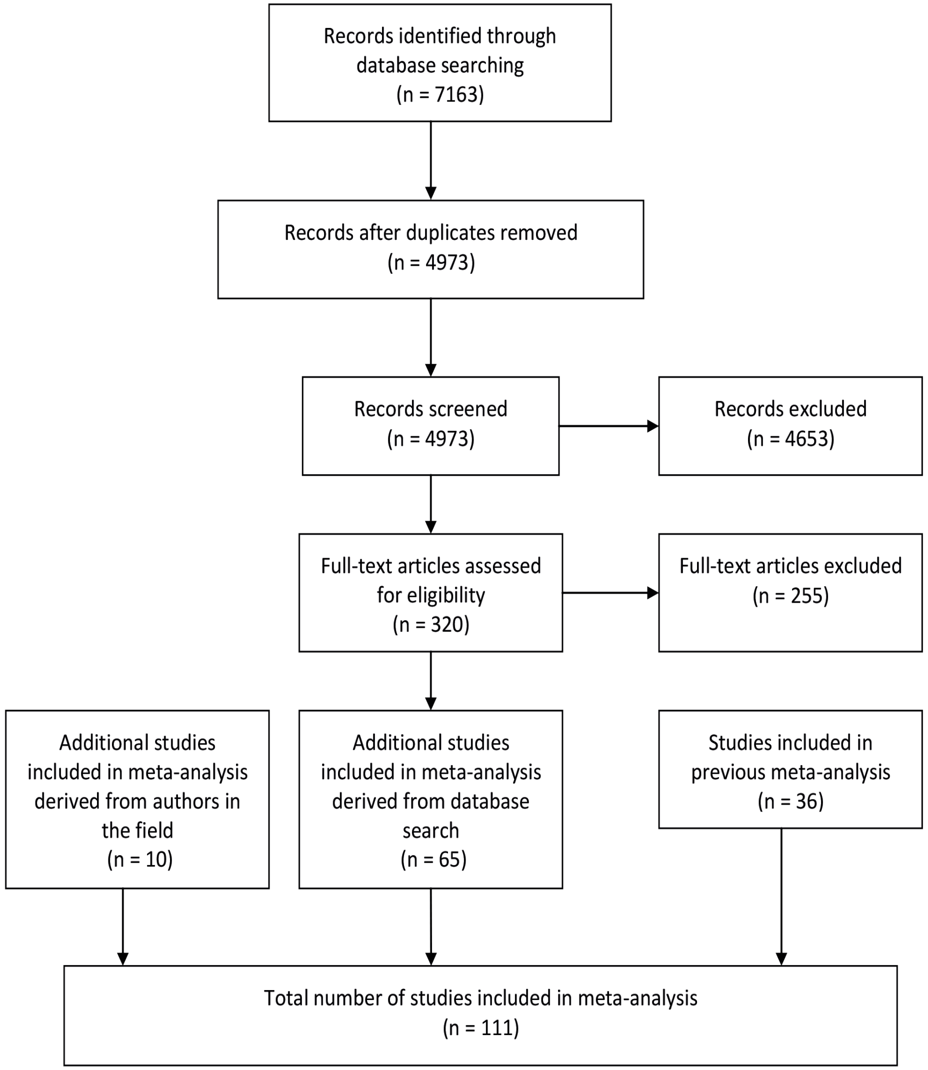
1. HIA and aggression were empirically assessed using standardized instruments.
    a. When studies distinguished between reactive- and proactive aggression, effect sizes were derived from the reactive aggression data only, since based on theory no relation between HIA and proactive aggression was expected.
    b. Studies that compared clinically aggressive children to other clinical groups, but not to nonaggressive controls were excluded since no reliable comparison could be made between clinical groups.
    c. Studies that used rejection as the only selection criterion were excluded. Studies that used both aggression and rejection as selection criterion were included.
    d. Studies that used social competence instead of aggression as a selection criterion were excluded. Low social competence and aggression are not opposite poles on a continuum and therefore low social competence was not considered as an indicator of aggression.
    e. Studies that used Attention-Deficit Hyperactivity Disorder (ADHD) as the main selection criterion were only included when the ADHD group demonstrated high aggression scores as well.
2. HIA and aggression were measured on the same time point. Studies that measured HIA and aggression on different time points were excluded since it is impossible to determine whether this relation would have been identical on the same time point (e.g., Fontaine et al., 2010; Godleski & Ostrov, 2010).
3. HIA was operationalized as specific cognitions about a presented social situation. Thus, studies that assessed hostility as a general pattern of cognitions or personality trait were excluded (e.g., Rubio-Garay et al., 2016).
4. HIA was not measured following experimental manipulation. It is impossible to determine the effect of the experimental manipulation on the relation between HIA and aggression. Thus, with regard to studies that used experimental manipulations such as the induction of emotions (e.g., De Castro et al., 2003; Reijntjes et al. 2011) or treatment (e.g., Stoltz et al., 2013) effect sizes were derived from pre-manipulation data only.
5. The presented social situations were standardized social interactions with peers. Studies that presented social situations concerning social interactions with solely adults were excluded. In studies that used social interactions with peers and adults and reported a composite score, effect sizes were based on this composite score. We decided to focus on interactions with peers only because of the presumed role of peer rejection as a cause for hostile attributions (Dodge, 2006) and the fact that almost every study on HIA and childhood aggression used social situations with peers to measure HIA. Studies that used unstandardized stimulus materials were not included since unstandardized stimulus materials prohibit to make between study comparisons.
6. Part of the stimulus materials were required to be ambiguous. Studies that solely presented non-ambiguous social situations were excluded. Regarding studies that used a mixture of ambiguous- and non-ambiguous social situations and reported a composite score of HIA, effect sizes were based on this composite score.

To derive reliable estimates of true effect sizes and to minimize the possibility of publication bias, multiple authors in the field were contacted for unpublished data. In addition, for studies that measured HIA and aggression but did not report sufficient information to calculate effect sizes, authors were contacted for additional information. The previous meta-analysis of De Castro et al. (2002) included 41 studies, however, one study (Dodge & Price, 1994) needed to be excluded from the present meta-analysis since it used a measure of behavioral competence instead of aggression. In addition, the previous meta-analysis treated different samples tested in the same study (Crick & Dodge, 1996; Lochman & Dodge, 1994) as independent studies, however, these were treated as from the same study in the present meta-analysis. From the 36 independent studies included in the previous meta-analysis, 51 effect sizes were derived, and the new search resulted in an additional 75 studies (68%) and 168 effect sizes (77%). Thus, the present meta-analysis included 111 studies and 219 effect sizes in total. An overview of the included studies and effect sizes in this meta-analysis is provided in the supplementary materials (See Table S24).

Coding

To examine whether specific variables influenced the relation between HIA and aggression, child characteristics and methodological characteristics were coded for each effect size.

Figure 1 *Flow diagram of search and identification of studies*



Methodological Characteristics

Methodological characteristics that were hypothesized to influence the relation between HIA and aggression were operationalized in following manner:

Type of Stimulus Presentation. Type of stimulus presentation was used as an indicator of the level of emotional engagement and coded categorically. Categories consisted of hypothetical stories read by the participant, hypothetical stories read to the participant (e.g., read by experimenter, played from audiotape), video-taped hypothetical stories, hypothetical stories presented through pictures, cartoons or illustrations, hypothetical stories presented through both audio and pictures, cartoons or illustrations, hypothetical stories presented through doll-play, real-time computerized interactions between the participant and a presumed peer or real-time interactions between the participant and a real peer.

Provocateur's Status. Provocateur's status was coded categorically. Categories consisted of the provocateur in the presented social situation being an unknown peer, a boy or girl from the neighborhood or school, a classmate, a friend, or an enemy of the participant.

Child Characteristics

Child characteristics that were hypothesized to influence the relation between HIA and aggression were operationalized in following manner:

Sociometric Status. Sociometric status was coded categorically. Categories consisted of effect size was based on an aggressive-rejected sample (samples consisting of aggressive-rejected children) or an aggressive sample (samples where only aggression was measured).

Function of Aggressive Behaviors. Function of aggressive behaviors was coded categorically. Categories consisted of aggression was measured as reactive aggression or aggression measured as a general construct.

Proportion of ADHD in the Sample. Proportion of ADHD in the sample was coded as a continuous variable representing the proportion of Attention Deficit Hyperactivity Disorder (ADHD) diagnoses in the sample.

Additional Moderators

The additional moderators were coded as in the 2002 meta-analysis. Details are provided in the supplementary materials.

Inter-rater Agreement

To make sure all studies were coded consistently, the studies included in the original meta-analysis were recoded for the present analysis.

To determine inter-rater agreement, 41 randomly selected studies (out of 111 studies; 37%) were coded by a second rater. In case of rater disagreement, the two raters discussed the discrepancy and tried to solve this by consensus. In rear cases where no consensus could be achieved, a third rater was asked to solve the discrepancy. Cohens kappa's for categorical variables were calculated and satisfying, ranging from 0.74. to 1.00 (M = 0.83, Mdn = 0.80). Inter-rater reliability of the coding of continuous variables was examined with a two-way random-effect model, absolute agreement, average-measures intra-class correlations (ICCs). ICCs were good ranging from 0.66 to 0.90 (M = 0.79, Mdn = 0.84, SD = 0.11). Frequency distributions of child- and methodological characteristics are reported in Table 1.

Table 1 *Moderators of Effect Size by Severity Classification*

	AGG Severity															
	All				Non-referred general				Non-referred extremes				Clinically-referred			
Characteristic and Level	**No. of studies**	**No. of ES**	**N**	**d**	**No. of studies**	**No. of ES**	**N**	**d**	**No. of studies**	**No. of ES**	**N**	**d**	**No. of studies**	**No. of ES**	**N**	**d**
Child characteristics																
Sociometric status																
Aggressive	97	202	28002	.30	58	124	22269	.28	28	50	4285	.31	14	28	1448	.48
Aggressive-Rejected	15	17	1270	.61	1	2	80	.51	14	15	1190	.62	—			
AGG function																
General	98	193	22502	.33	50	111	17259	.26	38	55	3877	.41	13	27	1366	.42
Reactive	18	26	6770	.36	11	15	5147	.40	7	10	1541	.27	1	1	82	.62
% ADHD	7	22	919	β0 = .39 β1 = -.00 *t* = -.05 *p* = .958	—				2	7	249	β0 = -.62 β1 = .01 *t* = 1.86 *p* = .122	5	15	670	β0 = .53 β1 = .00 *t* = .14 *p* = .890
Methodological characteristics																
Type of stimulus																
Self-reading	17	30	6363	.44	12	22	5787	.38	5	8	605	.64	—			
Pictures	6	7	1340	.25	3	3	501	.28	2	3	775	.11	1	1	64	.37
Audio	42	82	10152	.36	19	50	7809	.24	17	21	1795	.45	7	11	548	.58
Audio Pictures	14	32	2756	.27	8	15	2027	.27	6	15	689	.26	1	2	40	.17

Table 1 *Moderators of Effect Size by Severity Classification (continued)*

	AGG Severity															
	All				Non-referred general				Non-referred extremes				Clinically-referred			
Characteristic and Level	**No. of studies**	**No. of ES**	**N**	**d**	**No. of studies**	**No. of ES**	**N**	**d**	**No. of studies**	**No. of ES**	**N**	**d**	**No. of studies**	**No. of ES**	**N**	**d**
Video	19	35	3674	.23	11	16	2447	.20	5	7	595	.23	4	12	632	.39
Real-time (physical)	2	3	57	1.33	—				2	3	57	1.33	—			
Doll-play	1	3	98	.27	1	3	98	.27	—				—			
Real-time (computerized)	1	2	75	.36	—				1	2	75	.36	—			
Unclear	10	25			5	17			3	6			2	2		
Provocateur's status																
Don't know each other	31	54	6303	.29	15	24	4009	.25	13	18	1803	.33	4	12	491	.39
From neighborhood/ school	37	76	8369	.35	18	46	6427	.29	17	25	1722	.41	3	5	220	.51
Classmate	19	31	4214	.41	9	13	2982	.26	7	11	984	.58	3	7	248	.62
Friend	7	10	1396	.25	5	8	1143	.26	1	1	112	-.13	1	1	141	.35
Enemy	3	4	724	.26	3	4	724	.23	—				—			
Unclear	22	44			15	31			4	10			3	3		
Total	111	219	29272	.33	59	126	22349	.27	41	65	5475	.39	14	28	1448	.48

Statistical Analysis

All study outcomes were transformed into Fisher Z. Fisher Z is similar to a correlation coefficient, but corrects for nonlinearity of extreme correlation coefficients. Fisher Z calculations were derived from reported test statistics and if required test statistics were derived from reported means and standard deviations. Subsequently, Fisher Z scores were re-transformed into Cohen's *d* to facilitate interpretation. According to Cohen (1988), a Cohen's *d* of 0.3, 0.5, and 0.8 represents respectively a small, medium, and large effect size.

We applied a multi-level modeling approach using the "*metafor*" package (Viechtbauer, 2010) of the R Statistical Software version 3.0.2. A multi-level modeling approach allows to derive multiple effect sizes from each study by modeling dependency in effect sizes (Van den Noortgate et al., 2013). To account for dependency in effect sizes, a three-level meta-analytic model was estimated. A three-level meta-analytic model estimates sample variance for each effect size on level 1, variance in effect sizes within studies on level 2, and variance in effect sizes between studies on level 3 (Hox, 2002; Wibbelink & Assink, 2015). The standard errors of the coefficients in the three-level meta-analytic models were estimated with the Knapp & Hartung method (2003). Parameters were estimated using Restricted Maximum Likelihood estimation (Wibbelink & Assink, 2015). Analyses were conducted in four steps.

1. We first tested whether the overall mean effect size significantly deviated from zero.
2. Two log-likelihood ratio tests were used to evaluate whether estimating within-study variability (level 2) and between-study variability (level 3) in effect sizes significantly improved model fit. Subsequently, the Higgins and Thompson method (2002) was used to demonstrate how much variance in effect sizes was due to sampling variability (level 1), within-study variability (level 2), and between-study variability (level 3).
3. The influence of multiple moderators on the relation between HIA and aggression was analyzed using a multilevel mixed-effect model. Since including multiple moderators in one model inflates the type II error rate, separate 3-level mixed-effect models were fitted for each moderator separately. Subsequently, significant moderators were fitted in a 3-level mixed-effect model to address possible confounding among moderators. A multi-model inference approach was used to fit each possible model including none, one, and up to all of the selected moderators to the data and compare the goodness of fit of each model using Akaike information criterion values (AICs, see Burnham & Anderson, 1998). This method allows to examine the relative importance of each predictor when taking all possible models into consideration. Dependence in study characteristics prohibited to examine higher order interaction effects, as several combinations of child- and methodological characteristics often occurred and others rarely or never occurred.

4. Fourth, since the previous meta-analysis showed a significant effect of aggression severity on the relation between HIA and aggression, and to avoid confounding between aggression severity and other moderators, subset analyses were run for each of the three aggression severity groups separately (i.e., non-referred children with normal aggression scores, non-referred children with extreme aggression scores, clinically-referred aggressive children). Findings from these subset-analyses corresponded to the main study findings and are therefore only reported in the supplementary materials.

Publication Bias

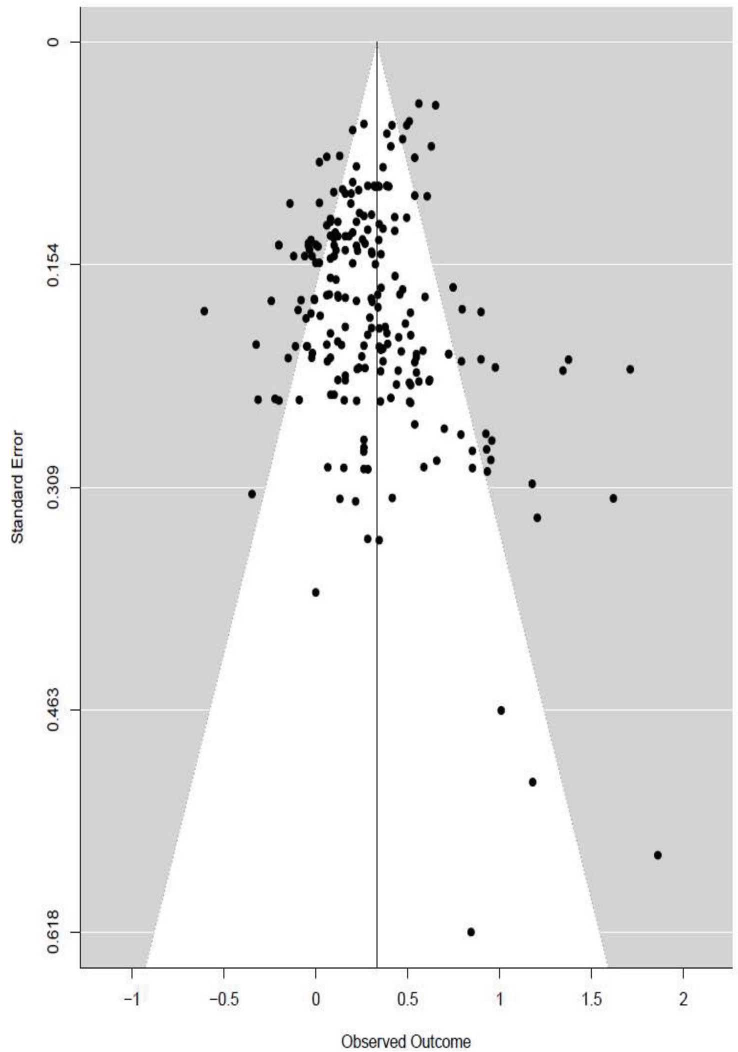
The fail-safe N method is frequently used in meta-analyses (e.g., in the 2002 HIA meta-analysis), but has been criticized for not providing a valid assessment of publication bias and its statistical weakness (e.g., Becker 2005; McDaniel et al., 2005). It is unclear whether a funnel plot, weighted Egger's test, and the trim and fill method are informative indicators of publication bias in heterogeneous data-sets (e.g., Van Assen et al., 2014; Coburn & Vevea, 2015).

To handle publication bias we therefore tried to include as many effect sizes derived from unpublished data as possible. This effort resulted in 66 effect sizes derived from unpublished data of 219 effect sizes total (30.1%). Unpublished data was not only operationalized as each effect size derived from unpublished studies, but also as each effect size derived from published studies where additional information needed to be provided by the authors. If publication bias was present it would be expected that effect sizes derived from unpublished data were smaller than effect sizes based on published data. However, results showed that effect sizes derived from unpublished data were actually larger than effect sizes derived from published data (d = 0.40 vs d = 0.31 vs, p = .128) and thereby indicated no effect of publication bias towards null-findings. In addition, using a strict criterion where unpublished data was operationalized as each effect size derived from unpublished studies (e.g., dissertations) showed no indication of publication bias towards null-findings. This strict criterion resulted in 16 effect sizes derived from unpublished studies (7.3%) and results demonstrated that effect sizes derived from unpublished data were significantly larger than effect sizes derived from published data (d = 0.54 vs d = 0.31, p = .014).

Funnel Plot

Figure 2 shows a funnel plot of the effects. Although this was not used as an indicator of publication bias, it allows to evaluate whether there is a pattern in the data. A weighted Egger' test demonstrated that effect sizes were not distributed in symmetrical manner across the funnel (r_t = 0.16, p < .001). Larger studies were mainly distributed around the overall mean effect size whereas smaller studies were more spread across the funnel. Moreover, the funnel plot demonstrated multiple datapoints fall outside of the funnel, indicating these datapoints show significant heterogeneity in effect size relative to its standard error. However, examining the leverage values and Cook's distance of the data-points demonstrated none should be considered as outliers or indicate excessive influence on the results. In addition, the funnel plot showed a gap on the bottom left, indicating that relatively large positive effect sizes combined with a large standard error were more often observed than negative effect sizes with a large standard error. A plausible explanation might be that larger positive effect sizes were derived from clinically-referred aggressive samples which in general showed larger effects (d = .48) and consisted of a smaller sample (mean N = 103) than studies with non-referred aggressive samples (respectively d = .27 and mean N = 379).

Figure 2. *Funnel plot with Fisher's z transformed Cohen's d. On the y-axis are the standard errors of the effect sizes, with smaller standard errors representing larger sample sizes. On the x-axis are the associations between childhood HIA and aggression*



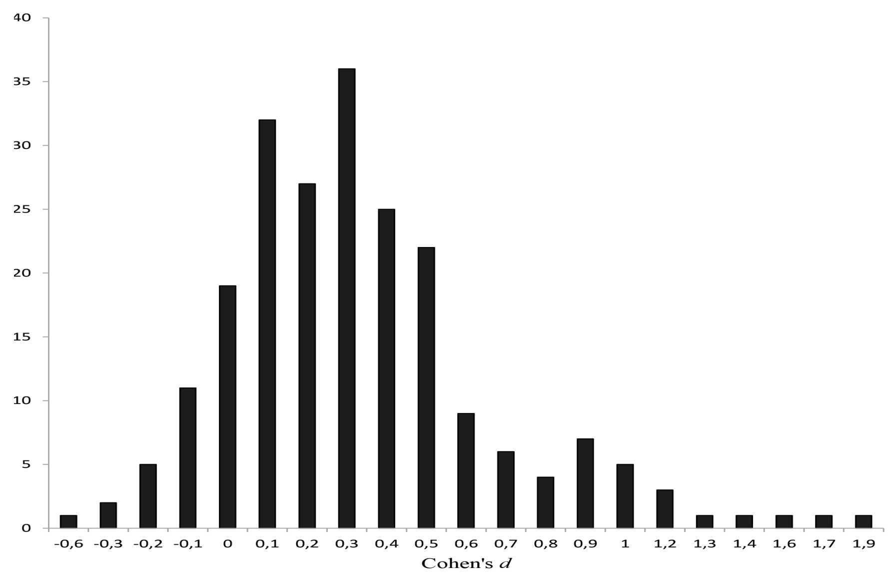
RESULTS

Overall Effect Size

219 effect sizes from 111 studies with 29.272 participants were included in this meta-analysis. Figure 3 shows the distribution of effect sizes. 186 out of 219 effects were in the hypothesized direction. The overall weighted mean effect size was d = 0.33, which significantly deviated from zero, SE = 0.03, $t(218)$ = 12.16, p < .001, 95% CI [0.28-0.39]. Thus, overall results demonstrated a robustly significant, modest positive association between childhood aggression and HIA.

However, this mean effect size should be interpreted with care, because effect sizes varied significantly between studies. The test for residual heterogeneity of the main-effect model showed there was significant heterogeneity in effect sizes not explained by the model ($Q(218)$ = 748.57, p < .001). In addition, two likelihood ratio tests demonstrated that effect sizes differed significantly within, $\chi^2(1)$ = 7.68, p = .006, and between studies, $\chi^2(1)$ = 48.57, p < .001. Subsequently, the distribution of the total variance in effect sizes across the three levels was examined. The percentage of the variance in effect sizes explained by sampling variability was 23.68%. The percentage of the variance in effect sizes explained by differences within studies (within-study variability) was 7.42%. The percentage of the variance in effect sizes explained by differences between studies (between study variability) was 68.90%. The two likelihood ratio tests and test for (residual) heterogeneity indicated that specific child- and methodological characteristics could possibly explain the variability in effect sizes. Therefore planned univariate moderator analyses were conducted.

Figure 3 *Distribution of effect sizes*



Moderator Analyses

The statistics for the test of the moderators (Q_M) and statistics for the test of residual heterogeneity (Q_E) are reported in the supplementary materials (See Table S2). For all the moderators the test of residual heterogeneity was significant, demonstrating there was still unexplained variance in effect sizes beyond each moderator.

Emotional Involvement

To examine whether effect sizes were dependent on emotional involvement, moderation by type of stimulus presentation was tested. Mean effect sizes derived from self-read (*d* = 0.44), audiotaped (*d* = 0.36), pictorial (*d* = 0.25), audiotaped and pictorial (*d* = 0.27), videotaped hypothetical stories (*d* = 0.23) and real-time interactions with a real peer (*d* = 1.33) significantly deviated from zero. The mean effect sizes derived from real-time computerized interactions with a presumed peer (*d* = 0.36) and hypothetical stories presented through doll-play (*d* = 0.27) did not deviate from zero, indicating there was no relation between HIA measured through these types of stimulus presentation and aggression. The mean effect size of HIA measured through real-time interactions with a real peer was significantly larger than the mean effect sizes of all other types of stimulus presentation (vs self-read, *p* = .013; vs audiotaped, *p* = .006; vs pictorial, *p* = .004; vs audiotaped and pictorial, *p* = .003, vs videotaped, *p* = .002; vs real-time computerized interactions with a presumed peer, *p* = .033; vs doll-play, *p* = .016). The mean effect size of HIA measured through self-read hypothetical stories was significantly larger than the mean effect size derived from videotaped hypothetical stories (*p* = .024). The coefficients for the type of stimulus presentation are reported in the supplementary materials (See Table S3). Thus, in line with our hypothesis, results on the type of stimulus presentation indicate that the strength of the relation between HIA and aggression increased with the level of emotional involvement the social situations elicited.

HIA towards Familiar vs Unfamiliar Others

To examine whether the relation between HIA and aggression is present in situations with both familiar and unfamiliar others, but stronger in situations with disliked others encountered in previous problematic social encounters, moderation by provocateur's status was tested. Results showed that the relation between HIA and aggression significantly deviated from zero for all types of provocateur's status (*d* = 0.25-0.41). However, no differences between types of provocateur's status were found (*p* = .539). Thus, contrary to our hypothesis, results on the provocateur's status indicate that the strength of the relation between HIA and aggression was not dependent on the familiarity of peers.

HIA in Aggressive-rejected and Aggressive Samples

To examine whether the relation between HIA and aggression is present in aggressive-rejected and aggressive samples, moderation by sociometric status was tested. The mean effect sizes of aggressive-rejected samples (*d* = 0.61) and aggressive samples (*d* = 0.30) both significantly deviated from zero. Results showed that in both types of samples there was a small to moderate positive association between HIA and aggression. In addition, the mean effect size of aggressive-rejected samples was significantly larger than the mean effect size of aggressive samples (*p* < .001). The coefficients for sociometric status are reported in the supplementary materials (See Table S4). Thus, in line with our hypothesis, results indicate that the relation between HIA to peers and aggression existed irrespective of the sociometric status of participants, and was stronger for children who are both aggressive and rejected.

HIA and Reactive Aggression

To examine whether the relation between HIA and aggression is stronger for reactive aggression, moderation by function of aggression was tested. Results showed that the relation between HIA and aggression significantly deviated from zero for both reactive aggression (*d* = 0.36) and aggression measured as a general construct (*d* = 0.33). However, no differences between the types of aggression function were found (*p* = .602). Thus, contrary to our hypothesis, results indicate that the relation between HIA and aggression was not stronger for reactive aggression than for aggression in general.

HIA and Proportion of ADHD Diagnoses in the Sample

To examine whether the strength of the relation between HIA and aggression increased with the proportion of ADHD diagnoses in the aggressive sample, moderation by ADHD comorbidity was tested. The association between HIA and aggression was not dependent on the percentage of ADHD diagnoses in the sample (*p* = .958). Thus, contrary to our hypothesis, results indicate that the strength of the relation between HIA and aggression did not increase with the proportion of ADHD comorbidity in the aggressive sample.

Exploratory Analyses of Moderators

Consistent with the findings in the meta-analysis of De Castro and colleagues (2002), effect sizes in the current meta-analysis were larger in samples with more severe behavioral problems. Moreover, aggression assessed by a staff-member was associated with higher effect sizes than all other types of informants, except for aggression assessed by an observer. In addition, results demonstrated that effect sizes were larger when more reliable HIA measures were used. For the other exploratory moderators no effects were found. For details see supplementary materials.

Multi-model Inference: Selection of Moderators

To examine whether moderators explained significant variance in effects size over and above the effects of other moderators, we used a multi-model inference approach. This procedure resulted in 74 effect sizes (out of 219) used for estimating all possible models. Results demonstrated that moderators were too confounded to distinguish unique effects of moderators when multiple models were taken into account (see supplementary materials for details).

DISCUSSION

Social-cognitive theories propose a relation between HIA and aggression in children and specific moderators of this relation. This meta-analysis found an overall modest positive association between childhood HIA and aggression (mean effect size *d* = 0.33). However, this mean effect size should be interpreted with care, because effect sizes varied significantly between studies. As expected, the relation between HIA and aggressive behavior was found to be stronger in emotionally engaging situations, and not to be limited to interactions with known peers, nor to rejected-aggressive children, nor to reactive aggression, nor to a comorbid ADHD diagnosis. In line with the previous meta-analysis (De Castro et al., 2002), results showed that the association between childhood HIA and aggression is stronger in more severely aggressive samples. In addition, the exploratory moderator analyses demonstrated that the strength of the association between HIA and aggression was dependent on the reliability of the HIA measures and the type of informant to assess aggression.

We tested specific hypotheses about moderators of the relation between HIA and aggression in children. The first hypothesis stated that the relation between HIA and aggression is stronger in emotionally-engaging situations. In line with our hypothesis, effect sizes derived from real-time interactions with a real peer were very large (*d* = 1.33), and significantly larger than for other types of stimulus presentation. However, it should be mentioned that only three effect sizes derived from two different studies concerned real-time interactions with a real peer. Almost 98% of the effect sizes were derived from studies using hypothetical stories to measure HIA. Although hypothetical stories were presented in different formats (e.g., self-read, audiotaped, pictorial, videotaped), their effect sizes were relatively small (*d* = 0.23 to 0.44). The findings seem to be in line with SIP models that postulate that HIA in aggressive children is particularly present in personally-involving and emotionally-engaging situations (Dodge, 1991).

Methodologically, it is important to note that results only showed a large effect for real-time interactions with a real peer and not computerized real-time interactions with a presumed peer. A plausible explanation could be the lack of observations for computerized real time interactions (two effect sizes from one study), which could have resulted in an unreliable estimate of the true effect size. Another explanation could be that this study assessed computerized real-time interactions with a presumed peer through a race-car game (Yaros & Lochman, 2014). This type of stimulus presentation might not have elicited sufficient levels of emotional engagement to evoke strong HIA, because the peer's behavior may have been considered legitimate in the gaming context. In sum, the findings on the type of stimulus presentation suggest that particularly social interactions that evoke sufficient emotional engagement elicit the emotional processes that activate HIA. This finding has implications for clinical practice, since it implies that cognitive-behavioral interventions should assess and target HIA in emotionally engaging situations.

The second hypothesis stated that the relation between HIA and aggression is present in social situations with both unfamiliar and familiar others. In addition, we expected this relation to be stronger in situations with disliked others encountered in previous problematic social situations. Results demonstrated that the relation between HIA and aggression was present irrespective of the provocateur's familiarity. Results did not show that the relation between HIA and aggression was stronger in social situations with disliked others who children had encountered in previous problematic social situations. This finding might suggest that HIA is not context-specific. However, another explanation could be the lack of observations (4 effect sizes from 3 studies) on HIA towards disliked others encountered in previous problematic social interactions, which could have resulted in unreliable estimates. Nonetheless, the findings seem to be in line with social-cognitive theory that proposes that the tendency to attribute hostile intent others derives from a general cognitive disposition towards both known and unknown others. For clinical practice this implies that CBT interventions could target a general cognitive disposition to establish significant and prolonged changes in SIP and subsequent behaviors across a wide range of contexts.

The third hypothesis stated that the relation between HIA and aggression is present irrespective of the sociometric status of participants, yet would be particularly pronounced in aggressive-rejected samples. Results showed support for this hypothesis and demonstrated that the relation between HIA and aggression was present in both aggressive-rejected and generally aggressive samples, but was stronger in aggressive-rejected samples. This finding supports the assumption that HIA derives from a general cognitive disposition that guides information processing across a broad range of contexts. In addition, since our meta-analysis only included studies that used social situations with peers to measure HIA, the finding that the relation between HIA and aggression was stronger in aggressive-rejected samples might indicate that the relation between HIA and aggression is stronger in situations that match specific memories of rejection by peers. For clinical practice this implies that cognitive-behavioral interventions could possibly be more effective when HIA is targeted in contexts similar to specific memories of aversive social experiences.

The fourth hypothesis stated that the relation between HIA and aggression is stronger when aggression is operationalized as reactive aggression. Results did not support this hypothesis and demonstrated no difference in effect sizes based on aggression measured as reactive aggression or as a general construct. An explanation could be the method used for the coding of this variable. Since empirical research suggests that the majority of aggressive children to some extent engage in reactive aggressive behaviors (Dodge et al., 1997), it may well be true that a substantial part of the samples where aggression was measured as a general construct, were primarily reactive- or reactive-proactive samples. This could have caused the null-result for this hypothesis. Another explanation could be that the relation between reactive HIA and aggression was based on 26 effect sizes and only one of these effect sizes was derived from clinically-referred aggressive samples. Since aggression severity seems to contribute to the strength of the relation between HIA and aggression it would be expected that the relation between HIA and reactive aggression is particularly strong in clinically-referred aggressive samples. Although the one effect size derived from clinically-referred aggressive samples was relatively large (*d* = 0.62), a lack of observations prohibits from drawing firm conclusions.

The fifth hypothesis stated that the relation between HIA and aggression is stronger in aggressive samples consisting of children with attention-deficit hyperactivity disorder (ADHD). Results did not support this hypothesis and demonstrated no effect of ADHD on the relation between HIA and aggression. However, only 22 effect sizes (10%) were based on samples where the presence of a ADHD diagnosis was measured and the majority of these samples were not full-ADHD samples. The lack of observations on ADHD comorbidity could have caused a lack of power to detect true effects and thereby the null-findings for this moderator. Another explanation could be that deficits in cognitive capacities of ADHD children are similar to deficits in cognitive capacities of aggressive children.

Exploratory analyses showed that the strength of the association between HIA and aggression significantly increased with higher Cronbach's α reliability. Cronbach's α's were reported for only 97 out of 219 effect sizes and ranged from .37 to .94, with a mean of .73. In addition, since more than half of all effect sizes were derived from studies that did not report a Cronbach's α for the HIA measure, it is unclear how the reliability of the HIA measure influenced effect sizes in these studies. It could be that at least several studies that did not report a Cronbach's α for the HIA measure, used an unreliable instrument to measure HIA and thereby reduced effect sizes. Thus, despite emphasis put on the importance of reliability of HIA measures in the previous HIA meta-analysis (De Castro et al., 2002), still less than half of the studies included reported a Cronbach's α. This seems cause for worry, as clinical decision making should not depend on unreliable measures or idiosyncrasies of particular vignettes chosen to assess HIA. The finding that larger effect sizes were associated with a higher Cronbach's α, emphasizes the importance for clinicians and researchers to only use reliable instruments to adequately measure HIA.

Exploratory analyses also demonstrated that the type of informant to assess aggression in children moderated the association between aggression and HIA. Results showed that aggression assessed by a staff-member yielded larger effect sizes than aggression assessed by all other type informants, except for aggression assessed by an observer. The latter might be due to a lack of observations (*k* = 2). A plausible explanation for the fact that effect sizes were larger in studies where aggression was assessed by a staff-member might be that all these studies (*k* = 5) were performed in clinically-referred aggressive samples. Since results demonstrated that the severity of aggressive behavioral problems contributes to the strength of the association between childhood aggression and HIA, the larger effect sizes for aggression assessed by a staff-member might be explained by the severity of aggressive behavioral problems for this subgroup.

Although the univariate moderator analyses demonstrated that several moderators influenced the relation between childhood HIA and aggression, a multi-model inference approach to combine these moderators was not feasible. An explanation might be that there was a strong interdependence between child- and methodological characteristics, where specific combinations of child- and methodological characteristics frequently, rarely, or never occurred (e.g., real-time interactions for clinically-referred aggressive samples). As a result, moderators were too confounded to distinguish unique effects of moderators when taking multiple models into account. Moreover, results demonstrated that the predictors that yielded the largest effect sizes consisted of relatively few observations. For example, only 28 effect sizes (12.8%) were derived from clinically-referred aggressive samples, 17 effect sizes (7.8%) from aggressive-rejected samples and only 3 effect sizes (1.4%) from real-time interactions with a real peer. The lack of observations on the strongest predictors could also be an explanation for the fact that a model without moderators included best fitted the data.

The large amount of residual heterogeneity seems to suggest that we did not capture important moderators of effect yet. Perhaps surprisingly, SIP theory is more specific about moderators of HIA performance than current research methods capture. For example, this meta-analysis did not examine the effect of several demand characteristics of HIA tasks that are implied by SIP theory. Cognitive capacities are considered key moderators of SIP (e.g., Dodge & Pettit, 2003) and tasks to measure HIA may inadvertently differ in the cognitive capacities they require for children. For example, to understand the task and to indicate that they do not interpret intentions as hostile (e.g., by requiring complex words like 'accidental' or 'unintended') or the amount of working memory understanding a task requires (e.g., remembering that you were the actual target child in the vignette while watching a video). In the current meta-analysis too few studies assessed executive functioning (e.g., working memory) and this prohibited from adequately testing the effect of this moderator. Therefore this meta-analysis used IQ as an indicator of cognitive abilities. However, this moderator did not show an effect. Nonetheless, given that children differ greatly in cognitive abilities, the presumed role of cognitive abilities in SIP, and the methods used to measure HIA varied considerably between studies, it could be that this influenced the results. Systematically studying (and varying) such test characteristics would be highly informative in understanding the roles of cognitive functioning in HIA.

Another moderator that was not measured in this meta-analysis was social desirability. Since 98% of the effect sizes were based on paper-pencil hypothetical stories to measure HIA, it could be that social desirability influenced participants' responses in studies using hypothetical stories. More specifically, it could be that using a paper-pencil format in an individual or group-based setting reminds children of an exam or test and therefore children may feel more reluctant to give socially undesirable answers. Another moderator that was not measured and could have influenced results is socio-economic status (SES). Research indicates that low SES is associated with chronic stressors such as parental psychopathology, deprived neighborhoods and social isolation (Baum et al., 1999; Pinderhughes et al., 2001). From a schema-theory perspective it can be assumed that these chronic stressors contribute to the development and maintenance of hostile schemata and thereby HIA (Nas et al., 2005). In this meta-analysis, 137 effect sizes (63%) from 69 studies (62%) were based on samples from the United States, a nation with large socio-economic inequalities (e.g., gini index; Central Intelligence Agency, 2009). It could be that effect sizes depend on the magnitude of variance in SES both within and between samples. Unfortunately, an insufficient number of studies (k = 5) included in the current meta-analysis measured SES and this prohibited from adequately testing the effect of this moderator.

Strengths and Limitations

An important strength of this meta-analysis is that it included studies from over 40 years of research on the relation between childhood HIA and aggression, and applied a multi-level modeling approach to analyze results. Multi-level model analyses allow to correct for dependency in effect sizes within studies and thereby allows to derive multiple effect sizes per study (Van den Noortgate et al., 2013). This resulted in 219 effect sizes based on the relation between aggression in children and HIA. In addition, this meta-analysis not only examined the overall relation between childhood HIA and aggression, but also examined specific theory-driven moderators of this relation. Thus, we obtained findings that inform our understanding of when and how HIA is related to aggression, with clear implications for the nature of HIA.

An important limitation of this meta-analysis is the strong interdependence between study characteristics. In other words, many studies used similar methodologies to measure HIA and aggression. As a consequence, specific combinations of child- and methodological characteristics frequently, rarely, or never occurred. The lack of observations for various specific combinations of child- and methodological characteristics might have contributed to confounding of moderators when included in one model. This made it impossible to disentangle specific effects of certain child- and methodological characteristics. A second limitation is that publication bias was only addressed through one method. This method yielded no indication for publication bias towards null-findings, and the fact that effect sizes from unpublished data were larger than effect sizes from published data could suggest true effect sizes in this meta-analysis were actually underestimated. More certainty about publication biases could be attained when multiple methods for testing publication bias become available for multi-level meta analyses.

Future Recommendations and Implications

The significant amount of residual heterogeneity emphasizes the need for theory development and research on the effects of specific combinations of child- and methodological characteristics on the relation between childhood HIA and aggression. Therefore, future research may focus on testing a variety of child- and methodological characteristics that are not frequently measured to date. To examine the effect of emotional engagement, researchers could manipulate the level of emotional engagement across presented social situations and directly compare HIA in real-time interactions and HIA as assessed through hypothetical stories using a within-subjects design.

In addition, context specificity of HIA seems to deserve more attention because of its relevance to intervention. To further examine the effect of social experiences on SIP in different contexts, future studies may link experiences in specific contexts (e.g., harsh parenting and peer rejection) prospectively to HIA in the same and differing contexts (e.g., with peers or adults) and manipulate the provocateur's status (e.g., unknown, friend, enemy) and type of context (e.g., provocation, peer entry, expectation, failure, unjust punishments). This would allow to evaluate whether the relation between HIA and aggression is stronger when the current social situation matches specific memories of previous aversive social experiences.

Last but not least, the current analysis did not address malleability of HIA and its effects on aggressive behavior. Experimental research on moderators of the relation between HIA and aggression may go hand in hand with experimental micro trials testing specific ways to reduce HIA. Recent studies suggest that HIA may be reduced with relatively simple means, such as implicit cognitive bias modification (Penton Voak et al., 2013) or parental instructed story reading (Van Dijk et al., 2018). Such experimental manipulation of HIA may help understand the dynamics of HIA and simultaneously inform effective intervention.

Conclusion

In sum, the meta-analytical findings indicate that HIA is a general cognitive disposition that guides information processing across a broad variety of contexts, including interactions with unknown peers. The relation between HIA and aggression is stronger in social situations that elicit sufficient emotional engagement and for more severely aggressive children. In addition, the relation between HIA and aggression depends on the reliability of HIA measures, but is not stronger for reactive aggression or proportion of ADHD diagnoses in the samples. Future research will further our understanding of this key variable in the development of aggressive behavior.

SUPPLEMENTARY MATERIALS

Coding of Additional Moderators

The additional moderators were coded as in the 2002 meta-analysis. Frequency distributions of child- and methodological characteristics are reported in Table S23.

Child Characteristics

Additional child characteristics that were examined were operationalized as follows:

Aggression Severity. Aggression severity was coded categorically. Categories consisted of the relation between HIA and aggression was measured in a non-referred general sample, non-referred extreme sample, or clinically-referred sample. Effect sizes derived from a non-referred general sample were either based on a correlation between HIA and aggression in a non-referred sample with normal aggression scores or a comparison of HIA scores between a group scoring above the mean and below the mean on the aggression measure. Effect sizes derived from a non-referred extreme sample were based on a comparison of HIA scores between non-referred groups that differed extremely in aggression scores (e.g., one group of children scoring more than 1 *SD* below the sample mean vs one group of children scoring more than 1 *SD* above the sample mean). Effect sizes derived from a clinically-referred sample were based on a comparison of HIA between a clinically-referred aggressive sample (e.g., special education, psychiatric care, prison) and non-referred control sample.

Participant's Age. Participant's age was coded categorically. Categories consisted of 3-6 years, 6-12 years, 9-12 years, and > 12 years old. The reason that categories overlap is that a large proportion of studies contained children from a large age-range. When the age-range of a sample overlapped with two categories, the decision for the coding was based on which of the categories showed the most overlap with the age-range in the sample (e.g., a sample with an age-range of 8-12 years was coded as 9-12 years, a sample with an age-range of 7-10 years was coded as 6-12 years).

Gender. Gender was coded categorically. Categories consisted of effect was based on boys sample, girls sample, or a mixed sample.

Form of Aggressive Behaviors. Form of aggressive behaviors was coded categorically. Categories consisted of aggression was measured as overt aggression (aggression measured as open confrontational aggression such as physical fighting or verbal threats), physical aggression (aggression measured as physical aggressive acts such as starting fights), indirect aggression (aggression measured as non-open forms of aggressive behaviors such as spreading rumors or criticizing other's performance), relational aggression (aggression measured as aggressive behaviors aimed at damaging other's social status or relations), and physical and relational aggression (composite measure of relational and physical aggressive behaviors).

Intelligence. Intelligence was coded categorically. Categories either consisted of effect size was controlled for (verbal) intelligence (studies that matched groups on intelligence or effect size was based on analyses that included intelligence as a covariate) or effect size was not controlled for intelligence (studies that not matched groups on intelligence or effect size was based on analyses that did not include intelligence as a covariate).

Proportion of ODD or CD in the Sample. Proportion of ODD or CD in the sample was coded as a continuous variable representing the proportion of Oppositional Deviant Disorder (ODD) or Conduct Disorder (CD) diagnoses in the sample.

Methodological Characteristics

Additional methodological characteristics that were examined were operationalized as follows:

The Participant's Perspective. The participant's perspective was coded categorically. Categories were participant's perspective was a first-person perspective (e.g., you are hit by a ball in the back), non-first person perspective where participants were instructed being the protagonist (e.g., a child is hit by a ball in the back, imagine you are this child) or a non-first person perspective where participants were instructed to evaluate the protagonist's perspective (e.g., another child is hit by a ball in the back).

Type of Context. Type of context coding was based on Dodge et al. (1985) taxonomy of social problem situations. Categories consisted of presented social situations were provocative social situations (e.g., participant is victim of a provocative act), peer entry situations (e.g., participant is rejected to join a peer group, activity or play), social expectation situations (e.g., participant needs help from a peer), social failure situations (e.g., participant's is being outperformed by a peer), or conflicts with adults (e.g., being falsely accused by a parent or teacher).

Type of Informant. Type of informant was coded categorically. Categories consisted of aggression in children was measured through parents, teachers, peers, observer, self-report, staff member, or more than one informant.

Type of Measure to Assess Aggression in Children. Type of measure to assess aggression in children was coded categorically. Categories consisted of aggression in children was measured through a checklist or rating scale, sociometric nomination (e.g., nomination by peers), observational rating, or two or more of the previous.

The Number of Presented Social Situations. The number of presented social situations was coded continuously and then transformed into a categorical variable. Categories consisted of less than six or more than five presented social situations.

Ambiguity of Social Situation. Ambiguity of social situation was coded continuously as the proportion of ambiguous social situations and then transformed into a categorical variable. Categories consisted of presented social situations were solely ambiguous or presented social situations were a mixture of ambiguous- and non-ambiguous social situations (e.g., hostile intention, ambiguous intention, benign intention).

Peers in Social Situations. Peers in social situations was coded continuously as the proportion of peers in the presented social situations and then transformed into a categorical variable. Categories consisted of the provocateurs in the presented social situations were solely peers or a mixture of both peers and adults.

Type of Setting of the Experiment. The type of setting of the experiment was coded categorically. Categories consisted of assessment was conducted in an individual- or group-based setting.

Type of Response Format. Type of response format was coded categorically. Categories consisted of HIA was measured through an open response format (e.g., why do you think he/she acted that way?), multiple choice (e.g., the participant is asked to select a hostile, benign, or neutral interpretation), or rating scale format (e.g., the participants is asked to rate how strongly he/she thinks the other person did it on purpose/to be mean), or a composite score of open- and multiple choice or rating format.

Type of HIA Scoring. Type of HIA scoring was coded categorically. Categories consisted of HIA was calculated as either the number or proportion of hostile attributions, or hostile minus benign attributions, or average of attributions made on a rating scale ranging from hostile to benign.

Reliability of the HIA Measure. Reliability of the HIA measure was coded as a continuous variable representing the reported Cronbach's α. For interpretational purposes, the reported Cronbach's α were standardized with a mean of zero.

Exploratory Analyses of the Other Child-and Methodological Characteristics

The Q_M statistics for the test of moderators and Q_E statistics for the test of residual heterogeneity are reported in Table 5 and Table 7.

Aggression Severity

For aggression severity, the mean effect sizes of non-referred general samples, non-referred extreme samples and clinically-referred samples all significantly deviated from zero. Moreover, the mean effect size of the clinically-referred samples (d = 0.48) was significantly larger than the mean effect size of non-referred general samples (d = 0.27). In addition, the mean effect size of non-referred samples with extreme aggression scores (d = 0.39) was significantly larger than the mean effect size of non-referred general samples. The coefficients for aggression severity are reported in Table 6.

Type of Informant to Assess Aggression

For type of informant to assess aggression, mean effect sizes based on aggression assessed by parents, teachers, peers, more than one informant, self-report and staff-member significantly deviated from zero, with a small to large positive effect. The mean effect size based on aggression assessed by an observer did not significantly deviate from zero, indicating there was no relation between aggression assessed by an observer and HIA. In addition, the mean effect size of aggression assessed by a staff-member (d = 0.74) was significantly larger than the mean effect sizes of aggression assessed by all other types of informants (d = 0.29-0.37), except for aggression assessed by an observer (d = 0.31). The coefficients for the type of informant to assess aggression are reported in Table 8.

Reliability of the HIA Measure

For reliability of the HIA measure a significant overall effect on the relation between HIA and aggression was found. The strength of the association between HIA and aggression significantly increased with a higher Cronbach's α. The coefficients for the reliability of the HIA measure are reported in Table 9.

Other Moderators

The association between (childhood) HIA and aggression was not dependent on participant's perspective, type of context, type of aggression measure, number of presented social situations, ambiguity of social situations, proportion of peers in the social situations, type of setting, type of response format, type of HIA scoring, age-group, gender, form of aggressive behavioral problems, intelligence, and proportion of ODD or CD diagnoses in the sample.

Meta 2002 vs Meta 2017. The coefficients for effect sizes from the meta 2002 vs additional effect sizes included in the meta 2017 are reported in Table 10. The overall mean effect size based on studies included in the previous meta-analysis (d = 0.46) was significantly larger than effect sizes derived from additional studies (d = 0.28). An explanation could be the fact that the majority of studies in the original meta-analysis used non-referred samples with extreme aggression scores (57%) whereas the majority of additional studies used non-referred samples with normal scores (67%).

A Multimodel Inference Approach

A multimodel inference approach demonstrated that the main-effects model and a model with sociometric status as a moderator yielded the best goodness-of-fit. Based on parsimony, the model with no moderators included was selected as the best model to fit the data. All other possible models contained values that were outside a range of 2 AIC-units, which implies these models were substantially less plausible. The AIC-values of the foremost models are depicted in Table 11 and Figure 4.

Subsequently, the relative importance of each predictor was examined when taking all possible models into consideration. The importance value of each predictor equals the sum of the probabilities for all models in which the predictor is included. Results showed that none of the selected moderators were of substantial importance. The relative importance of each moderator over all models is depicted in Figure 5. The vertical line is drawn at 0.8 and can be used as a cutoff to differentiate between important and non-important moderators. It is clearly shown that none of the predictors are close to this cut-off and thereby of predictive importance. Results again demonstrated that the best model is the model without any predictors included. Moreover, results showed that when including all of these moderators in the same model, none of the moderators remained significant and there was still a significant amount of residual heterogeneity (Q_E (df = 60) = 154.52, p < .001). This implies that when the moderators were controlled for each other's influence, none of the significant moderators had a significant effect on the association between childhood HIA and aggression, and that a significant amount of variability in effect sizes is not explained by the model. The latter suggests that moderators that were not measured caused this unexplained heterogeneity in effect sizes.

Subset Analyses

As mentioned in the method section, subset analyses were run for each of the three aggression severity groups separately (i.e., non-referred children with normal aggression scores, non-referred children with extreme aggression scores, clinically-referred aggressive children).

Non-referred Samples with Normal Aggression Scores

Child Characteristics. The Q_M statistics for the test of the moderators and Q_E statistics for the test of residual heterogeneity are reported in Table 12.

Age-group. For age-group, the mean effect sizes in non-referred normal aggressive samples of 3-6, 6-12, 9-12 and over 12 year-old children all significantly deviated from zero. For all age-groups in non-referred normal aggressive samples, there was a small to medium positive association between HIA and aggression. In addition, the mean effect size of over 12 year old children (*d* = 0.43) was significantly larger in non-referred normal aggressive samples than the mean effect size of 3-6 year (*d* = 0.22) and 6-12 year (*d* = 0.17) old children. The test of residual heterogeneity was significant, demonstrating there was still unexplained variance in effect sizes beyond this moderator. The coefficients for age-group are reported in Table 13.

Other Child Characteristics. The association between HIA and aggression in non-referred normal samples was not dependent on the gender, form or function of aggressive behavioral problems and sociometric status. For intelligence, proportion of ADHD diagnoses in the sample, and proportion ODD or CD diagnoses in the sample there were no observations in non-referred samples with normal aggression scores.

Methodological Characteristics. The Q_M statistics for the test of the moderators and Q_E statistics for the test of residual heterogeneity are reported in Table 14.

The Type of Measure to Assess Aggression. For the type of measure to assess aggression, the mean effect sizes in non-referred normal samples of aggression assessed through a checklist or rating scale, and sociometric nomination (nomination by peers) significantly deviated from zero. This indicated that for these type of measures to assess aggression there was a significant positive association with HIA in non-referred samples with normal aggression scores. The mean effect size of aggression assessed through observation or two or more types of measures did not significantly deviate from zero, indicating there was no relation between aggression assessed through observation or two or more types of measures and HIA in non-referred normal aggressive samples. In addition, the mean effect size of aggression assessed through a checklist or rating scale (*d* = 0.33) was significantly larger in non-referred normal aggressive samples than the mean effect sizes of aggression assessed through sociometric nomination (*d* = 0.16) or two or more types of measures (*d* = -0.07). The test of residual heterogeneity was significant, demonstrating there was still unexplained variance in effect sizes beyond this moderator. The coefficients for the type of measure to assess aggression are reported in Table 15.

Reliability of the HIA Measure. For the reliability of the HIA measure, the strength of the association between HIA and aggression in non-referred normal samples significantly increased with a higher Cronbach's α. The test of residual heterogeneity was significant, demonstrating there was still unexplained variance in effect sizes beyond this moderator. The coefficients for the reliability of the HIA measure are reported in Table 16.

Other Methodological Characteristics. The association between HIA and aggression in non-referred normal samples was not dependent on the type of stimulus presentation, participant's perspective in the presented social situation, provocateur's status, type of context, number of presented social situations, type of informant to assess aggression, ambiguity of presented social situations, proportion of peers in presented social situations, type of setting, type of response format, and type of HIA scoring.

Non-referred Samples with Extreme Aggression Scores

Child Characteristics. The Q_M statistics for the test of the moderators and Q_E statistics for the test of residual heterogeneity are reported in Table 17. Results showed that sociometric status was a significant moderator of effect in the relation between HIA and (childhood) aggression in non-referred extreme samples.

Sociometric Status. In non-referred samples with extreme aggression scores, the mean effect sizes of aggressive-rejected samples and aggressive samples both significantly deviated from zero. This indicated that for both type of samples there was a small to medium positive association between HIA and aggression. In addition, the mean effect size of non-referred aggressive-rejected samples with extreme aggression scores (*d* = 0.62) was significantly larger than the mean effect size of non-referred aggressive samples with extreme aggression scores (*d* = 0.31). The test of residual heterogeneity was significant, demonstrating there was still unexplained variance in effect sizes beyond this moderator. The coefficients for sociometric status are reported in Table 18.

Other Child Characteristics. The association between HIA and aggression in non-referred extreme samples was not dependent on age-group, gender, form or function of aggressive behavioral problems, intelligence, proportion of ADHD diagnoses in the sample, and ODD or CD diagnoses in the sample.

Methodological Characteristics. The Q_M statistics for the test of the moderators and Q_E statistics for the test of residual heterogeneity are reported in Table 19.

Type of Stimulus Presentation. For type of stimulus presentation, the mean effect sizes in non-referred extreme samples derived from self-read, audiotaped and audiotaped/pictorial hypothetical stories and real-time interactions with a real peer significantly deviated from zero. This indicated that in non-referred extreme samples, there was a significant positive association between HIA and aggression for these types of stimulus presentation. The effect sizes derived from videotaped hypothetical stories and real-time computerized interactions with a presumed peer did not deviate from zero, indicating there was no relation between HIA measured through these type of stimulus presentation and aggression in non-referred extreme samples. In addition, the mean effect size of HIA measured through real-time interactions with a real peer (*d* = 1.33) was significantly larger in non-referred extreme samples than the mean effect sizes of all other types of stimulus presentation (*d* = 0.11-0.45), except for self-read hypothetical stories (*d* = 0.64). Moreover, the mean effect size of HIA measured through self-read hypothetical stories (*d* = 0.64) was significantly larger in non-referred extreme samples than the mean effects derived from videotaped (*d* = 0.23), pictorial (*d* = 0.11) and audiotaped/pictorial hypothetical stories (*d* = 0.26). The test of residual heterogeneity was significant, demonstrating there was still unexplained variance in effect sizes beyond this moderator. The coefficients for the type of stimulus presentation are reported in Table 20.

Other Methodological Characteristics. The association between aggression and HIA in non-referred extreme samples was not dependent on the participant's perspective in the presented social situation, provocateur's status, type of context, number of presented social situations, type of informant to assess aggression, type of measure to assess aggression, ambiguity of presented social situations, proportion of peers in presented social situations, type of setting, type of response format, type of HIA scoring and reliability of HIA measure.

Clinically-referred Samples

Child Characteristics. Results demonstrated that no child characteristics influenced the association between HIA and aggression in clinically-referred samples. The association between HIA and aggression in clinically-referred samples was not dependent on age-group, gender, form or function of aggression, intelligence or proportion of ADHD diagnoses or ODD/CD diagnoses in the sample. The Q_M statistics for the test of the moderators and Q_E statistics for the test of residual heterogeneity are reported in Table 21.

Methodological Characteristics. Results demonstrated that no methodological characteristics influenced the association between (childhood) HIA and aggression in clinically-referred samples. The association between HIA and aggression in clinically-referred samples was not dependent on the type of stimulus presentation, participant's perspective in the presented social situation, provocateur's status, type of context, number of presented social situations, type of informant to assess aggression, type of measure to assess aggression, ambiguity of presented social situations, type of response format, type of HIA scoring and reliability of HIA measure. For type of setting and proportion of peers in the presented social situations there were no observations on both levels. The Q_M statistics for the test of the moderators and Q_E statistics for the test of residual heterogeneity are reported in Table 22.

Table S2 *All the Q_M statistics for the moderator analyses, and the Q_E statistics for the test of residual heterogeneity for the moderators in the main-analyses*

Moderator	QM (df)	p	QE (df)	p
Type of stimulus presentation	2.12 (7)	.044	540.85 (186)	< .001
Provocateur's status	0.78 (4)	.539	503.83 (170)	< .001
Sociometric status	12.07 (1)	<.001	724.68 (217)	< .001
AGG function	0.27 (1)	.602	740.13 (217)	< .001
% ADHD	0.00 (1)	.958	62.62 (20)	< .001

Table S3 *Coefficients for the univariate moderator analysis with type of stimulus presentation*

	ES	SE	t	95% CI	p
Self-reading>0	0.44	0.07	6.71	0.31-0.58	>.001
Pictures>0	0.25	0.12	2.06	0.01-0.49	.041
Audio>0	0.36	0.05	7.80	0.27-0.45	<.001
AudioPictures>0	0.27	0.07	3.62	0.12-0.42	<.005
Video>0	0.23	0.07	3.33	0.09-0.36	<.002
Real-time (physical) >0	1.33	0.35	3.83	0.65-2.02	<.001
Doll-play>0	0.27	0.27	1.02	-0.25-0.79	.311
Real-time (computerized)>0	0.36	0.29	1.22	-0.22-0.94	.223
Difference Self-reading-Pictures	0.19	0.14	1.40	-0.08-0.47	.164
Difference Self-reading-Audio	0.09	0.08	1.11	-0.07-0.25	.270
Difference Self-reading-AudioPictures	0.17	0.10	1.73	-0.02-0.37	.085
Difference Self-reading-Video	0.22	0.10	2.28	0.03-0.41	.024
Difference Self-reading-Real-time (physical)	-0.89	0.35	-2.51	-1.59--0.19	.013
Difference Self-reading-Doll play	0.17	0.27	0.64	-0.37-0.71	.524
Difference Self-reading- Real-time (computerized)	0.09	0.30	0.29	-0.51-0.68	.773
Difference Pictures-Audio	-0.10	0.13	-0.81	-0.36-0.15	.422
Difference Pictures-AudioPictures	-0.02	0.14	-0.14	-0.30-0.26	.887
Difference Pictures-Video	0.02	0.14	0.17	-0.25-0.30	.867
Difference Pictures-Real-time (physical)	-1.08	0.37	-2.93	-1.81--0.35	.004
Difference Pictures-Doll play	-0.02	0.29	-0.07	-0.60-0.56	.949
Difference Pictures-Real-time (computerized)	-0.11	0.32	-0.34	-0.73-0.52	.737
Difference Audio-AudioPictures	0.08	0.09	0.96	-0.09-0.26	.337

Table S3 *Coefficients for the univariate moderator analysis with type of stimulus presentation (continued)*

	ES	SE	t	95% CI	p
Difference Audio-Video	0.13	0.08	1.55	-0.03-0.29	.122
Difference Audio-Real-time (physical)	-0.98	0.35	-2.78	-1.67--0.28	.006
Difference Audio-Doll play	0.09	0.27	0.32	-0.45-0.62	.751
Difference Audio-Real-time (computerized)	-0.00	0.30	-0.01	-0.59-0.58	.995
Difference AudioPictures-Video	0.04	0.10	0.43	-0.16-0.24	.669
Difference AudioPictures-Real-time (physical)	-1.06	0.36	-2.98	-1.76--0.36	.003
Difference AudioPictures-Doll play	0.00	0.28	0.01	-0.54-0.55	.996
Difference AudioPictures-Real-time (computerized)	-0.09	0.30	-0.29	-0.68-0.51	.775
Difference Video-Real-time (physical)	-1.10	0.35	-3.11	-1.80--0.40	.002
Difference Video-Doll play	-0.04	0.27	-0.15	-0.58-0.50	.878
Difference Video-Real-time (computerized)	-0.13	0.30	-0.43	-0.72-0.46	.666
Difference Real-time (physical)-Doll play	1.06	0.44	2.43	0.20-1.92	.016
Difference Real-time (physical)-Real-time (computerized)	0.97	0.45	2.14	0.08-1.87	.033
Difference Doll-play-Real-time (computerized)	-0.09	0.40	-0.22	-0.87-0.69	.824

Table S4 *Coefficients for the univariate moderator analysis with sociometric status*

	ES	SE	t	95% CI	p
Aggressive>0	0.30	0.03	11.17	0.25-0.36	<.001
Aggressive-Rejected>0	0.61	0.08	7.21	0.44-0.78	<.001
Difference Aggressive-Aggressive-Rejected	-0.31	0.09	-3.48	-0.48--0.13	<.001

Table S5 *All the Q_M statistics for the moderator analyses, and the Q_E statistics for the test of residual heterogeneity for the child characteristics in the exploratory analyses*

Moderator	QM (df)	p	QE (df)	p
AGG Severity	3.91 (2)	.022	748.26 (216)	< .001
Age group	1.59 (3)	.192	646.94 (215)	< .001
Gender	0.80 (2)	.451	709.81 (216)	< .001
AGG form	1.26 (4)	.298	240.13 (56)	< .001
Intelligence	0.09 (1)	.766	745.41 (217)	< .001
% ODD/CD	0.11 (1)	.742	63.61 (26)	< .001

Table S6 *Coefficients for the univariate moderator analysis with aggression severity*

	ES	SE	t	95% CI	p
Non-referred general>0	0.27	0.04	7.75	0.20-0.34	<.001
Non-referred extreme>0	0.39	0.05	8.44	0.30-0.48	<.001
Clinically-referred>0	0.48	0.08	5.95	0.32-0.64	<.001
Difference non-referred general-non referred extreme	-0.12	0.06	-2.09	-0.23--0.01	.038
Difference non-referred general-clinically-referred	-0.21	0.09	-2.36	-0.38--0.04	.019
Difference non-referred extreme-clinically-referred	-0.09	0.09	-1.00	-0.27-0.09	.320

Table S7 *All the Q_M statistics for the moderator analyses, and the Q_E statistics for the test of residual heterogeneity for the methodological characteristics in the exploratory analyses*

Moderator	QM (df)	p	QE (df)	p
Participant's perspective	2.85 (2)	.060	499.50 (175)	< .001
Type of context (provocation)	1.07 (1)	.302	694.74 (214)	< .001
Type of context (peer-entry)	0.37 (1)	.542	620.35 (211)	< .001
Type of context (expectations)	0.26 (1)	.613	620.65 (211)	< .001
Type of context (failure)	0.08 (1)	.773	620.96 (211)	< .001
Type of context (adult)	1.27 (1)	.261	584.74 (212)	< .001
Number of presented social situations	0.13 (1)	.715	742.39 (213)	< .001
Type of aggression informant	2.34 (6)	.033	616.83 (212)	< .001
Type of aggression measure	1.72 (3)	.164	686.52 (215)	< .001
Ambiguity of social situations	0.19 (1)	.665	740.38 (217)	< .001
Proportion of peers in social situations	1.07 (1)	.302	585.62 (212)	< .001
Type of setting	0.01 (1)	.933	603.27 (198)	< .001
Type of response format	0.17 (2)	.841	717.96 (216)	< .001
Type of HIA scoring	2.23 (1)	.137	732.87 (211)	< .001
Reliability of HIA measure	6.68 (1)	.011	340.48 (95)	< .001

Table S8 *Coefficients for the univariate moderator analysis with type of aggression informant*

	ES	SE	t	95% CI	p
Parent>0	0.29	0.06	5.10	0.18-0.40	< .001
Teacher>0	0.32	0.04	8.12	0.24-0.40	< .001
Peer>0	0.29	0.05	6.02	0.19-0.38	< .001
Observer>0	0.31	0.25	1.24	-0.18-0.79	.216
More than one Informant>0	0.32	0.06	5.39	0.20-0.44	< .001
Self-report>0	0.37	0.04	8.58	0.29-0.46	< .001
Staff-member>0	0.74	0.13	5.69	0.48-0.99	< .001
Difference Parent-Teacher	-0.03	0.06	-0.58	-0.15-0.08	.564
Difference Parent-Peer	-0.00	0.07	-0.02	-0.14-0.13	.980
Difference Parent-Observer	-0.02	0.25	-0.08	-0.52-0.48	.938
Difference Parent-More than one Informant	-0.03	0.08	-0.40	-0.19-0.13	.687
Difference Parent-Self-report	-0.09	0.06	-1.52	-0.20-0.03	.131
Difference Parent-Staff-member	-0.45	0.14	-3.24	-0.73--0.18	.001
Difference Teacher-Peer	0.03	0.05	0.64	-0.07-0.13	.524
Difference Teacher-Observer	0.02	0.25	0.06	-0.48-0.51	.954
Difference Teacher-More than one Informant	0.00	0.07	0.02	-0.14-0.14	.981
Difference Teacher-Self-report	-0.05	0.05	-1.08	-0.15-0.04	.281
Difference Teacher-Staff-member	-0.42	0.13	-3.20	-0.67--0.16	.002
Difference Peer-Observer	-0.02	0.25	-0.07	-0.51-0.48	.943
Difference Peer-More than one Informant	-0.03	0.08	-0.41	-0.18-0.12	.683
Difference Peer-Self-report	-0.09	0.06	-1.53	-0.20-0.02	.127
Difference Peer-Staff-member	-0.45	0.14	-3.30	-0.72--0.18	.001

Table S8 *Coefficients for the univariate moderator analysis with type of aggression informant (continued)*

	ES	SE	t	95% CI	p
Difference Observer-More than one Informant	-0.01	0.25	-0.05	-0.51-0.49	.960
Difference Observer-Self-report	-0.07	0.25	-0.27	-0.56-0.43	.785
Difference Observer-Staff-member	-0.43	0.28	-1.55	-0.98-0.12	.123
Difference More than one Informant Self-report	-0.06	0.07	-0.78	-0.20-0.09	.437
Difference More than one Informant-Staff-member	-0.42	0.14	-2.94	-0.70--0.14	.004
Difference Self-report-Staff-member	-0.36	0.14	-2.69	-0.63--0.10	.008

Table S9 *Coefficients for the univariate moderator analysis with reliability of the HIA measure*

	ES	SE	t	95% CI	p
Intercept	0.26	0.03	7.97	0.20-0.32	<.001
Slope	0.09	0.04	2.59	0.02-0.16	.011

Table S10 *Coefficients for the univariate moderator analysis with Meta2002-Meta2017*

	ES	SE	t	95% CI	p
Meta2002	0.46	0.05	9.11	0.36-0.56	<.001
Meta2017	0.28	0.03	9.22	0.22-0.34	<.001
Slope Meta2002-Meta2017	0.18	0.06	3.06	0.06-0.30	.003

Table S11 *AICs of foremost models*

Best model	Model	AICs
1	Main-effects model (no moderators)	-128.35
2	Sociometric status	-128.35
3	Reliability of HIA measure	-126.18
4	Sociometric status + Reliability of HIA measure	-126.18
5	Aggression severity	-124.13
6	Aggression severity + Sociometric status	-124.13
7	Aggression severity + Reliability of HIA measure	-121.85
8	Aggression severity + Sociometric status + Reliability of HIA measure	-121.85
9	Type of Stimulus presentation	-119.30
10	Type of Stimulus presentation + Sociometric status	-119.30

Table S12 *All the Q_M statistics for the moderator analyses, and the Q_E statistics for the test of residual heterogeneity for the child characteristics in the subset analyses of non-referred samples with normal aggression scores*

Moderator	QM (df)	p	QE (df)	p
Sociometric status	0.41 (1)	.522	477.25 (124)	< .001
AGG function	3.33 (1)	.071	463.89 (124)	< .001
% ADHD	-			
Age group	3.08 (3)	.030	349.51 (122)	< .001
Gender	1.16 (2)	.317	442.87 (123)	< .001
AGG form	1.10 (4)	.369	195.34 (45)	< .001
Intelligence	-			
% ODD/CD	-			

Table S13 *Coefficients for the univariate moderator analysis with age group of non-referred samples with normal aggression scores*

	ES	SE	t	95% CI	p
3-6>0	0.22	0.08	2.90	0.07-0.36	.004
6-12>0	0.17	0.06	2.88	0.05-0.29	.005
9-12>0	0.30	0.04	6.93	0.21-0.38	<.001
>12>0	0.43	0.07	6.21	0.29-0.57	<.001
Difference 3-6-6-12	0.05	0.10	0.48	-0.14-0.23	.629
Difference 3-6-9-12	-0.08	0.09	-0.94	-0.25-0.09	.351
Difference 3-6->12	-0.22	0.10	-2.11	-0.42--0.01	.037
Difference 6-12-9-12	-0.13	0.07	-1.81	-0.26-0.01	.072
Difference 6-12->12	-0.26	0.09	-2.87	-0.44--0.08	.005
Difference 9-12->12	-0.13	0.08	-1.65	-0.30-0.03	.101

Table S14 *All the Q_M statistics for the moderator analyses, and the Q_E statistics for the test of residual heterogeneity for the methodological characteristics in the subset analyses of non-referred samples with normal aggression scores*

Moderator	QM (df)	p	QE (df)	p
Type of stimulus presentation	0.67 (5)	.651	312.95 (103)	< .001
Provocateur's status	0.08 (4)	.988	269.19 (90)	< .001
Participant's perspective	1.84 (2)	.164	278.66 (98)	< .001
Type of context (provocation)	1.26 (1)	.264	434.66 (124)	< .001
Type of context (peer-entry)	0.02 (1)	.878	367.55 (121)	< .001
Type of context (expectations)	0.11 (1)	.738	367.59 (121)	< .001
Type of context (failure)	3.83 (1)	.053	362.00 (121)	< .001
Type of context (adult)	3.77 (1)	.055	331.48 (122)	< .001
Number of presented social situations	1.01 (1)	.316	470.82 (120)	< .001
Type of aggression informant	1.94 (5)	.093	337.28 (120)	< .001
Type of aggression measure	5.63 (3)	.001	375.08 (122)	< .001
Ambiguity of social situations	0.11 (1)	.743	463.93 (124)	< .001
Proportion of peers in social situations	3.74 (1)	.055	332.08 (122)	< .001
Type of setting	0.02 (1)	.877	352.28 (110)	< .001
Type of response format	0.91 (2)	.404	434.32 (123)	< .001
Type of HIA scoring	0.76 (1)	.384	469.20 (124)	< .001
Reliability of HIA measure	6.74 (1)	.012	286.08 (71)	< .001

Table S15 *Coefficients for the univariate moderator analysis with type of aggression measure of non-referred samples with normal aggression scores*

	ES	SE	t	95% CI	p
Checklist/Rating scale>0	0.33	0.03	10.56	0.27-0.40	<.001
Observation>0	0.31	0.22	1.40	-0.13-0.74	.165
Sociometric nomination >0	0.16	0.05	2.88	0.05-0.26	.005
Two or more types of measures >0	-0.07	0.13	-0.53	-0.34-0.19	.597
Difference Checklist/Rating scale-Observation	0.02	0.22	0.13	-0.41-0.47	.899
Difference Checklist/Rating scale-Sociometric nomination	0.18	0.06	3.12	0.07-0.29	.002
Difference Checklist/Rating scale-Two or more types of measures	0.40	0.14	2.93	0.13-0.68	.004
Difference Observation-Sociometric nomination	0.15	0.23	0.66	-0.30-0.60	.510
Difference Observation-Two or more types of measures	0.38	0.26	1.47	-0.13-0.89	.145
Difference Sociometric nomination-Two or more types of measures	0.23	0.14	1.57	-0.06-0.51	.120

Table S16 *Coefficients for the univariate moderator analysis with reliability of the HIA measure of non-referred samples with normal aggression scores*

	ES	SE	t	95% CI	p
Intercept	0.26	0.04	6.33	0.18-0.34	<.001
Slope	0.12	0.05	2.60	0.03-0.22	.012

Table S17 *All the Q_M statistics for the moderator analyses, and the Q_E statistics for the test of residual heterogeneity for the child characteristics in the subset analyses of non-referred samples with extreme aggression scores*

Moderator	QM (df)	p	QE (df)	p
Sociometric status	9.21 (1)	.004	162.82 (63)	< .001
AGG function	1.46 (1)	.231	187.01 (63)	< .001
% ADHD	3.46 (1)	.122	1.72 (5)	.887
Age group	0.91 (3)	.444	183.61 (61)	< .001
Gender	1.17 (2)	.317	162.06 (62)	< .001
AGG form	1.74 (4)	.259	9.98 (6)	.125
Intelligence	3.83 (1)	.055	178.33 (63)	< .001
% ODD/CD	1.53 (1)	.284	0.56 (6)	.968

Table S18 *Coefficients for the univariate moderator analysis with sociometric status of non-referred samples with extreme aggression scores*

	ES	SE	t	95% CI	p
Aggressive>0	0.31	0.05	5.84	0.20-0.41	< .001
Aggressive-Rejected>0	0.62	0.09	6.93	0.44-0.80	< .001
Difference Aggressive-Aggressive-Rejected	-0.31	0.10	-3.04	-0.52--0.11	.004

Table S19 *All the Q_M statistics for the moderator analyses, and the Q_E statistics for the test of residual heterogeneity for the methodological characteristics in the subset analyses of non-referred samples with extreme aggression scores*

Moderator	QM (df)	p	QE (df)	p
Type of stimulus presentation	2.69 (6)	.024	135.15 (52)	< .001
Provocateur's status	2.13 (3)	.108	162.79 (51)	< .001
Participant's perspective	2.14 (2)	.128	160.17 (54)	< .001
Type of context (provocation)	0.26 (1)	.614	173.54 (61)	< .001
Type of context (peer-entry)	1.63 (1)	.206	167.12 (61)	< .001
Type of context (expectations)	0.09 (1)	.764	173.54 (61)	< .001
Type of context (failure)	0.03 (1)	.869	172.83 (61)	< .001
Type of context (adult)	0.04 (1)	.850	172.53 (61)	< .001
Number of presented social situations	0.56 (1)	.455	190.16 (63)	< .001
Type of aggression informant	0.76 (4)	.558	187.02 (60)	< .001
Type of aggression measure	1.75 (2)	.183	182.30 (62)	< .001
Ambiguity of social situations	0.06 (1)	.809	190.48 (63)	< .001
Proportion of peers in social situations	0.04 (1)	.850	172.53 (61)	< .001
Type of setting	0.76 (1)	.387	169.01 (59)	< .001
Type of response format	1.26 (2)	.292	179.20 (62)	< .001
Type of HIA scoring	2.53 (1)	.117	178.46 (59)	< .001
Reliability of HIA measure	2.58 (1)	.125	32.58 (18)	.019

Table S20 *Coefficients for the univariate moderator analysis with type of stimulus presentation of non-referred samples with extreme aggression scores*

	ES	SE	t	95% CI	p
Self-reading>0	0.64	0.13	4.88	0.38-0.90	<.001
Pictures>0	0.11	0.20	0.58	-0.28-0.50	.566
Audio>0	0.45	0.08	5.67	0.29-0.62	<.001
AudioPictures>0	0.26	0.11	2.31	0.03-0.49	.025
Video>0	0.23	0.14	1.67	-0.05-0.51	.101
Real-time (physical)>0	1.33	0.35	3.83	0.63-2.03	<.001
Doll-play>0	-				
Real-time (computerized)>0	0.36	0.29	1.24	-0.22-0.94	.222
Difference Self-reading-Pictures	0.53	0.23	2.26	0.06-1.00	.028
Difference Self-reading-Audio	0.18	0.15	1.20	-0.12-0.49	.234
Difference Self-reading-AudioPictures	0.38	0.17	2.19	0.03-0.73	.033
Difference Self-reading-Video	0.41	0.19	2.12	0.02-0.79	.039
Difference Self-reading-Real-time (physical)	-0.69	0.37	-1.86	-1.43-0.06	.069
Difference Self-reading-Doll play	-				
Difference Self-reading-Real-time (computerized)	0.28	0.32	0.88	-0.36-0.92	.381
Difference Pictures-Audio	-0.34	0.21	-1.64	-0.76-0.08	.108
Difference Pictures-AudioPictures	-0.15	0.22	-0.66	-0.60-0.30	.511
Difference Pictures-Video	-0.12	0.24	-0.51	-0.60-0.36	.612
Difference Pictures-Real-time (physical)	-1.22	0.40	-3.06	-2.01--0.42	.004
Difference Pictures-Doll play	-				
Difference Pictures-Real-time (computerized)	-0.25	0.35	-0.71	-0.95-0.45	.483
Difference Audio-AudioPictures	0.19	0.14	1.41	-0.08-0.47	.166

2

Table S20 *Coefficients for the univariate moderator analysis with type of stimulus presentation of non-referred samples with extreme aggression scores (continued)*

	ES	SE	t	95% CI	p
Difference Audio-Video	0.22	0.16	1.37	-0.10-0.54	.177
Difference Audio-Real-time (physical)	-0.87	0.36	-2.45	-1.59--0.16	.018
Difference Audio-Doll play	-				
Difference Audio-Real-time (computerized)	0.10	0.30	0.32	-0.51-0.70	.751
Difference AudioPictures-Video	0.03	0.18	0.15	-0.33-0.39	.883
Difference AudioPictures-Real-time (physical)	-1.07	0.37	-2.93	-1.80--0.34	.005
Difference AudioPictures-Doll play	-				
Difference AudioPictures-Real-time (computerized)	-0.10	0.31	-0.32	-0.72-0.53	.754
Difference Video-Real-time (physical)	-1.09	0.37	-2.92	-1.85--0.34	.005
Difference Video-Doll play	-				
Difference Video-Real-time (computerized)	-0.12	0.32	-0.39	-0.77-0.52	.700
Difference Real-time (physical)-Doll play	-				
Difference Real-time (physical)-Real-time (computerized)	0.97	0.45	2.14	0.06-1.88	.037
Difference Doll-play-Real-time (computerized)	-				

Table S21 *All the Q_M statistics for the moderator analyses, and the Q_E statistics for the test of residual heterogeneity for the child characteristics in the subset analyses of clinically-referred samples*

Moderator	QM (df)	p	QE (df)	p
Sociometric status	-			
AGG function	0.17 (1)	.681	77.87 (26)	< .001
% ADHD	0.02 (1)	.890	42.67 (13)	< .001
Age group	0.74 (3)	.542	79.65 (24)	< .001
Gender	2.31 (2)	.120	66.28 (25)	< .001
AGG form	-			
Intelligence	0.17 (1)	.685	79.88 (26)	< .001
% ODD/CD	0.10 (1)	.759	61.29 (20)	< .001

Table S22 *All the Q_M statistics for the moderator analyses, and the Q_E statistics for the test of residual heterogeneity for the methodological characteristics in the subset analyses of clinically-referred samples*

Moderator	QM (df)	p	QE (df)	p
Type of stimulus presentation	0.24 (3)	.865	60.29 (22)	< .001
Provocateur's status	0.11 (3)	.952	65.13 (21)	< .001
Participant's perspective	0.02 (1)	.878	55.92 (18)	< .001
Type of context (provocation)	0.10 (1)	.756	79.74 (25)	< .001
Type of context (peer-entry)	0.36 (1)	.553	77.24 (25)	< .001
Type of context (expectations)	0.14 (1)	.716	71.33 (25)	< .001
Type of context (failure)	0.18 (1)	.674	71.06 (25)	< .001
Type of context (adult)	-			
Number of presented social situations	0.28 (1)	.601	78.93 (26)	< .001
Type of aggression informant	1.58 (3)	.219	51.99 (24)	< .001
Type of aggression measure	1.46 (1)	.238	78.06 (26)	< .001
Ambiguity of social situations	0.38 (1)	.544	74.26 (26)	< .001
Proportion of peers in social situations	-			
Type of setting	-			
Type of response format	0.32 (2)	.730	60.89 (25)	< .001
Type of HIA scoring	0.11 (1)	.744	70.88 (24)	< .001
Reliability of HIA measure	0.22 (1)	.688	6.37 (2)	.041

Table S23 *Moderators of Effect Size by Severity Classification*

	AGG Severity															
	All				Non-referred general				Non-referred extremes				Clinically-referred			
Characteristic and Level	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d
Child characteristics																
Sociometric status																
Aggressive	97	202	28002	.30	58	124	22269	.28	28	50	4285	.31	14	28	1448	.43
Aggressive-Rejected	15	17	1270	.61	1	2	80	.51	14	15	1190	.62				
AGG function																
General	98	193	22502	.33	50	111	17259	.26	38	55	3877	.41	13	27	1366	.42
Reactive	18	26	6770	.36	11	15	5147	.40	7	10	1541	.27	1	1	82	.62
% ADHD	7	22	919	$\beta0$ = .39 $\beta1$ = -.00 t = -.05 p = .958					2	7	249	$\beta0$ = -.62 $\beta1$ = .01 t = 1.86 p = .122	5	15	670	$\beta0$ = .53 $\beta1$ = .00 t = .14 p = .890
Age group																
3-6	14	20	2856	.23	11	16	2371	.22	2	3	365	.26	1	1	120	.34
6-12	28	69	5882	.28	13	34	3680	.17	10	19	1564	.30	6	16	638	.62
9-12	52	102	12936	.37	26	63	9780	.30	22	30	2718	.48	6	9	438	.35
>12	19	28	7713	.39	10	13	6636	.43	8	13	805	.38	2	2	272	.12
Gender																
Boys	34	56	3923	.38	10	17	1535	.25	17	20	1507	.50	9	19	881	.34
Both	74	148	24819	.32	48	101	20535	.22	22	40	3727	.34	5	7	557	.53
Girls	7	15	736	.31	3	8	397	.13	3	5	239	.30	1	2	100	.93

Table S23 *Moderators of Effect Size by Severity Classification (continued)*

	AGG Severity															
	All				Non-referred general				Non-referred extremes				Clinically-referred			
Characteristic and Level	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d
AGG form																
Overt	4	8	486	.37	3	6	283	.39	1	2	203	.10				
Physical	12	23	4883	.27	10	20	4609	.24	3	3	274	.63				
Indirect	1	3	193	.45	1	2	114	.37	1	1	79	.75				
Relational	12	22	4872	.31	10	20	4594	.29	2	2	278	.06				
Physical & Relational	4	5	713	.10	2	2	291	.08	2	3	422	.09				
Intelligence																
Controlled for	7	20	1094	.33					1	2	248	-.14	7	18	846	.48
Not controlled for	104	199	28191	.37	59	126	22467	.28	40	63	5122	.41	7	10	602	.38
% ODD/CD	12	28	1171	$\beta0$ = .68 $\beta1$ = -.00 t = -.33 p = .741					2	6	146	$\beta0$ = .77 $\beta1$ = -.01 t = -1.24 p = .284	10	22	1025	$\beta0$ = .75 $\beta1$ = -.00 t = -.31 p = .759
Methodological characteristics																
Type of stimulus																
Self-reading	17	30	6363	.44	12	22	5758	.38	5	8	605	.64				
Pictures	6	7	1340	.25	3	3	501	.28	2	3	775	.11	1	1	64	.37
Audio	42	82	10152	.36	19	50	7809	.24	17	21	1795	.45	7	11	548	.58
AudioPictures	14	32	2756	.27	8	15	2027	.27	6	15	689	.26	1	2	40	.17

Table S23 *Moderators of Effect Size by Severity Classification (continued)*

	AGG Severity															
	All				Non-referred general				Non-referred extremes				Clinically-referred			
Characteristic and Level	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d
Video	19	35	3674	.23	11	16	2447	.20	5	7	595	.23	4	12	632	.39
Real-time (physical)	2	3	57	1.33					2	3	57	1.33				
Doll-play	1	3	98	.27	1	3	98	.27								
Real-time (computerized)	1	2	75	.36					1	2	75	.36				
Unclear	10	25			5	17			3	6			2	2		
Provocateur's status																
Don't know eachother	31	54	6303	.29	15	24	4009	.25	13	18	1803	.33	4	12	491	.39
From neighborhood/ school	37	76	8369	.35	18	46	6427	.29	17	25	1722	.41	3	5	220	.51
Classmate	19	31	4214	.41	9	13	2982	.26	7	11	984	.58	3	7	248	.62
Friend	7	10	1396	.25	5	8	1143	.26	1	1	112	-.13	1	1	141	.35
Enemy	3	4	724	.26	3	4	724	.23								
Unclear	22	44			15	31			4	10			3	3		
Participant's perspective																
First person	62	125	16346	.36	31	77	13235	.29	28	41	2793	.46	5	7	318	.45

Table S23 *Moderators of Effect Size by Severity Classification (continued)*

	AGG Severity															
	All				Non-referred general				Non-referred extremes				Clinically-referred			
Characteristic and Level	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d
Non-first person: Imagine	20	43	4868	.20	10	16	2438	.14	7	14	1300	.22	4	13	1130	.39
Non-first person: Observer	8	10	1258	.29	6	8	905	.30	2	2	353	.21				
Unclear	21	41			12	25			4	8			5	8		
Type of context																
Provocation	102	173	27004	.33	53	95	20382	.29	40	59	5347	.38	12	19	1275	.43
Peer-entry	58	109	15073	.31	38	74	12104	.27	15	26	2432	.31	6	9	537	.51
Expectation	9	13	1429	.28	7	9	1123	.30	1	1	51	.26	1	3	255	.38
Failure	3	5	535	.36	1	1	60	.93	1	1	220	.43	1	3	255	.37
Adult	8	13	4408	.42	7	11	4286	.42	1	2	122	.43				
Unclear	2	3							1	2			1	1		
Type of AGG informant																
Parent	15	27	3915	.29	9	14	3152	.23	6	9	663	.25	1	4	100	.72
Teacher	39	63	7829	.32	22	38	5119	.29	14	20	2389	.34	4	5	321	.51
Peer	20	47	5399	.29	15	38	4839	.24	5	9	560	.50				
Observer	2	2	133	.31	2	2	133	.31								
>1Informant	25	36	4417	.32	5	6	2090	.13	15	17	1664	.48	5	13	663	.15

2

Table S23 *Moderators of Effect Size by Severity Classification (continued)*

	AGG Severity															
	All				Non-referred general				Non-referred extremes				Clinically-referred			
Characteristic and Level	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d
Self	26	38	12636	.37	20	28	11952	.36	6	10	684	.37				
Staff-member	5	6	494	.74									5	6	494	.63
Type of AGG measure																
Checklist/ rating-scale	75	151	22581	.36	45	89	18320	.33	21	39	3101	.31	11	23	1160	.51
Observation	2	2	133	.31	2	2	133	.31								
Sociometric nomination	17	40	3741	.23	12	31	3181	.16	5	9	560	.50				
>2	21	26	2830	.30	3	4	833	-.07	15	17	1709	.51	3	5	288	.17
Nr. of presented social situations																
<6	64	123	16976	.34	30	62	13518	.25	24	38	2537	.43	11	23	921	.41
>5	47	92	12142	.32	29	60	8782	.31	17	27	2833	.35	4	5	527	.50
Unclear	1	4			1	4										
All ambiguous social situations																
Yes	102	203	27430	.33	55	118	21234	.29	36	59	4931	.39	13	26	1265	.41
No	9	16	1960	.38	4	8	1233	.24	5	6	544	.43	1	2	183	.69
All peers in social situations																
Yes	98	197	22171	.31	49	109	15414	.25	40	61	5329	.38	13	27	1428	.44
No	10	17	4922	.40	9	15	4800	.40	1	2	122	.43				
Unclear	3	5			1	2			1	2			1	1		

Table S23 *Moderators of Effect Size by Severity Classification (continued)*

	AGG Severity															
	All				Non-referred general				Non-referred extremes				Clinically-referred			
Characteristic and Level	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d
Type of setting																
Individual	87	176	18628	.32	43	95	12934	.27	34	54	4224	.36	13	27	1470	.44
Group	17	24	6711	.32	12	17	5708	.25	5	7	1003	.50				
Unclear	7	19			4	14			2	4			1	1		
Type of response format																
Open	42	81	6945	.35	15	32	3426	.26	23	38	2895	.38	6	11	624	.42
MC/rating-scale	65	127	21734	.33	42	92	18666	.30	15	20	2173	.48	7	15	895	.37
Both	7	11	909	.28	2	2	375	.05	4	7	407	.21	2	2	127	.62
Type of HIA scoring																
Hostile	81	171	22524	.31	48	111	18101	.27	26	44	3540	.33	10	16	883	.49
Hostile-benign	26	42	6567	.41	11	15	4366	.34	13	17	1764	.51	3	10	437	.43
Unclear	3	6							2	4			1	2		
Reliability of HIA measure	48	97	16091	$\beta0$ = .26 $\beta1$ = .09 *t* = 2.59 *p* = .011	37	73	13310	$\beta0$ = .26 $\beta1$ = .12 *t* = 2.60 *p* = .012	10	20	2266	$\beta0$ = .24 $\beta1$ = .09 *t* = 1.61 *p* = .125	3	4	515	$\beta0$ = .36 $\beta1$ = .09 *t* = -.46 *p* = .688
Meta2002-Meta2017																
Meta2002	36	51	5791	.46	9	14	2057	.31	22	29	3027	.51	6	8	707	.56
Meta2017	75	168	23581	.28	50	112	20292	.28	19	36	2448	.29	9	20	841	.36

Table S23 *Moderators of Effect Size by Severity Classification (continued)*

	AGG Severity															
	All				Non-referred general				Non-referred extremes				Clinically-referred			
Characteristic and Level	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d	No. of studies	No. of ES	N	d
Total	111	219	29272	.33	59	126	22349	.27	41	65	5475	.39	14	28	1448	.48

Note. No. of studies = number of studies; No. of ES = number of effect sizes; N = number of participant's; *d* = effect size; Aggression Severity = severity of aggressive behavioral problems (non-referred general = non-referred sample scoring in the normal aggression range; non-referred extreme = non-referred sample scoring in the extreme aggression range; clinically-referred = clinically-referred aggressive sample); Sociometric status = sociometric status of the sample (aggressive = sample consisted of aggressive children; Aggressive-Rejected = sample consisted of aggressive-rejected children); Aggression function = function of aggressive behavior in sample (General = aggression was measured as a general construct; Reactive = aggression was measured specifically as reactive aggression); % ADHD = the proportion of ADHD diagnoses in the sample; Age group = Age group of the children in the sample (3-6 = between 3 and 6 years old; 6-12 = between 6 and 12 years old; 9-12 = between 9 and 12 years old; > 12 = older than 12 years); Gender = Gender of the sample (Boys; sample consisted of solely boys; Both = sample consisted of both boys and girls; Girls = sample consisted of solely girls); Aggression form = form of aggressive behavior in sample (Overt = overt aggression; Physical = physical aggression; Indirect = indirect aggression; Relational = relation aggression; Physical & Relational = physical and relational aggression); Intelligence = effect size was controlled for (verbal) intelligence (Controlled = effect size was controlled for intelligence; Not controlled = effect size not controlled for intelligence); % ODD/CD = the proportion of ODD/CD diagnoses in the sample; Type of stimulus presentation = type of stimulus materials used for presentation of social situations (Self-reading; hypothetical stories read by the participant; Pictures = hypothetical stories presented through pictures/cartoons/illustrations; Audio = hypothetical stories presented through audiotape/read by experimenter; AudioPictures = hypothetical stories presented through audiotape/read by experimenter and pictures/cartoons/illustration; Video = hypothetical stories presented through videotape; Real-time (physical) = real time interactions with a real peer; Doll-play = hypothetical stories presented through doll-play; Real-time (computerized) = real time interactions with a presumed peer); provocateur's status = the status of the provocateur in the presented social situation (Don't know each other = the provocateur in the presented social situation was an unknown peer; From neighborhood/school = the provocateur in the presented social situation was a boy/girl from the neighborhood or school; classmate = the provocateur in the presented social situation was a classmate; Friend = the provocateur in the presented social situation was a friend; Enemy = the provocateur in the presented social situation was an enemy); Participant's perspective = the perspective of the participant in the presented social situation (First person = participant's perspective was a first-person perspective (e.g., you are hit by a ball in the back) ; Non-first person; Imagine = participant's perspective was a non-first person perspective where participants were instructed being the protagonist (e.g., a child is hit by a ball in the back, imagine you are this child); Non-first person; Observer = participant's perspective was a non-first person perspective where participants were instructed to evaluate the protagonist's perspective (e.g., a child is hit by a ball in the back); Pro; Type of context = type of social problem situation (Provocation = presented social situations were provocative social situations (participant is victim of a provocative physical act); Peer-Entry = presented social situations were peer entry situations (e.g., participant is rejected to join a peer group, activity or play); Expectation = presented social situations were peer expectation situations (e.g., participant needs help from a peer); Failure = presented social situations were failure situations (e.g., participant's is being outperformed by a peer); Adult = presented social situations were conflicts with adults (e.g., being falsely accused by a parent or teacher); Number of presented social situations = the number of presented social situations (< 6 = less than 6 presented social situations; > 5 = more than 5 presented social situations); AGG informant = type of informant to assess aggression in participants (Parent = aggression in sample was measured through parents; Teacher = aggression in sample was measured through teachers; Peer = aggression in sample was measured through peers; Observer = aggression in sample was measured through observers; >1 Informant = aggression in sample was measured through more than one informant; Self = aggression in sample was measured through self-report; Staff-member = aggression in sample was measured through a staff-member); AGG measure = type of measure to assess aggression in participants (Checklist/Rating-scale = aggression in children was measured through a checklist/rating scale; Observation = aggression in children was measured through observation; Sociometric nomination = aggression in children was measured through nomination by peers; > 1 measure = aggression in children was measured by more than one type of measure); All ambiguous social situations= ambiguity of presented social situations (Yes = presented social situations were all ambiguous; No = presented social situations were a mixture of ambiguous and non-ambiguous social situations); All peers in social situations = proportion of peers in presented social situations (Yes = provocateurs in the presented social situations were solely peers; No = provocateurs in the presented social situations were a mixture of both peers and adults); Type of response format = type of response format HIA was measured through (Open = HIA was measured through an open response format; Mc-rating = HIA was measured through multiple choice/rating scale format; Both = HIA was measured through a composite of open- and multiple choice/rating scale format); Type of setting = the type of setting the assessment had been conducted in (Individual = assessment was conducted in an individual setting; Group = assessment was conducted in an group-based setting); Type of HIA scoring = type of scoring used to calculate HIA (Hostile = HIA was calculated as the number/proportion of hostile attributions; Hostile-Benign = HIA was calculated as the number of hostile minus benign attributions, from a rating scale ranging from hostile to benign or other method to calculate HIA); Reliability of HIA measure; standardized with mean Cronbach α of the HIA-scale; Meta2002-Meta2017 = effect size was derived from studies included in the original meta-analysis or the update meta-analysis (Meta2002 = effect size was derived from studies included the original meta-analysis (2002); Meta2017: effect size was derived from studies included the update meta-analysis (2017).

Table S24 *Overview of included studies and effect sizes*

Study	Country	Number of effect sizes	N
Dodge (1980) *	USA	1	90
Dodge & Frame (1982) *	USA	1	81
Steinberg & Dodge (1983) *	USA	2	39
Milich & Dodge (1984) *	USA	2	98
White (1984) *	USA	1	18
Dodge et al. (1986) *	USA	1	79
Dodge & Coie (1987) *	USA	2	80
Dodge & Somberg (1987) *	USA	1	65
Dodge & Tomlin (1987) *	Not reported	1	74
Fesbach (1989) *	USA	2	59
Guerra & Slaby (1989) *	USA	1	48
Sancillio et al. (1989) *	USA	1	74
Graham et al. (1992) *	USA	1	88
Quiggle et al. (1992) *	USA	1	220
Weiss et al. (1992) *	USA	1	248
Hart (1993) *	USA	1	74
Hudley & Graham (1993) *	USA	1	102
Graham & Hudley (1994) *	USA	1	52
Lochman & Dodge (1994) *	USA	4	296
Vlerick (1994) *	Belgium	1	48
Crick (1995) *	USA	2	239
Katsurada (1995) *	USA	2	53

Table S24 *Overview of included studies and effect sizes (continued)*

Study	Country	Number of effect sizes	N
Barrett et al. (1996) *	Australia	1	53
Bickett et al. (1996) *	USA	1	50
Crick & Dodge (1996) *	USA	2	590
Erdley & Asher (1996) *	USA	1	781
Waldman (1996) *	Canada	1	81
Cuperus et al. (1997) * same study as Matthys et al. (1999)	The Netherlands	1	141
Dodge et al. (1997) *	USA	2	487
Gibbins (1997) *	Canada	1	148
Hyatt (1998) *	USA	2	112
Williams (1998) *	USA	1	87
Brendgren et al. (1999)*	Canada	1	322
Matthys et al. (1999) same study as Cuperus et al. (1997)	The Netherlands	6	114
Webster-Straton (1999)	USA	1	120
Zelli et al. (1999)*	USA	3	279
Gomez & Gomez (2000)	Australia	1	84
Yoon et al. (2000)*	USA	2	153
De Castro et al. (2001)*	The Netherlands	1	82
Hubbard et al. (2001)	USA	1	66
Crick et al. (2002)	USA	3	127
Dodge et al. (2002)	USA	2	322
Muris et al. (2003)	The Netherlands	3	103
De Castro et al. (2003)	The Netherlands	2	57

Table S24 *Overview of included studies and effect sizes (continued)*

Study	Country	Number of effect sizes	N
Dodge et al. (2003)	USA	4	837
Frick et al. (2003)	USA	2	73
MacBrayer et al. (2003)	USA	4	100
Manel (2003)	Canada	2	40
Schultz & Shaw (2003)	USA	1	178
Heidgerken et al. (2004)	USA	2	239
Crain et al. (2005)	USA	2	125
Lavallee (2005)	USA	1	288
Burgess et al. (2006)	USA	4	161
Kempes et al. (2006)	USA	1	78
Reid et al. (2006)	Australia	1	192
Van Nieuwenhuijzen et al (2006)	The Netherlands	2	130
Werner et al. (2006)	USA	1	67
Bowker et al. (2007)	USA	2	385
Garner & Lemerise (2007)	USA	1	92
Halligan et al. (2007)	UK	2	134
Meeks-Gardner et al. (2007)	Jamaica	1	202
Peets et al. (2007)	Finland	6	137
Runions & Keating (2007)	Australia	9	193
Yeung & Leadbeater (2007)	Canada	2	140
Crozier et al. (2008)	USA	1	411
Gentile & Gentile (2008)	USA	10	411

Table S24 *Overview of included studies and effect sizes (continued)*

Study	Country	Number of effect sizes	N
Mikami et al. (2008)	USA	3	203
Nelson et al. (2008)	USA	8	353
Peets et al (2008)	Finland	3	209
Teisl & Cicchetti (2008)	UK	1	267
Arsenio et al. (2009)	USA	2	100
Ellis et al. (2009)	USA	1	83
Orue & Calvete (2009)	Spain	1	807
Stickle et al. (2009)	USA	1	150
Calvete & Orue (2010)	Spain	1	1371
Halligan & Philips (2010)	UK	1	910
Hyde et al. (2010)	USA	2	230
Meece & Mize (2010)	USA	1	128
Schultz et al. (2010)	USA	1	125
Freeman et al. (2011)	UK	2	144
Goldweber et al. (2011)	USA	1	429
Kupfersmidt et al. (2011)	USA	1	244
Mathieson et al. (2011)	USA	1	635
Van Nieuwenhuijzen et al (2011)	The Netherlands	2	64
Calvete & Orue (2012)	Spain	2	650
Chaux et al. (2012)	Colombia	1	1235
Helmsen et al. (2012)	Germany	1	193
Ogelman & Seven (2012)	Turkey	1	60

Table S24 *Overview of included studies and effect sizes (continued)*

Study	Country	Number of effect sizes	N
Werner (2012)	USA	8	91
Ziv (2012)	Israel	1	256
Choe et al. (2013)	USA	1	231
Goraya & Kazim (2013)	Pakistan	1	106
Stolz et al. (2013a)	The Netherlands	3	73
Stolz et al. (2013b)	The Netherlands	1	206
Cilessen et al. (2014)	The Netherlands	2	366
Gentile et al. (2014)	USA	2	2253
Laible & Murphy et al. (2014)	USA	1	148
Laible & McGinley et al. (2014)	USA	1	1013
Vassilopoulos et al. (2014)	Greece	1	52
Yaros et al. (2014)	USA	2	75
Choe et al. (2014)	USA	1	370
Dodge et al. (2015) same study as Di Giunta et al. (2015)	USA	2	1244
Di Giunta et al (2015) same study as Dodge et al. (2015)	Italy	2	541
Healy et al. (2015)	UK	3	98
Helseth et al. (2015)	USA	4	46
Vassilopoulos et al. (2015)	Greece	1	34
Williams et al. (2015)	USA	1	51
Yaros et al. (2016)	USA	2	122
Ziv et al. (2016)	Israel	2	218
Van Dijk et al. (2017)	The Netherlands	1	283

Table S24 *Overview of included studies and effect sizes (continued)*

Study	Country	Number of effect sizes	N
Nelson et al. (2017)	USA	4	222
Singh (2017)	India	2	126
Wang & Dix (2017)	USA	1	1082

Note. Studies with an asterisk (*) are studies included in the 2002 meta-analysis.

Figure S4 *AIC support for each model.*

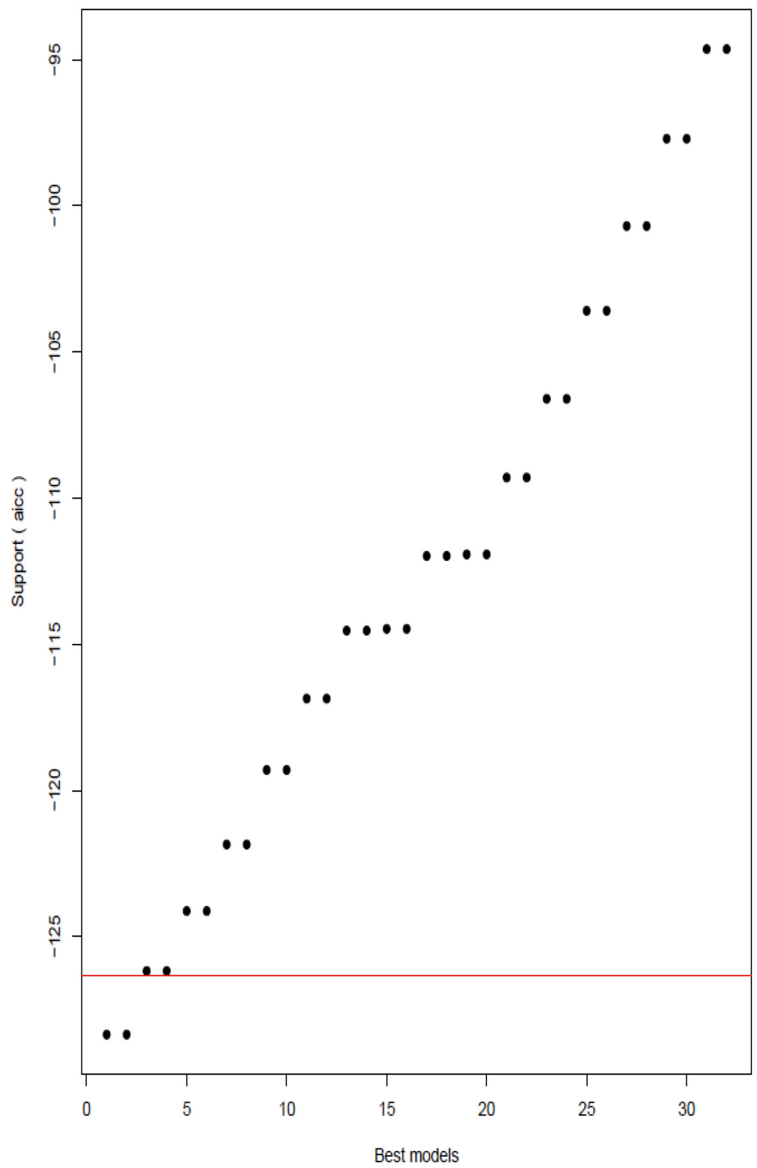
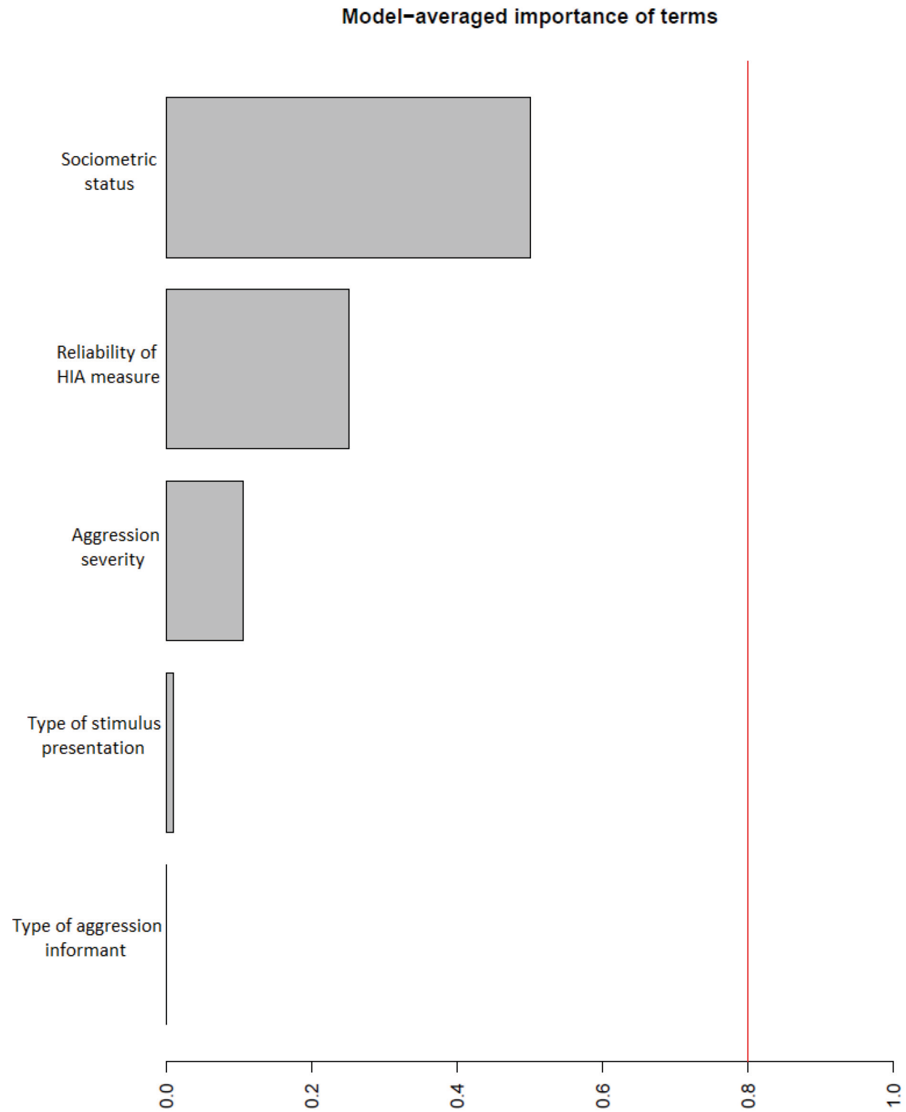


Figure S5 *Model-averaged importance of predictors*



CHAPTER 3

Interactive Virtual Reality Assessment of Aggressive Social Information Processing in Boys with Behavior Problems: A Pilot Study

This chapter is published as:
Verhoef, R. E. J., Van Dijk, A., Verhulp, E. E., & De Castro, B. O. (2021). Interactive virtual reality assessment of aggressive social information processing in boys with behaviour problems: A pilot study. Clinical Psychology & Psychotherapy, 28(3), 489-499. https://doi.org/10.1002/cpp.2620

The data and syntax that support the findings of this study are available through the Open Science Framework at https://doi.org/10.17605/OSF.IO/CS3AK

Author contributions:
RV conceptualized the study, AD, EV, and BC gave advice and feedback; RV and trained undergraduate students collected the data; RV analyzed the data and drafted the manuscript; AD, EV, and BC provided feedback on the analyses and manuscript.

ABSTRACT

Children's aggressive behavior is partly determined by how they process social information (e.g., making hostile interpretations, or aiming to seek revenge). Such aggressive social information processing (SIP) may be most evident if children are emotionally engaged in actual social interactions. Current methods to assess aggressive SIP, however, often ask children to reflect on hypothetical vignettes. This pilot study therefore examined a new method that actually involves children in emotionally engaging social interactions: interactive Virtual Reality (VR). We developed a virtual classroom where children could play games with virtual peers. A sample of boys (*N* = 32; ages 8-13) from regular and special education reported on their SIP in distinct VR contexts (i.e., neutral, instrumental gain, and provocation). They also completed a standard vignette-based assessment of SIP. Results demonstrated good convergent validity of interactive VR assessment of SIP, as indicated by significant moderate to large correlations of VR-assessed SIP with vignette-assessed SIP for all SIP variables except anger. Interactive VR showed improved measurement sensitivity (i.e., larger variances in SIP compared to vignettes) for aggressive responding, but not for other SIP variables. Discriminant validity (i.e., distinct SIP patterns across contexts) of interactive VR was supported for provocation contexts, but not for instrumental gain contexts. Last, children were more enthusiastic about the VR assessment compared to the vignette-based assessment. These findings suggest that interactive VR may be a promising tool, allowing for the assessment of children's aggressive SIP in standardized yet emotionally engaging social interactions.

Keywords: Social information processing, aggression, behavior problems, children, Virtual Reality, pilot study

INTERACTIVE VIRTUAL REALITY ASSESSMENT OF AGGRESSIVE SOCIAL INFORMATION PROCESSING IN BOYS WITH BEHAVIOR PROBLEMS: A PILOT STUDY

Children frequently encounter challenging social situations such as being laughed at, losing a game, or being excluded. How children mentally process such situations influences their subsequent behavior (Anderson & Bushman, 2002; Crick & Dodge, 1994; Lemerise & Arsenio, 2000). The Social Information Processing (SIP) model distinguishes several internal processing steps children engage in before responding to social events: (1) encoding and (2) interpreting social cues, (3) setting interactional goals, (4) generating and (5) evaluating responses, and (6) enacting a selected response (Crick & Dodge, 1994). Over the past decades, this SIP model has been shown to provide a convincing theoretical framework for the understanding, prevention, and treatment of aggressive behavior problems (for a review, see: De Castro & Van Dijk, 2017). Children's aggression has been shown to derive from deviations in each of these SIP steps, such as perceiving more threatening cues, attributing more hostile intentions to others, pursuing revenge- or instrumental goals more often, generating more aggressive responses, and evaluating aggressive responses more positively (for reviews, see: De Castro & Van Dijk, 2017; Dodge, 2011). Moreover, intervention studies have shown that changing children's SIP can reduce aggression (Lochman et al., 2017; Lochman et al., 2019; Maixner-Schindel & Shechtman, 2021; Wilson & Lipsey, 2006). Given the important role of SIP underlying children's aggression, valid assessment of SIP is essential. The present pilot study examines a new method to assess children's SIP in an ecologically valid manner: interactive Virtual Reality (VR).

Current methods to assess aggressive SIP have important shortcomings. Until now, most studies have assessed children's SIP using hypothetical stories, where children are asked to imagine that a hypothetical social event is actually happening to them and to reflect on their SIP in response to this hypothetical event. Using such hypothetical vignettes limits the ecological validity of SIP assessment, especially because many children may only show aggressive SIP when they are emotionally engaged in actual social events (Anderson & Bushman, 2002). Strong emotions such as anger, embarrassment or excitement may trigger aggressive cognitions that would not be triggered when children feel calm (Lemerise & Arsenio, 2000). For example, children may only attribute hostile intent to others when they feel frustrated, or may only pursue instrumental goals when they strongly desire an object. The relevance of assessing children's aggressive SIP in emotionally engaging social situations is emphasized by empirical work showing that inducing negative emotions elicits more aggressive SIP and behavior (e.g., Caporaso & Marcovitch, 2021; De Castro et al., 2003; Reijntjes et al., 2011). Thus, an ecologically valid assessment of children's aggressive SIP requires the use of emotional engaging social situations.

A few earlier attempts to promote ecological validity have used staged real-time conflicts with (alleged) peers or child actors (Hubbard et al., 2001; Kempes et al., 2008;

Steinberg & Dodge, 1983; Van Dijk et al., 2019). A meta-analysis has demonstrated such studies found stronger associations between hostile intent attribution and aggression (d = 1.33) than studies using vignettes (d = 0.23 to 0.44; Chapter 2 of this dissertation). This suggests that ecologically valid methods may improve the assessment of children's aggressive SIP. Research using of staged conflicts, however, can be ethically challenging and difficult to standardize. First, staged conflicts are prone to escalation, complicating adherence to ethical guidelines. Second, when staging real-time conflicts between children it is difficult to ensure that child actors behave identically with each participant, limiting standardization. As such, there is a need for innovative methods to assess children's aggressive SIP that can combine highly emotionally engaging, realistic social interactions with adequate standardization and adherence to ethical guidelines.

Interactive VR may provide a viable solution for limitations encountered by previous research. VR technology is already used for the assessment and treatment of various forms of psychopathology in adults (for reviews, see: Carl et al., 2019; Emmelkamp & Meyerbröker, 2021; Freeman et al., 2017). For children, though, research using VR is relatively limited. VR has been utilized for the treatment of autism and attention deficit-hyperactivity disorder, and for teaching emotion regulation skills to prevent risk taking behavior in adolescents (Hadley et al., 2019; Mesa-Gresa et al., 2018; Shema-Shiratzky et al., 2018). In addition, one study assessing SIP in children with autism spectrum disorder has used non-interactive VR, in which children navigated an avatar through a simulated 3D environment by selecting response options through a computer menu (Russo-Ponsaran et al., 2018). This method, however, may be less suitable to assess children's aggressive SIP because not being able to respond through actual behavior in the VR-environment may lower children's emotional engagement. To our knowledge, interactive VR has not been previously used to assess children's aggressive SIP.

Interactive VR may have several benefits for the assessment of children's aggressive SIP. First, it enhances ecological validity by immersing children in an emotionally engaging environment where they can interact with, and possibly aggress against, virtual peers. Second, interactive VR allows for rigorous experimental control. By controlling the course and content of social events in VR, researchers can standardize scenarios between participants and adhere to ethical guidelines. Third, VR can be flexibly used to present children with various contexts, enabling researchers to assess individual differences in aggressive behavior and associated SIP patterns. For the present study, we developed an interactive VR environment, aiming to optimize these benefits to provide an ecological valid assessment of children's aggressive SIP. In this first pilot study, we targeted school-aged boys with different levels of behavior problems to maximize potential variance stemming from differences in aggression. As such, we could examine whether our interactive VR would be a valid assessment method for boys across the whole spectrum from non-aggressive children to children with severe aggressive behavior problems.

First, to promote ecologically validity, we designed the VR environment to be interactive and realistic. Participants are visually completely immersed in a virtual classroom that, just like the real world, responds naturally to each single motion (Figure 1). Participants can freely walk around in the virtual classroom (in reality, they walk around in a demarcated space in an empty room at their school with the VR glasses on). They interact with virtual peers in similar fashion as in real life: through verbal and physical behaviors. They use controllers that mimic their hands in virtual reality, allowing them to use objects and play games. The virtual peers are manually controlled by the experimenter using standardized speech options and physical actions. This VR environment allows for various engaging interactions to assess children's SIP, such as building a 2-meter-high block tower that is being bumped over by a virtual peer (i.e., an ambiguous provocation).

A particularly sensitive aspect of ecological validity is the assessment of participants' aggressive behavior in the VR environment. Aggressive behavior is defined as "any behavior directed towards another individual with the intent to cause harm" (Anderson & Bushman, 2002, p. 27). Thus, to ensure ecological validity it is important that children believe that their aggressive behavior in VR does actual harm to the virtual peer. This is not self-evident, as many children play digital games where they use violence against characters they know do not exist. Therefore, we presented our virtual classroom as an actual classroom where participants allegedly met with real children from other schools who also participated in our study and were simultaneously logged on to the VR environment.

Second, to promote experimental control, we scripted all social interactions between the participant and virtual peers (Figure 2). The responses of virtual peers were controlled by an experimenter, using default movements and pre-recorded verbal responses. These standardized responses were designed to respond naturally to participants' behavior, thus facilitating participants' immersion in the social interactions.

Third, to assess individual differences in children's aggressive SIP, we designed different social scenarios to assess both reactive and proactive aggression (Dodge, 1991). Reactive aggression is defined as an impulsive aggressive response to perceived threat or provocation, whereas proactive aggression is defined as planned aggressive behavior aimed at obtaining a desired outcome (Dodge, 1991; Hubbard et al., 2010; Van Dijk et al., 2021). This suggests that different social contexts are needed to assess these types of aggression and their underlying SIP patterns, such as peer provocation for reactive SIP and the opportunity to obtain instrumental gain for proactive SIP. Yet, most previous studies on children's SIP only used provocation scenarios, perhaps because vignettes seem less suitable to provide children with an 'opportunity' to aggress than to present them with a provocative event (for reviews, see: Hubbard et al., 2010; Martinelli et al., 2018). As provocation scenarios seem theoretically more relevant to assess SIP underlying children's reactive aggression such as hostile intent attributions and revenge goals, earlier studies may have missed out on SIP underlying proactive aggression, such as instrumental goals (Hubbard et al., 2010). We therefore designed both provocation and instrumental gain scenarios in our VR, which we based on taxonomies of problematic situations for children with aggressive behavior problems (Dodge et al., 1985; Matthys et al., 2001). We used two scenarios to cover the context of provocation: being refused to join (i.e., social provocation) and participants' game being ruined (i.e., object provocation). Similarly, we used two scenarios to cover the context of instrumental gain: having the opportunity to steal (i.e., object acquisition) and having the opportunity to cheat (i.e., competition).

In sum, we designed a new interactive VR environment to assess children's aggressive SIP, aiming to accommodate for shortcomings of current assessment methods by immersing children in standardized, emotionally engaging social interactions. Therefore, in line with methodological guidelines (Boateng et al., 2018), we conducted a first-phase pilot study to test whether our VR measure demonstrates sufficient convergent validity, measurement sensitivity, and discriminant validity. We included a sample of boys recruited from both regular and special education to maximize variance in aggressive SIP. We also administered a traditional vignette-based assessment (De Castro et al., 2005). For both VR and vignettes, we assessed children's anger, intent attributions, goals, and responses. First, regarding convergent validity, we expected that the SIP assessment in VR would be positively associated with SIP assessed using vignettes. Second, regarding measurement sensitivity, we expected that the VR assessment would yield larger variances in children's SIP than the vignette assessment. Third, regarding discriminant validity, we expected that the provocation scenarios would elicit more anger, hostile intent attributions, and revenge goals than instrumental gain and neutral scenarios, and more aggressive responses than neutral scenarios. We further expected that the instrumental gain scenarios would elicit more instrumental goals than the provocation and neutral scenarios, and more aggressive responses than neutral scenarios. Last, in support of potential utility of interactive VR for assessment and intervention in clinical practice, we expected that children would be more enthusiastic about participating in the VR than the vignette assessment.

METHOD

Participants

Thirty-two boys ages 8 to 13 years (*M* = 10.34; *SD* = 1.36) were recruited from primary schools in the Netherlands. Children were from ethnically diverse backgrounds (34.4% Turkish/Moroccan, 15.6% Surinamese/Antillean, 50% Caucasian). To maximize variance in children's aggressive SIP, we created our sample by including boys from special education selected on aggressive behavior problems by their teacher (*n* = 14) and a random selection of boys from regular education (*n* = 18). In special education, children were excluded if they had an Autism Spectrum Disorder reported in their casefiles, had a clinical score on the teacher-rated Social Responsiveness Scale (Dutch translation; Roeyers et al., 2011), or had an IQ below 80 reported in their casefiles. Parents gave written consent for their child's participation in the study. All children who received consent participated in this study (*N* = 32). This pilot study was approved by the Medical Ethics Committee of University Medical Center Utrecht.

Procedure

Participants were individually tested in a silent room at their school. We informed participants that the study is about peer interactions and that they would listen to stories and would enter a virtual classroom where they could interact with peers from other schools. They completed the VR- and vignette-based SIP assessments on two different days with approximately a week in between (the order was counterbalanced across participants). We emphasized that no wrong answers could be given and assured participants of the confidentiality of their responses. Each assessment lasted approximately 45 minutes. The VR assessment was conducted by the first author (controlling the VR) and a trained graduate student (noting participants' responses). The vignette assessment was conducted by trained graduate students. At the end of each assessment and when they had completed both assessments, participants rated how enthusiastic they were about the VR and vignettes. We debriefed participants on the second assessment day, explaining that we wanted to examine how children interact with real peers rather than computer-controlled characters. Last, they received a small gift for their participation.

Interactive Virtual Reality

Development

We based the content of the VR scenarios and formulation of the SIP questions on the extant literature on SIP assessment (Chapter 2 of this dissertation). Both were discussed in multiple feedback rounds with colleagues knowledgeable in SIP research. We conducted early try-outs of the VR with 18 children (of the same age range as our participants) to ensure the intentions of the virtual peers were perceived as ambiguous, and that the games were not too difficult but challenging enough to evoke sufficient engagement.

VR Environment

The VR environment consisted of a virtual classroom, built by CleVR (Figure 1). We introduced the classroom to participants as an actual classroom with specific behavior rules (i.e., having respect for other children, being friendly to other children). Participants wore a HTC Vive with a combined resolution of 2160×1200, with an approximate diagonal field of view of 110°, and support for 6DOF tracking. They could walk around freely (in a 4×4 space), talk with virtual peers, and play games with them. Virtual peers were boys from the same age-range and average height for their age. Each scenario included virtual peers that differed slightly in haircut, facial features, and print of their clothing. The verbal responses of virtual peers were pre-recorded by 12 children from theatre schools. The experimenter controlled these pre-recorded responses, which included standardized responses used for all participants (Figure 2) and general statements allowing for a natural response to the participant (e.g., "I am 10. What's your age?"). During the VR scenarios, virtual peers' emotional expressions were neutral, however when the participant aggressed, the virtual peers' expression changed to upset.

The VR environment included two games: (1) building a tower of blocks as high as possible, and (2) using five balls to hit as many cans from a table as possible. We integrated high scores and bonuses in both games to increase participants' emotional engagement and to provide experimental control over gains and losses. The assessment scenarios were designed around these games, allowing participants to engage in aggression aimed at the virtual peer (e.g., hitting, name calling) as well as at the virtual peer's property (e.g., knocking over his tower). The instructions, game rules and score count were provided on a digital whiteboard and verbally explained using standard instructions recorded by a female experimenter (Figure 2).

Figure 1a & 1b *Virtual Classroom with the Tower Game or with the Cans Game*

VR Scenarios

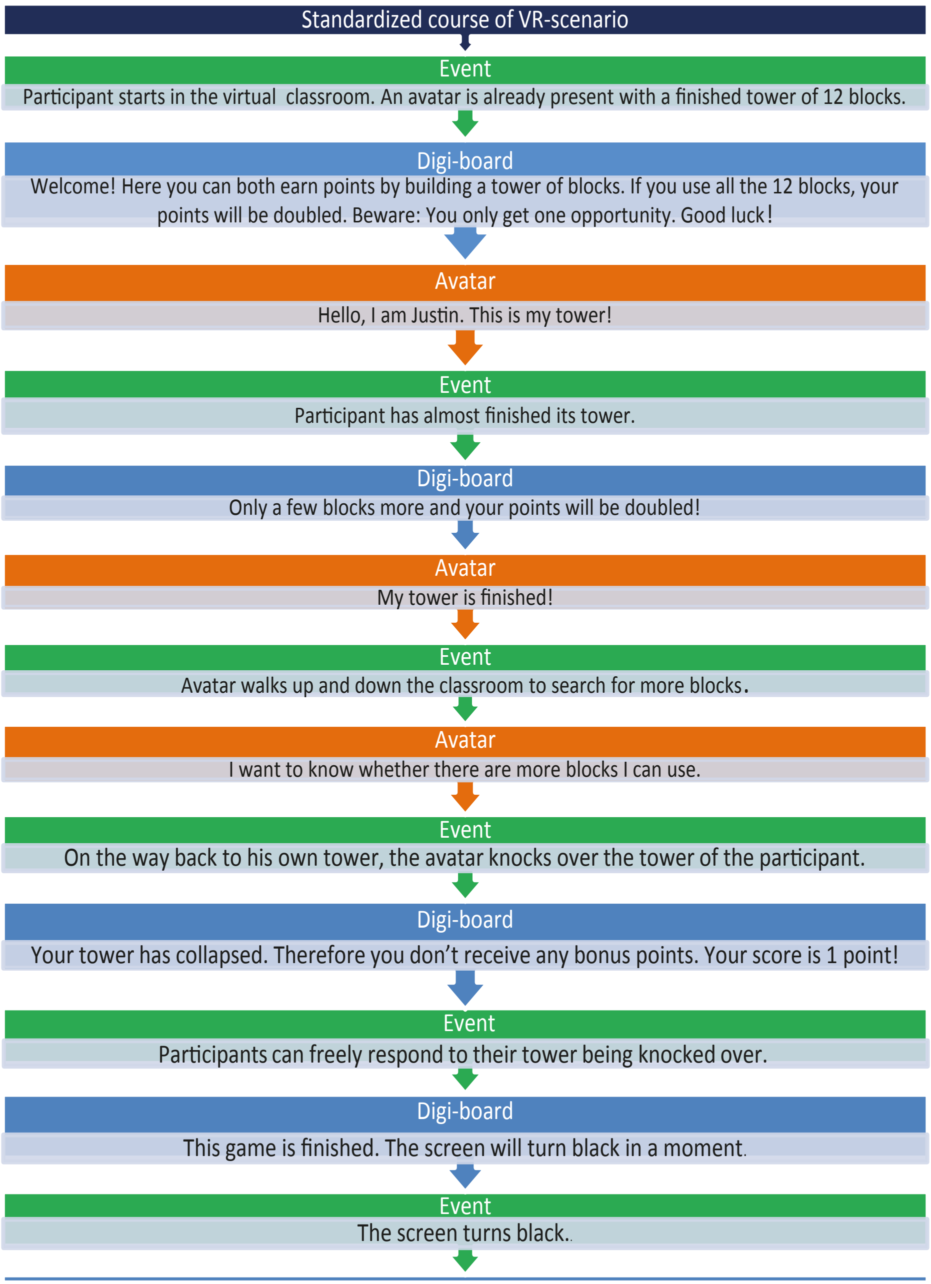
Each scenario followed a standardized course, designed around the game played by the participant and virtual child (Figure 2). Each scenario included specific social events presented in fixed order. The exact timing of events depended on the individual participant's behavior and progress in the game. At the end of each scenario, the experimenter presented the specific contextual event: provocation (e.g., participants' game being ruined by the virtual character) or instrumental gain (e.g., the virtual character is winning the game).

We developed six VR scenarios: one practice, one neutral, two instrumental gain, and two provocation scenarios. The practice scenario served to familiarize participants with the VR environment by letting them play the game without any virtual characters. The neutral scenario served to assess participants' SIP in a situation with no engaging events (i.e., during small talk with a virtual peer). The two instrumental gain scenarios assessed SIP in response to object acquisition (i.e., participants could choose to steal a block or ball from the virtual peer to obtain additional points) and competition (i.e., participant could win the game by sabotaging the virtual peer's game). The two provocation scenarios assess SIP in response to social provocation (i.e., participants were refused to join a game by two virtual peers) and object provocation (i.e., participants'

game was ruined by the virtual peer). The six VR scenarios were presented in fixed order: practice scenario, neutral scenario, object acquisition, competition, social provocation, and object provocation. We expected the provocation scenarios to elicit the strongest emotions, and therefore presented them last to prevent carry-over effects.

Participants completed all six scenarios for the same game (i.e., the tower or cans game), which was randomly assigned. As such, differences in SIP between scenarios reflected scenario effects rather than game effects. A description of each scenario per type of game is provided in the Supplementary Materials (Table S1).

Figure 2 *Example of a Standardized Course in a VR-scenario for Object Provocation (Tower Game)*



Measures

SIP Assessment using VR

We assessed participants' aggressive responding by observing their actual behavioral response in the VR scenario. We assessed their SIP by asking several questions directly after each scenario. During these 1-minute assessments between scenarios, participants kept the VR-glasses on. We assessed participants' anger, intent attributions, interaction goals, aggressive response generation, and evaluation of their actual aggressive response in VR. For the present study, however, we excluded aggressive response generation because too few children mentioned additional aggressive responses in addition to their actual response, and we excluded evaluation of aggression because we could only assess this variable if children responded aggressively, resulting in too few observations to analyze.

3

Anger. Participants' anger in VR was assessed with one item: "The other boy did [behavior other boy]. How angry did this make you feel, on a scale from 1-10?" Anger scores were averaged, creating separate scores for provocation (2 items; r = .78), instrumental gain (2 items; r = .53), and neutral contexts (1 item).

Hostile Intent Attribution. Participants' intent attributions were assessed using two items following each VR scenario: "The other boy did [behavior other boy]. To what extent did he try to be mean, on a scale from 1-10?" and "To what extent did he try to hinder you, on a scale from 1-10?". Correlations between the two items within each VR scenario were acceptable (M = .73, Mdn = .77, range = .34-.91) and therefore averaged to create a single hostile intent attribution score for each scenario. Next, hostile intent attribution scores were averaged, creating separate scores for provocation (2 items; r = .53), instrumental gain (2 items; r = .64), and neutral contexts (1 item).

Goals. Participants' goals were assessed using one open-ended question following each VR scenario: "When the other boy did [behavior other boy], you did [behavior participant]. What was the reason you did this?" During the assessment, a trained graduate student directly wrote down all participants' answers. Afterwards, we coded these responses in line with previous research (De Castro et al., 2012) into the categories *no goals* (e.g., "I don't know," "I had no goal"), *revenge goals* (e.g., "to retaliate," "because I was angry"), *instrumental goals* (e.g., "to win the game," "to show him who's boss"), and *goals underlying non-aggressive behavior* (e.g., "to become friends," "to avoid problems"). A second rater also coded 50% of the transcriptions. Inter-rater reliability was good, with κ ranging from 0.88-1.00 (M = .93, Mdn = .88). Scores for revenge goals were created by assigning 1 to *revenge goals* codes and 0 to other codes, and then averaged to create separate scores for provocation (2 items; τ = .53), instrumental gain (2 items; τ = .56) and neutral contexts (1 item). Similarly, scores for instrumental goals were created by assigning 1 to *instrumental goals* codes and 0 to other codes, and averaged to create separate scores for provocation (2 items; τ = .47), instrumental gain (2 items; τ = .68), and neutral contexts (1 item).

Behavioral Responses. Behavioral responses in VR were assessed through observation of participants' behavior during each scenario. A trained graduate student directly wrote down participants' behavior. We coded this behavior afterwards using standard procedures (De Castro et al., 2005) into the categories 0 for *non-aggressive behavior* (e.g., prosocial, avoidance), 1 for *mild aggressive behavior* (e.g., coercion, verbal aggression), and 2 for *severe aggressive behavior* (e.g., physical aggression, destructive aggression). A second rater also coded 50% of the behavioral descriptions. Inter-rater reliability was good, with κ ranging from 0.87-1.00 (*M* = .97, *Mdn* = 1.00). Aggressive response scores were averaged, creating separate scores for provocation (2 items; *r* = .55), instrumental gain (2 items; *r* = .92), and neutral contexts (1 item).

SIP Assessment using Vignettes

We used a validated vignette measure to assess participants' SIP (De Castro et al., 2005). This measure—as most standard SIP measures—only includes provocation stories. Participants were presented with 5 audiotaped vignettes describing ambiguous peer provocations, such as losing a computer game through fault of a peer (De Castro et al., 2005). We informed participants that they would listen to vignettes about daily social events and asked them to imagine each story was actually happening to them. Participants first practiced with one vignette, so that the experimenter could check whether they understood the procedure (all participants did). Next, following each vignette, we assessed children's SIP using the same questions and coding schemes as used for the VR assessment, except for two minor modifications. First, we formulated SIP vignette questions as hypothetical ("What would you…?") instead of actual ("What did you…?"). Second, we assessed aggressive responding using an open-ended question (i.e., "What would you do if [peer provocation]?") instead of observation. Hostile intent attribution items were correlated within each vignette and therefore averaged (*M* = .76, *Mdn* = 80; range = .66-.83). Inter-rater reliability (κ) was based on 50% of transcriptions and was acceptable for both interaction goals (*M* = .80, *Mdn* = .77, range = .66-1.00) and aggressive responding (*M* = .79, *Mdn* = .78, range = .77-.86). We averaged participants' responses across the five vignettes, creating single scores for anger (α = .61), hostile intent attribution (α = .72), revenge goals (α = .78), and aggressive responding (α = .74).

Enthusiasm about the VR and Vignettes Assessment

We assessed children's enthusiasm about the VR and vignette assessments using five items at the end of each assessment (e.g., "How much did you like the VR/vignettes?," "How much would you like to do the VR/vignettes again?"). Children responded on a rating scale from 1 (*not at all*) to 10 (*very much*). We averaged across the five items to create enthusiasm scores for both VR (α = .87) and vignettes (α = .86). To capture children's explicit comparison, we also asked them to rate how much they liked the VR and vignette assessment on a scale from 1-10 after they had completed both assessments.

Statistical Analyses

We had four main goals. First, we examined the convergent validity of SIP assessment in VR by calculating correlations between VR- and vignette-assessed SIP variables, using Pearson's *r* and Kendalls τ. In these analyses, we only included the VR provocation scenarios to relate to the vignette scores, because the vignettes only covered the domain of provocation. We analyzed correlations for anger, hostile intent attribution, revenge goals, and aggressive responding (i.e., the SIP variables relevant to provocation contexts). Second, we examined measurement sensitivity by comparing the variances of VR- and vignette-assessed SIP variables (i.e., anger, hostile intent attribution, revenge goals, and aggressive responding). To this end, we used the Pittman-Morgan test based on Spearman's rank correlations (McCulloch, 1987). Third, to test the discriminant validity of SIP assessment in VR, we conducted planned comparisons of participants' SIP between provocation, instrumental gain, and neutral contexts, using paired *t*-tests. Fourth, we examined whether children were more enthusiastic about our SIP assessment in VR than with vignettes, using paired *t*-tests. Given the small sample size and non-normal distribution of the variables, we conducted these analyses using bootstrapped bias-corrected accelerated (BCa) 95% confidence intervals (CI) based on 5000 resamples.

RESULTS

Convergent Validity: Association between SIP in VR versus Vignettes

Table 1 presents the zero-order correlations between VR- and vignette-assessed SIP variables. Confidence intervals that exclude the value of 0 signify that the correlation was significant. Supporting convergent validity, we found small to high significant correlations between VR- and vignette-assessed anger (*r* = .37, BCa 95% CI: .02-.65), hostile intent attribution (*r* = .56, BCa 95% CI: .28-.79), revenge goals (τ = .67, BCa 95% CI: .46-.84), and aggressive responding (*r* = .73, BCa 95% CI: .50-.89). These results indicate that children's SIP assessed with interactive VR corresponds with their SIP assessed through a traditional validated vignette-based measure.

Table 1 *Bivariate Correlations between SIP Variables in Provocation VR-Contexts and Vignettes*

	Range	M (SD)	2.	3.	4.	5.	6.	7.	8.
VR: Anger	2.50-10.00	7.09 (2.60)	73*	.48*	.49*	.37*	.49*	.53*	.58*
VR: Hostile intent attribution	1.25-10.00	6.76 (2.54)		.41*	.52*	.13	.56*	.38*	.47*
VR: Revenge goals	0.00-1.00	0.38 (0.42)			.87*	.21	.25	.67*	.72*
VR: Aggressive responding	0.00-2.00	0.88 (0.87)				.35*	.37*	.60*	.73*
Vignette: Anger	2.80-10.00	6.88 (1.86)					.34*	.37*	.48*
Vignette: Hostile intent attribution	1.00-8.90	3.72 (1.99)						.35*	.45*
Vignette: Revenge goals	0.00-1.00	0.28 (0.33)							.93*

Table 1 *Bivariate Correlations between SIP Variables in Provocation VR-Contexts and Vignettes (continued)*

	Range	M (SD)	2.	3.	4.	5.	6.	7.	8.
Vignette: Aggressive responding	0.00-1.80	0.45 (0.53)							

Note. All correlations including revenge goals are calculated using Kendall's τ, other correlations used Pearson's *r*.

*Indicates significance at .05, as the bootstrap 95% confidence interval did not include zero.

Measurement Sensitivity: Variances of SIP in VR versus Vignettes

To examine whether VR captured more individual differences in SIP than the vignettes, we compared variances of SIP variables between VR and vignettes. Table 1 shows the standard deviations for VR- versus vignette-assessed SIP variables, with larger standard deviations signifying larger variances. Results revealed significantly larger variances in SIP using VR versus vignettes for aggressive responding, $t(30) = 4.09$, $p < .001$, but not for anger, $t(30) = 1.43$, $p = .163$, hostile intent attribution, $t(30) = 1.40$, $p = .173$, and revenge goals, $t(30) = 2.01$, $p = .053$. So, only for aggressive responding, we found larger variances, meaning that interactive VR is more sensitive to capture individual differences in children's aggressive responding compared to vignettes.

Discriminant Validity: SIP Outcomes across VR Contexts

Table 2 presents the descriptive statistics of the VR-assessed SIP variables for the provocation, instrumental gain and neutral contexts separately, as well as the significance levels for our planned comparisons. Supporting discriminant validity, we found that the provocation context elicited significantly more anger ($d = 1.56$), hostile intent attributions ($d = 1.62$), and revenge goals ($d = 0.83$) than the instrumental gain context. As predicted, it also elicited significantly more anger ($d = 1.89$), hostile intent attributions ($d = 1.89$), revenge goals ($d = 0.71$), and aggressive responses ($d = 0.84$) than the neutral context. However, contrary to expectations, the instrumental gain context did not elicit significantly more instrumental goals than the provocation context ($d = 0.31$) or the neutral context ($d = 0.29$), nor did it elicit more aggressive responses than the neutral context ($d = 0.36$). Taken together, these findings provide partial support for the use of distinct VR contexts to capture distinct SIP patterns in children.

Table 2 *Descriptive Statistics of SIP Variables per Context and Significance Levels for Planned Comparisons Between Contexts Based on Bootstrap 95% Confidence Intervals*

	Provocation context	Instrumental gain context	Neutral context	Provocation vs. Neutral	Provocation vs. Instrumental	Instrumental vs. Neutral
	Mean (SD)	Mean (SD)	Mean (SD)	p	p	p
Anger	7.09 (2.60)	3.13 (2.80)	1.81 (2.11)	< .001	< .001	
Hostile intent attribution	6.76 (2.54)	2.48 (2.43)	2.05 (1.86)	< .001	< .001	
Revenge goals	0.38 (0.42)	0.06 (0.21)	0.03 (0.18)	< .001	= .001	
Instrumental gain goals	0.06 (0.21)	0.17 (0.35)	0.06 (0.25)		= .113	= .114
Aggressive responding	0.88 (0.87)	0.47 (0.84)	0.19 (0.59)	= .001		= .068

Note. Cells are empty if we had no hypotheses about this comparison.

Enthusiasm about the VR and Vignettes Assessment

We asked children to rate their enthusiasm directly after each assessment, and after completing both assessments. As predicted, children were more enthusiastic about the VR assessment (*M* = 8.54, *SD* = 1.98) than the vignette assessment directly after completing each assessment (*M* = 6.94, *SD* = 2.17) directly after each assessment, $p = .001$, $d = 0.72$. They also gave higher ratings to the VR (*M* = 9.06, *SD* = 1.95) than the vignettes (*M* = 6.78, *SD* = 2.56) after completing both assessments, $p < .001$, $d = 1.08$.

DISCUSSION

This pilot study examined whether interactive Virtual Reality (VR) provides a valid assessment of children's aggressive social information processing (SIP). We developed a virtual classroom where children played games with virtual peers. Children reported on their SIP in three distinct VR contexts (i.e., neutral, instrumental gain, provocation) and also completed a vignette-based assessment of SIP. Supporting convergent validity, results showed positive associations between VR- and vignette-based assessments of children's anger, hostile intent attributions, revenge goals, and aggressive responding. Supporting measurement sensitivity, results showed larger variances in of SIP assessment in VR for aggressive responding in VR versus vignettes. However, variances did not differ between VR and vignettes for anger, hostile intent attributions, and revenge goals. Supporting discriminant validity, results showed that the provocation context elicited more anger, hostile intent attributions, and revenge goals than the instrumental gain and neutral contexts, and more aggressive responses than the neutral context. However, one aspect of discriminant validity was not supported: The instrumental gain context did not elicit more instrumental goals than provocation and neutral contexts, nor more aggressive responses than the neutral context. Last but not least, results showed that children were more enthusiastic about participating in VR than completing a vignette-based measure of SIP.

Several findings stand out, warranting further discussion. We found significant correlations between SIP in VR and vignettes, supporting the convergent validity of our new measure. Yet, although correlations were moderate to large for hostile intent attributions, revenge goals, and aggressive responses, they were small for anger. This may be caused by the potentially limited reliability of using vignettes to assess children's anger. Participants may have found it difficult to report on their anger in hypothetical scenarios—in fact, people generally struggle to report on anticipated negative affective states (for a review, see: Robinson & Clore, 2002). Perhaps, interactive VR could provide a more reliable assessment of children's anger than current instruments do—a direction for future research worth investigating.

Another notable finding is that the improved measurement sensitivity of our interactive VR assessment was supported for aggressive responding, but not for other SIP variables. This could be due to the small sample size of our pilot study: variances were larger in VR for all SIP variables, but these differences were not significant. Still, the larger variance for aggressive responding in VR suggests that VR may be more sensitive to assess individual differences in aggressive responding than current vignette-based methods. As interactive VR immerses children in actual social interactions and allows them to actually aggress against virtual peers—an important difference with vignettes, which ask children to reflect on their hypothetical aggressive responses—it may have triggered aggressive responses in some children that were not triggered by vignettes. This notion aligns with theoretical work suggesting that many children may only respond aggressively when they are emotionally engaged (Anderson & Bushman, 2002; Lemerise & Arsenio, 2000). For clinical practice, this finding suggests that interactive VR assessment may detect individual differences in children's aggressive responding that would remain undetected when using measures that ask children to reflect on their hypothetical responses.

We developed distinct VR scenarios, aiming to assess the distinct SIP patterns that may underlie children's reactive and proactive aggression (Hubbard et al., 2010). Our findings showed that VR scenarios including peer provocation elicited more anger, hostile intent attributions, and revenge goals than the instrumental gain and neutral contexts, and more aggressive responses than the neutral context. These findings align with social-cognitive theory suggesting that contexts where the participant is provoked, frustrated or threatened should evoke SIP patterns related to reactive aggression (Dodge, 1991). They also align with empirical research suggesting that children's aggressive SIP patterns are context-dependent, as we found that the same children showed aggressive SIP in some VR scenarios but not others (De Castro & Van Dijk, 2017). These findings underscore the relevance of using distinct contexts to validly assess children's aggressive SIP patterns—and interactive VR may be an engaging and flexible method to do so.

Yet, what stood out is that our instrumental gain scenarios—developed to assess SIP underlying proactive aggression—did not elicit more instrumental goals than the provocation or neutral contexts. An explanation for this finding could be that proactive aggression is relatively rare (Dodge et al., 1997; Thomson & Centifanti, 2018). Indeed, few children in our sample displayed instrumental SIP in our VR scenarios, reducing statistical power to find significant differences between contexts. This idea is reflected in our data: although non-significant, on average, children showed more instrumental goals in the instrumental gain versus other contexts. Research in a larger sample of children is needed to examine whether instrumental gain contexts indeed elicit more instrumental tendencies than other contexts. This would be worthwhile, because VR—due to its realistic nature—seems more suitable to present instrumental gain scenarios than hypothetical vignettes (in fact, vignette-based SIP assessments have rarely included instrumental gain scenarios).

Lastly, children reported to be more enthusiastic about participating in VR than completing a vignette-based measure of SIP. This finding may have important implications for clinical practice. VR could be an attractive option for psychological assessment and intervention, because it may increase children's motivation to participate. If future research supports this idea, this may be particularly relevant for children with aggressive behavior problems because they are often less motivated to engage in therapy (Frick, 2012).

This study had several strengths. To our knowledge, it is the first study that used interactive VR to assess children's aggressive SIP. Moreover, we maximized variance in children's SIP by recruiting children from both regular and special education for children with behavior problems. The use of interactive VR in a sample with substantial variance in SIP allowed us to examine individual differences in children's aggressive SIP in an ecologically valid, experimentally controlled, theoretically comprehensive, and engaging context.

This study also had several limitations. First, the small sample size of this pilot study limits generalization of the findings. Moreover, it prevented us from running additional analyses, for instance to test whether the order in which children completed the VR- and vignette-assessments, despite counterbalancing, may have affected the results. Second, the provocation and instrumental gain contexts were assessed using only two VR scenarios each, reducing reliability of SIP measurements within each context. Relatedly, since research demonstrated that children display aggression in different contexts (e.g., Matthys et al., 2001), using two scenarios for each context may not have covered the broad range of social situations known to evoke aggression in children. Third, we only ran convergent validity analyses for the provocation VR scenarios because—to our knowledge—no well-established instrumental gain vignettes exist. Fourth, although the interactive nature of VR might have enhanced children's engagement in the actual interactions, it also caused the VR scenarios to slightly differ between children.

The findings of this pilot study open up valuable opportunities for both research and clinical practice. For research, an important next step is to examine to what extent SIP and behavior assessed in VR predict real-life aggression, using parent- or teacher-report questionnaires, peer nomination, and observation (Boateng et al., 2018). Moreover, the greater experimental control over social stimuli provided in the VR environment (e.g., nonverbal behaviors, emotion expressions) will allow researchers to test more specific hypotheses about causal effects in children's social interactions. For clinical practice, interactive VR—if it is further validated by future research—may provide a more attractive, flexible and valid method to assess children's aggressive SIP. Ultimately, interactive VR may also provide inroads for intervention. Clinicians could use the flexible VR environment to create engaging exercises tailored to individual clients, with precise control to adapt difficulty and complexity during the intervention.

In sum, this pilot study suggests that interactive VR is a promising tool to assess children's aggressive SIP. The use of VR allows researchers and clinicians to assess aggressive SIP in an emotionally engaging, ecologically valid context that is interactive and realistic. In the future, VR may further our understanding of SIP patterns underlying children's aggression and be used to enhance assessment and intervention for children with aggressive behavior problems.

SUPPLEMENTARY MATERIALS

Table S1 *Description of VR-Scenarios per Game; Italics Indicate Differences Between Games*

VR-scenario	Tower game	Cans game
Practice	The behavioral rules in the classroom are explained and the participant practices the game. The participant chats with a computerized avatar to practice talking in VR.	The behavioral rules in the classroom are explained and the participant practices the game. The participant chats with a computerized avatar to practice talking in VR.
Neutral	The participant *builds a tower of blocks*. During the game a virtual child engages in small-talk with the participant.	The participant *throws cans from the table*. During the game a virtual child engages in small-talk with the participant.
Object acquisition	The participant *builds a tower of blocks as high* as possible, however is one *block* short to finish the *tower* and earn bonus-points. During the game a virtual child is also *building a tower*, but then leaves the room and asks the participant if he could watch *his tower and blocks* until he returns.	The participant *throws as much cans from a table* as possible, however is one *ball* short to finish the *game* and earn bonus-points. During the game a virtual child is also *throwing cans from a table*, but then leaves the room and asks the participant if he could watch *his cans and balls* until he returns.
Competition	The participant and a virtual child both *build a tower of blocks as high* as possible. They are instructed that the player with the *highest tower* earns bonus-points, and that none of the players get bonus-points when it turns out to be a draw. The virtual child *has finished his tower* before the participant does, and announces that he is winning.	The participant and a virtual child both *throw as much cans from a table* as possible. They are instructed that the player with the *most cans thrown from the table* earns bonus-points, and that none of the players get bonus-points when it turns out to be a draw. The virtual child *has almost thrown all his cans from the table* before the participant does, and announces that he is winning.
Social provocation	The participant sees two virtual children in the classroom who are busy *building a tower of blocks*. The participant is prompted by the digital whiteboard to ask if he can join the game. Upon asking this question, the virtual children tell the participant he cannot join the game.	The participant sees two virtual children in the classroom who are busy *throwing cans from a table*. The participant is prompted by the digital whiteboard to ask if he can join the game. Upon asking this question, the virtual children tell the participant he cannot join the game.
Object provocation	The participant *builds a tower of blocks as high* as possible. A virtual child is already present with *a finished tower*. When the participant has *finished his tower*, the virtual child *walks up and down the classroom in search of more blocks. On his way back, the virtual child knocks over the tower*, thus ruining the participant's game.	The participant *throws as much cans from a table* as possible. A virtual child is already present with *his own balls*. When the participant has *one ball left to throw*, the virtual child *picks up the participant's last ball and drops it. The ball rolls out of reach*, thus ruining the participant's game.

Note. In all scenarios participants can earn points by playing the game: 1 point for each block built or each can thrown over, and bonus points for completing the game.

3

CHAPTER 4

Interactive Virtual Reality versus Vignette-Based Assessment of Children's Aggressive Social Information Processing

This chapter is published as:
Verhoef, R. E. J., Verhulp, E. E., Van Dijk, A., & De Castro, B. O. (2021). Interactive virtual reality versus vignette-based assessment of children's aggressive social information processing. Research on Child and Adolescent Psychopathology. Advance online publication. https://doi.org/10.1007/s10802-021-00879-w

The data and syntax that support the findings of this study are available through the Open Science Framework at https://doi.org/10.17605/OSF.IO/7SA6M

Author contributions:
RV conceptualized the study, EV, AD, and BC gave advice and feedback; RV and trained graduate students collected the data; RV analyzed the data and drafted the manuscript; EV, AD, and BC provided feedback on the analyses and manuscript.

ABSTRACT

This study examined whether interactive Virtual Reality (VR) provides a more ecologically valid assessment of children's aggressive social information processing (SIP) and aggressive responses than a standard vignette-based assessment. We developed a virtual classroom where children could meet and play games with virtual peers. Participants were boys (*N* = 184; ages 7-13) from regular education and special education for children with disruptive behavior problems. They reported on their SIP in four scenarios (i.e., two instrumental gain and two provocation scenarios) presented through both interactive VR and vignettes. Teachers reported on children's real-life aggressive behavior and reactive and proactive motives for aggression. Results demonstrated that children found the interactive VR assessment more emotionally engaging and immersive than the vignette-based assessment. Moreover, compared to vignettes, the interactive VR assessment evoked higher levels of aggressive SIP and responses in provocation scenarios only. Results supported the enhanced predictive validity of the interactive VR assessment of children's aggressive SIP and responses, which predicted children's real-life aggression above and beyond the vignette-based assessment with 2 to 12% additional explained variance. Similar results were found for children's real-life reactive and proactive motives for aggression, with 3 to 12% additional variance explained by interactive VR above and beyond vignettes. Interactive VR did not, however, evoke larger individual differences (i.e., variances) in children's aggressive SIP and responses than vignettes. Together, these findings suggest that interactive VR provides a more ecologically valid method to assess children's aggressive SIP and responses than hypothetical vignettes.

Keywords: Social information processing, aggression, children, Virtual Reality, reactive and proactive motives

INTERACTIVE VIRTUAL REALITY VERSUS VIGNETTE-BASED ASSESSMENT OF CHILDREN'S AGGRESSIVE SOCIAL INFORMATION PROCESSING

Children are often confronted with challenging social situations, such as not being allowed to join a peer group or being reprimanded by their teachers or parents. Such situations are likely to elicit strong emotions, which may affect children's thinking and responding in these situations (Caporaso & Marcovitch, 2021; Lemerise & Arsenio, 2000; Reijntjes et al., 2011). In many children, strong emotions such as anger, frustration, desire, or jealousy may trigger aggressive cognitions that would not have been triggered without these emotions. For instance, children may only interpret others' behavior as hostile when they feel frustrated, or may only justify stealing when they strongly desire an object. Thus, to better understand, predict, and treat children's aggressive behavior, we need to assess how children think in social situations *when they are emotionally engaged*. Yet traditional methods to assess children's social information processing (SIP) often use hypothetical stories (i.e., vignettes) that are unlikely to elicit strong emotions. We have therefore developed an interactive Virtual Reality (VR) environment to assess children's aggressive SIP and responses. The present study examines whether our VR-based assessment of children's SIP and responses better predicts their real-life aggressive behavior compared to a standard, vignette-based assessment.

Our interactive VR assessment is based on the SIP model (Crick & Dodge, 1994: Lemerise & Arsenio, 2000). This SIP model proposes that children's behavioral responses to social situations result from a sequence of mental processing steps: (1) encoding of social cues, (2) representation of social cues, (3) specification of interactional goals, (4) generation of responses, (5) evaluation of responses, and (6) enactment of a selected response. Children's aggressive behavior has been associated with deviations in each of these SIP steps, such as biased encoding, making hostile intent attributions, setting interactional goals directed at revenge or instrumental gain, generating more aggressive responses, and evaluating aggressive responses and their outcomes more positively (for reviews, see: De Castro & Van Dijk, 2017; Dodge, 2011). Moreover, children with aggressive behavior problems are more likely to experience anger (De Castro & Van Dijk, 2017), and research suggests that their SIP is more strongly affected by negative emotions (De Castro et al., 2003).

Previous work has shown that children's SIP patterns explain substantial variance in their concurrent and future aggressive behavior (e.g., De Castro & Van Dijk, 2017; Lansford et al., 2006; Chapter 2 of this dissertation). Nonetheless, findings vary considerably between studies and SIP measures used. A meta-analysis (Chapter 2 of this dissertation) revealed that the association between aggressive behavior and children's hostile intent attributions was stronger in studies using actual social interactions (d = 1.33) than in studies using vignettes (d = 0.23 to 0.44) or video-game tasks (d = 0.36; Yaros et al., 2014). The small to moderate effect sizes for vignettes and video games may be due to a lack of emotional engagement (i.e., vignettes may not evoke strong emotions) or limited ecological validity (i.e., video games may not resemble real-life social interaction). These findings align with theoretical work suggesting that strong emotions such as anger or frustration may trigger aggressive SIP patterns that are not triggered when children are calm (Anderson & Bushman, 2002; Lemerise & Arsenio, 2000; Chapter 6 of this dissertation). However, few studies exist that used actual social interactions to assess children's SIP—possibly because using this method is challenging in terms of standardization and ethics (Underwood, 2005).

Ideally, SIP assessment would combine highly emotional engaging, realistic social interactions with adequate standardization and ethically and practically feasible methodology. To attain this goal, we developed an interactive VR classroom where children can walk around freely, talk to virtual peers, and play games, allowing us to present standardized social events within an engaging environment. As children are fully immersed in the VR environment, the peer interactions they have (e.g., their game being ruined by a peer) may evoke substantial levels of anger, frustration, or jealousy. A recent pilot study revealed that our interactive VR assessment evoked larger individual differences in aggressive responses than a vignette-based assessment (Chapter 4 of this dissertation), suggesting that interactive VR may also enhance the prediction of individual differences in real-life aggressive behavior. The present study capitalizes on these findings by examining whether our interactive VR assessment of children's SIP indeed is (1) more immersive and emotionally engaging, and (2) more strongly associated with children's real-life aggression, compared to a vignette-based assessment of children's SIP.

Another advantage of using interactive VR may be that it allows for more precise assessment of distinct SIP patterns underlying reactive and proactive aggression (Dodge, 1991). Reactive aggression—an impulsive aggressive response to perceived threat or provocation (Dodge, 1991)—may stem from SIP characterized by excessive anger, heightened sensitivity to threatening cues, a tendency to attribute hostile intent to others, and goals directed at self-defense or taking revenge (e.g., Hubbard et al., 2010; Martinelli et al., 2018). Such reactive SIP patterns may particularly be triggered in provocation contexts (Hubbard et al., 2010) where children are refused to join a peer group (i.e., social provocation) or a peer damages their property (i.e., object provocation). In contrast, proactive aggression—planned aggressive behavior aimed at obtaining a desired outcome (Dodge, 1991)—may stem from SIP characterized by instrumental goals, positive outcome expectations of aggression, and positive evaluations of aggression (e.g., Hubbard et al., 2010). Such proactive SIP patterns may particularly be triggered in instrumental gain contexts (Hubbard et al., 2010), where children have the opportunity to steal something (i.e., object acquisition) or win a game by (i.e., competition). Although reactive and proactive motives for aggression can be mixed (e.g., taking revenge to show who is the boss; Bushman & Anderson, 2001), there is ample empirical work to suggest that they often occur in isolation (Polman et al., 2007; Van Dijk et al., 2021). Earlier studies, however, have not always found clearly delineated reactive versus proactive SIP patterns; possibly because their vignette-based assessment did not evoke the specific emotions underlying real-life reactive and proactive aggression (e.g., Crick & Dodge, 1996; Dodge et al., 1997; Oostermeijer et al., 2016; Stoltz et al., 2013).

Our interactive VR may address this issue by immersing children in engaging social interactions with virtual peers, where they are actually (not just hypothetically) provoked or tempted to use aggression. They may experience anger, frustration, or jealousy, activating the unique SIP patterns underlying reactive- and proactive aggression. Interactive VR then allows children to actually aggress against virtual peers instead of reporting on their hypothetical aggressive responses as with vignettes. Consequently, interactive VR permits an assessment of children's outcome expectancies and evaluations regarding their actual behavior instead of presenting them with hypothetical response options they might never carry out in real life.

In sum, the present study examines whether interactive VR provides a better assessment of children's aggressive SIP and responding than a standard, vignette-based assessment. We chose vignettes for this comparison because they are the standard method to assess children's SIP, and have been shown to yield similarly modest associations with children's real-life aggression as other methods, such as video-game tasks (Chapter 2 of this dissertation). Children completed both an interactive VR-based and a hypothetical vignette-based assessment of SIP, and teachers reported on their aggressive behavior. We had three main goals. First, we tested whether interactive VR, compared to vignettes, would elicit higher levels of emotional engagement (1a) and immersion (1b). Consequently, we expected that interactive VR would trigger aggressive SIP and response patterns that are not triggered when children are calm. This should result in larger individual differences (i.e., variances) in SIP and aggressive responses (1c), and higher scores on aggressive SIP and aggressive responses (1d). Moreover, it should result in more congruent SIP and response patterns, visible as stronger correlations between all SIP and aggressive response variables in each scenario (1e). Second, we examined whether interactive VR explained additional variance in children's real-life aggressive behavior reported by teachers, above and beyond the vignette-based assessment. We examined this both for the assessment of children's aggressive SIP (2a) and children's aggressive responses (2b). Third, we examined whether interactive VR explained additional variance in teacher-reported reactive and proactive motives for aggression, above and beyond the vignette-based assessment—again, both for aggressive SIP (3a) and aggressive responses (3b).

METHOD

Participants

Participants were 184 Dutch boys ages 7 to 13 years (*M* = 10.22; *SD* = 1.30). They were recruited from 18 Dutch primary schools. Schools were from neighbourhoods representative of the Dutch population, with on average 9% inhabitants with a Western migration background (*SD* = 3%), 13% with a non-Western migration background (*SD* = 9%), 21% with a lower educational level (*SD* = 4%), and with 7% of the households having a low-income (*SD* = 3%) (Statistics Netherlands, 2018, 2019). To maximize variance in aggressive behavior, boys high on disruptive behavior problems were oversampled by including boys from special education for disruptive behavior problems (*n* = 118) and a random sample of boys from regular education (*n* = 66). In the Netherlands, special education for children with disruptive behavior problems and/or psychiatric problems is reserved for children whose behavior problems are so severe that they require extra support that cannot be provided in regular education. In our study, boys from special education were nominated by their teacher for frequently showing aggressive behavior problems. Boys were excluded if they had an IQ below 80 or an Autism Spectrum Disorder (ASD) according to their casefiles, or had a clinical score on ASD symptoms on the teacher-rated Social Emotional Questionnaire (SEQ; Scholte & Van der Ploeg, 2007). Schools sent parents an information letter in which the study was explained. All parents provided written consent for their child's participation in the study by signing the attached informed consent form and returning it to their child's teacher. Boys provided verbal assent.

Procedure

Participants were individually tested in a silent room at their school by trained graduate students or the first author. Graduate students were trained in multiple sessions by the first author and were supervised during the first two assessments to ensure assessment fidelity. The interactive VR- and vignette-based SIP assessments both lasted 45 minutes and were completed on two different days with approximately one week in between. We counterbalanced the order of these assessments across participants to control for order effects. At the end of each assessment, boys reported on their emotional engagement and immersion during the assessment. Boys received a small monetary reward (€5) for their participation. Teachers reported on boys' aggressive behavior and filled out the SEQ through online questionnaires (response rate = 98%). The study was approved by the Medical Ethics Committee of University Medical Center Utrecht.

Materials

Interactive Virtual Reality Environment

Participants wore VR glasses to immerse them in the VR environment. They could walk around freely (in a demarcated 4 × 4 meter space), use controllers that mimicked their hands, and respond in similar fashion as in real life: through verbal and physical behavior. The interactive VR environment was designed as a virtual school classroom where participants could interact and play games with virtual peers (for a detailed description of the interactive VR environment, see Chapter 3 of this dissertation). We presented the virtual classroom to participants as an actual classroom where standard behavior rules applied (e.g., respecting other children) and where they would meet real children from other schools who were also participating in the study. In reality, virtual peers were controlled by the experimenter through default movement options and standardized verbal responses.

Participants could play two games: (1) building a tower of blocks as high as possible, and (2) throwing five balls to hit as many cans from a table as possible. We designed our VR assessment around these games to allow for both peer-directed aggression (e.g., hitting, name calling) and property-directed aggression (e.g., knocking over the peer's tower). To increase participants' emotional engagement and to provide experimental control over gains and losses, we included high scores and bonuses for participants' performance during the games (e.g., building a high tower). The instructions, game rules, and score count were displayed on a digital school board, which also explained these matters through standardized verbal instructions.

Virtual Reality Scenarios

Participants were presented with six VR scenarios in a fixed order: (1) practice scenario, (2) neutral scenario, (3) object acquisition, (4) competition, (5) social provocation, and (6) object provocation—all centering around one of the games (i.e., the tower or cans game; randomly assigned). The practice VR scenario served to familiarize participants with the VR environment and game rules by practicing the game without any virtual characters present. The neutral scenario served to familiarize participants with the SIP questions by having them play the game while engaging in neutral small talk with a virtual peer, and asking the SIP questions afterwards. Next, participants completed the four experimental scenarios, which we based on taxonomies of problematic situations for children with aggressive behavior problems (Matthys et al., 2001). The first two scenarios involved instrumental gain. In the object acquisition scenario, participants had the opportunity to steal a block or ball from the virtual peer, which would earn them additional points in the game. In the competition scenario, they could win the game and thus earn additional points by sabotaging the virtual peer's progress in the game (i.e., by knocking over the peer's tower, ruining the virtual peer's balls). The last two scenarios involved provocation. In the social provocation scenario, participants were refused to join the game by two virtual peers. In the object provocation scenario, their game was ruined by a virtual peer. As such, the provocations caused them to earn no points. In the two provocation scenarios, participants could not obtain any points by responding aggressively. We expected these provocation scenarios to elicit the strongest emotions, and therefore presented them last to prevent carry-over effects.

Hypothetical Vignettes

For the vignette-based SIP assessment, we developed audiotaped vignettes with the exact same content as the VR scenarios (e.g., describing how participants would gain or lose high scores and bonuses), allowing for a clean comparison between assessment methods. We counterbalanced the type of game across participants (i.e., participants who received the tower game in interactive VR, received the cans game with vignettes, and vice versa). As in most vignette procedures, participants were told that they would listen to stories about everyday social situations with peers and were asked to imagine that each story actually happened to them (Chapter 2 of this dissertation).

Measures

Emotional Engagement

We assessed children's emotional engagement during the assessment in two ways. First, we used two items immediately after each assessment to directly capture children's emotional engagement during the assessment, aiming to minimize the effect of memory on their ratings (i.e., "How angry did you feel when something bad happened to you in VR/vignettes?" and "How much did you care when something bad happened to you in VR/vignettes?"). Children responded on a rating scale from 1 (*not at all*) to 10 (*very*). We averaged the two items to create emotional engagement scores for both interactive VR (r = .83) and vignettes (r = .67). Second, to allow children to make a comparison between the VR- and vignette-based assessment, we again administered these two items after they had completed both assessments, but then phrased in comparative form (e.g., for the first item: "You have completed both the VR and the stories. How angry did you feel when something bad happened to you in the VR? And in the stories?;" question order was counterbalanced). We again averaged the two items to create emotional engagement scores for interactive VR (r = .74) and vignettes (r = .74).

Immersion

We assessed children's immersion during the assessment in two ways. First, we used six items immediately after each assessment, which were adapted from the Dutch translation of the Igroup Presence Questionnaire (Schubert et al., 1999). Two of the six items had low factor loadings (i.e., below 0.60) and were excluded. The four items used were: 1) "I was totally caught up by the events in VR/vignettes;" 2) "I had the feeling that the events in VR/vignettes were actually happening to me;" 3) "During the VR/vignettes it felt like I was actually experiencing the events;" and 4) "The events in VR/vignettes seemed almost real." Participants rated the items on a scale from 1 (*strongly disagree*) to 5 (*strongly agree*). We averaged across items to create immersion scores for both interactive VR (α = .78) and vignettes (α = .81).

Second, to allow children to make a comparison between the VR- and vignette-based assessment, we administered one item after they had completed both assessments, but then phrased in comparative form (i.e., "You have participated in both the VR- and vignette-based assessment. How much did you have the feeling that the events in VR were actually happening to you? And in the stories?"). Children responded on a rating scale from 0 (*not at all*) to 10 (*very*).

Aggressive SIP and Responses

We assessed participants' aggressive SIP and responses in two provocation scenarios and two instrumental gain scenarios (both in interactive VR and with vignettes). Initially, we planned to create aggregate SIP and response variables for provocation and instrumental gain contexts. However, we found low correlations for SIP and response variables between the social provocation and object provocation scenario (i.e., ranging from .37-.60 in VR and from .27-.50 with vignettes) and between the object acquisition and competition scenario (i.e., ranging from .34-.58 in VR and from .35-.48 with vignettes), suggesting that aggressive SIP and behavior may be highly situation specific (Dodge et al., 1985; Matthys et al., 2001). Hence, we decided to create variables for children's SIP and aggressive responses for each scenario separately.

Interactive VR Assessment. We assessed participants' aggressive responses through observation of their behavior in VR, and used self-report to assess their anger, intent attributions, goals, outcome expectancies, and response evaluations at the end of each VR-scenario. In between scenarios, participants kept their VR-glasses on while replying verbally to the experimenter's questions. For procedural clarity, we assessed all SIP questions following all scenarios, even though we were only interested in proactive SIP in instrumental scenarios (i.e., instrumental goals, outcome expectancies, and response evaluation) and reactive SIP in provocation scenarios (i.e., anger, hostile intent attribution, and revenge goals).

Anger. Anger was assessed using one item following each VR-scenario: "The other boy did [behavior of other boy]. How angry did this make you feel, on a scale from 1, meaning *not at all*, to 10, meaning *very*?".

Hostile Intent Attribution. Intent attributions were assessed using two items following each VR-scenario: "The other boy did [behavior of other boy]. To what extent did he try to be mean, on a scale from 1, meaning *not at all*, to 10, meaning *very*?" and "To what extent did he try to hinder you, on a scale from 1 to 10?" These two items were moderately to highly correlated within each of the four VR scenarios (M = .83, Mdn = .87, range = .67-.90) and were therefore averaged within each VR-scenario.

Interaction Goals. Interaction goals were assessed using one open-ended question following each VR-scenario: "When the other boy did [behavior of other boy], you did [behavior of participant]. What was the reason you did this?" In line with earlier research (De Castro et al., 2012), the first author coded each answer as *revenge goals* (e.g., "to retaliate," "because I was angry," "to defend myself"), *instrumental goals* (e.g., "to win the game," "to show him who's the boss"), *goals underlying non-aggressive behavior* (e.g., "to become friends," "to avoid problems"), or *no goals* (e.g., "I don't know"). A second rater also coded 35% of the transcriptions. Inter-rater reliability was excellent, with Cohen's κ ranging from .85-.96 across scenarios (M = .91, Mdn = .91). Scores for revenge goals were created by assigning 1 to *revenge goals* codes and 0 to other codes. Similarly, scores for instrumental goals were created by assigning 1 to *instrumental goals* codes and 0 to other codes.

Aggressive Responses. We assessed participants' behavioral responses in interactive VR through observation. A trained research assistant made detailed descriptions of participants' behavioral responses in each VR-scenario. The first author coded these descriptions into *non-aggressive behavior* (e.g., prosocial behavior, avoidance), *mild* *aggressive behavior* (e.g., coercion, verbal aggression), and *severe aggressive behavior* (e.g., physical aggression, destructive aggression) following standard coding procedures (De Castro et al., 2005). If multiple codes applied, the highest category was scored. A second rater also coded 35% of the behavioral descriptions. Inter-rater reliability was excellent, with κ ranging from .92-1.00 across scenarios (*M* = .97, *Mdn* = .98). Because frequencies of *mild aggressive behavior* were low or even absent (i.e., 0 to 17% across VR-scenarios and vignettes, *Mdn* = 2%), we created a dichotomous variable by coding *mild* and *severe aggressive behavior* as 1 and *non-aggressive behavior* as 0.

Outcome Expectancies. Outcome expectancies of aggression were assessed using one item following each VR-scenario: "What did you expect would happen when you [behavior of participant]?" We coded only answers of participants who had actually used aggression in that VR-scenario and assigned missing values to other answers. The first author coded each answer as *positive outcome expectancies of aggression* (e.g., "I would win the game"), or *no positive outcome expectancies of aggression* (e.g., "He would dislike me"). A second rater also coded 35% of the transcriptions. Inter-rater reliability was excellent, with κ being 1.00 for each scenario. Scores for positive outcome expectancies of aggression were created by assigning 1 to *positive outcome expectancies of aggression* and 0 to *no positive outcome expectancies of aggression.*

Response Evaluations. Positive evaluations of aggression were assessed using one item following each VR-scenario: "When the other boy did [behavior of other boy], you did [behavior of participant]. To what extent do you approve your behavior on a scale from 1, meaning *not at all*, to 10, meaning *very*?" We only used scores of children who had actually used aggression in that VR-scenario and coded other scores as missing.

Participants' outcome expectancies and response evaluations of aggression were only scored when they displayed aggressive responses, limiting the number of observations for these variables. Conversely, other SIP variables (i.e., anger, hostile intent attributions, revenge goals and instrumental goals) could be scored irrespectively of whether participants engaged in aggressive responses, yielding full data for these variables (see Table 1 for descriptive statistics of SIP and aggressive response variables).

Vignette Assessment. Children reported on their SIP following each vignette. We used the same questions and coding schemes as used for the interactive VR-assessment, except that we formulated the questions as hypothetical (e.g., "How angry would you feel…?") instead of actual (e.g., "How angry were you…?"). The two items assessing intent attributions were averaged within each vignette as they were highly correlated (*M* = .80, *Mdn* = .81, range = .68-.90). Inter-rater reliability (κ) for open-ended questions was based on 35% of transcriptions and was excellent for both interaction goals (range = .81-1.00, *M* = .91, *Mdn* = .91) and outcome expectancies (range = .83-1.00, *M* = .94, *Mdn* = 1.00). We assessed participants' anticipated behavioral responses for each vignette using an open-ended question (i.e., "What would you do if [social event]?"). Inter-rater reliability was based on 35% of the transcriptions and was excellent, with κ ranging from .91-1.00 (*M* = .94, *Mdn* = .93).

Real-Life Aggressive Behavior

Teachers completed two questionnaires to assess participants' aggressive behavior in real life. First, teachers filled out the Aggressive Behavior subscale of the Dutch version of the Teacher Report Form (TRF; Verhulst et al., 1997). They rated 20 items (e.g., "This child threatens others") on a 3-point Likert scale (*1 = not true for this child*, 2 = *somewhat true for this child*, or 3 = *very often true for this child*). Scores were averaged across items (α = .96). Second, they filled out the Instrument for Reactive and Proactive Aggression (IRPA; Polman et al., 2009). This instrument differentiates between the frequency of aggression on the one hand, and the motives underlying aggression on the other hand. We used the frequency scale to assess children's real-life aggressive behavior. Teachers rated the frequency of 7 distinct forms of aggressive behavior (i.e., kicking, pushing, hitting, name calling, arguing, gossiping, and doing sneaky things) in the previous month on a 5-point Likert scale (1 = *never,* 2 = *once,* 3 = *weekly,* 4 = *multiple times a week,* 5 = *daily*). Scores on these seven items were averaged (α = .90). IRPA frequency scores (*M* = 1.95, *SD* = 0.86) and TRF scores (*M* = 1.67, *SD* = 0.57) were highly correlated (*r* = .85). We therefore standardized and averaged them to create a single aggressive behavior score.

4

Reactive & Proactive Motives for Aggression

We assessed reactive and proactive motives for aggression by again using the IRPA (Polman et al., 2009), but this time the motive scales. For each form of aggression rated above 0, teachers rated 3 reactive motives items (e.g., "Because someone teased or upset him") and 3 proactive motives items (e.g., "To hurt someone or to be mean") on a 5-point Likert scale (0 = *never,* 1 = *rarely,* 2 = *sometimes,* 3 = *often,* 4 = *always*). For aggression frequency items rated 0, motives scores were missing by design. We calculated reactive and proactive motives scores by averaging across all reactive motives items (i.e., 3 items times 7 forms of aggression; α = .94) and all proactive motives items (α = .95), respectively. Thus, high scores on reactive (*M* = 2.75, *SD* = 0.94) or proactive (*M* = 2.04, *SD* = 0.84) motives indicate that if participants engaged in aggressive behavior, they often had reactive or proactive motives. The correlation between reactive and proactive motives was non-significant (*r* = .14, *p* = .075).

Statistical Analyses

To test our first hypothesis that interactive VR is more engaging than vignettes, we considered five aspects. First, we examined whether interactive VR yielded higher mean levels of emotional engagement than vignettes, using paired *t*-tests. Second, we examined whether participants' immersion was higher in VR versus vignettes, also using paired *t*-tests. Third, we examined whether VR elicited larger individual differences in aggressive SIP and aggressive responses than vignettes. To this end, we used an adaptation of the Pittman-Morgan test which replaces Pearson's *r* with Spearman's rank correlation to account for non-normal data (McCulloch, 1987). Fourth, we examined whether interactive VR yielded higher scores on aggressive SIP and aggressive responses than vignettes, using paired *t*-tests for continuous SIP variables and McNemar's tests for dichotomous SIP and response variables. Fifth, we examined whether VR yielded stronger correlations among SIP and aggressive responses than vignettes. To do so, we calculated correlations between all SIP and response variables for each scenario using Pearson's *r*, Pearson's π, and Point-Biserial correlations. Next, we tested for inequality of the obtained correlation matrices using Steiger's test (1980), which directly compares all elements of two dependent correlation matrices instead of comparing each correlation separately.

To test our second hypothesis that interactive VR assessment of aggressive SIP (2a) and responses (2b) better predicts children's aggressive behavior in real life compared to vignettes, we examined whether VR explained additional variance in real life aggression above and beyond vignettes, but not vice versa. For aggressive SIP, we conducted two hierarchical regression analyses: the first with vignette-assessed SIP entered at step 1 and VR-assessed SIP at step 2; the second with VR-assessed SIP at step 1 and vignette-assessed SIP at step 2. For aggressive responses, we repeated these analyses with VR- versus vignette-assessed aggressive responses as predictors.

To test our third hypothesis that interactive VR assessment of aggressive SIP (3a) and responses (3b) better predicts children's reactive and proactive motives underlying their aggressive behavior in real life compared to vignettes, we conducted the same hierarchical regression analyses as used for our second hypothesis, but then with reactive motives as dependent variables for the provocation scenarios and proactive motives as dependent variables for the instrumental gain scenarios.

RESULTS

Preliminary Analyses

Table 1 presents the descriptive statistics for all SIP variables in both VR and vignettes. As most SIP variables were skewed, we conducted our analyses using a bootstrapping procedure with bias-corrected accelerated (BCa) 95% confidence intervals (CI) based on 5000 resamples.

Our VR elicited aggressive responses in 23% to 58% of children, depending on the scenario (Table 1). However, few children who responded aggressively in the VR, also responded aggressively in the same scenario in vignettes (i.e., 9 to 32% across scenarios, $Mdn = 10\%$; see Supplementary Material Table S1). As a result, we had insufficient data to compare VR versus vignettes on SIP variables that were only assessed if children actually responded aggressively (i.e., positive outcome expectancies and positive evaluations of aggression). We therefore reported descriptive statistics for these two variables (see Table 1) but excluded them from our main analyses.

Children's Engagement in Interactive VR versus Vignettes

Emotional Engagement

As predicted, children reported feeling more emotionally engaged in VR ($M = 5.59$, $SD = 2.74$) than with vignettes ($M = 2.91$, $SD = 2.37$), BCa 95% CI [2.23, 3.10], $p < .001$, , $d = 0.92$. The same result was found when we asked children about their engagement after they had completed both assessments: their engagement was higher in VR ($M = 5.68$, $SD = 2.69$) than with vignettes ($M = 3.05$, $SD = 2.16$), BCa 95% CI [2.23, 3.04], $p < .001$, $d = 0.96$.

Immersion

As predicted, children reported feeling more immersed in VR ($M = 4.18$, $SD = 0.83$) than with vignettes ($M = 2.57$, $SD = 1.07$), BCa 95% CI [1.44, 1.77], $p < .001$, $d = 1.45$. They also reported feeling more immersed in VR ($M = 7.96$, $SD = 2.21$) than with vignettes ($M = 3.70$, $SD = 2.37$) after they had completed both assessments, BCa 95% CI [3.85, 4.64], $p < .001$, $d = 1.52$.

Variances in Aggressive SIP and Responses

We found limited support for our hypothesis that VR elicits larger variances in SIP variables and aggressive responses than vignettes. For the object acquisition and competition scenario respectively, we found no difference in variances between VR and vignettes for instrumental goals, $t(178) < 0.01$, $p = 1.000$, and $t(179) = 1.56$, $p = .120$, and aggressive responses, $t(178) < 0.01$, $p = 1.000$, and $t(179) = 0.21$, $p = .835$. In the social provocation scenario, we did find higher variances in VR versus vignettes for revenge goals, $t(178) = 4.87$, $p < .001$, and aggressive responses, $t(178) = 4.37$, $p < .001$, but not for anger, $t(178) = -0.28$, $p = .777$, and hostile intent attributions, $t(178) = 0.61$, $p = .544$. Last, in the object provocation scenario, we found no support for our hypothesis: there were no differences for anger, $t(177) = 0.91$, $p = .364$, revenge goals, $t(177) = 0.33$, $p = .738$, and aggressive responses, $t(177) = -0.02$, $p = .983$, and, contrary to our expectation, larger variances with vignettes versus VR for hostile intent attributions, $t(177) = -2.08$, $p = .039$.

Levels of Aggressive SIP and Responses

We tested whether VR yielded more aggressive SIP and responses than vignettes (Table 1). Details of the McNemar's test for dichotomous variables can be found in Supplementary Material Table S1. Our hypothesis was partly supported. For the object acquisition and competition scenarios respectively, we found no differences between VR and vignettes in instrumental goals, $p = 1.000$, $OR = 1.000$, $p = .154$, $OR = 1.67$, nor aggressive responses, $p = 1.000$, $OR = 1.00$, $p = .885$, $OR = 1.09$. For the social provocation and object provocation scenarios respectively, children showed more hostile intent attributions, BCa 95% CI [.00, .97], $p = .048$, $d = 0.15$, BCa 95% CI [.07, 1.29], $p = .025$, $d = 0.17$, revenge goals, $p < .001$, $OR = 4.67$, $p = .014$, $OR = 1.85$, and aggressive responses, *p* < .001, *OR* = 6.63, *p* = .001, *OR* = 2.35, in VR than with vignettes. However, we found no differences in anger between VR and vignettes in the social provocation scenario, BCa 95% CI [-.24, .69], *p* = .354, *d* = 0.07, and even higher mean levels of anger for vignettes versus VR in the object provocation scenario, BCa 95% CI [-1.21, -.28], *p* = .002, *d* = -0.24.

Correlations between Aggressive SIP Variables and Responses

We tested whether correlations among aggressive SIP and response variables were stronger in VR versus vignettes. Table 2 presents all correlations between these variables for each scenario separately. Steiger's test to compare correlation matrices showed that support for our hypothesis was limited. Steiger's test revealed that the correlation matrix of aggressive SIP and response variables was significantly higher for VR versus vignettes for the competition scenario, $\chi^2(1) = 23.33$, $p < .001$, but did not significantly differ between VR and vignettes for the object acquisition scenario, $\chi^2(1) = 0.03$, $p = .862$, social provocation scenario, $\chi^2(6) = 6.58$, $p = .361$, and object provocation scenario, $\chi^2(6) = 6.46$, $p = .374$.

In sum, children reported more emotional engagement and immersion in VR than with vignettes. Partial support was found for VR outperforming vignettes on other aspects: It yielded more variance for 2 out of 12 results, higher levels of aggressive SIP and responses for 6 out of 12 results, and stronger correlations for 1 out of 4 results.

Table 1 *Descriptive Statistics of SIP Variables for each Scenario between VR and Vignettes (VIG)*

	Object Acquisition		**Competition**		**Social Provocation1**		**Object Provocation2**	
	VR	VIG	VR	VIG	VR	VIG	VR	VIG
Anger	1.75 (1.78)	1.65 (1.62)	3.13 (2.66)	4.98 (3.09)	5.36 (2.88)[1]	5.13 (2.94)[1]	6.79 (2.76)[2]	7.54 (2.58)[2]
Hostile Intent Attribution	1.47 (1.21)	1.59 (1.21)	2.17 (2.18)	5.00 (3.09)	5.23 (3.08)[1]	4.75 (2.94)[1]	7.35 (2.78)[2]	6.65 (3.32)[2]
Revenge Goals	1 (1)	1 (1)	6 (3)	15 (8)	51 (28)	19 (10)	92 (51)	69 (38)
Instrumental Goals	41 (23)	43 (24)	37 (20)	27 (15)	16 (9)	6 (3)	12 (7)	7 (4)
Aggressive Responses	42 (23)	44 (24)	43 (24)	42 (23)	69 (38)	25 (14)	105 (58)	77 (42)
Outcome Expectancies[a]	10 (24)	14 (32)	18 (42)	3 (7)	2 (3)	0 (0)	0 (0)	5 (6)
Positive Evaluations[a]	5.17 (3.56)	4.32 (2.77)	4.91 (3.71)	4.26 (3.25)	4.04 (3.24)	6.24 (3.38)	5.22 (3.40)	5.49 (3.16)

Note. Columns display means (*M*) and standard deviations (*SD*) of Anger, Hostile Intent Attribution, and Positive Evaluations, and the number (*n*) and proportion (%) of children displaying Revenge Goals, Instrumental Goals, Aggressive Responses, and Positive Outcome Expectancies. [a] Scores only apply to children who displayed an Aggressive Response for this scenario. [1] Based on *n* = 179 because 5 children had missing data in VR/vignettes; [2] Based on *n* = 178 because 6 children had missing data in VR/vignettes.

Table 2 *Bivariate Correlations of SIP and Response Variables in VR and Vignettes with Real-Life Aggressive Behavior and Reactive and Proactive Motives for Aggression in Instrumental Gain Scenarios (Object Acquisition Scenario above the Diagonal; Competition Scenario below) and Provocation Scenarios (Social Provocation Scenario above the Diagonal; Object Provocation Scenario below)*

Instrumental gain scenarios	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**				
VR: Instrumental Goals		.98*	.27*	.29*	.19*	.27*				
VR: Aggressive Responses	.91*		.26*	.28*	.20*	.27*				
Vignette: Instrumental Goals	.23*	.18*		.99*	.17*	.15				
Vignette: Aggressive Responses	.27*	.24*	.76*		.20*	.16				
Real-Life Aggressive Behavior	.26*	.29*	.17	.32*		.58*				
Real-Life Proactive Motives	.24*	.23*	.14	.20*	.58*					
Provocation scenarios	**1.**	**2.**	**3.**	**4.**	**5.**	**6.**	**7.**	**8.**	**9.**	**10.**
VR: Anger		.52*	.41*	.41*	.37*	.40*	.21*	.23*	.12	.06
VR: Hostile Intent Attribution	.55*		.45*	.36*	.29*	.40*	.10	.13	.05	.14
VR: Revenge Goals	.41*	.40*		.80*	.09	.17*	.16*	.19*	.34*	.30*
VR: Aggressive Responses	.36*	.40*	.86*		.07	.14	.16*	.23*	.35*	.29*
Vignette: Anger	.30*	.18*	.11	.10		.52*	.32*	.41*	.06	-.12
Vignette: Hostile Intent Attribution	.10	.12	.14	.11	.45*		.30*	.35*	-.05	-.21*
Vignette: Revenge Goals	.19*	.16*	.18*	.23*	.43*	.37*		.86*	.17*	.05
Vignette: Aggressive Responses	.19*	.19*	.24*	.28*	.39*	.39*	.91*		.23*	.09
Real-Life Aggressive Behavior	.20*	.29*	.27*	.32*	.15	.07	.14	.20*		.46*
Real-Life Reactive Motives	.19*	.21*	.29*	.27*	-.04	-.07	-.04	.01	.46*	

Note. Correlations of SIP and Responses in Instrumental Gain Scenarios are reported in the upper part of the Table (Object Acquisition Scenario above the Diagonal; Competition Scenario below) and correlations of SIP and Responses in Provocation Scenarios in the lower part of the Table (Social Provocation Scenario above the Diagonal; Object Provocation Scenario below). All correlations including Instrumental Goals, Revenge Goals or Aggressive Responses are point-biserial correlations, all correlations between Instrumental Goals, Revenge Goals and Aggressive Responses used Pearson's π, and other correlations used Pearson's r. * Indicates that the bootstrap 95% confidence interval did not include zero.

Predicting Real-Life Aggressive Behavior

Table 3 and 4 present the results of the hierarchical regression analyses of aggressive behavior in real life regressed on aggressive SIP a) and aggressive responses b), first conducted with vignettes in Step 1 and VR in Step 2, and next with VR in Step 1 and vignettes in Step 2. Analyses were conducted for each scenario separately.

Aggressive SIP

Children's aggressive SIP in all four VR scenarios significantly predicted their real-life aggression, with explained variances at Step 1 ranging from 4 to 13% across scenarios. As expected, effects were weaker for vignettes. Children's aggressive SIP assessed with vignettes significantly predicted their real-life aggression at Step 1 in the object acquisition scenario (R^2 = .03) and social provocation scenario (R^2 = .05), but not in the competition (R^2 = .02) and object provocation scenario (R^2 = .04). Turning to the incremental value of VR, we found that VR entered at Step 2 explained significant variance over and above vignettes in all scenarios (i.e., 2% in object acquisition, 5% in competition, 12% in social provocation, and 9% in object provocation). As predicted, vignettes did not explain significant variance over and above VR in any scenario.

Aggressive Responses

Children's aggressive responses in all four VR scenarios significantly predicted their real-life aggression, with explained variances at Step 1 ranging from 4 to 12% across scenarios. Similar effects were found for vignettes, with explained variances at Step 1 ranging from 4 to 10%. Turning to the incremental value of VR, we found that VR entered at Step 2 explained significant variance over and above vignettes in all scenarios (i.e., 2% in object acquisition, 5% in competition, 9% in social provocation, and 7% in object provocation). However, we also found that vignettes at Step 2 explained significant variance over and above VR in three scenarios, with higher levels of explained variance in the competition scenario (i.e., 6%), but lower levels in in social provocation and object provocation scenarios (i.e., 3% and 2%, respectively).

In sum, all eight hierarchical regression analyses regarding children's real-life aggression supported the incremental value of VR over vignettes, whereas only three analyses supported the reverse.

Table 3 *Hierarchical Regression Analyses of Real-Life Aggression Regressed both on Instrumental Goals and Aggressive Responses*

		Object acquisition scenario						Competition scenario					
Step	Predictor	β	β SE	95% CI	ΔR^2	*df*	*F* change	β	β SE	95% CI	ΔR^2	*df*	*F* change
1	Vignette: Instrumental Goals	.37	.20	-.02-.77	.03	1,174	4.54*	.40	.25	-.07-.91	.02	1,175	3.75
2	Vignette: Instrumental Goals	.27	.19	-.10-.65	.02	1,173	4.38*	.23	.25	-.26-.75	.05	1,174	10.27**
	VR: Instrumental Goals	.37	.20	-.01-.75				.58**	.18	.21-.94			
1	VR: Instrumental Goals	.44*	.20	.04-.82	.04	1,174	6.66*	.63**	.19	.27-1.01	.07	1,175	13.02***
2	VR: Instrumental Goals	.37	.20	-.01-.75	.01	1,173	2.30	.58**	.19	.22-.95	.01	1,174	1.18
	Vignette: Instrumental Goals	.27	.19	-.10-.64				.23	.25	-.21-.72			
1	Vignette: Aggressive Responses	.44*	.20	.07-.83	.04	1,174	6.86*	.74***	.20	.36-1.15	.10	1,175	19.98***
2	Vignette: Aggressive Responses	.34	.19	-.02-.72	.02	1,173	4.12*	.61**	.19	.24-1.00	.05	1,174	9.19**
	VR: Aggressive Responses	.35	.19	-.00-.73				.50**	.17	.15-.84			
1	VR: Aggressive Responses	.45*	.20	.08-.84	.04	1,174	7.09**	.66***	.18	.31-1.00	.08	1,175	16.07***
2	VR: Aggressive Responses	.35	.19	-.00-.72	.02	1,173	3.90	.50**	.17	.17-.83	.06	1,174	12.39***
	Vignette: Aggressive Responses	.34	.19	-.01-.71				.61**	.19	.24-.97			

Note. Hierarchical Regression Analyses were run for the Two Instrumental Gain Scenarios separately, both with Vignettes and VR Entered First. Model output is based on a non-bootstrapped procedure whereas output on separate predictors is based on a bootstrapping procedure.
$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

Table 4 *Hierarchical Regression Analyses of Real-Life Aggression Regressed both on Reactive SIP and Aggressive Responses*

		Social Provocation						Object Provocation					
Step	Predictor	β	β SE	95% CI	ΔR^2	*df*	*F* change	β	β SE	95% CI	ΔR^2	*df*	*F* change
1	Vignette: Anger	.01	.03	-.05-.08	.05	3,172	2.80*	.05	.03	-.01-.11	.04	3,171	2.55
	Vignette: Hostile Intent Attribution	-.04	.03	-.10-.02				-.00	.02	-.05-.04			
	Vignette: Revenge Goals	.67*	.31	.09-1.28				.22	.17	-.10-.55			
2	Vignette: Anger	.02	.03	-.05-.08	.12	3,169	7.85***	.04	.03	-.01-.10	.09	3,168	5.73**
	Vignette: Hostile Intent Attribution	-.05	.03	-.11-.02				-.01	.02	-.05-.03			
	Vignette: Revenge Goals	.50	.28	-.03-1.04				.14	.17	-.18-.47			
	VR: Anger	.01	.03	-.06-.07				-.01	.03	-.08-.06			
	VR: Hostile Intent Attribution	-.03	.04	-.10-.04				.07*	.03	.01-.14			
	VR: Revenge Goals	.79***	.19	.42-1.13				.35*	.16	.02-.65			
1	VR: Anger	.01	.03	-.06-.08	.13	3,172	8.67***	.01	.03	-.06-.07	.11	3,171	7.14***
	VR: Hostile Intent Attribution	-.04	.03	-.11-.03				.07*	.03	.01-.13			
	VR: Revenge Goals	.83***	.19	.47-1.18				.35*	.15	.05-.63			
2	VR: Anger	.01	.03	-.06-.07	.03	3,169	2.15	-.01	.03	-.08-.06	.02	3,168	1.31
	VR: Hostile Intent Attribution	-.03	.04	-.10-.04				.07*	.03	.01-.13			
	VR: Revenge Goals	.79***	.19	.42-1.15				.35*	.16	.03-.63			
	Vignette: Anger	.02	.03	-.05-.08				.04	.03	-.01-.09			
	Vignette: Hostile Intent Attribution	-.05	.03	-.12-.02				-.01	.02	-.05-.03			
	Vignette: Revenge Goals	.50	.28	-.04-1.10				.14	.17	-.19-.49			

4

Table 4 *Hierarchical Regression Analyses of Real-Life Aggression Regressed both on Reactive SIP and Aggressive Responses (continued)*

		Social Provocation						Object Provocation					
Step	Predictor	β	β SE	95% CI	ΔR^2	*df*	*F* change	β	β SE	95% CI	ΔR^2	*df*	*F* change
1	Vignette: Aggressive Responses	.68*	.26	.19-1.21	.06	1,174	11.12**	.44**	.15	.16-74	.05	1,173	9.80**
2	Vignette: Aggressive Responses	.48	.25	-.02-1.00	.09	1,173	18.07***	.30*	.14	.02-.58	.07	1,172	14.04***
	VR: Aggressive Responses	.60***	.15	.33-.89				.53***	.13	.28-.79			
1	VR: Aggressive Responses	.68***	.15	.38-.99	.12	1,174	23.86***	.62***	.13	.35-.87	.10	1,173	19.80***
2	VR: Aggressive Responses	.60***	.15	.32-.90	.03	1,173	5.78*	.53***	.13	.27-.79	.02	1,172	4.39*
	Vignette: Aggressive Responses	.48	.25	-.01-.98				.30*	.14	.02-.58			

Note. Hierarchical Regression Analyses were run for the Two Provocation Scenarios separately, both with Vignettes and VR Entered First. Model output is based on a non-bootstrapped procedure whereas output on separate predictors is based on a bootstrapping procedure.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Predicting Reactive & Proactive Motives

Next, we conducted the same set of hierarchical regression analyses as for children's real-life aggressive behavior, in this case predicting children's reactive and proactive motives for aggression. Detailed results of these analyses are provided in the Supplementary Materials (Table S2 and S3).

Aggressive SIP

As predicted, children's aggressive SIP in all four VR scenarios significantly predicted their reactive and proactive motives in real life, with explained variances at Step 1 ranging from 6 to 10% across scenarios. Effects were less pronounced for vignettes. Children's aggressive SIP assessed with vignettes significantly predicted their reactive and proactive motives in the object acquisition scenario (R^2 = .03) and social provocation scenario (R^2 = .06), but not in the competition (R^2 = .02) and object provocation scenario (R^2 < .01). Turning to the incremental value of VR, we found that VR entered at Step 2 explained significant variance over and above vignettes in all scenarios (i.e., 6% in object acquisition, 5% in competition, 12% in social provocation, and 11% in object provocation). In contrast, we found that vignettes at Step 2 explained significant variance over and above VR only in the social provocation scenario (i.e., 8%).

Aggressive Responses

Children's aggressive responses in all four VR scenarios significantly predicted their reactive and proactive motives in real life, with explained variances at Step 1 ranging from 5 to 9% across scenarios. Effects were weaker for vignettes. Children's aggressive responses assessed with vignettes significantly predicted their reactive and proactive motives in the object acquisition scenario (R^2 = .03) and competition scenario (R^2 = .05), but not in the social provocation (R^2 = .01) and object provocation scenario (R^2 < .01). Turning to the incremental value of VR, we found that VR entered at Step 2 explained significant variance over and above vignettes in all scenarios (i.e., 5% in object acquisition, 3% in competition, 8% in social provocation, and 8% in object provocation). In contrast, we found that vignettes at Step 2 explained significant variance over and above VR only in the competition scenario (i.e., 3%).

In sum, all eight hierarchical regression analyses regarding children's reactive and proactive motives supported the incremental value of VR over vignettes, whereas only two analyses supported the reverse.

DISCUSSION

This study tested whether interactive Virtual Reality (VR) provides a more ecologically valid assessment of social information processing (SIP) underlying aggressive behavior in children than a standard vignette-based assessment. In line with expectations, children reported that the interactive VR assessment was more emotionally engaging and immersive than the vignette-based assessment. Moreover, the assessment of children's aggressive SIP and responses in VR predicted their real-life aggressive behavior and reactive and proactive motives for aggression, above and beyond the vignette assessment.

Interactive VR immerses children in emotionally engaging social interactions and enables them to actually aggress against virtual peers — an important difference with vignettes, which ask children to consider their hypothetical aggressive responses. Accordingly, interactive VR has evoked higher levels of aggressive SIP and responses in children in provocation scenarios, and improved the predictive validity of their assessed aggressive SIP and responses. These findings support the proposition that emotional engagement influences SIP and consequent behavior (Lemerise & Arsenio, 2000; Anderson & Bushman, 2002). Thus, the emotionally engaging nature of our interactive VR assessment seems to have triggered aggressive SIP patterns and responses that may only occur with sufficient emotional engagement.

We expected that the engaging nature of interactive VR would also evoke larger individual differences in children's aggressive SIP and responses, and stronger correlations between children's aggressive SIP and responses compared to vignettes. However, interactive VR and vignettes generally evoked similar variances in children's aggressive SIP and responses, and similar correlations between aggressive SIP steps and responses. Perhaps, our vignettes validly assessed individual differences in children's "calm" SIP; that is, the way they would reflect on social situations when they do not experience strong emotions. Such "calm" SIP may also differ between children and show similar correlations between children's SIP and responses as their emotional SIP, but would be less suitable to predict children's real-life aggression. Indeed, our findings showed that interactive VR yielded incremental predictive value above and beyond the vignette-based assessment in all four scenarios, both for the prediction of children's real-life aggression (i.e., 2 to 12% additional explained variance) and underlying reactive and proactive motives (i.e., 3 to 12% additional explained variance).

One unexpected pattern in our findings was that interactive VR seemed to outperform vignettes more so for provocation scenarios than for instrumental gain scenarios: the incremental predictive value of VR versus vignettes was larger in provocation scenarios (with 7 to 12% increases in explained variance in children's real-life aggression) than in instrumental gain scenarios (with 2 to 5% increases in explained variance in children's real-life aggression), and only in provocation scenarios children showed more aggressive SIP and responses in VR versus vignettes. Although we expected that the engaging nature of interactive VR would enhance children's proactive aggressive tendencies in instrumental gain scenarios as well (e.g., because the stakes are higher, so they would experience more jealousy or desire), children did not show more proactive SIP and responses in VR versus vignettes. Possibly, the provocation scenarios were more salient than the instrumental gain scenarios, because the instrumental gain of points in the VR constituted no actual gain outside of the game in the real world. As such, it makes sense that the incremental value of interactive VR was the largest for children's aggressive SIP and aggressive responses in provocation scenarios. In sum, interactive VR seems to yield an improved assessment of both children's reactive SIP and proactive SIP patterns and responses compared to vignettes, but the difference in favor of interactive VR is the largest when measuring children's reactive SIP patterns and aggression.

Several findings on separate SIP steps warrant further discussion. First, children's interactional goals were the strongest SIP predictor of their real-life aggression and underlying motives, and yielded the largest effect sizes for levels of aggressive SIP in VR versus vignettes. Moreover, as children's revenge goals were strongly correlated with their anger and hostile intent attributions, they were the only significant SIP step predicting children's real-life aggression and reactive motives for aggression in most analyses. Although such overlap among predictors (i.e., multicollinearity) may seem problematic from a statistical point of view, it does make sense conceptually, because children's interactional goals seem to be most proximal to their (aggressive) behavior and may often derive from preceding SIP steps such as anger and hostile intent attributions (Crick & Dodge, 1994).

Second, contrary to our predictions, children reported similar levels of anger in interactive VR and vignettes, and even more anger with vignettes in the object provocation scenario. This finding contrasts with our finding that VR is more emotionally engaging than vignettes. However, it may also reveal a potential limitation of vignettes: asking children to reflect on their anticipated anger in a hypothetical situation could lead them to overestimate how they would actually feel. Indeed, research has shown that individuals generally find it difficult to report on anticipated negative affective states and tend to overestimate them (Robinson & Clore, 2002). Although we do not know whether this was actually the case, the stronger correlations of VR- versus vignette-assessed anger with children's real-life aggression indicate that children are more accurate when reporting on their anger in interactive VR. Perhaps, as in interactive VR children are actually (not just hypothetically) provoked or tempted to use aggression, they may experience emotions more similar to daily life than the anticipated emotions assessed with vignettes.

This study had several strengths. To our knowledge, it is the first empirical study that used interactive VR to assess children's aggressive SIP and responses and compared its external validity directly to a standard vignette-based assessment of children's aggressive SIP and responses. Moreover, we maximized clinically meaningful variance in children's SIP by recruiting boys from both regular and special education for disruptive behavior problems. The use of interactive VR in a sample with substantial variance in children's SIP allowed us to test important hypotheses concerning the validity of VR-based assessment of SIP.

Our study also had its limitations. First, as few children responded with aggression in the same scenarios in both VR and vignettes, we were not able to analyze whether interactive VR provides an improved assessment of children's positive outcome expectancies and response evaluations of aggression, as these were assessed only if children had actually aggressed. Consequently, we tested our hypotheses on children's proactive SIP and responses in instrumental gain scenarios using two variables only (i.e., instrumental goals and aggressive responses). Second, children's responses were coded for reliability by the first author, who may have been biased because he was aware of the research questions. However, inter-rater agreement with a second coder who was blind to the research questions was excellent, suggesting that this bias was limited. Third, as interactive VR is obviously more time-consuming and costly to develop than vignettes, we were only able to include four assessment scenarios. Given that children may show aggression in various contexts (De Castro & Van Dijk, 2017), it can be assumed that using only four scenarios involving playing games with peers in a school-setting did not cover the broad range of social situations known to evoke aggression in children. Fourth, as children's SIP and responses were only weakly to moderately associated across scenarios, we conducted our analyses for each scenario separately. Although this finding aligns with empirical research demonstrating that children's aggression is situation-specific (e.g., De Castro & Van Dijk, 2017; Matthys et al., 2001), it prohibited us from testing how reliable our SIP measurements were per type of scenario. Last, since our study included only boys between 7-13 years with limited diversity in cultural and socio-economic background, findings cannot directly be generalized to girls, older, or younger children, or children from other cultural or socio-economic backgrounds than our sample.

There are both advantages and disadvantages of using interactive VR to assess children's aggressive SIP and responses. One important disadvantage is that interactive nature of VR makes establishing ambiguity of social situations more difficult than with vignettes. This interactive nature might enhance the experience of an actual social interaction, however it might also affect ambiguity to some extent (e.g., children who talked a lot with the virtual peer during the interaction might be prone to attribute non-hostile intent). Moreover, interactive VR is obviously costly and time-consuming to develop, and so it is relevant to directly compare this method to other assessment methods besides vignettes, such as video game tasks (e.g., Yaros et al., 2014).

That said, VR has multiple advantages over the use of vignettes. In interactive VR, children are actually provoked, tempted to use aggression, and able to aggress against virtual peers, and may therefore experience similar emotions as in real-life (e.g., anger, frustration), activating similar SIP patterns and responses as in real-world interactions. As such, researchers may examine the effect of a broad range of emotions on children's SIP; that is, not only anger or frustration, but also shame, guilt, fear, desire or sadness. Relatedly, since children actually 'respond' in VR, it is possible to include physiological indicators of children's arousal, permitting researchers to test more specific hypotheses on the role of emotional arousal in children's SIP and responses. Moreover, the large experimental control over social stimuli provided by interactive VR (e.g., control over virtual peers' nonverbal behaviors and emotional expressions) allows researchers to test more specific hypotheses about causal effects of subtle social cues on children's SIP and responses than has been feasible thus far. In addition, interactive VR may allow researchers to use a broad variety of emotionally engaging contexts known to evoke aggression in children. For example, researchers may present children with more

salient cues to evoke proactive SIP and aggression (e.g., by allowing children to obtain actual gains outside of the VR environment). Relatedly, researchers may also examine children's SIP in other settings than playing games with peers, such as settings with parents, settings which do not involve play, settings that allow for the assessment of relational aggression (e.g., spreading rumors), or settings that better allow to examine cooperative behaviors. Last, using interactive VR may minimize cognitive load), increasing the validity of children's reported SIP.

In sum, this empirical study demonstrated that interactive VR is an improved method to assess children's aggressive SIP and behavior compared to a standard vignette-based assessment. The use of VR allows researchers and practitioners to assess aggressive SIP patterns in an emotionally engaging, ecologically valid context that is truly interactive and realistic. Ultimately, interactive VR may also facilitate interventions with children, because it allows for extensive practice with the specific situations relevant to their individual needs, with precise control to adapt difficulty and complexity during the intervention. Moreover, practitioners may use cooperative contexts that yield rewards for specific desirable behaviors (e.g., prosocial), reinforcing these behaviors repeatedly through operant conditioning. As such, interactive VR may further our understanding of the SIP mechanisms underlying aggressive behavior problems in children and may enhance assessment and intervention for children with aggressive behavior problems.

SUPPLEMENTARY MATERIALS

Table S1 *Cross-tabulations of Dichotomous SIP and Response Variables per Scenario for VR and Vignettes*

		Object Acquisition		Competition		Social Provocation		Object Provocation	
		Vignette		Vignette		Vignette		Vignette	
	VR	No	Yes	No	Yes	No	Yes	No	Yes
Revenge Goals	No	177 (98.9%)	1 (0.6%)	161 (89.4%)	13 (7.2%)	119 (66.5%)	9 (5.0%)	62 (34.8%)	26 (14.6%)
	Yes	1 (0.6%)	0 (0.0%)	5 (2.8%)	1 (0.6%)	42 (23.5%)	9 (5.0%)	48 (27.0%)	42 (23.6%)
Instrumental Goals	No	115 (64.2%)	23 (12.8%)	129 (71.7%)	15 (8.3%)	159 (88.8%)	4 (2.2%)	159 (89.3%)	7 (3.9%)
	Yes	23 (12.8%)	18 (10.1%)	25 (13.9%)	11 (6.1%)	14 (7.8%)	2 (1.1%)	12 (6.7%)	0 (0.0%)
Aggressive Responses	No	114 (63.7%)	23 (12.8%)	115 (63.9%)	23 (12.8%)	102 (57.0%)	8 (4.5%)	55 (30.9%)	20 (11.2%)
	Yes	23 (12.8%)	19 (10.6%)	25 (13.9%)	17 (9.4%)	53 (29.6%)	16 (8.9%)	47 (26.4%)	56 (31.5%)
Positive Outcomes of Aggression	No	10 (52.6%)	4 (21.1%)	9 (52.9%)	0 (0.0%)	15 (93.8%)	0 (0.0%)	52 (92.9%)	4 (7.1%)
	Yes	4 (21.1%)	1 (5.3%)	8 (47.1%)	0 (0.0%)	1 (6.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)

Note. Odds Ratios in the Main Manuscript are Calculated through the Ratio of Discordant Cells.

Table S2 *Hierarchical Regression Analyses of Proactive Motives in Real Life Regressed on both Instrumental Goals and Aggressive Responses*

		Object acquisition						Competition					
Step	Predictor	β	β SE	95% CI	ΔR2	df	F change	β	β SE	95% CI	ΔR2	df	F change
1	Vignette: Instrumental Goals	.33	.17	.00-.68	.03	1,157	4.59*	.33	.22	-.08-.76	.02	1,157	3.00
2	Vignette: Instrumental Goals	.22	.17	-.11-.55	.06	1,156	10.06**	.20	.22	-.21-.64	.05	1,156	7.65**
	VR: Instrumental Goals	.50**	.17	.16-.84				.45**	.18	.10-.78			
1	VR: Instrumental Goals	.55**	.16	.23-.86	.08	1,157	12.90***	.49**	.17	.16-.83	.06	1,157	9.63**
2	VR: Instrumental Goals	.50**	.17	.17-.81	.01	1,156	1.95	.45*	.18	.10-.80	.01	1,156	1.14
	Vignette: Instrumental Goals	.22	.17	-.10-.55				.20	.22	-.20-.64			
1	Vignette: Aggressive Responses	.34*	.17	.01-.69	.03	1,157	4.96*	.43**	.15	.15-.72	.05	1,157	7.92**
2	Vignette: Aggressive Responses	.22	.17	-.11-.57	.05	1,156	9.25**	.34*	.15	.06-.64	.03	1,156	5.74*
	VR: Aggressive Responses	.47**	.17	.16-.80				.36*	.16	.05-.68			
1	VR: Aggressive Responses	.53**	.16	.20-.86	.07	1,157	12.28**	.44**	.16	.14-.76	.05	1,157	8.86**
2	VR: Aggressive Responses	.47**	.17	.14-.81	.01	1,156	2.11	.36*	.16	.05-.69	.03	1,156	4.82*
	Vignette: Aggressive Responses	.22	.17	-.09-.54				.34*	.16	.05-.65			

Note. Hierarchical Regression Analyses were run for the Two Instrumental Gain Scenarios separately, both with Vignettes and VR Entered First. Model output is based on a non-bootstrapped procedure whereas output on separate predictors is based on a bootstrapping procedure.

* $p < .05$; ** $p < .01$; *** $p < .001$.

4

Table S3 *Hierarchical Regression Analyses of Reactive Motives in Real Life Regressed on both Reactive SIP and Aggressive Responses*

		Social Provocation						Object Provocation					
Step	Predictor	β	β SE	95% CI	ΔR^2	*df*	*F* change	β	β SE	95% CI	ΔR^2	*df*	*F* change
1	Vignette: Anger	-.03	.03	-.09-.04	.06	3,154	3.18*	.00	.04	-.06-.07	< .01	3,153	0.21
	Vignette: Hostile Intent Attribution	-.06*	.03	-.12--.01				-.02	.03	-.07-.03			
	Vignette: Revenge Goals	.46	.24	.00-.93				-00	.18	-.34-.34			
2	Vignette: Anger	-.03	.03	-.09-.04	.12	3,151	7.55***	-.01	.03	-.08-.05	.11	3,150	6.41***
	Vignette: Hostile Intent Attribution	-.09**	.03	-.15--.04				-.02	.02	-.07-.03			
	Vignette: Revenge Goals	.37	.24	-.08-.85				-.10	.18	-.43-.23			
	VR: Anger	-.00	.03	-.07-.06				.02	.03	-.05-.08			
	VR: Hostile Intent Attribution	.04	.03	-.02-.10				.04	.03	-.03-.11			
	VR: Revenge Goals	.58**	.18	.22-.94				.48**	.16	.16-.80			
1	VR: Anger	-.04	.03	-.09-.03	.10	3,154	5.65**	.01	.03	-.05-.07	.10	3,153	5.87**
	VR: Hostile Intent Attribution	.01	.03	-.04-.06				.04	.03	-.03-.10			
	VR: Revenge Goals	.66**	.18	.32-1.03				.46**	.16	.15-.77			
2	VR: Anger	-.00	.03	-.07-.06	.08	3,151	5.04**	.02	.03	-.04-.08	.01	3,150	0.79
	VR: Hostile Intent Attribution	.04	.03	-.02-.10				.04	.04	-.03-.11			
	VR: Revenge Goals	.58**	.18	.22-.94				.48**	.16	.18-.79			
	Vignette: Anger	-.03	.03	-.09-.04				-.01	.03	-.07-.05			
	Vignette: Hostile Intent Attribution	-.09**	.03	-.15--.04				-.02	.03	-.07-.03			
	Vignette: Revenge Goals	.37	.24	-.08-.85				-.10	.18	-.43-.22			
1	Vignette: Aggressive Responses	.30	.20	-.09-70	.01	1,156	2.01	.07	.15	-.23-.36	< .01	1,154	0.18
2	Vignette: Aggressive Responses	.15	.22	-.28-.59	.08	1,155	12.91***	-.10	.16	-.42-.22	.08	1,151	13.20***
	VR: Aggressive Responses	.54**	.15	.24-.84				.57**	.17	.24-.89			
1	VR: Aggressive Responses	.56***	.15	.27-.84	.09	1,156	14.06***	.54**	.15	.24-.84	.08	1,154	13.04***
2	VR: Aggressive Responses	.54**	.16	.23-.84	< .01	1,155	0.51	.57**	.16	.24-.89	< .01	1,151	0.40
	Vignette: Aggressive Responses	.15	.21	-.29-.59				-.10	.16	-.41-.19			

Note. Hierarchical Regression Analyses were run for the Two Provocation Scenarios separately, both with Vignettes and VR Entered First. Model output is based on a non-bootstrapped procedure whereas output on separate predictors is based on a bootstrapping procedure.
$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$.

4

CHAPTER 5

Detecting Social Information Processing Profiles of Boys with Aggressive Behavior Problems: An Interactive Virtual Reality Approach

This chapter is submitted as:
Verhoef, R. E. J., Van Dijk. A., Thomaes, S., Verhulp, E. E., Van Rest, M. M., & De Castro, B. O. (2022). Detecting social information processing profiles of boys with aggressive behavior problems: An interactive virtual reality approach. Manuscript submitted for publication.

Upon publication the data and syntax that support the findings of this study will be available at: https://osf.io/ja9uw/

Author contributions:
RV conceptualized the study, AD, ST, EV, MR, and BC gave advice and feedback; RV and trained graduate students collected the data; RV analyzed the data and drafted the manuscript; AD, ST, EV, MR, and BC provided feedback on the analyses and manuscript.

ABSTRACT

Children with aggressive behavior problems may aggress for very different reasons, requiring tailored assessment and treatment. The primary aim of this study was to test whether it is possible to detect distinct social information processing (SIP) profiles among boys with aggressive behavior problems. For this purpose, we conducted Latent Profile Analyses (LPA) on boys' SIP patterns assessed in interactive Virtual Reality (VR). In addition, we examined the discriminant validity of these SIP profiles by comparing them on theoretically relevant child characteristics (i.e., temperament, executive functioning, aggressive belief systems, punishment insensitivity, and sensation seeking). We presented boys (*N* = 181; ages 7-13) with a virtual classroom where they could play games with virtual peers. They reported on their SIP in four VR scenarios, designed to assess both reactive and proactive aggressive SIP. Results of the LPA revealed four distinct SIP profiles: a general reactive SIP profile, a situation-specific reactive SIP profile, a mixed reactive-proactive SIP profile, and a nonaggressive SIP profile. Planned contrasts revealed that boys with these SIP profiles differed in temperament, aggressive belief systems, and punishment insensitivity, but not in executive functioning and sensation seeking. Overall, findings suggest that boys may differ in the exact SIP patterns underlying their aggressive behavior, providing inroads to tailor interventions to children's individual needs.

Keywords: Social information processing, Aggression, Children, Virtual Reality, Latent profile analysis

DETECTING SOCIAL INFORMATION PROCESSING PROFILES OF BOYS WITH AGGRESSIVE BEHAVIOR PROBLEMS: AN INTERACTIVE VIRTUAL REALITY APPROACH

Children may engage in aggressive behavior for very different reasons. Some children may be easily angered, or prone to take revenge when they feel provoked. Other children may carefully plan their aggressive behavior, hoping to instrumentally benefit from it (for reviews, see: De Castro & Van Dijk et al., 2017; Dodge, 2011). It makes sense, then, to try to identify profiles of children with aggressive behavior problems, based on how they process social information. Yet, detecting distinct social information processing (SIP) profiles of children with aggressive behavior problems is difficult, especially when using conventional SIP assessment methods that ask children how they *would* think or feel in hypothetical situations (e.g., vignette-based measures). Such methods may insufficiently tap aggressive SIP as it often occurs in vivid, emotionally-laden contexts (Anderson & Bushman, 2002; Lemerise & Arsenio, 2000), thereby underestimating individual differences in SIP patterns between children. In the current study, we therefore assessed children's aggressive SIP using an interactive Virtual Reality (VR) environment in which children played games with virtual peers—a context that resembles real-world interaction and evokes strong emotions in children (Chapter 4 of this dissertation). The primary aim of our study was to distinguish SIP profiles of children with aggressive behavior problems by conducting latent profile analyses (LPA) on their aggressive SIP patterns assessed in interactive VR. Detecting distinct SIP profiles in children with aggressive behavior problems may uncover new possibilities to tailor cognitive-behavior interventions to the needs of individual children.

Social Information Processing in Children with Aggressive Behavior Problems

The SIP model (Crick & Dodge, 1994) offers a useful framework to try to distinguish SIP profiles in children with aggressive behavior problems. The model postulates that children's aggressive behavior in social situations derives from deviancies in a sequence of SIP steps: (1) encoding of social cues, (2) mental representation of social cues, (3) setting interactional goals, (4) generation of behavior options, (5) evaluation of behavior options, and (6) behavior enactment. Emotional processes are often implicated in each of these SIP steps (Lemerise & Arsenio, 2000). Over the past decades, researchers have gained considerable understanding of the SIP deviancies contributing to children's aggressive behavior, such as encoding social cues in a hostile manner, making hostile intent attributions, setting interactional goals directed at revenge or instrumental gain, generating more aggressive responses, and evaluating aggressive responses and their outcomes more positively (for reviews, see: De Castro & Van Dijk, 2017; Dodge, 2011; Chapter 2 of this dissertation). However, there appear to be striking differences between children in which SIP steps are implicated in their aggressive behavior. In the present study, we examined whether clusters of children with distinct SIP patterns can be discerned.

One relevant dimension to distinguish SIP profiles of children with aggressive behavior problems may be the well-known distinction between reactive and proactive aggression (e.g., Dodge, 1991; Hubbard et al., 2010; Kempes et al., 2005). Reactive aggression refers to hot-blooded, defensive, uncontrolled aggressive behavior triggered by perceived threat, provocation, or frustration. Proactive aggression, in contrast, refers to cold-blooded, offensive, controlled aggressive behavior driven by a desired goal such as acquiring material gain or social dominance. Children may differ in which type of aggression they display most. Research has typically identified predominantly reactive subgroups and mixed reactive-proactive subgroups of children, and sometimes also a predominantly proactive subgroup (Carroll et al., 2018; Euler et al., 2017; Marsee et al., 2014; Muñoz et al, 2008; Smeets et al., 2017; Thomson & Centifanti, 2018; Van Dijk et al., 2021). Scholars have suggested that children displaying predominantly reactive versus proactive aggression may show deviancies in different SIP steps (for reviews, see: De Castro & Van Dijk, 2017; Hubbard et al., 2010; Merk et al., 2005; Vitaro et al., 2006). That is, reactively aggressive children may experience anger, attribute hostile intent, and set revenge goals. In contrast, proactively aggressive children may set instrumental goals, expect positive outcomes of their aggressive behavior, and evaluate the use of aggression positively.

The empirical evidence thus far, however, is not conclusive. First, only a part of the available studies has found that children's SIP steps are differentially associated with reactive versus proactive aggressive behavior, suggesting that it may be possible to detect distinct SIP profiles of children with aggressive behavior problems based on their reactive and proactive SIP patterns (e.g., Arsenio et al., 2009; Hubbard et al., 2001; Dodge et al., 1997; De Castro et al., 2005). Other studies did not replicate such findings (e.g., Crick & Dodge, 1996; Dodge et al., 1997; Oostermeijer et al., 2016; Stolz et al., 2013). These divergent findings may arise because previous studies used conventional vignette-based measures, which may not tap into vivid, emotionally-laden SIP, and therefore may be limited in their ability to detect individual differences in children's aggressive SIP (Chapter 2 of this dissertation). Second, to our knowledge, all previous SIP research used a variable-based approach, studying associations between SIP variables instead of clustering children into distinct groups based on their SIP patterns. Hence, even if our understanding of the SIP deviancies that underlie children's aggression is substantial, we do not yet know whether children with aggressive behavior problems systematically cluster in terms of their SIP characteristics. Thus, research is needed that uses a) emotional, engaged interactions to assess substantial variance in SIP and b) person-based analytical methods to cluster this variance within children.

A Person-Based Interactive VR Approach to Detect Distinct SIP Profiles

To address these issues, we used a person-based analytical approach to examine children's SIP patterns in the context of emotionally engaging interactive VR. Interactive VR allows to assess SIP in a vivid, emotionally arousing context—a context in which individual differences in aggressive SIP tend to become salient (Anderson & Bushman, 2002; Lemerise & Arsenio, 2000). Indeed, a meta-analysis showed that children's aggressive behavior is better predicted by SIP assessed in actual social interactions, as compared to hypothetical social interactions described in vignettes (Chapter 2 of this dissertation). Interactive VR provides a context in which children are immersed in actual social interactions with virtual peers—but then in a standardized manner, allowing for more accurate assessment of individual differences in children's aggressive SIP.

Another advantage of interactive VR is that it allows for creating distinct, theoretically relevant contexts of interaction. For example, to assess reactive SIP, children can be presented with provocation contexts that may elicit anger or frustration (e.g., being excluded from a game, or being hindered by a peer; Hubbard et al., 2010). To assess proactive SIP, children can be presented with instrumental gain contexts that may elicit envy and desire (e.g., competing against a peer, or having an opportunity to steal; Hubbard et al., 2010). In interactive VR, children actually enact these scenarios, making it likely that both provocation and instrumental gain contexts can be truly engaging for children. We thus expected that interactive VR would enable us to detect distinct reactive and proactive SIP profiles. Based on the literature, we expected four SIP profiles: a reactive SIP profile, a proactive SIP, a mixed reactive-proactive SIP profile, and a nonaggressive SIP profile (see Table 1 for our hypotheses on specific SIP patterns within each SIP profile).

5

Table 1 *Hypothesized Scores on SIP Variables for Boys in each SIP Profile*

SIP variables	SIP profile			
	Reactive	Proactive	Mixed	Nonaggressive
Anger	High	Low	High	Low
Hostile Intent Attributions	High	Low	High	Low
Revenge Goals	High	Low	High	Low
Instrumental Goals	Low	High	High	Low
Positive Outcomes of Aggression	Low	High	High	Low
Positive Evaluations of Aggression	Low	High	High	Low
Aggressive responding	High	High	High	Low

Discriminant Validity of SIP Profiles

A secondary aim of this study was to provide further validation of the existence of distinct SIP profiles in children with aggressive behavior problems. We therefore examined whether children with different SIP profiles differed on 1) teachers' impressions of children's reactive and proactive aggression, and 2) theoretically relevant child characteristics (i.e., temperament, executive functioning, aggressive belief systems, punishment insensitivity, and sensation seeking; De Castro & Van Dijk, 2017; Dodge & Pettit, 2003; Frick & Morris, 2004; Hubbard et al., 2010; Merk et al., 2005; Chapter 6 of this dissertation). We expected each SIP profile to stand out on unique characteristics.

Children with a reactive SIP profile may be characterized by a highly emotionally reactive temperament, which predisposes them to experience excessive anger or frustration in social interactions (for reviews, see: Bookhout et al., 2018; Moore et al., 2018). Moreover, these children may be prone to attribute hostile intent, which may stem from hostile memory structures and limited working memory capacities impeding their accurate processing of social events (for reviews, see: Dodge, 2006; de Castro & Van Dijk, 2017; Chapter 6 of this dissertation). Last, these children's tendency to promptly seek revenge may be rooted in inhibition deficits (Ellis et al., 2009; Thomson & Centifanti, 2018).

In contrast, children with a proactive SIP profile may be characterized by callous and unemotional (CU) traits. These traits may predispose them to value aggression as a useful strategy to obtain instrumental gain (for reviews, see: Frick et al., 2014a, 2014b; Frick & Morris, 2004). Moreover, these children may be prone to pursue instrumental goals and hold positive outcome expectancies of aggressive behavior, stemming from a moral belief system that justifies the use of aggression (Calvete & Orue, 2010; Zelli et al., 1999). Last, these children's positive expectations and evaluations of aggression may be rooted in their sensation seeking tendencies and insensitivity to punishment. They seek for thrills (for a review, see: Matthys et al., 2013), and are relatively unaffected by negative consequences of their behavior (for a review, see: Branje & Koot, 2018).

Last, children with a mixed reactive-proactive SIP profile would display characteristics of both groups, whereas children with a nonaggressive SIP profile would share none of these characteristics (see Table 2 for an overview of the expected differences between SIP profiles).

Table 2 *Theoretically Relevant Child Characteristics that Boys in the Reactive, Mixed, and Proactive SIP Profiles Are Expected to Stand Out For*

	Reactive & Mixed SIP Profile	Proactive & Mixed SIP Profile
Teachers' Impression	Reactive Motives	Proactive Motives
Child Characteristics	Anger-Frustration Temperament	Callous & Unemotional Traits
	Hostile Beliefs	Justification of Violence Beliefs
	Working Memory deficits	Sensation Seeking
	Inhibition deficits	Punishment Insensitivity

Note. We used planned contrasts to test if scores were higher for the two expected profiles (e.g., Reactive and Mixed) versus the other two profiles (e.g., Proactive and Nonaggressive).

The Present Study

Our study goals were to (1) detect distinct SIP profiles underlying children's aggressive behavior, and (2) validate these profiles against teachers' impressions and theoretically relevant child characteristics. We used interactive VR to assess emotionally-engaged SIP, presenting participants with an interactive VR classroom where they played games with virtual peers. To assess participants' reactive and proactive SIP patterns, we presented them with both provocation and instrumental gain contexts. To validate the obtained SIP profiles, we asked teachers to rate participants' reactive and proactive aggression at school, parents to rate their children's temperament, and participants themselves to complete questionnaires and tasks assessing their traits, beliefs, and executive functioning. Our study included only boys, because the development of VR is quite costly and assessing aggressive SIP in girls would have required us to develop additional VR scenarios relevant for girls' aggression (Ostrov & Godleski, 2010). We expected to find four distinct SIP profiles of boys with aggressive behavior problems (i.e., reactive, proactive, mixed, and nonaggressive; Table 1), differing on teachers' impressions and theoretically relevant child characteristics (Table 2).

METHOD

Participants

Participants were *N* = 181 boys ages 7 to 13 (*M* = 10.23; *SD* = 1.27), recruited from 18 Dutch primary schools. Schools were from neighbourhoods representative of the Dutch population, with on average mostly native Dutch middle class inhabitants, and a minority of inhabitants with a migration background (Western: 8.6%, *SD* = 2.5%; non-Western: 13.0%, *SD* = 9.5%), a lower educational level (20.9%, *SD* = 4.3%), or a low income (7.5% of households, *SD* = 3.15%) (Statistics Netherlands, 2018, 2019). To maximize variance in aggressive behavior, we included boys from special education for children with disruptive behavior problems (*n* = 115 boys), in addition to boys randomly selected from regular education (*n* = 66 boys). In the Netherlands, special education for children with disruptive behavior problems and/or psychiatric problems is offered to children whose problems are so severe that they require extra support. Our study included boys from special education who were nominated by their teacher for frequently showing aggressive behavior problems. Boys were excluded if they had an IQ below 80 or an Autism Spectrum Disorder (ASD) according to their casefiles, or showed Autism Spectrum Disorder symptomatology within the clinical range on the teacher-rated Social Emotional Questionnaire (Scholte & Van der Ploeg, 2007). Parents gave their written informed consent, boys themselves gave verbal assent. The study was conducted in accordance with the 2013 Helsinki Declaration and approved by the Dutch Medical-Ethical Testing Committee Utrecht (METC-Utrecht).

Procedure

Participants were individually tested in a silent room at their school by trained research assistants (i.e., undergraduate psychology students) or the first author. We tested participants in two 45-minute sessions, spaced about a week apart. Boys completed questionnaires and executive functioning tasks on a computer tablet in Session 1, and the VR-based SIP assessment in Session 2. Parents (95.0%) and teachers (98.3%) completed the questionnaires online, in their own time.

Interactive Virtual Reality Environment to Assess Aggressive SIP

Setting

The VR environment was built as a virtual school classroom. Participants could walk around freely, talk with virtual peers, and play games in the virtual classroom setting (see: <masked for blind review>). Participants played two games: (1) building a block tower, and (2) throwing cans from a table with a ball. We chose games to enhance emotional engagement, and augmented these games with high scores and bonuses to have experimental control over gain and losses. The games were created to allow participants to engage in aggression directed at the virtual peer (e.g., hitting, name calling) or the peer's property (e.g., knocking over the peer's tower). We provided the instructions, game rules, and score count within the VR, using a voice-over and a digital whiteboard. Before entering the VR environment, the experimenter explained participants that they would enter a classroom where specific behavior rules applied (i.e., be friendly to other children, have respect for others). We also told them that they would interact and play games in the VR environment with actual boys from other schools who simultaneously took part in the study. In reality, the experimenter controlled the virtual peers by activating default movements and standardized verbal responses.

VR Scenarios

Participants were presented with six VR scenarios in fixed order: (1) practice scenario, (2) neutral scenario, (3) object acquisition scenario, (4) competition scenario, (5) social provocation scenario, and (6) object provocation scenario. The practice scenario allowed participants to familiarize themselves with the VR environment, practice the games, and learn the classroom rules. The neutral scenario consisted of a brief interaction with an avatar, and was included to familiarize boys with answering our SIP questions. The next two scenarios (i.e., object acquisition, competition) covered the instrumental gain context. In the object acquisition scenario, participants could choose to steal a block or ball from the virtual peer, which would earn them additional points. In the competition scenario, participants could unfairly win the game by adjusting (i.e., setting back) the virtual peer's score. The last two scenarios (i.e., social and object provocation) covered the provocation context. In the social provocation scenario, participants were refused to join a game by two virtual peers. In the object provocation scenario, the game of the participant was ruined by a virtual peer (e.g., participant's

tower was knocked over by a peer). We presented the provocation scenarios last because we anticipated that these could elicit relatively strong emotion, potentially leading to carry-over effects if they would not be presented last. Participants completed all six scenarios for the same game (i.e., tower or cans). Games were randomly assigned to participants.

SIP Assessment using Interactive VR

At the end of each VR scenario, we assessed participants' anger, intent attributions, goals, outcome expectancies, evaluations of behavior, and behavioral responses. We emphasized that there were no wrong answers and that all responses would remain confidential. We analyzed participants' SIP for each scenario separately, and so we did not calculate aggregate scores (as is often done in vignette-based SIP research; Crick & Dodge, 1996; De Castro et al., 2005; Chapter 2 of this dissertation).

Anger. We assessed anger using a single item: "The other boy did [behavior other boy]. How angry did this make you feel on a scale from 1-10 (*not at all-very*)?"

Hostile Intent Attribution. We assessed intent attribution using two items: "The other boy did [behavior other boy]. To what extent was he trying to be mean, on a scale from 1-10 (*not at all-very*)?" and "To what extent was he trying to hinder you, on a scale from 1-10 (*not at all-very*)?" Within each VR scenario, the two items were highly correlated (M = .83, Mdn = .87, range = .67-.90). We averaged the items to create a single hostile intent attribution score for each VR scenario.

Goals. We assessed goals using a single open-ended question following each VR scenario: "When the other boy did [behavior other boy], you did [behavior participant]. What was the reason you did this?" Following existing guidelines (De Castro et al., 2012), we coded answers as *revenge-anger goals* (e.g., "because I was angry," "to retaliate," "to defend myself"), *instrumental goals* (e.g., "to win the game," "to show him who's boss"), *goals underlying nonaggressive behavior* (e.g., "to become friends," "to avoid problems"), or *no goals* (e.g., "I don't know," "I had no goal"). To test the inter-rater reliability of the scoring system, 35.4% of transcriptions were coded by a second rater. Inter-rater reliability was excellent, with κ ranging from .85-.96 (M = .91, Mdn = .91). Scores for revenge-anger goals were created by assigning 1 to *revenge-anger* codes and 0 to other codes. Similarly, scores for instrumental goals were created by assigning 1 to *instrumental* codes and 0 to other codes.

Behavioral Responses. We assessed behavioral responses in VR by observing participants' behavior in each VR scenario. We coded behavior afterwards using a well-established procedure (De Castro et al., 2005). We coded responses as *nonaggressive behavior* (e.g., prosocial, solution-focused, avoidance), *mild aggressive behavior* (e.g., coercion, verbal aggression), or *severe aggressive behavior* (e.g., physical aggression, destructive aggression). To test the inter-rater reliability of the scoring system, 35.4% of transcriptions were coded by a second rater. Inter-rater reliability was excellent, with κ ranging from .92-1.00 (M = .97, Mdn = .98). Because frequencies of *mild aggressive behavior* were low or even absent in all scenarios (i.e., 0.0 to 2.2%, Mdn = 0.6%), we

5

created a dichotomous variable for aggressive responding by scoring *mild* and *severe aggressive behavior* as 1 and *nonaggressive behavior* as 0.

Outcome Expectancies. We assessed outcome expectancies for aggression with a single item: "What did you expect would happen when you [behavior participant]?" Each answer was coded as: *positive instrumental outcomes of aggression* (e.g., "I would win the game") or *no positive instrumental outcomes of aggression* (e.g., "I would not receive bonus points"). To test the inter-rater reliability of the scoring system, 35.4% of transcriptions were coded by a second rater. Inter-rater reliability was excellent, with κ being 1.00. Scores for each outcome were created by assigning 1 to the specific outcome and 0 to other codes. Because we were interested in participants' outcome expectancies of *aggression*, we coded data of participants who displayed no aggression as missing.

Response Evaluations. We assessed positive evaluations of aggression using a single item: "When the other boy did [behavior other boy], you did [behavior participant]. To what extent do you approve this behavior on a scale from 1-10 (*not at all-very*)?". Again, we coded data of participants who displayed no aggression as missing.

Measures Used for Validation Purposes

Questionnaires

Teachers' Impressions of Reactive and Proactive Motives. We used the Instrument for Reactive and Proactive Aggression (IRPA) to assess teachers' impressions of boys' reactive and proactive motives (Polman et al., 2009). Teachers rated the frequency of 7 distinct forms of aggressive behavior (i.e., kicking, pushing, hitting, name calling, arguing, gossiping, and doing sneaky things) in the previous month on a 5-point Likert scale (0 = *never,* 1 = *once,* 2 = *weekly,* 3 = *multiple times a week,* 4 = *daily*). For aggression items rated above 0, teachers rated 6 items about the motives underlying boys' aggression on a 5-point Likert scale (0 = *never,* 1 = *rarely,* 2 = *sometimes,* 3 = *often,* 4 = *always*). For aggression frequency items rated 0, motives scores were missing by design. Three items described reactive motives (e.g., "because someone teased or upset him/her") and three items described proactive motives (e.g., "to hurt someone or to be mean"). We calculated reactive and proactive motive scores by averaging across all reactive motive items (i.e., 3 items for 7 forms of aggression; α = .81) and all proactive motive items (α = .83). Thus, high scores for reactive or proactive motive indicated that *if* children engaged in aggressive behavior, they often had reactive or proactive motives. The correlation between reactive and proactive motives was non-significant (r = .12, p = .146).

Anger-Frustration Temperament. We assessed anger-frustration temperament using the Dutch translation of the Anger/Frustration subscale of the Temperament for Middle Childhood Questionnaire (TMCQ: Simonds & Rothbart, 2004). Parents rated the extent to which 7 items applied to their child on a five-point Likert scale from 1 (*almost never true for my child*) to 5 (*almost always true for my child*), with "*does not apply*" as an additional option. A sample item is: "Has anger outbursts when he/she does not get what he/she wants." We calculated anger-frustration temperament scores as the average across items ($α$ = .84).

Hostile Beliefs. To assess hostile beliefs, we used 10 items derived from the Hostility subscale of the Child Automatic Thoughts Scale (CATS; Schniering & Rapee, 2002) and the Mistrust/Abuse subscale of the Schema Inventory for Children (SIC; Rijkeboer, & De Boo, 2009). Boys rated these items on a five-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*). A sample item is: "You can never trust someone". We calculated hostile belief scores as the average across items ($α$ = .85).

CU Traits. We assessed CU traits with the Callous & Unemotional Subscale of the Youth Psychopathic Inventory Child Version (YPI-CV; Van Baardewijk et al., 2009). Boys rated 15 items on a four-point Likert scale from 1 (*does not apply at all*) to 5 (*applies very well*). A sample item is: "When I have hurt other people's feelings, it doesn't really bother me." We calculated CU trait scores as the average across items ($α$ = .78).

Justification of Violence. We assessed justification of violence with 14 items derived from the How I Think Scale (HITS; Nas et al., 2008) and the The Irrational Beliefs Scale for Adolescents (IBSA; Cardeñoso & Calvete, 2004). Boys rated these items on a five-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*). A sample item is: "Sometimes you need to hurt or threaten someone to get what you want." We calculated justification of violence scores as the average across items ($α$ = .91).

Sensation Seeking. We assessed sensation seeking using the Brief Sensation Seeking Scale (BSSS; Dekkers et al., 2018). Boys rated 14 items on a five-point Likert scale from 1 (*strongly disagree*) to 5 (*strongly agree*). A sample item is: "If someone dares me to do something, I will do it." We calculated sensation seeking scores as the average across items ($α$ = .66).

Executive Functioning Tasks

Executive functioning tasks were presented in a game-based format on a tablet computer. The tasks were developed to be appealing for children (Van Rest et al., 2019).

Working Memory Deficits. We assessed visual working memory using a task based on the Klingberg principles for working memory (Klingberg, 2010; Klingberg et al., 2005). The task assesses boys' capacity to temporarily store and manipulate patterns of visual stimuli. The task consists of sequential trials in which participants are asked to replicate a visual pattern presented as a monkey or crocodile moving on a 4 × 4 check-like board. Participants listened to an instruction and conducted a sequence of four practice trials, before starting the test trials. These trials started easy (i.e., only two attached spaces) and increased stepwise in length (i.e., more spaces) and visual difficulty (i.e., detached spaces further apart). The monkey was presented for 1,000 milliseconds and disappeared for 750 milliseconds before appearing in another space. Participants were asked to replicate the monkey's movement pattern immediately after the monkey stopped moving. The test trials ended if boys had two consecutive incorrect responses on trials with the same length and difficulty level. Boys completed 12 trials, on average (range: 1-21). Next, boys took part in another round of trials, but this time they were asked to replicate the movement pattern of a crocodile in reversed order (i.e., starting with the last step of the crocodile, ending with the first step). Boys completed 9 of these trials, on average (range: 0-21). The number of correct trials was reverse-scored for each child so that higher scores represented more working memory deficits. We standardized and averaged scores on the backward and forward trials (r = .54) to create a single working memory deficits score.

Inhibition Deficits. We assessed response inhibition using a task based on the Go/No-go principle (Nigg, 1999). The task assesses children's ability to inhibit action tendencies, asking them to press a button as fast as they can when a stimulus is presented on screen. The task consists of two phases that each include an instruction and a sequence of practice trials. In the first (i.e., learning) phase, participants were asked to press an apple-shaped button as fast as they could when an elephant appeared on the screen. Participants were presented 52 trials where an elephant appeared on screen for a maximum of 800 milliseconds, until they responded. Each trial started with a fixation symbol that was presented for 1,000 milliseconds before the elephant emerged. Participants were instructed to not press the button during the presentation of the fixation symbol, but only when the elephant appeared on screen, requiring them to inhibit their response. The duration between each trial was 1,000 milliseconds. In the second (i.e., inhibition) phase, participants were again presented with 52 trials and instructed to respond as fast as they could when the elephant appeared on the screen. However, this time, they were instructed not to respond when the elephant was presented with a red cross through it. There were 39 trials including the elephant without a red cross (i.e., Go-trials) and 13 trials including the elephant with a red cross (i.e., No-go trials). Again, elephants (with or without the red cross) were presented on screen for 800 milliseconds, preceded by the fixation symbol. Presentation order of Go and No-go trials was fixed. We used the number of premature responses in the first (i.e., learning) phase and second (i.e., inhibition) phase, as well as the number of incorrect responses (i.e., pressing the button while the elephant was presented with a red cross through it) the second phase, as indicators of participants' inhibition abilities. These three scores were standardized and averaged (r ranging from .48-.49) to create a single inhibition deficits score.

Punishment Insensitivity. We assessed punishment insensitivity using a door-opening task used in previous studies (e.g., Matthys et al., 1998, Matthys et al., 2004), which is based on the card-playing task (Newman et al., 1987). In this task, participants were asked to open doors by pressing a button. With each door, participants could earn or lose 10 cent. Participants started with 0 cents and the task included 110 trials. The probability of winning decreased gradually by 10% with each 10 trials (i.e., from 100% in trials 1-10 to 0% in trials 100-110). Participants were instructed to win as much money as they could and were told they could stop playing at each trial by pressing a stop button. The order of winning and losing doors was fixed across participants. We used the elapsed time between a losing door and opening a next as indicator of participant's punishment insensitivity. We standardized and reverse-scored children's scores so that higher scores indicated more punishment insensitivity.

Statistical Analysis

Our primary aim was to detect SIP profiles of boys with aggressive behavior problems. Before we conducted our main analyses, we inspected multicollinearity between SIP variables. We found high correlations of aggressive responding with instrumental goals in the object acquisition (Pearson's π = .98) and competition scenario (Pearson's π = .91), and with revenge goals in the social provocation (Pearson's π = .80) and object provocation scenario (Pearson's π = .86). To avoid multicollinearity, we therefore only included revenge goals and instrumental goals in our analyses.

Next, we conducted Latent Profile Analyses (LPA) using Mplus (version 8.5). We included proactive SIP variables assessed in instrumental scenarios (i.e., instrumental goals, outcome expectancies, and response evaluation), and reactive SIP variables assessed in provocation scenarios (i.e., anger, hostile intent attribution, and revenge goals).[1] We first tested a single-profile model, and increased the number of profiles until the model no longer improved in terms of model fit, interpretability, and parsimony (McCutcheon, 2002). Regarding model fit, we aimed to select the model with the lowest Akaike information criterion (AIC), Bayesian information criterion (BIC), and sample size adjusted BIC (aBIC); sufficient entropy (i.e., entropy > .80); a non-significant Lo-Mendell-Rubin likelihood ratio test (LMRT; Lo et al., 2001) and bootstrapped likelihood ratio test (BLRT; McLachlan & Peel, 2000); and a sufficient number of observations in each profile (i.e., more than the number of parameters estimated).

Our secondary aim was to validate the obtained SIP profiles, using boys' most likely profile membership as independent variable. First, we compared the SIP profiles of the best-fitting model on teachers' impressions of boys' reactive and proactive motives for aggression. Second, we compared the SIP profiles of the best-fitting model on theoretically relevant child characteristics. We conducted planned contrasts to test our a priori hypotheses (Table 2). In addition, if we found that the ANOVA of SIP profile membership on a variable was significant, we explored all possible contrasts. Given the non-normal distribution of our variables, we conducted these analyses using bootstrapped bias-corrected accelerated (BCa) 95% confidence intervals (CI) based on 5000 resamples. We used pairwise deletion to deal with missing data (4.2% of which 2.3% missing by design on the IRPA).

1 We began running the LPA with reactive and proactive SIP variables assessed in both provocation and instrumental gain scenarios, but this model would not identify because it included too many parameters given the study sample size.

RESULTS

Descriptive Statistics

Interactive VR evoked aggressive responses in 38.3% to 58.3% of boys for the provocation scenarios, but only in 23.2% to 23.8% of boys for the instrumental gain scenarios (see Table S1 in the supplementary materials for descriptive statistics of SIP variables for each VR scenario separately). This frequency of aggression was lower than anticipated, and implied substantial missing data for those SIP variables that could only be assessed in the context of an aggressive response (i.e., positive outcome expectancies and positive evaluations of aggression). Accordingly, we excluded these variables from the LPA, and report descriptive statistics of these variables in the supplementary materials (Table S2).

Distinct SIP Profiles of Boys with Aggressive Behavior Problems

Table 3 shows the fit indices for the Latent Profile analyses. We selected the 4-profile model as the best-fitting model. This model had a better fit than the 3-profile model according to all three information criteria, entropy, and BLRT, although not according to VLMR, which was marginally non-significant. Although the 5-profile model fitted slightly better than the 4-profile model according to AIC and aBIC, entropy, and BLRT, it conceptually added little. That is, it showed a similar pattern as the 4-profile solution, with two profiles only slightly differing in mean scores and item probabilities of SIP variables. Thus, the 4-profile model provided a more parsimonious, and well-fitting solution.

Table 3 *Fit Indices for the LPA Models*

	AIC	BIC	aBIC	Entropy	VLMR p	BLRT p	n of smallest profile
1-profile model	4420.289	4458.671	4420.666				181
2-profile model	4137.596	4204.764	4138.256	0.866	0.0000	0.0000	67
3-profile model	4084.447	4180.402	4085.390	0.808	0.1815	0.0000	48
4-profile model	4045.023	4169.764	4046.249	0.823	0.0798	0.0000	27
5-profile model	4031.748	4185.276	4033.257	0.835	0.5243	0.0000	21
6-profile model*	4012.371	4194.686	4014.163	0.863	0.1482	0.0000	6

* The 6-profile solution yielded profiles that were too small (n = 6) given the number of free parameters to be estimated (q = 12).

Figure 1 shows a visual representation of the four SIP profiles based on their reactive and proactive SIP patterns per scenario (for descriptive statistics, see Table S2 in the supplementary materials). As predicted, we found evidence for a reactive SIP profile, mixed reactive-proactive SIP profile, and a nonaggressive SIP profile. Contrary to our predictions, we did not find a proactive SIP profile, and detected an additional situation-specific profile.

As expected, we found one "general reactive SIP" profile of n = 47 boys (26.0%), characterized by reactive SIP in both provocation scenarios. Boys with this profile showed very high levels[2] of anger (M = 8.38, SD = 1.48; M = 9.06, SD = 1.51), high levels of hostile intent attributions (M = 7.43, SD = 2.63; M = 8.70, SD = 1.88), and a moderate-to-high probability of displaying revenge goals (ρ = .426; ρ = .660) in the social provocation and object provocation scenario, respectively, and very low-to-zero probabilities of displaying instrumental goals in the object acquisition (ρ = .064) and competition scenario (ρ = .000).

In addition, and unexpectedly, we found a second reactive SIP profile of n = 59 boys (32.6%) who displayed reactive SIP only in the object provocation scenario. We refer to this group as the "situation-specific reactive SIP" profile. Boys with this profile showed high levels of anger (M = 6.84, SD = 2.20), very high levels of hostile intent attributions (M = 8.49, SD = 1.46), and a moderate probability of displaying revenge goals in the object provocation scenario (ρ = .517), but low levels of anger (M = 3.76, SD = 1.64), moderate levels of hostile intent attributions (M = 4.70, SD = 2.49) and a very low probability of displaying revenge goals (ρ = .138) in the social provocation scenario, as well as low probabilities of displaying instrumental goals in the object acquisition (ρ = .203) and competition scenario (ρ = .237).

5

We also found the expected "mixed reactive-proactive SIP" profile, consisting of n = 27 boys (14.9%) displaying reactive SIP in both provocation scenarios and proactive SIP in both instrumental gain scenarios. Boys with this profile showed high levels of anger (M = 7.67, SD = 2.08) and hostile intent attributions (M = 6.87, SD = 3.08), and a very high probability of displaying revenge goals in the social provocation scenario (ρ = .815), very high levels of anger (M = 8.19, SD = 2.19) and hostile intent attributions (M = 8.57, SD = 2.62), and a very high probability of displaying revenge goals (ρ = 1.00) in the object provocation scenario, as well as very high probabilities of displaying instrumental goals in the object acquisition (ρ = .889) and competition scenario (ρ = .852).

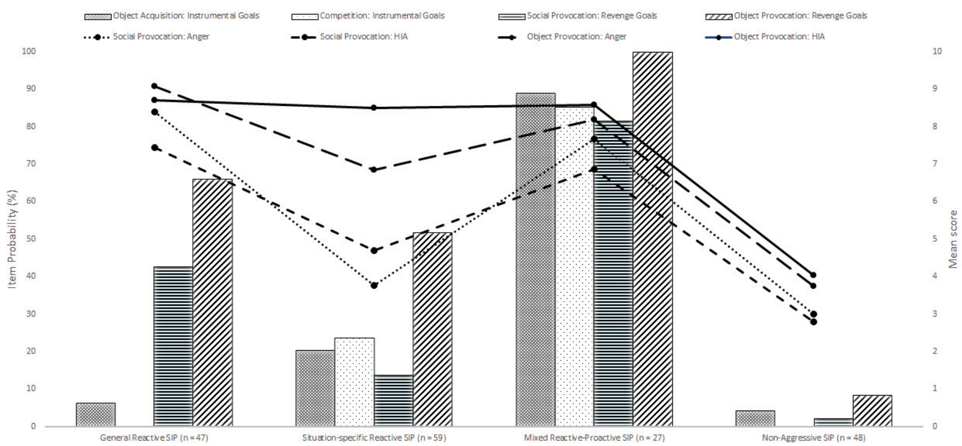
The remaining 26.5% of boys (n = 48) showed a "nonaggressive SIP" profile, characterized by nonaggressive SIP in all scenarios. Boys with this profile showed low levels anger (M = 3.00, SD = 1.62; M = 3.75, SD = 1.55), very low-to-low levels of hostile intent

2 We used 5 labels (i.e., very high, high, moderate, low, and very low) to describe levels of aggressive SIP. We applied these labels by dividing the scale of continuous SIP variables (i.e., 1-10) and item probabilities of dichotomous SIP variables (i.e., 0-1) by 5. Although arbitrary, these labels may help interpret differences between the SIP profiles in terms of aggressive SIP.

attributions (*M* = 2.79, *SD* = 1.86; *M* = 4.04, *SD* = 2.02), and a very low probability of displaying revenge goals (ρ = .021; ρ = .083) in the social- and object provocation scenario, respectively, as well as very low probabilities of displaying instrumental goals in the object acquisition (ρ = .042) and competition scenario (ρ = .000).

Taken together, the 4-profile solution suggests that distinct profiles exist of reactive SIP (i.e., general and situation-specific), mixed reactive-proactive SIP, and nonaggressive SIP. We hypothesized but did not find a proactive SIP profile. When we explored the data for children with scores fitting this profile (i.e., displaying instrumental goals but no revenge goals), we found only 6 boys that could potentially fit this profile—a group too small to be picked up by LPA.

Figure 1 *Latent SIP Profiles based on Boys' Reactive and Proactive SIP in each Virtual Reality Scenario*



Note. Bars refer to dichotomous SIP variables (i.e., instrumental goals, revenge goals) and correspond with item probabilities (%) displayed on the left vertical axis; lines refer to continuous SIP variables (i.e., anger, hostile intent attributions) and correspond mean scores displayed on the right vertical axis.

Discriminant Validity of SIP Profiles

Next, to validate the obtained SIP profiles, we compared them on 1) teacher's impression of boys' reactive and proactive motives for aggression and 2) theoretically relevant child characteristics. Table 4 shows descriptive statistics of the four SIP profiles on these variables and bootstrap 95% confidence intervals of the mean difference between SIP profiles of the planned contrasts, based on 5000 resamples. Bivariate correlations between these variables are reported in the supplementary materials (see Table S3).

Teacher-Reported Reactive and Proactive Motives for Boys' Aggression

Planned contrasts revealed that boys with a mixed reactive-proactive SIP profile showed more reactive motives for their teacher-reported aggression than those with a nonaggressive profile (*d* = 0.60) and more proactive motives than boys with a general-reactive SIP profile (*d* = 0.70), situation-specific SIP profile (*d* = 0.53) and nonaggressive SIP profile (*d* = 0.75). However, we found no significant differences in reactive motives between the reactive versus nonaggressive profiles.

Theoretically Relevant Child Characteristics

Planned contrasts yielded partial support for the distinctiveness of the reactive SIP profiles. Boys with a general reactive SIP profile displayed higher levels of anger-frustration temperament (*d* = 0.68) and hostile beliefs (*d* = 0.73) than boys with a nonaggressive SIP profile. Boys with a situation-specific reactive SIP profile reported more hostile beliefs (*d* = 0.40), but not more anger-frustration temperament, than those with a nonaggressive SIP profile. We found no significant differences in working memory and inhibition deficits between the reactive versus other SIP profiles.

We also found partial support for the distinctiveness of the mixed reactive-proactive SIP profile. Boys with this profile showed higher levels of hostile beliefs (*d* = 0.65), CU traits (*d* = 0.54), and justification of violence beliefs (*d* = 0.91) than those with a nonaggressive SIP profile. They also displayed higher levels of justification of violence beliefs than those with a situation-specific reactive SIP profile (*d* = 0.55), but not than those with a general reactive SIP profile. Furthermore, they showed higher levels of punishment insensitivity than those with a general reactive SIP profile (*d* = 0.50), but not than those with a situation-specific reactive or nonaggressive SIP profile. Contrary to our hypotheses, we found no differences between the mixed versus reactive SIP profiles on CU traits, nor any differences between SIP profiles in sensation seeking, working memory, or inhibition.

Last, explorative analyses revealed significant group differences for the variables anger-frustration temperament, $F(3,171) = 4.37$, p = 005, and justification of violence beliefs, $F(3, 180) = 5.29$, p = 002. Bootstrapped pairwise comparisons showed that boys with a general reactive SIP profile displayed higher levels of anger-frustration temperament than those with a mixed reactive-proactive SIP profile (*d* = 0.52) and a situation-specific reactive SIP profile (*d* = 0.53), and more justification of violence beliefs than those with a nonaggressive SIP profile (*d* = 0.64).

Table 4 *Descriptive Statistics of Theoretically Relevant Child Characteristics for Each SIP Profile and Bootstrap 95% Confidence Intervals of the Mean Difference between Planned Contrasts (Based on 5000 Resamples)*

	SIP Profiles				**Contrasts**					
	General Reactive (GR-SIP)	Situation-specific Reactive (SR-SIP)	Mixed Re- & Proactive (M-SIP)	Non-Aggressive- (NA-SIP)	GR-SIP vs. SR-SIP	GR-SIP vs. M-SIP	GR-SIP vs. NA-SIP	SR-SIP vs. M-SIP	SR-SIP vs. NA-SIP	M-SIP vs. NA-SIP
	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	*95% CI*	*95% CI*	*95% CI*	*95% CI*	*95% CI*	*95% CI*
Anger-frustration temperament	3.43 (0.84)^{a}	3.00 (0.79)^{b}	3.01 (0.71)^{b}	2.81 (0.94)^{b}	[.10, .76]	[.04, .79]	[.25, .98]	[-.38, .33]	[-.15, .51]	[-.18, .61]
Hostile beliefs	2.41 (0.87)^{a}	2.12 (0.71)^{a}	2.27 (0.66)^{a}	1.85 (0.66)^{b}	[-.02, .58]	[-.22, .48]	[.26, .86]	[-.49, .15]	[.01, .54]	[.12, .75]
Working Memory deficits	0.13 (0.95)^{a}	0.05 (0.98)^{a}	-0.04 (0.81)^{a}	-0.17 (0.68)^{a}	[-.31, .45]	[-.24, .60]	[-.08, .65]	[-.32, .49]	[-.12, .55]	[-.23, .50]
Inhibition deficits	0.10 (0.99)^{a}	-0.08 (0.65)^{a}	0.04 (1.04)^{a}	-0.03 (0.64)^{a}	[-.12, .51]	[-.47, .55]	[-.18, .45]	[-.63, .28]	[-.30, .19]	[-.31, .51]
Callous & Unemotional traits	2.03 (0.51)^{ab}	1.95 (0.48)^{ab}	2.09 (0.49)^{a}	1.85 (0.41)^{b}	[-.12, .26]	[-.32, .17]	[-.01, .37]	[-.37, .08]	[-.07, .27]	[.02, .45]
Justification of violence beliefs	2.15 (0.95)^{ab}	1.88 (0.78)^{ac}	2.34 (1.01)^{b}	1.65 (0.57)^{c}	[-.05, .59]	[-.67, .26]	[.19, .80]	[-.92, -.04]	[-.03, .48]	[.30, 1.09]
Sensation seeking tendencies	3.27 (0.57)^{a}	3.47 (0.51)^{a}	3.33 (0.44)^{a}	3.24 (0.56)^{a}	[-.41, .02]	[-.29, .16]	[-.19, .25]	[-.07, .35]	[.02, .42]	[-.14, .31]

Table 4 *Descriptive Statistics of Theoretically Relevant Child Characteristics for Each SIP Profile and Bootstrap 95% Confidence Intervals of the Mean Difference between Planned Contrasts (Based on 5000 Resamples) (continued)*

	SIP Profiles				**Contrasts**					
	General Reactive (GR-SIP)	Situation-specific Reactive (SR-SIP)	Mixed Re- & Proactive (M-SIP)	Non-Aggressive- (NA-SIP)	GR-SIP vs. SR-SIP	GR-SIP vs. M-SIP	GR-SIP vs. NA-SIP	SR-SIP vs. M-SIP	SR-SIP vs. NA-SIP	M-SIP vs. NA-SIP
	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	Mean (*SD*)	*95% CI*	*95% CI*	*95% CI*	*95% CI*	*95% CI*	*95% CI*
Punishment insensitivity	-0.30 (1.41)^{a}	-0.01 (0.89)^{ab}	0.31 (0.85)^{b}	0.14 (0.51)^{b}	[-.85, .16]	[-1.16, -.11]	[-.97, -.02]	[-.72, .10]	[-.45, .14]	[-.20, .52]
Teacher-rated reactive motives	2.81 (1.02)^{ab}	2.74 (0.89)^{ab}	3.03 (0.83)^{b}	2.49 (0.94)^{a}	[-.32, .47]	[-.69, .24]	[-.11, .74]	[-.70, .12]	[-.12, .61]	[.11, .95]
Teacher-rated proactive motives	1.91 (0.80)^{a}	2.02 (0.88)^{a}	2.48 (0.89)^{b}	1.90 (0.69)^{a}	[-.44, .21]	[-.98, -.16]	[-.30, .32]	[-.88, -.05]	[-.21, .45]	[.18, 1.00]

Note. Hypotheses for underlined variables were at least partly supported. SIP Profiles with different superscripts had a significant mean difference (i.e., the 95% CI does not include zero).

DISCUSSION

The present study used interactive VR methods to detect distinct SIP profiles of boys with aggressive behavior problems. Improved understanding of such profiles should help tailor cognitive-behavioral interventions to the needs of individual children. We found two reactive SIP profiles (i.e., a "general" and "situation-specific" reactive SIP profile), a mixed reactive-proactive SIP profile, and a nonaggressive SIP profile. We found no evidence for a proactive SIP profile. These findings demonstrate how different SIP patterns may underlie children's aggressive behavior problems. As such, our study extends previous work not only by showing that children's aggression may stem from deviations in distinct steps of the SIP model (De Castro & Van Dijk, 2017; Dodge, 2011), but also by demonstrating that these deviations can be used to demarcate subgroups of children with aggressive behavior problems.

Some of our findings were unexpected. First, while we did not anticipate to find two distinct reactive SIP profiles, this finding aligns with learning theory and research indicating that children's SIP may depend on conditioning of specific situational cues (De Castro & Van Dijk, 2017; Matthys et al., 2001; Dodge et al., 1985). Some children may be sensitized to specific situations, perhaps through past experience (Dodge, 2006; Matthys et al., 2001). For example, children whose belongings have been ruined by peers in the past will be likely to show aggressive SIP in similar situations (e.g., when their game is been ruined by a peer), but less so in other situations (e.g., when being excluded from a game by their peers). An alternative explanation for the emergence of a situation-specific profile may be sought with the fixed presentation order of our VR scenarios: it is possible that some children managed to regulate their anger up to a certain point in the first (social) provocation scenario, but were unable to regulate their anger in the face of yet another (object) provocation scenario, leading them to display aggressive SIP and behavior in this last scenario specifically (Kempes et al., 2008). However, such spill-over of anger seems unlikely given that boys had relatively low mean scores on anger in the first (social) provocation scenario.

Second, we did not find support for a proactive SIP profile. Prior research aiming to identify subgroups of children based on their reactive versus proactive motives for aggression has found mixed evidence for the existence of a predominantly proactive aggressive group, with some studies finding such a group (Carrol et al., 2018; Van Dijk et al., 2021), but not others (e.g., Euler et al., 2017; Marsee et al., 2014; Munoz et al., 2008; Smeets et al., 2017; Thomson & Centifanti, 2018). As previously detected proactive aggressive groups have been proportionally small (e.g., Van Dijk et al., 2021), it is possible that our study failed to detect such a group due to limited variance in proactive SIP (we used one dichotomous proactive SIP indicator only: instrumental goals). Alternatively, it is possible that proactively aggressive boys in our sample displayed aggressive SIP in the provocation scenarios (meant to assess reactive SIP) because of a carry-over effect, leading to a classification in the mixed reactive-proactive SIP profile. Indeed, the instrumental gain scenarios (meant to assess proactive SIP) that were presented first may have activated boys' aggressive responses, and perhaps facilitated the accessibility of aggressive response options in the subsequent provocation scenarios. Hence, although our findings suggest the absence of a proactive SIP profile, more research is needed to verify this finding.

Discriminant Validity of SIP Profiles

Our secondary aim was to provide further validation for the existence of distinct SIP profiles in boys with aggressive behavior problems by comparing them on 1) teachers' impressions of boys' reactive and proactive motives for aggression and 2) theoretically relevant child characteristics.

Teacher-reported Reactive and Proactive Motives for Boys' Aggression

Teachers' impressions partly corresponded with the obtained SIP profiles. They reported that boys with a mixed reactive-proactive profile displayed more reactive motives for their aggression than those with a nonaggressive profile, and more proactive motives than those with any other SIP profile. However, they did not report more reactive motives in boys with reactive versus other SIP profiles. One possible explanation is that, because boys with a mixed reactive-proactive SIP profile displayed the highest levels of aggressive SIP across all VR scenarios, their aggressive behavior and underlying motives are also more often observed by their teachers.

Theoretically Relevant Child Characteristics

For the general reactive SIP profile, we found partial support for unique child characteristics. Boys who showed this profile displayed higher levels of anger-frustration temperament and hostile beliefs than those with a nonaggressive SIP profile (Frick & Morris, 2004; Hubbard et al., 2010; Merk et al., 2005). Furthermore, exploratory analyses showed that boys with a general reactive SIP profile also displayed higher levels of anger-frustration temperament than those with a mixed reactive-proactive SIP profile or a situation-specific reactive SIP profile. These findings correspond with research indicating that children with an emotionally reactive temperament are prone to show reactive, but not proactive, SIP and aggression (Frick & Morris, 2004).

Unexpectedly, we found that boys who showed a general reactive SIP profile also reported more justification of violence beliefs than those with a nonaggressive SIP profile. It is possible that children displaying reactive SIP and aggression across provocation contexts view aggression an acceptable strategy to defend themselves or retaliate. Indeed, further analysis[3] revealed that boys with the general reactive SIP profile were particularly likely to endorse justification of violence items directly related to reactive aggression (e.g., "When somebody provokes you, it is normal to hit or threaten that

3 We identified items directly related to the justification of reactive (k = 3) and proactive aggression (k = 3), calculated an average score for each, and tested them against each other using a dependent t-test, $t(46)$, = 3.38, p = .001, d = 0.49.

person"). This finding aligns with studies showing that justification of violence beliefs are positively associated with reactive SIP and aggression (Calvete & Orue, 2010; Shu & Luo, 2021). In addition, we found that boys with a general reactive SIP profile were more sensitive to punishment than boys with a nonaggressive SIP profile and mixed reactive-proactive SIP profile. Although unexpected, this finding aligns with previous work proposing that children who engage in reactive aggression may be sensitized to negative stimuli (e.g., punishment and threat), possibly because they often grow up in harsh and punitive environments (Bubier & Drabick, 2009; Pederson et al., 2018).

The child characteristics of boys who showed a situation-specific profile were less pronounced than those of boys who showed a general reactive SIP profile. They displayed higher levels of hostile beliefs than boys with a nonaggressive SIP profile, but similar levels of anger-frustration temperament, inhibition and working memory. This may imply that boys with a situation-specific profile do not respond aggressively due to temperamental or inhibitory dispositions, but because of hostile schemas that are only activated in certain situations and predispose them to display reactive SIP in those situations (e.g., object provocations) but not others (e.g., social provocations; Chapter 6 of this dissertation).

For the mixed reactive-proactive SIP profile, we again found partial support for unique child characteristics. As predicted, boys with this profile more strongly justified violent beliefs than boys with a situation-specific reactive SIP and nonaggressive SIP profile. They also displayed CU traits more than boys with a nonaggressive SIP profile, but not more than boys with reactive SIP profiles. This last finding contradicts earlier work demonstrating that children who engage in both reactive and proactive aggression display higher levels of CU traits than children who solely engage in reactive aggression (Thomson & Centifanti, 2018). Moreover, we found that boys with a mixed reactive-proactive SIP profile were not less sensitive to punishment than those with a nonaggressive SIP profile. This contradicts earlier work suggesting that children who engage in proactive aggression may be less sensitive to punishment (Branje & Koot, 2018).

Last, we did not find any differences between SIP profiles for working memory and inhibition. This may be due to our measures. To match the emotionally engaging nature of interactive VR, we used standard executive functioning tasks presented in a game-based format designed to engage children. Nevertheless, these tasks may have evoked substantially less emotional arousal than our interactive VR assessment of SIP. It is possible that children's executive functioning assessed using "cool" tasks do not predict children's "hot" SIP assessed in emotionally engaging social interactions using interactive VR.

Strengths and Limitations

To our knowledge, our research is the first person-based study to distinguish between SIP profiles of children with aggressive behavior problems. We examined children's SIP patterns using interactive VR, which seems particularly suited to detect individual differences in children's SIP because it evokes relatively strong emotions in children, triggering aggressive SIP patterns that are not elicited when children are calm (Anderson & Bushman, 2002; Lemerise & Arsenio, 2000). We maximized clinically meaningful variance in children's SIP by recruiting boys from the entire spectrum of aggressive behavior problems, including children with severe aggressive behavior problems. This allowed us to detect four distinct SIP profiles, suggesting that different SIP patterns may underlie aggressive behavior in different children.

We acknowledge several limitations. First, the relatively small sample size of our study may have limited statistical power to find proportionally small SIP profiles, such as a proactive SIP profile. Relatedly, we were not able to identify enough children who displayed aggressive responses in instrumental gain scenarios. As such, we were not able to include children's positive outcome expectancies and response evaluations of aggression in the Latent Profile Analyses (LPA). Because our LPA was thus based on one proactive SIP indicator only (i.e., instrumental goals), the chance of finding a proactive SIP profile was further reduced.

Second, as interactive VR is relatively time-consuming and costly to develop, we were able to include four scenarios only. While we carefully chose the scenarios based on the literature and pilot work (Chapter 3 of this dissertation), they do not cover the broad range of social situations known to trigger aggressive SIP and behavior in children. Relatedly, all participants completed the VR scenarios in the same order. We presented the provocation scenarios last because those may arouse the strongest emotions. Nevertheless, we cannot rule out the possibility of potential order-effects affecting participants' SIP. Future studies could try to include a broader range of counterbalanced scenarios.

Third, our study included a relatively homogeneous sample of boys ages 7-13, with limited diversity in ethnic/cultural and socio-economic backgrounds. Future work is needed to test generalization to other subgroups of children.

Conclusion

This study shows, for the first time, that it is possible to detect distinct SIP profiles among children with aggressive behavior problems using interactive VR. Our findings advance our understanding of the SIP patterns contributing to children's aggressive behavior, and inform efforts to tailor cognitive-behavior interventions to individual children. We hope our findings will spur further work to more precisely delineate unique SIP profiles and experimentally test the effects of profile-tailored cognitive-behavioral interventions.

SUPPLEMENTARY MATERIALS

Table S1 *Means (and Standard Deviations) of Continuous SIP Variables (i.e., Anger, Hostile Intent Attribution, Positive Evaluations) and Number (and Proportions) of Dichotomous SIP variables (i.e., Revenge Goals, Instrumental Goals, Aggressive Responses, Outcome Expectancies) in Each Virtual Reality Scenario*

	Object Acquisition	Competition	Social Provocation	Object Provocation
Anger	1.75 (1.78)	3.13 (2.66)	5.35 (2.87)	6.80 (2.75)
Hostile Intent Attribution	1.47 (1.21)	2.17 (2.18)	5.23 (3.07)	7.37 (2.78)
Revenge Goals	1 (0.6%)	6 (3.3%)	51 (28.3%)	92 (51.1%)
Instrumental Goals	41 (22.7%)	37 (20.4%)	16 (8.9%)	12 (6.7%)
Aggressive Responses	42 (23.2%)	43 (23.8%)	69 (38.3%)	105 (58.3%)
Outcome Expectancies*	10 (23.8%)	18 (41.9%)	2 (2.9%)	0 (0.0%)
Positive Evaluations*	5.17 (3.56)	4.91 (3.71)	4.04 (3.24)	5.22 (3.40)

* Scores only apply to children who responded aggressively.

Table S2 *Means (M) and Standard Deviations (SD) of Continuous SIP Variables (i.e., Anger, Hostile Intent Attribution, Positive Evaluations) and Item Probabilities (p) and Number of Boys with Aggressive SIP (n) of Dichotomous SIP variables (i.e., Instrumental Goals, Revenge Goals, Outcome Expectancies) and Aggressive Responding for Each SIP Profile in Each Virtual Reality Scenario*

VR scenario	SIP variable	SIP Profile											
		General Reactive (*n* = 47)			Situation-Specific Reactive (*n* = 59)			Mixed Reactive-Proactive (*n* = 27)			Non-Aggressive (*n* = 48)		
		M/p	*SD*	*n*	*M/p*	*SD*	*n*	*M/p*	*SD*	*n*	*M/p*	*SD*	*n*
Object acquisition	Instrumental Goals	.064		3	.203		12	.889		24	.042		2
	Aggressive Responding	.064		3	.203		12	.889		24	.062		3
	Positive Evaluations*	8.67	2.31		4.75	3.22		5.00	3.67	24	4.67	4.73	
	Outcome Expectancies*	.667		2	.250		3	.167		4	.333		1
Competition	Instrumental Goals	.000		0	.237		14	.852		23	.000		0
	Aggressive Responding	.064		3	.254		15	.889		24	.021		1
	Positive Evaluations*	4.00	5.20		4.80	3.38		4.88	3.81		10.00	-	
	Outcome Expectancies*	.000		0	.400		6	.500		12	.000		0
Social provocation	Anger	8.38	1.48		3.76	1.64		7.67	2.08		3.00	1.62	
	Hostile Intent Attribution	7.43	2.63		4.70	2.49		6.87	3.08		2.79	1.86	
	Revenge Goals	.426		20	.138		8	.815		22	.021		1
	Aggressive Responding	.532		25	.293		17	.889		24	.062		3
Object provocation	Anger	9.06	1.51		6.84	2.20		8.19	2.19		3.75	1.55	
	Hostile Intent Attribution	8.70	1.88		8.49	1.46		8.57	2.62		4.04	2.02	
	Revenge Goals	.660		31	.517		30	1.00		27	.083		4
	Aggressive Responding	.702		33	.655		38	1.00		27	.146		7

*Scores only apply to children who responded aggressively.

Table S3 *Means* (M)*, Standard Deviations* (SD)*, and Number of Participants* (n)*, and Bivariate Correlations between Theoretically Relevant Child Characteristics*

	M	*SD*	*n*	2.	3.	4.	5.	6.	7.	8.	9.	10.
Anger-Frustration Temperament	3.06	0.86	172	.27*	.26*	.14	.18*	.21*	.07	.06	.30*	.07
Hostile Beliefs	2.15	0.76	181		.23*	.22*	.29*	.33*	-.03	.06	.25*	.09
Working Memory deficits	0.00	0.88	173			.32*	.12	.13	.21	.04	.27*	.14
Inhibition deficits	0.00	0.81	180				.11	.02	.04	-.02	-.03	.13
CU traits	1.96	0.48	181					.61*	.07	.10	.08	.22*
Justification of Violence Beliefs	1.96	0.85	181						.02	.05	.09	.24*
Punishment Insensitivity	0.00	1.00	165							-.03	.17*	.20*
Sensation Seeking	3.34	0.53	181								.09	.15
Teacher-rated Reactive Motives	2.74	0.94	160									.12
Teacher-rated Proactive Motives	2.03	0.83	160									

Note. Correlations were calculated using pairwise deletion. * Indicates that the bootstrap 95% confidence interval did not include zero.

CHAPTER 6

A Dual-Mode Social-Information-Processing Model to Explain Individual Differences in Children's Aggressive Behavior

This chapter is published as:
Verhoef, R. E. J., Van Dijk, A., & De Castro, B. O. (2021). A dual-mode social-information-processing model to explain individual differences in children's aggressive behavior. Clinical Psychological Science, 10(1), 41-57. https://doi-org/10.1177/21677026211016396

Author contributions:
RV, AD, and BC conceptualized the manuscript; RV drafted the manuscript; AD and BC provided feedback on the manuscript.

ABSTRACT

Children differ considerably in the social information processing (SIP) patterns underlying their aggressive behavior. To clarify these individual differences, we propose a dual mode SIP model that predicts which processing steps which children take, under which circumstances, and how this may lead to aggression. This dual mode SIP model distinguishes between an automatic and reflective processing mode. The automatic mode is characterized by fast automatic processing and impulsive behavioral responses, whereas the reflective mode is characterized by deliberate processing and controlled behavioral responses. Whether children use the automatic versus reflective processing mode is moderated by their level of arousal, which depends on an interplay between child-specific factors (i.e., emotional dispositions, motivational dispositions, and executive functioning) and dynamic factors (i.e., internal state and type of situation). The dual mode SIP model provides new insights into children's unique SIP styles and provides possibilities to tailor treatment to children's individual needs.

Keywords: Social information processing, aggression, children, dual mode, automatic processing, arousal

A DUAL-MODE SOCIAL-INFORMATION-PROCESSING MODEL TO EXPLAIN INDIVIDUAL DIFFERENCES IN CHILDREN'S AGGRESSIVE BEHAVIOR

Aggressive behavior problems in childhood have a detrimental impact on children, their environment, and society at large, and are among the most prevalent psychological problems in children (Dodge et al., 2006; Polanczyk et al., 2016; Romeo et al., 2006). Research has identified cumulative contextual and dispositional risk factors that predict the development of aggressive behavior problems (Lochman & Matthys, 2018). These distal risk factors may shape the way in which children process social information, so that aggressive behavior patterns will be maintained even when these risk factors are no longer present (Dodge & Pettit, 2003). For instance, children who are repeatedly victimized may generalize the expectation that others will victimize them to future, non-hostile, contexts, where this generalized expectation will maintain their aggressive behavior patterns (Perren et al., 2013). Furthermore, empirical work has demonstrated that children's deviant social information processing (SIP) patterns explain meaningful variance in their aggressive behavior, are associated with disruptive behavior disorders, and predict the development of future aggressive behavior patterns (e.g., De Castro & Van Dijk, 2017; Dodge et al., 1986; Lansford et al., 2006; Chapter 2 of this dissertation). Thus, research suggests that children's SIP plays a key role in the development and maintenance of their aggressive behavior.

Nonetheless, how well children's SIP explains their aggressive behavior varies considerably between children and studies. These divergent findings may reflect that relatively few empirical studies on children's SIP have considered automatic processes (Anderson & Bushman, 2002). Previous work has predominantly assessed reflective SIP in children, by explicitly asking them to reflect on hypothetical social events. Yet, many children may act aggressively *without* such reflection: When they are in a state of high arousal, an apparently hostile gesture of a peer may trigger a direct automatic aggressive response, with minimal cognitive control to guide this process reflectively. Distinct models have been developed that predict children's aggressive behavior based on their emotional dispositions (e.g., temperament), motivational dispositions (e.g., punishment and reward insensitivity), executive functioning capacities, or database of social experiences (e.g., Dodge, 2006; Guo & Mrug, 2017; Jarrett & Hilton, 2017; Matthys et al., 2013). However, each of these models does not explain how these factors then actually contribute to children's aggressive behavior through their automatic and reflective SIP. Therefore, we propose an overarching dual mode social information processing model that predict which processing steps which children take, under which circumstances, and how this may lead to aggression.

THE CURRENT SOCIAL INFORMATION PROCESSING MODEL

The current SIP model (Crick & Dodge, 1994) proposes that several social information processing steps take place between children's encounter with a social stimulus and their behavioral response: (1) encoding of social cues, (2) interpretation and mental representation of these cues, (3) setting of interactional goals, (4) generation of possible behavioral responses, (5) evaluation of these responses, and (6) enactment of the selected response. For each of these steps, children draw from their database of social knowledge to process the present situation (Crick & Dodge, 1994). Aggressive behavior may result from deviancies in each step of the SIP model. Research has shown that children with aggressive behavior problems encode more hostile cues and less non-hostile cues, make more hostile interpretations of others' behavior, set more interaction goals directed at revenge or instrumental gain, generate more aggressive responses, and evaluate aggressive responses and their outcomes less negatively than their less aggressive peers (for reviews, see: De Castro & Van Dijk, 2017; Dodge, 2011).

The current SIP model accommodates individual differences in the specific processing steps leading up to aggression. For instance, aggressive behavior may primarily stem from excessive anger and a tendency to attribute hostile intent to others in some children, but from the tendency to pursue instrumental goals and having positive outcome expectancies of aggression in others (Arsenio et al., 2009; De Castro et al., 2005; Crick & Dodge, 1996). From the onset, the model has also suggested that aggressive behaviors may be a consequence of skipping (part of) these deliberate SIP steps (Crick & Dodge, 1994). In everyday life, many children may not have the time or mental resources, or feel the need to engage in reflective processing of all SIP steps (Anderson & Bushman, 2002; De Castro, 2004). However, the SIP model does not describe how children who skip parts of these processing steps would actually engage in automatic processing, nor what would make children skip parts of the deliberate SIP steps. Our dual mode SIP model aims to explain what determines whether children process social information automatically or reflectively, and how these processes take place.

EVIDENCE FOR AUTOMATIC SIP

Several empirical findings support the necessity of incorporating automatic processing into the SIP model. Research has found a direct link between children's emotion regulation deficits and aggressive behavior, which could not be explained by their self-reported SIP (De Castro et al., 2005; Helmsen et al., 2012). This suggests that strong emotions, such as anger or frustration, may directly lead to aggression, without intermediate reflection or decision processes. Similarly, self-reports by children with aggressive behavior problems have revealed that they primarily explain their aggressive responses to peer provocations as being driven by strong emotions, whereas they seldom refer to goals or evaluations underlying their behavior (De Castro et al., 2012). Furthermore, in a detective game where children were asked to evaluate whether a peer acted with hostile intent or not, children with aggressive behavior problems made faster judgements, suggesting they spent less time to reflect on the peer's intentions (Dodge & Newman, 1981). Moreover, in a social problem solving task, aggressive-rejected children generated more conflict-escalating solutions (e.g., verbal or physical aggression) than their non-aggressive peers, but only when they were instructed to respond as quickly as possible with the first solution that came to mind—not when they were instructed to wait for 20 seconds and consider alternative solutions (Rabiner et al., 1990). These findings suggest that aggressive-rejected children may generate different responses under automatic versus reflective conditions.

Further indications for automatic SIP may stem from empirical studies using eye-tracking technology, which aimed to capture children's automatic encoding of social cues by assessing their eye movements. For instance, one study has shown that children with aggressive behavior problems attended more to non-hostile cues than their non-aggressive peers, but nonetheless recalled less non-hostile information (Horsley et al., 2010). The authors suggested that these children may have automatically encoded non-hostile cues—which are salient because they conflict with children's pre-existing hostile ideas—, without further reflecting on them. Another study also explained their seemingly contradictory findings by adopting a dual mode perspective. This study found that children who paid *less* attention to social threat cues were *more* likely to attribute hostile intent and exhibit aggressive behavior (Schippell et al., 2003). The authors proposed post hoc that children's encoding of social cues may consist of two steps, instead of one: 1) automatic encoding of cues, occurring before cues come into conscious awareness; and 2) deliberate encoding of cues, occurring after cues come into conscious awareness. Supporting this idea of deliberate encoding, another eye-tracking study in young adults has revealed that the number of attentional fixations on social cues was positively associated with the quality of moral decision-making justification (Garon et al., 2018). As such, empirical evidence suggests that children's encoding of social cues may be automatic as well as reflective.

6

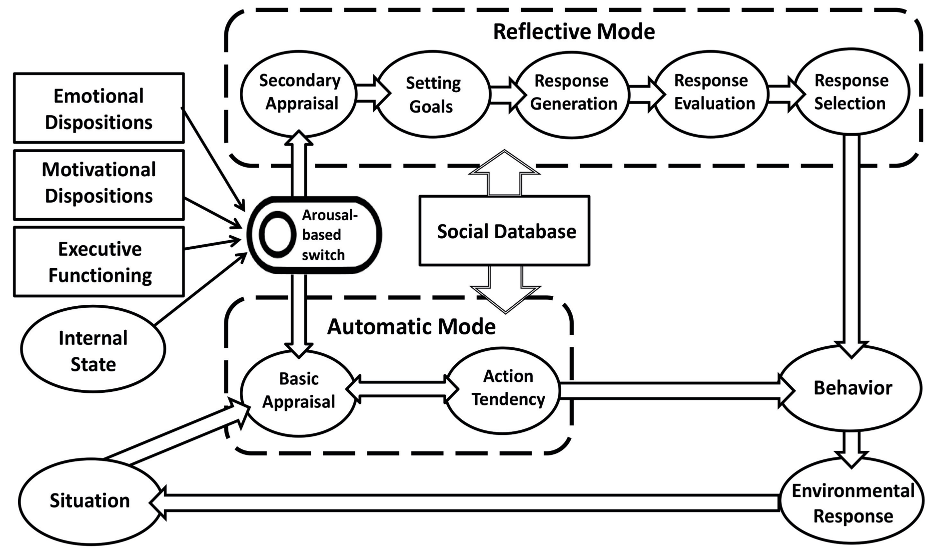
More support for automatic SIP stems from studies using experimental paradigms known to tap into automatic processes. For instance, research using a cued-recall paradigm has shown that individuals high on self-reported aggression spontaneously encoded and interpreted behavioral sentences as more aggressive than non-aggressive individuals, but this difference disappeared when they were asked to deliberately reflect on the motives of the actor in these sentences (Zelli et al., 1996). Relatedly, empirical work indicates that subliminal priming with aggressive concepts may predict aggressive behavioral tendencies in subsequent unrelated tasks (for a review, see: Todorov & Bargh, 2002). In addition, one study using an implicit association task has demonstrated that children's implicit aggressive tendencies explained additional variance in their aggressive behavior above and beyond their explicit aggressive tendencies, suggesting that automatic processes form a unique route to aggressive behavior in children (Grumm et al., 2011). Together, these findings illustrate that aggressive behavior may result from automatic processes. An adapted SIP model is therefore needed that incorporates both reflective and automatic SIP, and provides a fine-grained and testable description of which processing steps which children take, under which circumstances, and how this may lead to aggression.

THE DUAL MODE SOCIAL INFORMATION PROCESSING MODEL

We propose a dual mode SIP model that distinguishes between a reflective and automatic processing mode underlying children's social behavior (Figure 1). Our model is in line with a more general shift in cognitive psychology, where a distinction is made between a reflective and automatic processing mode in so-called dual mode processing models (e.g., Chaiken & Trope, 1999; Kahneman, 2011). Our dual mode SIP model combines insights from current SIP models (Crick & Dodge, 1994; Lemerise & Arsenio, 2000), the General Aggression Model (Anderson & Bushman, 2002), and dual mode processing theory (e.g., Frijda, 1993; Smith & DeCoster, 2000). Both modes may lead children to respond with aggressive behavior: an automatically enacted aggressive action tendency, or a deliberately selected aggressive response strategy. However, although all children may use both processing modes, some children may predominantly engage in aggression that derives from automatic SIP, whereas others more frequently engage in aggression that derives from reflective SIP. Distinguishing between processing modes may thus have important implications for intervention: Intervening on reflective SIP would not help children whose aggression is predominantly driven by automatic SIP. Conversely, intervening on automatic SIP would not help children whose aggression is predominantly driven by reflective SIP.

The dual mode SIP model incorporates several factors to explain which children will use which processing mode under which circumstances. We derived these factors from previous research on correlates of children's aggressive SIP and behavior (e.g., Bookhout et al., 2018; De Castro & Van Dijk, 2017; Jarrett & Hilton, 2017; Moore et al., 2018). As can be seen in Figure 1, we propose that whether children use the automatic or reflective processing mode is determined by an interplay between child-specific factors (i.e., children's emotional dispositions, motivational dispositions, and executive functioning) and dynamic factors (i.e., children's internal state and the type of situation). Moreover, similar to previous SIP models, we further propose that aggressive SIP patterns—in either processing mode—are explained by children's adverse learning histories, stored in the social database.

Figure 1 *The Dual Mode Social Information Processing Model*



Note. The figure depicts one SIP cycle: A situation activates the automatic and/or reflective mode, resulting in behavior, responded to by the environment, thereby creating a new situation. Circles represent dynamic processes that may differ between SIP cycles. Rectangles represent dispositions that are relatively stable across SIP cycles.

THE AUTOMATIC AND REFLECTIVE MODE

The dual mode SIP model proposes that, during any social interaction, children will process social information in either the automatic or the reflective mode. The automatic mode is characterized by fast implicit processing and consists of a basic appraisal and a dominant action tendency. Both derive from an associative network in memory that is part of children's social database. This network links specific situational triggers (e.g., a child's goal is blocked) to specific affective reactions (e.g., frustration) and behavioral responses (e.g., fight) through automatic if-then contingencies (e.g., if frustrated, then fight). Note that this fast process does not involve any deliberate thought, but rather is a direct emotional response to specific situational cues (Frijda, 1993). Thus, children in the automatic mode solely use their database of automated contingencies, without further consideration of situational cues or decision processes. This allows them to quickly process and respond to situational triggers while saving cognitive resources, but goes at the cost of careful decision making (Anderson & Bushman, 2002). The if-then contingencies in the automatic mode differ between children, depending on their learning histories stored in the social database, which is in line with the original SIP model (Crick & Dodge, 1994). In contrast to the original model, though, we propose that the automatic mode bypasses all SIP steps except the encoding of cues (i.e., it does not include representation of intent, goal setting and response decision processes; see Figure 1).

The reflective mode is characterized by slower explicit processing and consists of several sequential mental steps, resulting in a deliberate behavioral response. In contrast to the automatic mode, the reflective mode integrates information from the social database with situational cues (Huesmann & Guerra, 1997; Zelli et al., 1999). In the reflective mode, children re-consider their basic appraisal by allocating their attention to the situation and re-appraising the encoded cues, which may include attributions of intent or causality (e.g., “His face shows that he has harmed me on purpose”). Next, based on this secondary appraisal, children generate possible responses and evaluate those responses based on, for instance, potential outcomes (e.g., “If I punch him now, he won’t bother me anymore”), their self-efficacy to enact the response (e.g., “I am stronger than him”), and moral considerations (e.g., “Someone who provokes me deserves a beating”). Last, children select the response evaluated most positively to be enacted. The reflective mode has the benefit of better problem solving, but depletes cognitive resources and takes more processing time (Anderson & Bushman, 2002). This reflective mode includes all processing steps formulated in the original SIP model (Crick & Dodge, 1994).

The dual mode SIP model thus extends current SIP models by explaining not only deliberate, controlled aggression preceded by reflective processing, but also fast, emotion-driven aggression preceded by automatic processing. According to the dual mode SIP model, children will typically process social information in the automatic mode. From an evolutionary perspective, it is plausible that children will only engage in reflective processing when it is necessary. It would be maladaptive to spend cognitive resources to correct basic appraisals of situations that are familiar, satisfactory or concern irrelevant stimuli (Aron et al., 2012; Barlett & Anderson, 2011). Thus, our dual mode SIP model assumes that children will only switch to the reflective mode in situations that are unfamiliar, potentially threatening, or of other personal relevance (Frijda, 1993). For instance, children may use the automatic mode when they are engaging in small talk with the bus driver (i.e., a personally irrelevant situation), or when they are routinely playing with their sibling (i.e., a familiar and non-threatening situation). Yet, children may switch to the reflective mode when they encounter an unfamiliar peer at the schoolyard (i.e., an unfamiliar and potentially relevant or threatening situation). Our dual mode model further assumes that children who are using the reflective mode may switch back to the automatic mode if the situation becomes excessively threatening and requires a direct response (Kunimatsu & Marsee, 2012). For instance, children may reflectively process the encounter with an unfamiliar peer, but immediately switch to the automatic mode when they feel threatened.

How do children switch between modes? The dual mode SIP model proposes that children’s basic appraisal evokes a certain level of arousal that determines which processing mode they will use. This relation follows an inverted U-shaped function (Obradović*, 2016;* Yerkes & Dodson, 1908), such that the reflective mode will only be used when arousal levels are moderate, rather than too low or too high (cf. ‘window of tolerance’ as an optimal state of arousal to process information reflectively; Siegel, 1999). Thus, in situations of low arousal, children will use the automatic mode, saving cognitive resources. In situations of moderate arousal, children will use the reflective mode, allowing them to deliberately process and carefully respond to the situation. In situations of excessive arousal, children are forced to use the automatic mode, allowing them to quickly process and respond to the situation (Frijda, 1993). Whether a situation evokes low, moderate, or excessive arousal, will differ between children and situations, depending on their dispositions and dynamic factors (to be discussed later). These factors together determine children’s basic appraisal. Thus, children’s basic appraisal can be seen as an unconscious motivational heuristic that works as the ‘switch’ between the automatic and reflective processing mode by changing children’s level of arousal.

The biological underpinnings of children’s arousal-based switch between the automatic versus reflective mode may be found within the autonomic nervous system (ANS). When children encounter a social stimulus, the thalamus processes the stimulus and information is sent to the amygdala, leading to an initial appraisal in terms of emotional relevance. Next, the amygdala sends a signal to the hypothalamus, which in turn activates the ANS (Cunningham et al., 2007). The ANS has two branches. The sympathetic nervous system (SNS) incites physiological processes to prepare children for fight-or-flight responses (e.g., by increasing heart rate, blood pressure) through the release of neurotransmitters such as norepinephrine (McCorry, 2007). In contrast, the parasympathetic nervous system (PNS) enables children’s physiology to recover to a calm state (e.g., by decreasing heart rate, blood pressure) through the release of neurotransmitters such as acetylcholine (McCorry, 2007). The SNS and PNS generally operate simultaneously and have opposing functions: the SNS increases physiological arousal and prepares for action in the face of threat, whereas the PNS decreases physiological arousal and allows for greater attentional and cognitive capacity (Beauchaine, 2001). As such, a well-performing ANS allows for a flexible physiological response to diverse social situations. In terms of our model, this allows children to switch between the automatic and reflective mode, enabling them to adequately respond to social demands. Conversely, a dysfunctional ANS may trigger maladjusted behavior (Branje & Koot, 2018). For instance, high SNS reactivity in conjunction with PNS withdrawal may lead to high arousal levels, triggering exaggerated fight-or-flight responses (Branje & Koot, 2018). Indeed, autonomic processes such as increased heart-rate have been linked to more hostile intent attributions and aggressive behavior in children (Crozier et al., 2008; Williams et al., 2003). Thus, research suggests that children’s arousal levels modulate their social information processing—as we propose, through triggering their automatic versus reflective mode.

THE SOCIAL DATABASE

In line with previous SIP models, our dual mode SIP model proposes that aggressive SIP patterns—in both processing modes—are determined by children's social database. This database is built from children's social experiences, which accumulate into memory structures consisting of specific affective, cognitive, and behavioral tendencies, which in turn guide children's SIP in future social contexts (Anderson & Bushman, 2002; Crick & Dodge, 1994). For instance, children growing up in harsh environments may develop hostile memory structures predisposing them to processes social interactions in a hostile manner (e.g., by attributing hostile intent or having access to aggressive response options; Dodge, 2006; Dodge et al., 1997; Lansford et al., 2010), whereas children growing up in supportive environments may develop memory structures allowing them to trust and cooperate with others (Frankenhuis et al., 2016). Thus, memory structures, or schemas, guide children's perception, encoding, storage and retrieval of social information (Beck et al., 2004). They may also guide behavior, prompting children with action plans or scripts of how to react to schema-relevant events (Beck et al., 2004). In this way, memory structures allow children to quickly process and respond to relevant situational triggers, but may also induce errors or biases—especially in schema-relevant situations (Crick & Dodge, 1994).

Empirical work has identified several memory structures related to aggressive SIP (for a review, see: De Castro & Van Dijk, 2017). Longitudinal research has shown that children are more likely to display aggressive SIP and behavior if they hold hostile schemas (e.g., "Other people cannot be trusted;" Burks et al., 1999; Calvete & Orue, 2012) or believe that aggression is morally acceptable and instrumentally useful (e.g., "Sometimes you need to fight to get what you want;" Calvete, 2008; Calvete & Orue, 2012; Huesmann & Guerra, 1997; Zelli et al., 1999). Moreover, research has shown that narcissistic memory structures, including beliefs about grandiosity, self-entitlement and being superior to others (e.g., "Children like me deserve something extra"), are associated with aggressive SIP and behavior (Calvete & Orue, 2010; Calvete & Orue, 2012). Thus, empirical research suggests that children's social database may shape their aggressive SIP.

Our dual mode SIP model extends previous scientific work, proposing that both automatic and reflective SIP rely on children's social database, but in different ways. The automatic mode—being fast and implicit—fully relies on the database. Any situational input activates an associative memory network, where this input is directly linked to fully automatized affective and behavioral tendencies (Anderson & Bushman, 2002; Smith & DeCoster, 2000). A high- or low-arousal basic appraisal will directly trigger an associated emotional response and dominant action tendency. In low-arousal situations, children's associative memory structures will trigger action tendencies for familiar and habitual situations, which will often be nonaggressive (e.g., routinely playing with a sibling) but may also be aggressive (e.g., routinely bullying a classmate). Similarly, in high-arousal situations (e.g., being provoked by a peer), individual differences in children's associative memory structures will predict whether they respond in a nonaggressive or aggressive way (e.g., walking off versus starting a fight). Thus, children's social database directly determines their automatic SIP, leading them to respond without reflecting on the situation (e.g., by considering others' intentions or the consequences of their response).

The reflective mode, in contrast, does not solely rely on the social database, but explicitly integrates situational input with pre-existing memory structures. Thus, children will use knowledge from their database to consider situational cues and evaluate possible response options. In the reflective mode, children may use both associative memory structures (which are also used in the automatic mode) and rule-based memory structures. Associative memory structures may influence children's reflections because they trigger dominant emotional and behavioral tendencies (e.g., children may be more likely to consider aggressive responses when a hostile schema is activated; Anderson & Bushman, 2002). Rule-based memory structures are more complex, and include reasoning heuristics that require children to integrate situational input with pre-existing knowledge (Smith & DeCoster, 2000). Examples are: "When someone frustrates me, I first need to check his height, weight and reputation before I hit him," and "I first need to check his body language to evaluate whether his intentions are hostile or not." Individual differences in children's rule-based memory structures will predict whether they respond with aggression or not. Thus, children's social database indirectly influences their reflective SIP, triggering dominant tendencies and rule-based heuristics that may steer, but will not fully determine, how they process situational input and reflect on their response options.

In sum, children's social database determines whether their SIP is aggressive or not. The dual mode SIP model further proposes that some children are prone to use the automatic mode, whereas others are prone to use the reflective mode. Although both modes may predict aggressive behavior, it is important to understand what determines children's 'predominant' mode to help explain how their aggressive behavior typically originates. This has implications for intervention. Children whose aggressive SIP is predominantly automatic would require a different intervention approach (e.g., intervening on arousal regulation) than children whose aggressive SIP is predominantly reflective (e.g., changing deviant cognitions). The dual mode SIP model therefore describes factors influencing children's predominant processing mode.

EXPLAINING INDIVIDUAL DIFFERENCES IN CHILDREN'S PREDOMINANT SIP MODE

To our knowledge, no empirical research has been conducted on factors influencing whether children are prone to process social information in the automatic versus reflective mode. We considered work on reactive and proactive aggression (e.g., Dodge, 1991; Hubbard et al., 2010; Kempes et al., 2005; Merk et al., 2005), but found that this work may not directly concern our dual mode SIP model, as the reactive-proactive and automatic-reflective distinctions are not interchangeable (Bushman & Anderson, 2001).

It may seem intuitive to assume that reactive aggression, which is described as emotional and impulsive (Dodge, 1991), should derive from automatic SIP, whereas proactive aggression, described as unemotional and controlled (Dodge, 1991), should derive from reflective SIP. Yet, our dual mode SIP model predicts differently: Reactive aggression may result from automatic SIP (e.g., impulsively hitting back) as well as reflective SIP (e.g., deciding to take revenge), same as proactive aggression may result from automatic SIP (e.g., routinely bullying) as well as reflective SIP (e.g., planning to steal). There is empirical work to support this notion. Research has examined children's in-the-moment physiology and aggression during a game where a virtual peer provoked them, and found that children's reactive aggression may be accompanied by high arousal and weak regulation, as well as by low arousal and strong regulation (Moore et al., 2018). Similarly, a review on physiological predictors of children's aggressive behavior has illustrated that both reactive aggression and proactive aggression can be related to low as well as high physiological arousal (Fanti, 2018). This work suggests that there are both automatic and reflective routes to both reactive and proactive aggression. Last, there is a host of SIP research to support that reactive aggression is predominantly associated with early SIP deviancies (i.e., hostile encoding, hostile intent attributions), whereas proactive aggression is predominantly associated with late SIP deviancies (i.e., instrumental goals, positive evaluations of aggression; for reviews, see: Hubbard et al., 2010; Vitaro et al., 2006). Yet, as these studies have assessed both early and late SIP by explicitly asking children to reflect on hypothetical social events, these findings may actually suggest that reflective SIP predicts both reactive and proactive aggression.

In sum, although the reactive-proactive distinction is important to explain individual differences in children's aggression, empirical work on this distinction cannot be used as direct support for the automatic-reflective distinction made in our dual mode SIP model, nor to select factors that may explain individual differences in children's predominant SIP mode. We therefore selected factors for our dual mode SIP mode based on the following considerations: a) previous research has shown that the factor is related to aggression; b) the factor has been theoretically and empirically linked to children's SIP; and c) the factor is relevant for arousal regulation, and as such may contribute to whether children predominantly engage in automatic versus reflective SIP. In the next sections, we describe these factors, aiming to provide starting points for future research on individual differences underlying children's aggressive behavior.

Emotional Dispositions

A wealth of empirical research suggests that children's aggressive SIP and behavior are associated with specific emotional dispositions, assessed through temperament questionnaires as well as their physiological reactivity (for reviews, see: Bookhout et al., 2018; Branje & Koot, 2018; Frick & Morris, 2004; Moore et al., 2018). Overall, this work suggests that both children who display high (hyper-) emotional reactivity and children who display blunted (hypo-) emotional reactivity are prone to aggressive SIP and behavior. The dual mode SIP model proposes that children's emotional dispositions modulate their arousal across social situations (Rothbart & Derryberry, 1981) and thereby directly influence whether they are prone to use the automatic versus reflective processing mode. Hyper-emotional children will be prone to experience excessive arousal levels in stressful situations, forcing them to use the automatic mode more often than other children would. This may, for example, make them instantly respond with excessive anger and aggression, without any reflection such as making intent attributions, or considering their response options. Conversely, hypo-emotional children will experience lower arousal levels in the same situation. This may affect their SIP in two ways, depending on how stressful the situation is. In highly stressful situations, hypo-emotional children's lower arousal levels may enable them to use the reflective mode. They may, for example, exhibit moderate (instead of high) arousal levels when they are threatened, allowing them to reflect on the situation and carefully plan a retaliatory strike or reconciliatory attempt. However, in moderately stressful situations that would trigger the reflective mode in most children, hypo-emotional children may still use the automatic mode. They may, for example, exhibit low levels of arousal when they are bullying someone, prohibiting them from reflecting on the potential harm caused by their behavior. Thus, children's emotional dispositions may be an important determinant of their predominant processing mode, both for children who are hyper- and hypo-emotionally reactive.

Indirect support for the idea that children's emotional dispositions affect their predominant processing mode stems from research on children's temperament. This research demonstrated that children with a highly emotionally reactive temperament become highly aroused when confronted with peer provocation (Hessler & Katz, 2007), suggesting they are prone to engage in automatic processing. In contrast, children with callous-unemotional (CU) traits have been shown to exhibit blunted physiological arousal in challenging social situations (i.e., lower skin conductance reactivity, less heart-rate change, and lower cortisol reactivity; for a review see: Frick et al., 2014a). These children report to feel more alert and in control after a virtual fear induction (Thomson et al., 2020), suggesting that they indeed may use reflective SIP in highly stressful situations. Conversely, these children may use automatic SIP in situations that would trigger reflective SIP in others. For instance, research has shown that children with CU traits are insensitive to peers' expressions of fear and distress or signs of potential punishment (for a review see: Frick et al., 2014b). Relatedly, empirical work demonstrated that children who exhibit low autonomic arousal during hypothetical moral transgressions report lower levels of guilt and, in turn, display more aggressive behavior according to their caregivers (Colasante et al., 2021). The authors suggested that these children may aggress without reflecting on their moral transgressions or considering the other person's discomfort.

The idea that children's emotional dispositions affect their predominant mode is also indirectly supported on a biological level, where the same distinction between hyper- and hypo-emotional children is observed (e.g., Bookhout et al., 2018; Branje & Koot, 2018). Stress-regulating systems such as the Hypothalamic-Pituitary-Adrenal

(HPA) axis have been linked to aggression in children through both hyperactive and hypoactive HPA-axis functioning, leading to excessive or blunted stress responses, respectively (for reviews, see: Branje & Koot, 2018; Van Goozen et al., 2007). Similarly, empirical research has linked respiratory sinus arrhythmia (RSA) functioning—a physiological marker of children's emotion regulation capacities—to aggressive behavior in children through both low and elevated resting RSA (for reviews, see: Bookhout et al., 2018; Branje & Koot, 2018). These findings fit with the dual mode SIP model, which proposes that hyper-emotional children may be particularly prone to engage in aggression that derives from automatic SIP, whereas hypo-emotional children may be prone to engage in aggression that derives from either automatic SIP (i.e., when confronted with relatively mild social stressors) or reflective SIP (i.e., when confronted with more severe social stressors).

Motivational Dispositions

Children's predominant SIP mode may also be influenced by their motivational dispositions. Empirical work suggests that children's insensitivity to punishment and reward may predispose them to engage in aggressive behavior (for reviews, see: Matthys et al., 2012, Matthys et al., 2013; Weeland et al., 2015). Children who are insensitive to punishment and reward exhibit lower arousal levels when confronted with regular punishment and reward cues, as compared to other children (e.g., Matthys et al., 2013). The dual mode SIP model therefore proposes that punishment and reward insensitivity may predispose children to use the automatic mode when others would use the reflective mode, and use the reflective mode when others would use the automatic mode—in similar fashion as explained for children with CU-traits (who are also prone to experience lower arousal levels in response to social stressors). In fact, there is emerging evidence to suggest that punishment and reward insensitivity may be specific characteristics of children with CU-traits; Frick et al., 2014b).

Which mode these children will use, depends on the severity of the punishment and reward cues. Relatively mild cues that would trigger the reflective mode in most children, may activate the automatic mode in children who are insensitive to punishment and reward. For instance, these children may experience little arousal from being caught by their teacher, and therefore continue bullying a peer from the automatic mode. However, more severe punishment and reward cues that would make most children switch to the automatic mode because of high arousal levels, may activate the reflective mode in children who are insensitive to punishment and reward. For instance, these children may experience moderate arousal levels when they are caught for stealing a candy bar from the store, allowing them to carefully reflect on the situation and successfully get away with it.

In sum, our dual mode model proposes that children who are insensitive to punishment and reward are prone to engage aggression that derives from either automatic SIP (i.e., when they are confronted with relatively mild punishment and reward cues) or reflective SIP (i.e., when they are confronted with more severe punishment and reward cues).

Executive Functioning

There is much empirical research to suggest that children's capacities to consciously control thought and action, called executive functions (EF; Zelazo & Müller, 2002), are associated with less aggressive SIP (e.g., Ellis et al., 2009; Goldweber et al., 2011; Van Nieuwenhuijzen et al., 2017) and less aggressive behavior (for reviews, see: Jarrett & Hilton, 2017; Morgan & Lilienfeld, 2000; Ogilvie et al., 2011). The dual mode SIP model further proposes that children's EF capacities may help them regulate their arousal by facilitating reflective skills such as perspective-taking, problem-solving and judgement, and thereby affect whether they are prone to use the automatic or reflective mode. If arousal levels determine children's mode, how then can the reflective mode be used to regulate arousal? As mentioned, the relation between children's arousal level and activated SIP mode is thought to follow an inverted U-shaped function, by which non-arousing or highly arousing situations directly activate the automatic mode, whereas moderately arousing situations activate the reflective mode. However, in many situations children's arousal levels will increase gradually (e.g., when they are awaiting their turn to play), first residing with the automatic mode, then activating the reflective mode when arousal levels become moderate, and then again activating the automatic mode when arousal levels become excessively high.

We propose that EF may determine whether children manage to remain in the reflective mode or are forced to switch to the automatic mode. For instance, children high in EF may be able to make a secondary appraisal of their emotion (e.g., evaluating whether their anger is justified) and carry out a deliberately selected emotion-regulating response (e.g., counting to ten, taking multiple perspectives) that down-regulates their arousal, whereas children low in EF will not be able to do so, causing arousal levels to further increase and forcing them to use automatic SIP. In support of this notion, empirical work has shown that children high in EF exhibit greater regulation of their physiological arousal levels in emotionally demanding situations and have better emotion regulation skills than children low in EF (Obradović, 2016; Obradović & Finch, 2016).

Children's EF may not only affect their predominant processing mode, but may also influence the quality of their reflective SIP. If children low in EF use the reflective processing mode, they may exhibit more errors and biases compared to children high in EF, who will be more accurate in their reflective processing. For instance, children low in EF may have difficulties to hold multiple response options and outcomes in mind and evaluate them adequately, which may steer their decision processes towards aggression in situations where aggression is the most accessible response option. Or, as another example, these children may attribute hostile intent in clearly benign social interactions, because they fail to inhibit a schema-driven tendency to assume others cannot be trusted. Indeed, research suggests that low EF (i.e., low focused attention, working memory, and inhibition) is linked to SIP biases, such as the generation and positive evaluation of aggressive responses (Van Nieuwenhuijzen et al., 2017; Van Rest et al., 2019).

In sum, our dual mode SIP model proposes that in social situations where children's arousal increases gradually, children low in EF are particularly prone to switch to the automatic mode, whereas children high in EF are able to remain in the reflective mode. Moreover, when using the reflective mode, children low in EF are expected to produce errors or biases, whereas children high in EF are expected to be more accurate in their processing.

DYNAMIC FACTORS INFLUENCING CHILDREN'S SIP MODE

Internal State

We have just described how stable dispositions may predispose children to use either the automatic or reflective mode. In addition to these factors, children's processing mode is affected by dynamic processes that may vary across situations and SIP cycles (see Figure 1). The dual mode SIP model proposes that children's arousal level in any given situation is affected by their internal state—the arousal, affect and cognitions that were already activated just before this situation. This internal state can be influenced both by preceding events and by physiological factors such as temperature, stress, frustration, fatigue, hunger and pain (Anderson & Bushman, 2002). Obviously, children may be more likely to use the automatic mode when their level of arousal just prior to the social event was already high, for example when they were already moody or fatigued, or when hostile memory structures were already activated. Conversely, children may be prone to use the reflective mode when they were concentrated or rested just prior to the social event, or when goal-related memory structures were already activated. In line with this idea, research has shown that experimentally manipulating children's mood exacerbates their aggressive SIP and behavior in subsequent provocative situations (De Castro et al., 2003; Dodge & Somberg, 1987). These findings illustrate that children's internal state affects their subsequent SIP, but do not show which mode was activated. Empirical work in adults suggests that hot temperatures or feeling hungry may activate the automatic mode: these internal states produced increases in arousal levels, hostile affect, hostile cognitions and aggressive tendencies (Anderson et al., 1995; Bushman et al., 2014). In sum, the dual mode SIP model proposes that children's internal state affects the likelihood of using the automatic versus reflective mode in forthcoming situations.

Type of Situation

A second dynamic factor affecting whether children engage in automatic or reflective SIP is the type of situation. Some situations may be more arousing than others, making children prone to use the automatic mode. Empirical research has demonstrated that children may show aggressive behavior across various situations, such as being threatened, provoked or disadvantaged, having to cope with competition, and dealing with authority figures (Dodge et al., 1985; Matthys et al., 2001). Which specific situations elicit high arousal levels differs between children, depending on their social database. That is, children may be sensitized to certain situations by their past experiences. For example, children who have been bullied in the past will likely experience high arousal levels when they encounter a group of unknown peers at the schoolyard, making them use the automatic mode. Conversely, children who have never been bullied, would in the same situation experience a moderate level of arousal (because it is a personally relevant but not threatening situation), making them use the reflective mode. Thus, our dual mode SIP model proposes that specific social situations affect whether children's SIP and behavior is steered by the automatic or reflective processing mode.

AGGRESSIVE BEHAVIOR AND ENVIRONMENTAL RESPONDING

Thus far, we have identified dispositional and dynamic factors that may affect whether children engage in aggression stemming from automatic or reflective SIP, and have described how children's social database may contribute to their aggressive SIP and behavior. Importantly, children's aggressive SIP patterns (in either mode) are not just a product of their own dispositions, but are also shaped by their environment (Dodge & Pettit, 2003; Dodge, 2009). Children's aggressive behavior may, at least in some cases, evoke hostile responses from their social environment that may reinforce children's aggressive SIP, initiating a new cycle of aggressive behavior and environmental responding. Longitudinal research has supported this cyclical process, demonstrating that children's tendency to attribute hostile intent to others may not only be a result from peer rejection, but may also contribute to future peer rejection (Lansford et al., 2010). The dual mode SIP model proposes that this vicious cycle may be established through the automatic mode: Hostile responses from children's social environment may strengthen certain memory structures (e.g., hostile schemas) for specific situations (e.g., provocation by peers), causing these situations to trigger automatic SIP faster and faster (e.g., in increasingly benign interactions), leading children to respond aggressively time and again, thereby creating a vicious cycle beyond children's cognitive control.

CLINICAL IMPLICATIONS

Our dual mode SIP model provides novel insights for the treatment of children's aggressive SIP and behavior. It illustrates how clinicians may most effectively target aggressive SIP in which children, under which conditions, and through which factors. We have described that children's aggressive behavior may derive from hyper-aroused automatic SIP, hypo-aroused automatic SIP, or reflective SIP. Each of these SIP styles may require a different intervention approach, with a tailored combination of intervention techniques that may be presented within the context of applied effective intervention approaches such as Cognitive Behavior Therapy (for a review, see: Smeets et al., 2015) and Behavioral Parent Training (for reviews, see: Leijten et al., 2013; McCart et al., 2006).

For children whose aggressive behavior primarily derives from hyper-aroused automatic SIP (i.e., highly emotional aggression), interventions may focus on arousal regulation by children themselves and their environment. These children could be taught cognitive and behavioral arousal regulation skills, such as deep breathing, focusing on helpful or joyful thoughts, or taking a time-out. For the environment, it will be helpful to understand under which circumstances and in which internal state a child is more likely to 'flip the switch' to automatic processing. This way, teachers and parents can foresee when a child is moody, tired or frustrated and approach these situations differently than they would have if the child was calm or rested. For instance, instead of appealing to reason (e.g., "Why are you doing this?"), they could help the child calm down or prevent the situation from escalating.

For children whose aggressive behavior primarily derives from hypo-aroused automatic SIP (i.e., automatic, callous aggression), interventions may focus on following social rules in the absence of emotional urges to do so. As this population of children may be challenging to treat, interventions may include the entire system around the individual child to provide consequent guidance and supervision. An option could be to teach parents and teachers how they can repeatedly rehearse simple interactional rules (e.g., "Stop when others say no") and practice prosocial behavior skills with these children in problematic situations, so that prosocial behavioral strategies eventually become part of children's automatic tendencies.

For children whose aggressive behavior primarily derives from reflective SIP (i.e., deliberately selected aggression), interventions may focus on changing the content of children's reflective processing. This may prove challenging, because these children consciously value the use of aggression to achieve their goals and may not see the benefit of replacing this behavior with prosocial response options. Nevertheless, it has been suggested that children's moral disengagement may be influenced by others: in younger children more so by the moral values of adults, and in older children more so by the moral values of their peers (Caravita et al., 2013). An option may therefore be to identify specific role-models (e.g., popular peers or adults) or other persons these children might look up to (e.g., famous athletes or musicians) and have them elaborate on the negative consequences of aggressive behavior based on their personal experience. Another option may be to offer these children meaningful roles (e.g., representative of their school or class, captain of their football team) that may yield the same magnitude of rewards (e.g., material or social) as their aggressive behavior (Ellis et al., 2016).

Another implication of the dual mode SIP model is that social-cognitive interventions may be most effective when children's deviant SIP patterns are targeted in the processing mode that is also active when they actually engage in aggressive behavior. This implies that interventions should use techniques that elicit similar arousal levels as are present when children engage in aggression in real-life. For instance, for children who engage in automatic aggressive behavior as a result of high arousal levels, it seems

most effective to practice with social situations that also elicit high levels of arousal (e.g., actually being provoked by a peer in a real-time interaction, or, possibly, in a virtual reality environment). Similarly, for children who engage in deliberate aggression as a result of moderate arousal levels elicited by opportunities to obtain instrumental gain (e.g., cheating, stealing), it may be most effective to target their SIP in situations where actual instrumental gain could be acquired.

The dual mode SIP model also provides inroads to change children's SIP indirectly, through targeting factors that are expected to contribute to children's dominant processing mode. For example, children who engage in aggressive behavior to counteract their hypo-arousal, could be taught more adaptive ways to seek stimulation (e.g., physical exercise or extreme sports), children who easily become hyper-aroused when they are tired could improve their sleeping hygiene (Miadich et al., 2020), children who exhibit cognitive errors when reflecting on social situations could receive an executive function training (Diamond & Ling, 2016), and children who display context-specific aggression could practice alternative non-aggressive behavior responses in the specific contexts that are most problematic.

In sum, our dual mode SIP model yields valuable opportunities to tailor treatment to the predominant SIP mode of individual children, illustrating which children may benefit most from which approaches under which conditions.

FUTURE RESEARCH DIRECTIONS

In this theoretical paper, we have proposed a dual mode SIP model to explain individual differences in children's aggressive SIP and behavior. Our model yields new directions for future research.

First, more research is needed on children's automatic versus reflective SIP and behavior. Current research is limited, simply for the reason that assessing automatic SIP is challenging. Children are unaware of their automatic SIP and asking them about it will involuntary activate their reflective SIP. Nonetheless, future research may attempt to assess automatic SIP by including indirect indicators, such as physiological arousal (e.g., skin conductance, heart-rate variability), reaction times, eye-movements, and observation of children's emotions and behavior. Such measures would allow to distinguish between automatic versus reflective SIP (e.g., high versus moderate physiological arousal, fast versus slower reaction times, few versus many eye-movements, observation of strong versus mild emotions, and observation of impulsive versus deliberate behavior). Alternatively, researchers could examine children's associative knowledge structures (e.g., using implicit association tasks or sentence completion tasks, see: Burks et al., 1999; Grumm et al., 2011), and assess the extent to which children's *online* SIP in various social situations resembles their *offline* knowledge structures, as an index of automaticity. Such measures would further our understanding of how children's SIP contributes to their aggressive behavior.

Second, future research may examine our model's prediction that children's arousal levels determine whether their SIP derives from the automatic or reflective processing mode. To test this, researchers may assess children's automatic versus reflective SIP (e.g., using reaction times as indicator of SIP automaticity) after manipulating their arousal levels, for instance by using interactive Virtual Reality in which children play games with virtual peers that vary in stakes and time pressure. In addition, researchers may use physiological measures during such experiments to link children's ANS functioning to their automatic versus reflective SIP patterns.

Third, more research is needed on executive functioning (EF) in relation to SIP. The dual mode SIP model predicts that children's EF has two different functions: arousal regulation, preventing children from switching to the automatic mode; and cognitive control, preventing children from making SIP errors. To disentangle these functions, it seems valuable to assess children's EF under stressful conditions (arousal regulation) and non-stressful conditions (cognitive control). This issue is addressed by research on cool and hot EF (e.g., Peterson & Welsh, 2014; Zelazo & Müller, 2002). Yet, these studies often base the distinction between cool and hot EF on the specific type of EF or task, not on actual arousal levels (as noted by Schoorl et al., 2018). For example, working memory is often considered a cool EF, although it may affect children's SIP differently when they are aroused. Thus, research is needed that tests each EF under hot conditions (e.g., by using incentives to motivate children to do well on the task, or when they are frustrated by negative feedback) and cool conditions (e.g., when they are not).

Fourth, given the detrimental impact that childhood adversity may have on children's SIP (e.g., Lansford et al., 2010), longitudinal research would be valuable to investigate how specific childhood adversities may shape children's SIP styles. For instance, children who have been chronically victimized may develop an automatic aggressive SIP style because such aversive experiences may have impaired their executive functioning and emotion regulation skills (for a review, see: Pechtel & Pizzagalli, 2011) and facilitated the development of hostile memory structures (Dodge, 2006). Or, as another example, children who grew up in environments where they frequently witnessed violence, may develop a reflective aggressive SIP style because they may have developed memory structures linking aggression to positive outcomes (Bandura, 1978; Guerra et al., 2003).

Fifth, more research is needed on children's automatic versus reflective SIP in relation to their reactive and proactive motives for aggression. Our dual mode SIP model predicts that both reactive and proactive aggression may be preceded by automatic as well as reflective SIP. It would be interesting to investigate how strongly the two dimensions are actually related, and whether distinct subtypes of children may exist, such as children who tend to routinely bully others (automatic-proactive) or calmly plan their revenge (reflective-reactive).

Sixth, the dual mode SIP model may provide researchers with an explanatory framework to investigate clinical syndromes such as intermittent explosive disorder, mood-, or anxiety disorders. Our dual mode SIP model would predict that children who are predisposed towards hyper-emotional reactivity are more likely to be diagnosed with such disorders, especially if they have limited executive functioning capacities. Such children may frequently depend on the automatic mode in specific disorder-relevant contexts (e.g., social threat, personal failure, uncertainty), especially when they already were in a negative emotional state (e.g., depressed or irritable mood), triggering responses that can be recognized as clinical symptoms (e.g., tantrums, hostility, rumination, or anxiety). In addition, children who are predisposed towards hypo-emotional reactivity may be more likely to be diagnosed with conduct disorder, especially if they are insensitive to punishment or others' distress cues. Such children may depend on the automatic mode in situations that would trigger the reflective mode in others (e.g., victimizing others), or on the reflective mode in stressful situations that would trigger the automatic mode in others (e.g., being pressured by peers to engage in criminal activities), triggering responses that can be recognized as clinical symptoms (e.g., callous bullying, calculated burglary).

Last, future research may link children's automatic versus reflective SIP to brain functioning (Beauchamp & Anderson, 2010). It may be that activity in specific brain structures contribute to the dominance of either the automatic or reflective processing mode. For instance, activation in the amygdala has been associated with the emotional salience of initial, automatic appraisals, facilitating *fight*-or-*flight* tendencies, whereas activity in the prefrontal cortex (PFC) has been associated with the deliberate reprocessing of stimuli, facilitating cognitive and behavioral control (Cunningham et al., 2007; Cunningham & Zelazo, 2007). We would hypothesize that the interplay between the affective input from the amygdala and executive control of the PFC may determine whether children's SIP is predominantly automatic or reflective. Similarly, researchers may examine how the interaction between activity in specific brain areas and ANS functioning may predict children's automatic versus reflective SIP. Although research linking SIP to brain- and ANS functioning is still in its infancy, emerging evidence suggests that there is much to benefit from integrating these perspectives (e.g., for reviews, see: Adolphs, 2009; Insel & Fernald, 2004; Krain et al., 2006).

CONCLUSION

This article has presented a dual mode social information processing (SIP) model that predicts which processing steps which children take, under which circumstances, and how this may lead to aggression. This dual mode SIP model distinguishes between children's automatic versus reflective SIP, and describes how the dominance of either mode is determined by an interplay between child-specific factors (i.e., emotional dispositions, motivational dispositions, and executive functioning) and dynamic factors (i.e., internal state and type of situation). We hope this dual mode SIP model may further our understanding of children's deviant SIP underlying their aggressive behavior and will help to identify promising targets for theory development, empirical research, and intervention.

CHAPTER 7

Summary & General Discussion

CHAPTER 7: SUMMARY & GENERAL DISCUSSION

The **overarching aim** of this dissertation was to improve our understanding of the social information processing (SIP) patterns underlying children's aggressive behavior. Previous SIP studies only explained relatively small proportions of variation in children's concurrent and future real-life aggression (e.g., De Castro et al., 2002; Dodge et al., 1986; Lansford et al., 2006). This may have been due to three knowledge gaps in the current SIP literature: (1) the limited ecological validity of current SIP assessment methods; (2) the use of analytical methods that may not uncover the individual differences in children's aggressive SIP patterns because they study average SIP patterns over groups of children instead of clustering children based on their SIP patterns; (3) the current theoretical SIP models that only provide a clear explanation of how children's calm, deliberate SIP may contribute to aggression, but do not capture the emotional nature of children's SIP that may lead them to respond aggressively without any deliberate thought.

The current dissertation aimed to fill these knowledge gaps in the current SIP literature by answering three research questions:

1. Can interactive Virtual Reality (VR) provide a more ecologically valid assessment of children's aggressive SIP?
2. Is it possible to distinguish SIP profiles of children with aggressive behavior problems?
3. Can we specify an improved theoretical model to explain aggressive SIP and behavior?

The answers to these research questions may advance our understanding of how children's SIP contributes to their real-life aggressive behavior, allowing inroads for clinical practice to more effectively tailor cognitive-behavior interventions to individual children.

Summary of Findings

This dissertation started with a multi-level meta-analysis on the relation between one aspect of children's SIP (i.e., the tendency to attribute hostile to others) and their real-life aggression (using parent-report, teacher-report, self-report, observation, and peer nomination measures; **Chapter 2**). The aim of this meta-analysis was to acquire more insight in the child-specific and methodological determinants of the strength of this relation (i.e., moderators). Results revealed a positive association between children's tendency to attribute hostile intent to others and their real-life aggression, but the strength of this relation varied widely between studies. A key finding was that this relation appeared to be stronger for studies that used emotionally engaging SIP assessment methods. These meta-analytic findings indicate that emotionally engaging SIP assessment methods may yield a more ecologically valid assessment of children's aggressive SIP. Ideally, an ecologically valid SIP assessment would thus present children with highly emotional engaging, realistic, standardized social interactions.

Can Interactive Virtual Reality Provide a More Ecologically Valid Assessment of Children's Aggressive SIP?

To provide a more ecologically valid, standardized SIP assessment, we developed an interactive VR classroom where children could walk around freely, talk to virtual peers, and play games (i.e., building a tower of blocks, throwing cans from a table with five balls), allowing us to present standardized social events within an engaging, standardized environment. We extended previous SIP research by presenting children with distinct social events to assess their reactive and proactive aggressive SIP and behavior, allowing for more accurate assessment of individual differences in children's aggressive SIP and behavior. To this end, we presented children with four experimental scenarios, two scenarios included the opportunity to obtain instrumental gain (to assess children's proactive SIP and behavior) and two scenarios included being provoked by a peer (to assess children's reactive SIP and behavior). For a description of each scenario, see Table S1 in Chapter 3.

Chapter 3 includes a pilot study that tested whether our interactive VR environment provides a valid assessment of children's SIP. Our interactive VR assessment of SIP demonstrated good convergent validity (i.e., moderate to large correlations with SIP assessed through vignettes for all SIP variables except anger). Interactive VR also yielded improved measurement sensitivity (i.e., larger variances in SIP compared to vignettes) for aggressive responding, but not for other SIP variables. Discriminant validity (i.e., VR contexts evoked distinct SIP patterns) of interactive VR was supported for provocation contexts, but not for instrumental gain contexts. Taken together, our pilot study suggests that interactive VR is a promising tool to assess children's aggressive SIP and behavior.

7

Based on our findings, we made a few minor adaptations to our interactive VR assessment for our next studies (i.e., to our instructions for children, verbal responses of virtual characters). Although discriminant validity was not supported for the instrumental gain contexts, we did not make any changes to the instrumental gain contexts. We had two reasons: (1) instrumental SIP and aggression is relatively rare (Dodge et al., 1997; Thomson & Centifanti, 2018) while the sample size in our pilot study was relatively small; (2) children in our pilot study, on average, showed more instrumental SIP and aggression in instrumental gain contexts versus other contexts. Thus, the lack of discriminant validity for instrumental gain contexts seemed most likely to be due to a lack of power.

An important next step was to test whether our interactive VR assessment of children's SIP and behavior predicts their real-life aggression as reported by their teachers better than a standard vignette-based SIP assessment (Boateng, et al., 2018). **Chapter 4** therefore concerns an external validation study that directly tested whether interactive VR provides a more ecologically valid assessment of children's SIP than a standard vignette-based SIP assessment. For the vignette-based SIP assessment, we developed audiotaped vignettes with the exact same content as the VR scenarios (e.g., describing how participants would play games and gain or lose high scores and bonuses), allowing for a clean comparison between assessment methods. We counterbalanced the type of game across participants (i.e., participants who received the tower game in interactive VR, received the cans game with vignettes, and vice versa). Results demonstrated that, compared to vignettes, our interactive VR assessment evoked more emotional engagement and immersion, higher levels of aggressive SIP in provocation scenarios, and explained additional variance in children's teacher-reported aggression. Unexpectedly, interactive VR seemed to outperform vignettes more for provocation scenarios than for instrumental gain scenarios. Last, results revealed that interactive VR particularly yielded an improved assessment of children's interactional goals and behavior, but did not clearly outperform vignettes for anger and hostile intent attributions. Taken together, interactive VR seems to provide a more ecologically valid assessment of children's aggressive SIP patterns and behavior than a standard vignette-based assessment. For this reason, interactive VR may provide a more accurate assessment of individual differences in children's aggressive SIP and behavior than traditional vignettes.

Is it Possible to Distinguish SIP Profiles of Children with Aggressive Behavior Problems?

As children differ considerably in the SIP patterns leading up to their aggressive behavior (De Castro & Van Dijk, 2017; Dodge, 2011), it makes sense to try to identify SIP profiles in children with aggressive behavior problems. **Chapter 5** therefore included an empirical study that tested whether it is possible to distinguish SIP profiles among boys with aggressive behavior problems. Latent profile analyses revealed four distinct SIP profiles: a general reactive SIP profile, a situation-specific reactive SIP profile, a mixed reactive-proactive SIP profile, and a non-aggressive SIP profile. Planned contrasts demonstrated that boys with these SIP profiles differed in temperament, aggressive belief systems, and sensitivity to punishment, but not in executive functioning and sensation seeking. These findings seem to provide further validation for the existence of distinct SIP profiles in boys with aggressive behavior problems. In sum, this study demonstrated, for the first time, that distinct SIP profiles can be distinguished in children with aggressive behavior problems, allowing inroads for clinical practice to more effectively tailor cognitive-behavior interventions to individual children.

Can we Specify an Improved Theoretical Model to Explain Aggressive SIP and Behavior?

The current dissertation suggests that children's emotional engagement may strongly affect the nature of their SIP, but how precisely is not yet clearly specified in current SIP models. In **Chapter 6** we therefore proposed a new theoretical SIP model that specifies how and why children's emotional engagement has such a profound influence on their SIP. Our new SIP model explains how children's arousal levels may contribute to fast, emotion-driven aggression preceded by automatic SIP as well as deliberate, controlled aggression preceded by reflective SIP based on an interplay between child-specific (i.e., emotional dispositions, executive functioning, motivational dispositions) and situations-specific factors (i.e., internal state and type of situation). Our theoretical SIP model thus aims to provide a more ecologically valid explanation of children's unique SIP styles, allowing to more effectively tailor treatment to children's individual needs.

Unexpected Findings and Required Adaptations of our Interactive VR Measure to Assess Children's Aggressive SIP and Behavior

Our research on interactive VR assessment of children's SIP yielded several unexpected findings, indicating further improvements to our VR measure are required.

First, our interactive VR assessment particularly yielded an improved assessment of children's interactional goals and responses, but less improvement for children's anger and hostile intent attributions. For anger, we decided to rely on self-report, because we conducted the first empirical studies on the use of interactive VR to assess children's SIP. We therefore preferred to use similar SIP assessment methods as in most earlier SIP studies. It has been shown, though, that individuals may find it difficult to accurately report on their emotional states, which may have obscured our findings on children's anger (Robinson & Clore, 2002). To provide a more valid assessment of children's emotional states, the use of self-report could be combined with more objective measures, such as physiological indicators (e.g., skin conductance, heart rate; Hubbard et al., 2002, 2004).

For children's hostile intent attributions, our findings may be a result of suboptimal ambiguity of our VR scenarios: As the mean scores of children's hostile intent attributions in interactive VR were quite high, this may indicate that our VR scenarios may have been perceived as less ambiguous than intended. This may have led children to attribute hostile intent irrespective of their tendency to respond aggressively in real life. To improve the assessment of children's hostile intent attributions, it is important to subtly modify the scenarios to render the intentions of virtual characters more truly ambiguous. Although the content of the provocation scenarios in VR was identical to the vignettes (in Chapter 4), the ambiguity of the VR scenarios may have been complicated by specific characteristics of the VR environment. After being provoked by a virtual peer (which is the end of the scenario in vignettes), the VR scenario does not end directly but proceeds to allow children to respond to the provocation. To keep our provocation scenarios as ambiguous as possible, we decided, though, to have the virtual peer remain silent after the provocation, even if participants asked the virtual peer about what happened. This awkward silence may have facilitated children's hostile intent attributions. To deal with this unforeseen issue, we suggest testing the ambiguity of a set of virtual characters' responses following provocation and select the response options that are being perceived as ambiguous.

Second, our interactive VR assessment of children's SIP yielded larger incremental value for provocation scenarios compared to instrumental gain scenarios. This finding indicates that interactive VR provides an improved assessment of children's reactive aggressive SIP and behavior, which is particularly useful because it is the most prevalent form of aggression (e.g., Dodge et al., 1997; Van Dijk et al., 2021). On the other hand, the incremental value of interactive VR to assess children's proactive aggressive SIP and behavior was less pronounced. An explanation may be that our instrumental gain scenarios evoked less emotional engagement than our provocation scenarios, reducing the incremental value of the instrumental gain scenarios in interactive VR compared to vignettes. This would imply that stealing or cheating during a game in interactive VR is less arousing than being provoked. If so, more salient or tempting instrumental cues may be required when assessing children's proactive aggressive SIP and behavior using interactive VR (e.g., to have the instrumental gain in VR constitute actual gain in the real world). Alternatively, it could be that interactive VR simply has less incremental value compared to vignettes when assessing children's proactive aggressive SIP and behavior. Some children who use aggression to obtain instrumental gain may also report this in response to vignettes as they consciously value the use of aggression. To test the incremental value of interactive VR for assessing children's proactive SIP and behavior, researchers may recruit a large sample of children, specifically selected on their tendency to engage in proactive aggression, and use a broad variety of tempting instrumental cues.

Third, we expected that the engaging nature of interactive VR would also evoke larger individual differences in children's aggressive SIP and responses compared to vignettes. However, interactive VR and vignettes generally evoked similar variances in children's aggressive SIP and responses. Perhaps our vignettes validly assessed individual differences in children's calm SIP; that is, the way they would reflect on social situations when they are calm. Such calm SIP may differ as much between children as their emotional SIP, but would be less suitable to predict children's real-life aggression. Indeed, our findings showed that interactive VR yielded incremental predictive value above and beyond the vignette-based assessment in all four scenarios, both for the prediction of children's real-life aggression and underlying reactive and proactive motives.

Last, our latent profile analyses (LPA) yielded two unexpected findings. We did not find the expected proactive SIP profile. Prior research aiming to identify subgroups of children based on their reactive versus proactive motives for aggression has found mixed evidence for the existence of a predominantly proactive aggressive group, with some studies finding such a group (Carrol et al., 2018; Van Dijk et al., 2021), but others not (e.g., Euler et al., 2017; Marsee et al., 2014; Munoz et al., 2008; Smeets et al., 2017; Thomson & Centifanti, 2018). As previously detected proactive aggressive groups have been relatively small (e.g., Van Dijk et al., 2021), it is possible that our study failed to detect it because of the limited variance in proactive SIP (we had one dichotomous proactive SIP indicator only: instrumental goals).

Also unexpected was the obtained distinction between a general and situation-specific reactive aggressive SIP profile, but this aligns with empirical work that demonstrated children's SIP may depend on situational cues (De Castro & Van Dijk, 2017; Matthys et al., 2001; Dodge et al., 1985). Some children may be sensitized to specific situations, perhaps through past experience (Dodge, 2006; Matthys et al., 2001). For example, children whose property have been ruined by peers in the past will be likely to show aggressive SIP in similar situations (e.g., when their game is been ruined by a peer), but less so in other situations (e.g., when being excluded from a game by their peers). Alternatively, the design of our research may provide another explanation for the two unexpected findings in our LPA. Participants completed all scenarios in the same order, starting with two instrumental gain scenarios (used to assess proactive SIP), followed by two provocation scenarios (used to assess reactive SIP). We decided to present the provocation scenarios last because we expected these scenarios to evoke the strongest emotions, hoping to minimize carry-over effects. This decision seems justified because for each SIP profile, anger levels increased with each subsequent scenario. Nevertheless, we cannot rule out the possibility that earlier scenarios affected children's SIP and behavior in subsequent scenarios. If so, our findings concerning the existence of two reactive aggressive SIP profiles and the absence of a purely proactive aggressive SIP profile need to be interpreted cautiously, as they might partly be due to a methodological artefact.

Scientific Implications for the Development of Interactive VR to Assess Children's Aggressive SIP and Behavior

The effects of emotional engagement on children's SIP have been understudied. With interactive VR as new SIP assessment method and a new theoretical SIP model, important first steps have been taken to advance our understanding of SIP underlying children's aggressive behavior. By providing children with an emotionally engaging, realistic, flexible and standardized environment, interactive VR yields an improved assessment of their aggressive SIP and a better prediction of their real-life aggression than has been feasible thus far. Interactive VR opens up a broad variety of new opportunities for research. For instance, it may be used to study relational aggression in girls, bullying behavior in school-aged children, oppositional behavior towards authority figures, and delinquency in juveniles. We hope that our work may encourage scientists to use emotionally engaging, ecologically valid assessment methods to study human behavior. Several considerations warrant attention when researchers indeed decide to use interactive VR for research purposes.

First, when developing an interactive VR environment it is important to find a balance between on one hand the interactive nature of VR (which might enhance realism and engagement), and on the other hand the ambiguity of the virtual characters' behavior (which is important to assess children's hostile intent attributions) and the standardization of presented social events (which is important for adequate internal validity). To this end, researchers need to make a careful selection of the virtual characters' verbal and physical responses to be able to adequately respond to a broad range of participants' behaviors in interactive VR and at the same time maintain ambiguity and standardization across participants.

Second, to maximize participants' natural responses in the VR environment, they should —before being presented the assessment scenarios—be familiar with how they can respond in the interactive VR environment and how the VR environment responds to them. This can be achieved by making the VR environment as similar to the real world as possible and to have participants explore the VR environment in several practice scenarios. In addition, the virtual characters' responses in the VR environment should be adequately timed by the experimenter. For instance, some lag in the virtual characters' responses may unintentionally remind participants of interacting in a computerized environment rather than a real-world like context which often involves more fluent social interactions. Hence, experimenters should exhibit sufficient skills in handling the interactive VR interface, which requires practice.

Third, to ensure a valid SIP assessment, the VR environment needs to evoke sufficient emotional engagement in children. To attain this goal, certain incentives can be incorporated in the VR environment. In our interactive VR assessment of SIP, we therefore used games that are challenging but not too difficult, allowed participants' to obtain high scores for their performance during these games, and included a time-limit in each VR-scenario. Alternatively, for researchers assessing relational aggression, other incentives may be more appropriate. They may enhance children's emotional engagement by having participants' believe they will be evaluated by alleged peers based on their behavior in interactive VR and relate this to likes and dislikes using a social-media paradigm (Lee et al., 2020).

Fourth, an important consideration when using interactive VR to assess aggression is that participating children should believe that they are interacting with real other children who can actually be harmed. This is not self-evident, as many children who never engage in aggression in real life do play digital games where they aggress against characters they know do not exist. For example, in our studies we presented our virtual classroom as an actual classroom where participants allegedly met with real children from other schools who also participated in our studies and were simultaneously logged on to the VR environment.

Last, when presenting multiple VR-scenarios it is important to consider potential order-effects: children's emotional arousal and SIP in the scenarios presented first may be carried over to subsequent scenarios. For this purpose, the order of the scenarios can be counterbalanced or scenarios can be presented on different assessment days.

Clinical Implications of Using Interactive VR to Assess and Intervene on Children's Aggressive SIP and Behavior

For clinical practice, interactive VR—if it is further validated by future research—may provide a more attractive, flexible, and ecologically valid method to assess and treat children's aggressive SIP. Clinicians could use a flexible VR environment to create emotionally engaging exercises tailored to individual clients, with precise control to adapt difficulty and complexity during the intervention. For instance, for children who tend to show reactive SIP across a broad range of provocative social situations, interactive VR may be used to have them gradually practice with anger exposure and challenging hostile cognitions in a variety of provocative situations (e.g., by varying the type of the provocation and gradually increasing the severity of the provocations). For children who tend to show situation-specific reactive SIP, interventions using interactive VR may be even more specific, aiming to challenge children's reactive aggressive SIP patterns only in those situations that are most problematic to them. In addition, to treat children's proactive aggressive SIP, practitioners may use interactive VR to have children gradually practice with a variety of situations that tempt them to use aggression to obtain instrumental or relational gain (e.g., by gradually increasing the stakes) and repeatedly rehearse prosocial response strategies. Such application of interactive VR to intervention is currently being tested, based on our experience with the present VR (Alsem et al., 2021).

Our findings on the existence of distinct SIP profiles among children with aggressive behavior problems may also shed another light on standardized diagnostic assessment methods used in clinical practice. Given that children with aggressive behavior problems may show very different underlying SIP profiles, it does not seem most informative to use a standardized SIP assessment instrument that is identical for each child and evaluates to what extent a child deviates from the population norm on each SIP step. Rather than presenting children with a norm-based standardized SIP assessment instrument, practitioners may use a functional analysis to uncover the specific situational triggers and associated aggressive SIP and behavior patterns in specific children. Such a tailored diagnostic approach may provide a more accurate assessment of the unique SIP styles underlying children's aggressive behavior problems, allowing to more effectively tailor cognitive-behavioral interventions to the individual child.

Strengths and Limitations

The current dissertation had several strengths and limitations. Strengths of this dissertation primarily relate to its innovative nature. First, our multi-level meta-analysis yielded an extensive synthesis of empirical work on the relation between children's SIP and real-life aggression, including studies from over a timeframe of 40 years, using multiple effect sizes per study. This allowed us to reliably test novel hypotheses on child-specific and methodological determinants of the strength of this relation to explain the divergent findings between and within studies. Our meta-analytic findings provided valuable starting points for the development of our interactive VR assessment of children's SIP. Second, we conducted the first empirical studies thus far that used interactive VR to assess children's aggressive SIP and behavior. This yielded a) a more ecologically valid assessment of children's aggressive SIP and behavior than a standard vignette-based SIP assessment, and b) provided valuable insights into the advantages and disadvantages of using interactive VR to assess children's aggressive SIP and behavior. Third, we extended previous SIP research by presenting children with distinct social contexts to assess their reactive and proactive aggressive SIP and behavior, allowing for more accurate assessment of individual differences in children's aggressive SIP and behavior patterns. Fourth, we demonstrated, for the first time, that distinct SIP profiles can be distinguished among children with aggressive behavior problems, providing clinical practice with opportunities to more effectively tailor cognitive-behavioral interventions to individual children. Last, we proposed a new theoretical SIP model that provides a more ecologically valid, fine-grained and testable explanation of the SIP patterns leading up to children's aggressive behavior.

This dissertation also has its limitations. These relate to the generalizability of our findings, the selection and presentation order of VR scenarios, the measurement of SIP, and the assessment of emotional engagement. First, as we conducted the first empirical studies to examine whether interactive VR yields a valid SIP assessment, we aimed to include relatively homogenous samples (i.e., boys between 7-13 years, with limited diversity in cultural and socio-economic background). As we know from earlier empirical work that children's SIP patterns may be affected by their gender, age, and cultural or socio-economic background (Coie & Dodge, 1998; Dodge et al., 2003, Ostrov & Godleski, 2010, Lansford et al., 2006; Ziv & Sorongon, 2011), our findings therefore cannot be directly generalized to other subpopulations.

Second, as interactive VR is obviously time-consuming and costly to develop, we were only able to include four assessment scenarios. Given that children may show aggression in various contexts (De Castro & Van Dijk, 2017), it can be assumed that using only four scenarios involving playing games with peers in a school classroom did not cover the broad range of social situations known to evoke aggression in children. Moreover, as we presented the four VR-scenarios in the same order to all participants, we could not disentangle order-effects from scenario-effects. To minimize carry-over effects, we decided to present the scenarios that were expected to arouse the strongest emotions last. Indeed our data revealed children's anger levels increased with each subsequent scenario. In addition, despite potential carry-over effects, children's SIP and behavior in all VR scenarios predicted their real-life aggression above and beyond vignettes. Future research may extend our work by using other social contexts, for example aimed at assessing children's relational aggression (e.g., gossiping).

Third, we assessed children's SIP by explicitly asking them about each SIP step at the end of each VR scenario. This yielded an improved assessment of children's SIP compared to vignettes, possibly because with interactive VR children reflect on their actual SIP instead of on their hypothetical SIP. However, when children experience strong emotions in interactive VR, many of them may instantly respond with aggression without any reflection on intent attributions or decision processes (Chapter 6; De Castro 2005; 2012; Helmsen et al., 2012). For these children, aggression may derive from automatic SIP which is not captured when asking children to reflect on each SIP step, because they are unaware of their automatic SIP patterns. Consequently, our interactive VR assessment only captured children's reflective SIP. This may explain why, although interactive VR yielded an improved assessment of children's SIP and behavior compared to vignettes, there still remained relatively large proportions of unexplained variance in their real-life aggression.

Fourth, for our measurement of children's outcome expectancies and evaluations of the use of aggression, we asked children to report on their outcome expectancies and evaluations of their actual aggressive responses in VR. We decided to measure these variables in this way because it seems less valid to have children evaluate hypothetical (aggressive) responses they would never carry out in real life. Nevertheless, this method forced us to exclude these proactive SIP variables from our analyses because an insufficient number of children actually aggressed in the instrumental gain VR scenarios. As a result, we could test our hypotheses in the different studies for one proactive SIP variable only (i.e., instrumental goals), prohibiting us, for example, from testing whether interactive VR yields an improved assessment of the other two proactive SIP variables (outcome expectancies and response evaluations), and limiting variance in proactive SIP in our latent profile analyses.

Last, we assessed children's emotional engagement in interactive VR based on all scenarios together instead of for each scenario separately. This prohibited us from examining the unique effects of children's emotional engagement on their SIP in each provocation and instrumental gain scenario. Consequently, we were not able to test whether the larger incremental value of interactive VR for provocation scenarios compared to instrumental gain scenarios was due to differences in the level of emotional engagement both types of social contexts elicited in children.

Future Research Directions

This dissertation yields several future research directions that may further our understanding of children's aggressive SIP and behavior. First, interactive VR allows researchers to conduct an abundance of experiments to test hypotheses about how children's emotions may affect their SIP and behavior. For instance, children's emotional engagement in interactive VR may be manipulated by gradually adding incentives to test whether this exacerbates their aggressive SIP and behavior in a broad range of social situations including provocation as well as opportunities to obtain instrumental gain (De Castro et al., 2003; Dodge & Somberg, 1987; Reijntjes et al., 2011). In addition, interactive VR may be used to test the effect of a variety of emotions on children's SIP: not only anger or frustration, but also fear, sadness, shame, or guilt. For example, researchers may present children with specific social situations to induce feelings of guilt (e.g., by having children ruin a virtual characters' game) or shame (e.g., having children fail at a game in front of a group of virtual peers), and test whether children's SIP includes signs of worry about others' evaluations. Similarly, researchers may present children with social situations that evoke positive emotions such as happiness or joy (e.g., by having children succeed at a game or receive positive feedback from a virtual peer), and examine whether children's SIP and behavior becomes more prosocial and cooperative.

Second, the large experimental control over social stimuli provided by interactive VR allows researchers to test a wealth of hypotheses about the impact of specific social stimuli on children's SIP. Researchers may examine children's encoding of social cues by manipulating virtual characters' facial expressions (e.g., hostile, neutral, friendly). They may test, for example, if children attend more to schema-congruent or schema-incongruent cues and whether this is dependent on their arousal levels and subsequent automatic versus deliberate encoding. In addition, future studies may examine children's intent attributions and generation of responses by manipulating the ambiguity virtual characters' intentions. For instance, the virtual character's intentions could be manipulated into ambiguous versus non-ambiguous to test the idea that children's past social experiences only steers their intent attributions in ambiguous situations. Or as another example, experiments may induce hostile versus benign intent attributions to test whether children's aggressive behavior results from their inability to generate non-aggressive response options. The use of interactive VR would also allow to test hypotheses on the flexibility of children's SIP. For instance, the social bond with virtual characters may be manipulated (e.g., into disliked, unknown, or liked peer) to test whether children with aggressive behavior problems show aggressive SIP irrespective of their social bond with the interactional partner (indicating inflexible SIP), or only with disliked others (indicating flexible SIP). Relatedly, researchers may present children with several VR scenarios with the same virtual character whose intentions are ambiguous in the first scenario, benign in the second scenario, and again ambiguous in the third scenario. If SIP of children with aggressive behavior problems is less flexible, it would be expected that the difference between their intent attributions in the first versus third scenario is smaller compared to children without aggressive behavior problems. Interactive VR may also be used to test hypotheses on peer group dynamics. For example, researchers may test how peer group pressure may exacerbate children's aggressive SIP and bullying behavior while manipulating certain features of the virtual characters, such as social status (e.g., popular, unpopular) or demographics (i.e., age, gender, cultural background).

Third, researchers may investigate the validity of using interactive VR to directly assess children's temperament, motivational dispositions, and executive functioning, perhaps providing a more ecological valid assessment than questionnaires or computerized tasks. For instance, children's anger-frustration temperament may be assessed by measuring their anger and frustration while presenting them with challenging tasks that differ in difficulty (moderate, difficult, extremely difficult). Similarly, children's sensitivity to punishment and reward could be assessed by measuring children's arousal levels when presenting them with a sequence of trials including rewards and punishment (e.g., by awarding and deducting points during a game) that differ in magnitude (e.g., mild, moderate, severe). Or as another example, children's cool and hot executive functioning may be assessed by presenting them with a sequence of stressful and non-stressful social events in which they are asked to inhibit certain behavior responses (to assess inhibition or self-control), or to decide on risky versus non-risky response options (to assess risky decision making; Hobson et al., 2011). Relatedly, researchers could manipulate multiple social cues in a sequence of trials (e.g., virtual character's facial expression, posture, verbal speech) and instruct children to fixate on each of these cues for a few seconds under stressful and non-stressful conditions, and have them recall and integrate these cues at the end of each trial (to assess working memory).

Fourth, future studies may use interactive VR to test predictions of our theoretical SIP model. For instance, researchers may manipulate children's arousal levels in interactive VR (e.g., by having them play games with virtual peers that vary in stakes and time pressure) to test whether children engage in reflective SIP when arousal levels are moderate and engage in automatic SIP when arousal levels are low or high (Obradović, 2016; Yerkes & Dodson, 1908). Relatedly, interactive VR could be used to test the influence of children's dispositions (e.g., temperament, executive functioning, motivational dispositions) on their arousal levels and subsequent SIP patterns. For example, children with a highly emotional reactive temperament may be presented with VR scenarios that evoke different levels of arousal to test whether they exhibit more automatic SIP in highly stressful situations. Likewise, children with callous and unemotional traits may be presented with mild and severe social stressors (e.g., by having them be provoked by a group of older peers or by a single younger peer in interactive VR) to test its influence on their arousal levels and subsequent SIP. Such a study would allow, for example, to test our model's prediction that these children show blunted arousal levels and automatic SIP when being confronted with mild stressors, and moderate arousal levels and reflective SIP when being confronted with severe stressors.

Last, scientists may further specify and extend our new theoretical SIP model. For instance, researchers may examine how the interplay between specific contextual- (e.g. dysfunctional parenting or peer relations) and dispositional risk factors (e.g., temperament, executive functioning) affects children's SIP styles and subsequent aggressive behavior patterns over time. Further, we know that children's arousal levels and subsequent SIP patterns are modulated by biological and brain systems (Adolphs, 2009; Branje & Koot, 2018; Insel & Fernald, 2004; Krain et al., 2006). Our theoretical SIP model may therefore be extended by, for example, specifying how the interplay between the limbic system and prefrontal cortex influences children's automatic versus reflective SIP through their autonomic nervous system functioning.

General conclusion

This dissertation emphasizes the appreciation of emotional engagement as pivotal to our understanding of the SIP patterns leading up to children's real-life aggression. As interactive VR presents children with emotionally engaging, realistic, standardized social interactions, we are now able to better predict their real-life aggressive behavior. Given the advantages of interactive VR, it may open up a broad variety of new opportunities for SIP research and tailored interventions that has not been feasible thus far. In addition, our new theoretical SIP model may help scientists to better explain individual differences in children's aggressive SIP and behavior by providing a more ecologically valid, testable and fine-grained explanation of children unique SIP styles. In sum, we hope that this dissertation may inspire scientists and clinicians to use interactive VR and continue with enhancing our knowledge of children's SIP underlying their real-life aggression.

APPENDICES

References
Summary in Dutch (Samenvatting in het Nederlands)
Acknowledgements (Dankwoord)
Scientific Publications

REFERENCES

Adolphs, R. (2009). The social brain: neural basis of social knowledge. Annual Review of Psychology, 60(1), 693-716. http://doi.org/10.1146/annurev.psych.60.110707.163514

Alsem, S. C. Van Dijk, A., E. E. Verhulp, & De Castro, B. O. (2021). Using virtual reality to treat aggressive behavior problems in children: A feasibility study. Clinical Child Psychology and Psychiatry. Advance online publication. https://doi.org/10.1177/13591045211026160

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. Annual Review of Psychology, 53(1), 27-51. http://doi.org/10.1146/annurev.psych.53.100901.135231

Anderson, C. A., Deuser, W. E., & DeNeve, K. M. (1995). Hot temperatures, hostile affect, hostile cognition, and arousal: Tests of a general model of affective aggression. Personality and Social Psychology Bulletin, 21(5), 434-448. http://doi.org/10.1177/0146167295215002

Aron, E. N., Aron, A., & Jagiellowicz, J. (2012). Sensory processing sensitivity: A review in the light of the evolution of biological responsivity. Personality and Social Psychology Review, 16(3), 262-282. http://doi.org/10.1177/1088868311434213

Arsenio, W. F., Adams, E., & Gold, J. (2009). Social information processing, moral reasoning, and emotion attributions: Relations with adolescents reactive and proactive aggression. Child Development, 80(6), 1739-1755. https://doi.org/10.1111/j.1467-8624.2009.01365.x

Bandura, A. (1978). Social learning theory of aggression. Journal of Communication, 28(3), 12-29. https://doi.org/10.1111/j.1460-2466.1978.tb01621.x

Barlett, C. P., & Anderson, C. A. (2011). Reappraising the situation and its impact on aggressive behavior. Personality and Social Psychology Bulletin, 37(12), 1564-1573. http://doi.org/10.1177/0146167211423671

Barrett, P. M., Rapee, R. M., Dadds, M. M., & Ryan, S. M. (1996). Family enhancement of cognitive style in anxious and aggressive children. Journal of Abnormal Child Psychology, 24(2), 187-203. https://doi.org/10.1007/bf01441484

Baum, A., Garofalo, J. P., & Yali, A. M. (1999). Socioeconomic status and chronic stress: Does stress account for SES effects on health? Annals of the New York Academy of Sciences, 896(1), 131-144. https://doi.org/10.1111/j.1749-6632.1999.tb08111.x

Beauchaine, T. P. (2001). Vagal tone, development, and Gray's motivational theory: Toward an integrated model of autonomic nervous system functioning in psychopathology. Development and Psychopathology, 13(2), 183-214. https://doi.org/10.1017/S0954579401002012

Beauchamp, M. H., & Anderson, V. (2010). SOCIAL: An integrative framework for the development of social skills. Psychological Bulletin, 136(1), 39-64.https://doi.org/10.1037/a0017768

Beck, A. T., Freeman, A., & Davis, D. D. (2004). Cognitive Therapy of Personality Disorders (2nd ed.). Guilford Press.

Becker, B. J. (2005). Failsafe N or file-drawer number. In H. R. Rothsetin, A. J. Sutton, & M. Borenstein (Eds.). Publication bias in meta-analysis: Prevention, assessment and adjustments (pp. 111-125). John Wiley & Sons. https://doi.org/10.1002/0470870168.ch7

Bickett, L. R., Milich, R., & Brown, R. T. (1996). Attributional styles of aggressive boys and their mothers. Journal of Abnormal Child Psychology, 24(4), 457-472. https://doi.org/10.1007/BF01441568

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. Frontiers in Public Health, 6, 149. https://dx.doi.org/10.3389/fpubh.2018.00149

Bookhout, M. K., Hubbard, J. A., & Moore, C. C. (2018). Emotion regulation. In J. E., Lochman, & W. Matthys (Eds.), The Wiley handbook of disruptive and impulse- control disorders (pp. 221-236). Wiley Blackwell.

Bowker, J. C., Rubin, K. H., Rose-Krasnor, L., & Booth-LaForce, C. (2007). Good friendships, bad friends: Friendship factors as moderators of the relation between aggression and social information processing. European Journal of Developmental Psychology, 4(4), 415-434. http://doi.org/10.1080/17405620701632069

Branje, S., & Koot, H. M. (2018). Psychophysiology of aggression. In T. Malti & K. H. Rubin (Eds.), Handbook of child and adolescent aggression (pp. 84-106). The Guilford Press.

Brendgen, M., Bowen, F., Rondeau, N., & Vitaro, F. (1999). Effects of friends' characteristics on children's social cognitions. Social Development, 8(1), 41-51. http://doi.org/10.1111/1467-9507.00079

Bubier, J. L., & Drabick, D. A. (2009). Co-occurring anxiety and disruptive behaviour disorders: the roles of anxious symptoms, reactive aggression, and shared risk processes. Clinical Psychology Review, 29(7), 658–669. http://doi.org/10.1016/j.cpr.2009.08.005

Burgess, K. B., Wojslawowicz, J. C., Rubin, K. H., Rose-Krasnor, L., & Booth-LaForce, C. (2006). Social information processing and coping strategies of shy/withdrawn and aggressive children: Does friendship matter? Child Development, 77(2), 371-383. http://doi.org/10.1111/j.1467-8624.2006.00876.x

Burks, V. S., 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. http://doi.org/10.1111/1467-9507.00092

Burnham, K. P., & Anderson, D. R. (1998). Practical use of the information theoretic approach. In K. P. Burnham & D. R. Anderson (Eds.), Model selection and inference (pp. 75-117). Springer. https://doi.org/10.1007/978-1-4757-2917-7_3

Bushman, B. J., & Anderson, C. A. (2001). Is it time to pull the plug on hostile versus instrumental aggression dichotomy? Psychological Review, 108(1), 273-279. http://doi.org/10.1037/0033-295X.108.1.273

Bushman, B., Dewall, C., Pond, R., & Hanus, M. (2014). Low glucose relates to greater aggression in married couples. Proceedings of the National Academy of Sciences, 111(17), 6254-6257. http://doi.org/10.1073/pnas.1400619111

Calvete, E. (2008). Justification of violence and grandiosity schemas as predictors of antisocial behavior in adolescents. Journal of Abnormal Child Psychology, 36(7), 1083-1095. http://doi.org/10.1007/s10802-008-9229-5

Calvete, E., & Orue, I. (2010). Cognitive schemas and aggressive behavior in adolescents: The mediating role of social information processing. The Spanish Journal of Psychology, 13(1), 190-201. http://doi.org/10.1017/s1138741600003772

Calvete, E. & Orue, I. (2012). Social information processing as a mediator between cognitive schemas and aggressive behavior in adolescents. Journal of Abnormal Child Psychology, 40(1), 105-117. http://doi.org/10.1007/s10802-011-9546-y

Caravita, S. C. S., Sijtsema, J. J., Rambaran, A. J., & Gini, G. (2014). Peer influences on moral disengagement in late childhood and early adolescence. Journal of Youth and Adolescence, 43(2), 193-207. https://doi.org/10.1007/s10964-013-9953-1

Cardeñoso, O., & Calvete, E. (2004). Desarrollo del Inventario de Creencias Irracionales para Adolescentes (Development of the adolescents' irrational beliefs inventory). Psicología Conductual, 12(2), 289–304.

Caporaso, J. S., & Marcovitch, S. (2021). The effect of taxing situations on preschool children's responses to peer conflict. Cognitive Development, 57(6), 100989. https://doi.org/10.1016/j.cogdev.2020.100989

Carl, E., Stein, A. T., Levihn-Coon, A., Pogue, J. R., Rothbaum, B., Emmelkamp, P., Asmundson, G. J. G., Carlbring, P., & Powers, M. B. (2019). Virtual reality exposure therapy for anxiety and related disorders: A meta-analysis of randomized controlled trials. Journal of Anxiety Disorders, 61, 27-36. https://doi.org/10.1016/j.janxdis.2018.08.003

Carroll, A., McCarthy, M., Houghton, S., Sanders O'Connor, E., & Zadow, C. (2018). Reactive and proactive aggression as meaningful distinctions at the variable and person level in primary schoolaged children. Aggressive Behavior, 44(5), 431–441. https://doi.org/10.1002/ab.21763

Chaiken, S., & Trope, Y. (1999). Dual-process theories in social psychology. Guilford Press.

Chaux, E., Arboleda, J., & Rincón, C. (2012). Community violence and reactive and proactive aggression: The mediating role of cognitive and emotional variables. Revista Colombiana de Psicologia, 21, 233-251. Retrieved from https://www.revistas.unal.edu.co/index.php/psicologia/article/view/28511

Choe, D. E., Lane, J. D., Grabell, A. S., & Olson, S. L. (2013). Developmental precursors of young school-age children's hostile attribution bias. Developmental Psychology, 49(12), 2245-2256. https://doi.org/10.1037/a0032293

Choe, D. E., Shaw, D. S., & Forbes, E. E. (2015). Maladaptive social information processing in childhood predicts young men's atypical amygdala reactivity to threat. Journal of Child Psychology and Psychiatry, 56(5), 549-557. https://doi.org/10.1111/jcpp.12316

Cillessen, A. H. N., Lansu, T. A. M., & Van den Berg, Y. H. M. (2014). Aggression, hostile attributions, status, and gender. Development and Psychopathology 26(3), 635-644. https://doi.org/10.1017/s0954579414000285

Coburn, K. M., & Vevea, J. L. (2015). Publication bias as a function of study characteristics. Psychological Methods, 20(3), 310-330. https://doi.org/10.1037/met0000046

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Coie, J. D., & Dodge, K. A. (1998). Aggression and antisocial behavior. In W. Damon & N. Eisenberg (Ed.), Handbook of child psychology: Social, emotional, and personality development (pp. 779-862). John Wiley & Sons.

Colasante, T., Jambon, M., Goa, X., & Malti, T. (2021). A process model linking physiological arousal and fear recognition to aggression via guilt in middle childhood. Development and Psychopathology, 33(1), 109-121. https://doi.org/10.1017/S0954579419001627

Cornacchio, D., Bry, L. J., Sanchez, A. L., Poznanski, B., & Comer, J. S. (2018). Psychosocial treatment and prevention of conduct problems in early childhood. In J. E. Lochman & W. Matthys (Eds.), The Wiley handbook of disruptive and impulse-control disorders (pp. 433-449). Wiley Blackwell.

Crain, M. M., Finch, C. L., & Foster, S. L. (2005). The relevance of the social information processing model for understanding relational aggression in girls. Merrill-Palmer Quarterly, 51(2), 213-249. http://doi.org/10.1353/mpq.2005.0010

Crick, N. R. (1995). Relational aggression: The role of intent attributions, feelings of distress, and provocation type. Development and Psychopathology, 7(2), 313-322. http://doi.org/10.1017/s0954579400006520

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. http://doi.org/10.1037//0033-2909.115.1.74

Crick, N. R., & Dodge, K. A. (1996). Social information processing deficits in reactive and proactive aggression. Child Development, 67(3), 993-1002. http://doi.org/10.2307/1131875

Crick, N. R., Grotpeter, J. K., & Bigbee, M. A., (2002). Relationally and physically aggressive children's intent attributions and feelings of distress for relational and instrumental peer provocations. Child Development, 73(4), 1134-1142. http://doi.org/10.1111/1467-8624.00462

Crozier, J. C., Dodge, K. A., Fontaine, R. G., Lansford, J. E., Bates, J. E., Pettit, G. S., & Levenson, R. W. (2008). Social information processing and cardiac predictors of adolescent antisocial behavior. Journal of Abnormal Psychology, 117(2), 253-267. https://doi.org/10.1037/0021-843X.117.2.253

Cunningham, W. A., & Zelazo, P. D. (2007). Attitudes and evaluations: A social cognitive neuroscience perspective. Trends in Cognitive Sciences, 11(3), 97-104. http://doi.org/10.1016/j.tics.2006.12.005

Cunningham, W. A., Zelazo, P. D., Packer, D. J., & Van Bavel, J. J. (2007). The iterative reprocessing model: A multilevel framework for attitudes and evaluation. Social Cognition, 25(5), 736-760. http://doi.org/10.1521/soco.2007.25.5.736

Cuperus, J. M. (1997). Sociale probleemoplossing bij kinderen met gedragsstoornissen [Social problem solving in children with behavior disorders]. Unpublished doctoral dissertation, Utrecht University, The Netherlands.

De Castro, B. O. (2004). The development of social information processing and aggressive behaviour: Current issues. European Journal of Developmental Psychology, 1(1), 87-102. http://doi.org/10.1080/17405620444000058

De Castro, B. O., 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. http://doi.org/10.1207/s15374424jccp3401_10

De Castro, B. O., Slot, N. W., Bosch, J. D., Koops, W., & Veerman, J. W. (2003). Negative feelings exacerbate hostile attributions of intent in aggressive boys. Journal of Clinical Child and Adolescent Psychology, 31(1), 56-65. http://doi.org/10.1207/s15374424jccp3201_06

De Castro, B. O., Thomaes, S., & Reijntjes, A. (2015). Using experimental designs to understand the development of peer relations. Journal of Research on Adolescence, 25(1), 1-13. https://doi.org/10.1111/jora.12103

De Castro, B. O., & Van Dijk, A. (2017). "It's gonna end up with a fight anyway": Social cognitive processes in children with disruptive behavior disorders. In: Lochman, J. E., & Matthys, W. (Eds.), The Wiley handbook of disruptive and impulse-control disorders (pp. 237-253). Wiley Blackwell. http://doi.org/10.1002/9781119092254.ch15

De Castro, B. O., Veerman, J.W., Koops, W., Bosch, J. D., & Monshouwer, H.J. (2002). Hostile attribution of intent and aggressive behavior: A meta-analysis. Child Development, 73(3), 916-934. http://doi.org/10.1111/1467-8624.00447

De Castro, B., Verhulp, E. E. & Runions, K. (2012). Rage and revenge: Highly aggressive boys' explanations for their responses to ambiguous provocation. European Journal of Developmental Psychology, 9(3), 331-350. https://doi.org/10.1080/17405629.2012.680304

Dekkers, T. J., van Bergen, N. R. J., & Jansen, B. R. J. (2019). An assessment of the psychometric properties of the Brief Sensation Seeking Scale for children. Journal of Personality Assessment, 101(4), 446-451. https://doi.org/10.1080/00223891.2018.1468336

Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. Developmental Cognitive Neuroscience, 18, 34-48. http://doi.org/10.1016/j.dcn.2015.11.005

Di Giunta, L., Iselin, A. R., Eisenberg, N., Pastorelli, C., Gerbino, M., Lansford, J. E., Dodge, K. A....,Thartori, E. (2017). Measurement invariance and convergent validity of anger and sadness self-regulation among youth from six cultural groups. Assessment, 24(4), 484-502. http://doi.org/10.1177/1073191115615214

Dodge, K. A. (1980). Social cognition and children's aggressive behavior. Child Development, 51(1), 162-170. http://doi.org/10.2307/1129603

Dodge, K. A. (1991). The structure and function of reactive and proactive aggression. In D. Pepler & K. H. Rubin (Eds.), The development and treatment of childhood aggression (pp. 201-218). Lawrence Erlbaum Associates, Inc.

Dodge, K. A. (1991). Emotion and social information processing. In J. Garber & K. A. Dodge (Eds.), Cambridge studies in social and emotional development. The development of emotion regulation and dysregulation (pp. 159-181). Cambridge University Press.

Dodge, K. A. (2006). Translational science in action: Hostile attributional style and the development of aggressive behavior problems. Development and Psychopathology, 18(3), 791-814. http://doi.org/10.1017/S0954579406060391

Dodge, K. A. (2009). Mechanisms of gene–environment interaction effects in the development of conduct disorder. Perspectives in Psychological Science, 4(4), 408-414. http://doi.org/10.1111/j.1745-6924.2009.01147.x

Dodge, K. A. (2011). Social information processing patterns as mediators of the interaction between genetic factors and life experiences in the development of aggressive behavior. In P. Shaver & M. Mikulincer (Eds.), Human aggression and violence: Causes, manifestations, and consequences (pp. 165-185). American Psychological Association.

Dodge, K. A., & Coie, J. D. (1987). Social information processing factors in reactive and proactive aggression in children's peer groups. Journal of Personality and Social Psychology, 53(6), 1146-1158. http://doi.org/10.1037//0022-3514.53.6.1146

Dodge, K. A., Coie, J. D., & Lynam, D. (2006). Aggression and antisocial behavior in youth. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), Handbook of child psychology: Social, emotional, and personality development (pp. 719-788). John Wiley & Sons.

Dodge, K. A., & Frame, C. L. (1982). Social cognitive biases and deficits in aggressive boys. Child Development, 53(3), 620-635. http://doi.org/10.2307/1129373

Dodge, K. A., Laird, R., Lochman, J. E., & Zelli, A.(2002). Multidimensional latent-construct analysis of children's social information processing patterns: correlations with aggressive behavior problems, Psychological Assessment, 14(1), 60-73. http://doi.org/10.1037//1040-3590.14.1.60

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. http://doi.org/10.1111/1467-8624.7402004

Dodge, K. A., Lochman, J. E., Harnish, J. D., Bates, J. E., & Pettit, G. S. (1997). Reactive and proactive aggression in school children and psychiatrically impaired chronically assaultive youth. Journal of Abnormal Psychology, 106(1), 37-51. http://doi.org/10.1037/0021-843x.106.1.37

Dodge, K. A., Malone, P. S., Lansford, J. E., Sorbring, E., Skinner, A. T., Tapanya, S., … Pastorelli, A. (2015). Hostile attributional bias and aggressive behavior in global context. Proceedings of the National Academy of Sciences of the United States of America, 112(30), 9310-9315. http://doi.org/10.1073/pnas.1418572112

Dodge, K. A., McClaskey, C. L., & Feldman, E. (1985). Situational approach to the assessment of social competence in children. Journal of Consulting and Clinical Psychology, 53(3), 344-353. http://doi.org/10.1037/0022-006x.53.3.344

Dodge, K. A, & Newman, J. P. (1981). Biased decision-making processes in aggressive boys. Journal of Abnormal Psychology, 90(4), 375-379. http://doi.org/10.1037/0021-843X.90.4.375

Dodge, K. A., & Pettit, G. S. (2003). A biopsychosocial model of the development of chronic conduct problems in adolescence. Developmental Psychology, 39(2), 349-371. http://doi.org/10.1037/0012-1649.39.2.349

Dodge, K. A., Pettit, G. S., Bates, J. E., & Valente, E. (1995). Social information-processing patterns partially mediate the effect of early physical abuse on later conduct problems. Journal of Abnormal Psychology, 104(4), 632-643. http://doi.org/10.1037/0021-843x.104.4.632

Dodge, K., Pettit, G., McClaskey, C., Brown, M., & Gottman, J. (1986). Social competence in children. Monographs of the Society for Research in Child Development, 51(2), 1-85. http://doi.org/10.2307/1165906

Dodge, K. A., & Price, J. M. (1994). On the relation between social information processing and socially competent behavior in early school-aged children. Child Development, 65(5), 1385-1397. http://doi.org/10.2307/1131505

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(1), 213-224. http://doi.org/10.2307/1130303

Dodge, K. A., & Tomlin, A. M. (1987). Utilization of selfschemas as a mechanism of interpretational bias in aggressive children. Social Cognition, 5(3), 280-300. http://doi.org/10.1521/soco.1987.5.3.280

Doyle A. E. (2006). Executive functions in attention-deficit/ hyperactivity disorder. Journal of Clinical Psychiatry 67(8), 21-26.

Ellis, B. J., Volk, A. A., Gonzalez, J. M., & Embry, D. D. (2016). The meaningful roles intervention: An evolutionary approach to reducing bullying and increasing prosocial behavior. Journal of Research on Adolescence, 26(4), 622-637. https://doi.org/10.1111/jora.12243

Ellis, M. L., Weiss, B., & Lochman, J. E. (2009). Executive functions in children: Associations with aggressive behavior and appraisal processing. Journal of Abnormal Child Psychology, 37(7), 945-956. http://doi.org/10.1007/s10802-009-9321-5

Emmelkamp, P. M. G., & Meyerbröker, K. (2021). Virtual Reality therapy in mentalhealth. Annual Review of Clinical Psychology. Advance online publication. https://doi.org/10.1146/annurev-clinpsyc-091219-115923

Erdley, C. A., & Asher, S. R. (1996). Children's social goals and self-efficacy perceptions as influences on their responses to ambiguous provocation. Child Development, 67(4), 1329-1344. http://doi.org/10.2307/1131703

Euler, F., Steinlin, C., & Stadler, C. (2017). Distinct profiles of reactive and proactive aggression in adolescents: Associations with cognitive and affective empathy. Child and Adolescent Psychiatry and Mental Health, 11(1), 1–14. https://doi.org/10.1186/s13034-016-0141-4

Fanti, K. A. (2018). Understanding heterogeneity in conduct disorder: A review of psychophysiological studies. Neuroscience and Biobehavioral Reviews, 91, 4-20. https://doi.org/10.1016/j.neubiorev.2016.09.022

Feshbach, L. E. (1989). Aggression-conduct problems, attention- deficits hyperactivity, play, and social cognition in four-year-old boys. Unpublished doctoral dissertation, University of Washington, Seattle.

Fontaine, R. G., Tanha, M., Yang, C., Dodge, K. A., Bates, J. E., & Pettit, G. S. (2010). Does response evaluation and decision (RED) mediate the relation between hostile attributional style and antisocial behavior? Journal of Abnormal Child Psychology, 38(5), 615-626. http://doi.org/10.1007/s10802-010-9397-y

Frankenhuis, W., Panchanathan, K., & Nettle, D. (2016). Cognition in harsh and unpredictable environments. Current Opinion in Psychology, 7(2016), 76-80. http://doi.org/10.1016/j.copsyc.2015.08.011

Freeman, K., Hadwin, J. A., & Halligan, S. L. (2011). An experimental investigation of peer influences on adolescent hostile attributions. Journal of Clinical Child and Adolescent Psychology, 40(60, 897-903. http://doi.org/10.1080/15374416.2011.614582

Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., & Slater, M. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychological Medicine, 47(14), 1-8. http://dx.doi.org/10.1017/S003329171700040X

Frick, P. J. (2012). Developmental pathways to conduct disorder: Implications for future directions in research, assessment, and treatment. Journal of Clinical Child & Adolescent Psychology, 41(3), 378-389. https://doi.org/10.1080/15374416.2012.664815

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. http://doi.org/10.1023/a:1023899703866

Frick, P. J., Cornell, A. H., Bodin, S. D., Dane, H. E., Barry, C. T., & Loney, C. T. (2003). Callous–unemotional traits and developmental pathways to severe conduct problems. Developmental Psychology, 39(2), 246-260. http://doi.org/10.1037/0012-1649.39.2.246

Frick, P. J., & Morris, A. S. (2004). Temperament and developmental pathways to conduct problems. Journal of Clinical Child and Adolescent Psychology, 33(1), 54-68. http://dx.doi.org/10.1207/S15374424JCCP3301_6

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014a). Annual research review: A developmental psychopathology approach to understanding callous-unemotional traits in children and adolescents with serious conduct problems. Journal of Child Psychology and Psychiatry, 55(6), 532-548. http://doi.org/10.1111/jcpp.12152

Frick, P. J., Ray, J. V., Thornton, L. C., & Kahn, R. E. (2014b). 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. http://doi.org/10.1037/a0033076

Frijda, N.H. (1993). The place of appraisal in emotion. Cognition and Emotion, 7(3–4), 357-387. http://doi.org/10.1080/02699939308409193

Garner, P. W. & Lemerise, E. A. (2007). The roles of behavioral adjustment and conceptions of peers and emotions in preschool children's peer victimization. Development and Psychopathology, 19(1), 57-71. http://doi.org/10.1017/S0954579407070046

Garon, M., Lavallée, M. M., Vera Estay, E., & Beauchamp, M. H. (2018). Visual encoding of social cues predicts sociomoral reasoning. PLoS ONE, 13(7), Article e0201099. https://doi.org/10.1371/journal.pone.0201099

Gentile, D. A., & Gentile, J. R. (2008). Violent video games as exemplary teachers: A conceptual analysis. Journal of Youth and Adolescence, 9(2), 127-141. http://doi.org/10.1007/s10964-007-9206-2

Gentile, D.A., Li, D., Khoo, A., Prot, S., & Anderson, C.A. (2014). Mediators and moderators of long-term effects of violent video games on aggressive behavior practice, thinking, and action. JAMA Pediatrics, 168(5), 450-457. https://doi.org/10.1001/jamapediatrics.2014.63

Gibbins, C., & Craig, W. (1997, April). Mapping the path to aggression: Validation of a social-cognitive model of childhood aggression. Poster presented at the biennial meeting of the Society for Research in Child Development, Washington, DC.

Godleski, S. A., & Ostrov, J. M. (2010). Relational aggression and hostile attribution biases: testing multiple statistical methods and models. Journal of Abnormal Child Psychology, 38(4), 447-458. https://doi.org/10.1007/s10802-010-9391-4

Goldweber, A., Bradshaw, C. P., Goodman, K., Monahan, K., & Cooley-Strickland, M. (2011). Examining factors associated with (in)stability in social information processing among urban school children: A latent transition analytic approach. Journal of Clinical Child and Adolescent Psychology, 40(5), 715-729. https://doi.org/10.1080/15374416.2011.597088

Gomez, R., & Gomez, A. (2000). Perceived maternal control and support as predictors of hostile-based attribution of intent and responsive selection in aggressive boys. Aggressive Behavior, 26(2), 155-168. https://doi.org/10.1002/(sici)1098-2337(2000)26:2<155::aid-ab2>3.3.co;2-b

Goraya, F., & Kazim, S. S. (2013). Social information processing as a mediator between parenting and children's behavioral problems. Journal of Behavioural Sciences, 23(1), 39-61.

Graham, S., & Hudley, C. (1994). Attributions of aggressive and nonaggressive African-American male early adolescents: A study of construct accessibility. Developmental Psychology, 30(3), 365-373. https://doi.org/10.1037/0012-1649.30.3.365

Graham, S., Hudley, C., & Williams, E. (1992). Attributional and emotional determinants of aggression among African-American and Latino young adolescents. Developmental Psychology, 28(4), 731-740. https://doi.org/10.1037/0012-1649.28.4.731

Granic, I. (2014). The role of anxiety in the development, maintenance and treatment of childhood aggression. Development and Psychopathology, 26(4, Pt 2), 1515-1530. https://doi.org/10.1017/s0954579414001175

Grumm, M., Hein, S., & Fingerle, M. (2011). Predicting aggressive behavior in children with the help of measures of implicit and explicit aggression. International Journal of Behavioral Development, 35(4), 352-357. https://doi.org/10.1177/0165025411405955

Guerra, N. G., Huesmann, L. R., & Spindler, A. (2003). Community violence exposure, social cognition, and aggression among urban elementary school children. Child Development, 74(5), 1561-1576. https://doi.org/10.1111/1467-8624.00623

Guerra, N. G., & Slaby, R. G. (1989). Evaluative factors in social problem solving by aggressive boys. Journal of Abnormal Child Psychology, 17(3), 277-289. https://doi.org/10.1007/bf00917399

Guo, J., & Mrug, S. (2017). Temperament. In: Lochman, J. E., & Matthys, W. (Eds.), The Wiley handbook of disruptive and impulse-control disorders (pp. 237-253). Wiley Blackwell. http://doi.org/10.1002/9781119092254.ch11

Hadley, W., Houck, C., Brown, L. K., Spitalnick, J. S., Ferrer, M., & Barker, D. (2019). Moving beyond role-play: Evaluating the use of virtual reality to teach emotion regulation for the prevention of adolescent risk behavior within a randomized pilot trial. Journal of Pediatric Psychology, 44(4), 425-435. https://doi.org/10.1093/jpepsy/jsy092

Halligan, S. L., Cooper, P. J., Healy, S. J. & Murray, L. (2007). The attribution of hostile intent in mothers, fathers and their children. Journal of Abnormal Child Psychology, 35(4), 594-604. https://doi.org/10.1007/s10802-007-9115-6

Halligan, S. L., & Philips K. J. (2010). Are you thinking what I'm thinking? Peer group similarities in adolescent hostile attribution tendencies. Developmental Psychology, 46(5), 1385-1388. https://doi.org/10.1037/a0020383

Hart, M. T. (1993). Social-cognitive processing in aggressive/ withdrawn, aggressive/ nonwithdrawn, and nondeviant children. Unpublished doctoral dissertation, Texas A&M University, College Station.

Healy, S. J., Murray, L., Cooper, P. J., Hughes, C., & Halligan, S. L. (2015). A longitudinal investigation of maternal influences on the development of child hostile attributions and aggression. Journal of Clinical Child & Adolescent Psychology, 44(1), 80-92. https://doi.org/10.1080/15374416.2013.850698

Heidgerken, A. D., Hughes, J. N., Cavell, T. A., & Willson, V. L. (2004). Direct and indirect effects of parenting and children's goals on child aggression. Journal of Clinical Child and Adolescent Psychology, 33(4), 684-693. https://doi.org/10.1207/s15374424jccp3304_4

Helmsen, J., Koglin, U., & Petermann, F. (2012). Emotion regulation and aggressive behavior in preschoolers: The mediating role of social information processing. Child Psychiatry and Human Development, 43(1), 87-101. http://doi.org/10.1007/s10578-011-0252-3

Helseth, S. A., Waschbusch, D. A., King, S., & Willoughby, M. T. (2015). Aggression in children with conduct problems and callous-unemotional traits: Social information processing and response to peer provocation. Journal of Abnormal Child Psychology, 43(8), 1503-1514. http://doi.org/10.1007/s10802-015-0027-6

Hessler, D. M., & Fainsilber Katz, L. (2007). Children's emotion regulation: Self-report and physiological response to peer provocation. Developmental Psychology, 43(1), 27-38. http://doi.org/10.1037/0012-1649.43.1.27

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. Statistics in Medicine, 21(11), 1539–1558. http://doi.org/10.1002/sim.1186

Hobson, C. W., Scott, S., & Rubia, K. (2011). Investigation of cool and hot executive function in ODD/CD independently of ADHD. Journal of Child Psychology and Psychiatry, 52(10), 1035-1043. https://doi.org/10.1111/j.1469-7610.2011.02454.x

Horsley, T. A., Castro, B. O. de, & 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. https://doi.org/10.1007/s10802-009-9361-x

Hox, J. J. (2002). Multilevel analysis: Techniques and applications (1st ed.). Erlbaum. http://doi.org/10.4324/9781410604118

Hubbard, J. A., Dodge, K. A., Cillessen, A. H. N., 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), 268-280. https://dx.doi.org/10.1037/0022-3514.80.2.268

Hubbard, J. A., McAuliffe, M. D., Morrow, M. T., & Romano, L. J. (2010). Reactive and proactive aggression in childhood and adolescence: Precursors, outcomes, processes, experiences, and measurement. Journal of Personality, 78(1), 95-118. http://doi.org/10.1111/j.1467-6494.2009.00610.x

Hubbard, J. A., Parker, E. H., Ramsden, S. R., Flanagan, K. D., Relyea, N., Dearing, K. F., Smithmyer, C. M., Simons, R. F., & Hyde, C. T. (2004). The relations among observational, physiological, and self-report measures of children's anger. Social Development, 13(1), 14–39. https://doi.org/10.1111/j.1467-9507.2004.00255.x

Hubbard, J. A., Smithmyer, C. M., Ramsden, S. R., Parker, E. H., Flanagan, K. D., Dearing, K. E., et al. (2002). Observational, physiological, and self-report measures of children's anger: Relations to reactive versus proactive aggression. Child Development, 73(4), 1101-1118. https://doi.org/10.1111/1467-8624.00460

Hudley, C., & Graham, S. (1993). An attributional intervention to reduce peer-directed aggression among African-American boys. Child Development, 64, 124-138. https://doi.org/10.2307/1131441

Huesmann, L. R., & Guerra, N. G. (1997). Children's normative beliefs about aggression and aggressive behavior. Journal of Personality and Social Psychology, 72(2), 408-419. http://doi.org/10.1037/0022-3514.72.2.408

Hummer, T. A., Kronenberger, W. G., Wang, Y., Dunn, D. W., Mosier, K. M., Kalmin, A. J., & Mathews, V. P. (2010). Executive functioning characteristics associated with ADHD comorbidity in adolescents with Disruptive Behavior Disorders. Journal of Abnormal Child Psychology, 39(1), 11-19. http://doi.org/10.1007/s10802-010-9449-3

Hyatt, L. F. (1998). Moderating influences on the attributional bias towards hostile orientation in aggressive boys. Unpublished doctoral dissertation, Hofstra University, Hempstead, NY.

Hyde, L. W., Shaw, D. S., & Moilanen, K. L. (2010). Developmental precursors of moral disengagement and the role of moral disengagement in the development of antisocial behavior. Journal of Abnormal Child Psychology, 38(2), 197-209. http://doi.org/10.1007/s10802-009-9358-5

Insel, T. R. & Fernald, R. D. (2004). How the brain processes social information: searching for the social brain. Annual Review of Neuroscience, 27(1), 697-722.https://doi.org/10.1146/annurev.neuro.27.070203.144148

Jarret, M. A, & Hilton, D. C. (2017). Cognitive functions. In: Lochman, J. E., & Matthys, W. (Eds.), The Wiley handbook of disruptive and impulse-control disorders (pp. 237-253. Wiley Blackwell. http://doi.org/10.1002/9781119092254.ch11

Kahneman D. (2011). Thinking, fast and slow. Macmillan.

Katsurada, E. (1995). Preschoolers' hostile attribution, aggressive behavior, and relationships with their mothers' attributional style, parenting behavior, and affect. Unpublished doctoral dissertation, Oregon State University, Corvallis.

Kempes, M. M., De Castro, B.O, & Sterck, E. H. M. (2008). Conflict management in 6–8 year old aggressive Dutch boys: Do they reconcile? Behaviour, 145(11), 1701-1722. https://doi.org/10.1163/156853908786131306

Kempes, M., Matthys, W., De Vries, H., & Engeland, V. H. (2005). Reactive and proactive aggression in children: A review of theory, findings and the relevance for child and adolescent psychiatry. European Child & Adolescent Psychiatry, 14(1), 11-19. http://doi.org/10.1007/s00787-005-0432-4

Kempes, M, Matthys, W., Maassen, G., Van Goozen, S., & Van Engeland, H. (2006). A parent questionnaire for distinguishing between reactive and proactive aggression in children. European Journal of Developmental Psychology, 15(1), 38-45. https://doi.org/10.1007/s00787-006-0502-2

King, S., & Waschbusch, D. A. (2010). Aggression in children with attention deficit/hyperactivity disorder. Expert Review of Neurotherapeutics, 10(10), 1581-1594. https://doi.org/10.1586/ern.10.146

Klingberg, T. (2010). Training and plasticity of working memory. Trends in Cognitive Sciences, 14(7), 317–324. http://doi.org/10.1016/j.tics.2010.05.002

Klingberg, T., Fernell, E., Olesen, P. J., Johnson, M., Gustafsson, P., Dahlstrom, K., Gillberg, C. G., Forssberg, H., & Westerberg, H. (2005) Computerized training of working memory in children with ADHD—a randomized, controlled trial. Journal of the American Academy of Child & Adolescent Psychiatry, 44(2),177–186. https://doi.org/10.1097/00004 583-200502000-00010

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with asingle covariate. Statistics in Medicine, 22(17), 2693-2710. https://doi.org/10.1002/sim.1482

Krain, A. L., Wilson, A. M., Arbuckle, R., Castellanos, F. X. & Milham, M. P. (2006). Distinct neural mechanisms of risk and ambiguity: A meta-analysis of decision-making. NeuroImage 32(1), 477-484. http://doi.org/10.1016/j.neuroimage.2006.02.047

Kunimatsu, M. M., & Marsee, M. A. (2012). Examining the presence of anxiety in aggressive individuals: The illuminating role of fight-or-flight mechanisms. Child and Youth Care Forum, 41(3), 247-258. https://doi.org/10.1007/s10566-012-9178-6

Kupersmidt, J. B., Stelter, R., & Dodge, K. A. (2011). Development and validation of the social information processing application: A Web-based measure of social information processing patterns in elementary school-age boys. Psychological Assessment, 23(4), 834-847. https://doi.org/10.1037/a0023621

Laible, D. J., McGinley, M., Carlo, G., Augustine, M., & Murphy, T. (2014). Does engaging in prosocial behavior make children see the world through rose-colored glasses? Developmental Psychology, 50(3), 872-880. http://doi.org/10.1037/a0033905

Laible, D. J., Murphy, T. P., & Augustine, M. (2014). Adolescents' aggressive and prosocial behaviors: Links with social information processing, negative emotionality, moral affect, and moral cognition. The Journal of Genetic Psychology: Research and Theory on Human Development, 175(3), 270-286. http://doi.org/10.1080/00221325.2014.885878

Lansford, J. E., Malone, P. S., Dodge, K. A., Crozier, J. C., Pettit, G. S., & Bates, J. E. (2006). A 12-year prospective study of patterns of social information processing problems and externalizing behaviors. Journal of Abnormal Child Psychology, 34(5), 709-718. http://doi.org/10.1007/s10802-006-9057-4

Lansford, J. E., Malone, P. S., Dodge, K. A., Pettit, G. S., & Bates, J. E. (2010). Developmental cascades of peer rejection, social information processing biases, and aggression during middle childhood. Development and Psychopathology, 22(3), 593- 602. http://doi.org/10.1017/s0954579410000301

Lavallee, K. L., Bierman, K. L., & Nix, R. L. (2005). The impact of first-grade "friendship group" experiences on child social outcomes in the Fast Track Program. Journal of Abnormal Child Psychology, 33(3), 307-324. http://doi.org/10.1007/s10802-005-3567-3

Lee, H. Y., Jamieson, J. P., Reis, H. T., Beevers, C. G., Josephs, R. A., Mullarkey, M. C., O'Brien, J. M., & Yeager, D. S. (2020). Getting fewer "likes" than others on social media elicits emotional distress among victimized adolescents. Child Development, 91(6), 2141–2159. https://doi.org/10.1111/cdev.1342

Leijten, P., Raaijmakers, M., de Castro, B., & Matthys, W. (2013). Does socioeconomic status matter? A meta-analysis on parent training effectiveness for disruptive child behavior. Journal of Clinical Child & Adolescent Psychology, 42(3), 384–392. https://doi.org/10.1080/15374416.2013.769169

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. http://doi.org/10.1111/1467-8624.00124

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. Biometrika, 88(3), 767-778. https://doi.org/10.1093/biomet/88.3.767

Lochman, J. E., Boxmeyer, C. L., Jones, S., Qu, L., Ewoldsen, D., & Nelson, W. M. (2017). Testing the feasibility of a briefer school-based preventive intervention with aggressive children: A hybrid intervention with face-to-face and internet components. Journal of School Psychology, 6, 33-50. http://doi.org/10.1016/j.jsp.2017.03.010

Lochman, J. E., Boxmeyer, C. L., Kassing, F. L., Powell, N. P., & Stromeyer, S. L. (2019). Cognitive behavioral intervention for youth at risk for conduct problems: Future directions. Journal of Clinical Child and Adolescent Psychology, 48(5), 799-810. http://doi.org/10.1080/15374416.2019.1567349

Lochman, J. E., & Dodge, K. A. (1994). Social-cognitive processes of severely violent, moderately aggressive, and nonaggressive boys. Journal of Consulting and Clinical Psychology, 62(2), 366-374. http://doi.org/10.1037/0022-006x.62.2.366

Lochman, J. E., & Dodge, K. A. (1998). Distorted perceptions in dyadic interactions of aggressive and nonaggressive boys: Effects of prior expectations, context, and boys' age. Development and Psychopathology, 10(3), 495-512. http://doi.org/10.1017/S0954579498001710

Lochman, J. E., & Matthys, W. (2018). The Wiley handbook of disruptive and impulse-control disorders. Wiley Blackwell.

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. http://doi.org/10.1017/s0954579402004157

MacBrayer, E. K., Milich, R., Hundley, M. (2003). Attributional biases in aggressive children and their mothers. Journal of Abnormal Psychology, 112(4), 698-708. http://doi.org/10.1037/0021-843x.112.4.598

Maixner-Schindel, K., & Shechtman, Z. (2021). The impact of reappraisal skills on aggressive children. Aggressive Behavior, 47(2), 205-2014. https://doi.org/10.1002/ab.21939

Manel, W. (2003). Aggressive boys' attributions of emotions and intentions in social situations. Electronic Theses and Dissertations. Paper 916.

Marsee, M. A., Frick, P. J., Barry, C. T., Kimonis, E. R., Centifanti, L. C. M., & Aucoin, K. J. (2014). Profiles of the forms and functions of self-reported aggression in three adolescent samples. Development and Psychopathology, 26(3), 705–720. https://doi.org/10.1017/S0954579414000339

Martinelli, A., Ackermann, K., Bernhard, A., Freitag, C. M., & Schwenck, C. (2018). Hostile attribution bias and aggression in children and adolescents: A systematic literature review on the influence of aggression subtype and gender. Aggression and Violent Behavior, 39, 25-32. https://doi.org/10.1016/j.avb.2018.01.005

Mathieson, L. C., Murray-Close, D., Crick, N. R., Woods, K. E., Zimmer-Gembeck, M., Geiger, T. C., & Morales, J. R. (2011). Hostile intent attributions and relational aggression: The moderating roles of emotional sensitivity, gender, and victimization. Journal of Abnormal Child Psychology, 39(7), 977-987. https://doi.org/10.1007/s10802-011-9515-5

Matthys, W., Cuperus, J. M., & van Engeland, H. (1999). Deficient social problem-solving in boys with ODD/CD, with ADHD, and with both disorders. Journal of the American Academy of Child and Adolescent Psychiatry, 38(3), 311–321. http://doi.org/10.1097/00004583-199903000-00019

Matthys, W., Maassen, G. H., Cuperus, J. M., & van Engeland, H. (2001). The assessment of the situational specificity of children's problem behaviour in peer–peer context. Journal of Child Psychology and Psychiatry, 42(3), 413-420. http://doi.org/10.1111/1469-7610.00734

Matthys, W., Vanderschuren, L. J., & Schutter, D. J. (2013). The neurobiology of oppositional defiant disorder and conduct disorder: Altered functioning in three mental domains. Development and Psychopathology, 25(1), 193-207. http://doi.org/10.1017/S0954579412000272

Matthys, W., Vanderschuren, L. J. M. J., Schutter, D. J. L. G., & Lochman, J. E. (2012). Impaired neurocognitive functions affect social learning processes in oppositional defiant disorder and conduct disorder: Implications for interventions. Clinical Child and Family Psychology Review, 15(3), 234-246. https://doi.org/10.1007/s10567-012-0118-7

Matthys, W., Van Goozen, S. H. M., De Vries, H., Cohen-Kettenis, P., & Van Engeland, H. (1998). The dominance of behavioural activation over behavioural inhibition in conduct disordered boys with and without attention deficit hyperactivity disorder. Journal of Child Psychology and Psychiatry, 39(5), 643–651. https://doi.org/10.1111/1469-7610.00364

Matthys, W., Van Goozen, S. H. M., Snoek, H., & Van Engeland, H. (2004). Response perseveration and sensitivity to reward and punishment in boys with oppositional defant disorder. European Child and Adolescent Psychiatry, 13(6), 362–364. https://doi.org/10.1007/s00787-004-0395-x

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), 527–543. https://doi.org/10.1007/s10802-006-9031-1

McCorry L. K. (2007). Physiology of the autonomic nervous system. American journal of pharmaceutical education, 71(4), 78.

McCulloch, C. E. (1987). Tests for equality of variances with paired data. Communications in Statistics - Theory and Methods, 16(5), 1377-1391. http://doi.org/10.1080/03610928708829445

McCutcheon, A. L. (2002). Basic concepts and procedures in single-and multiple-group Latent class analysis. In J. A. Hagenaars & A. L. McCutcheon (Eds.), Applied latent class analysis (pp. 56-88). Cambridge University Press

McDaniel, M. A., Rothstein, H. R., & Whetzel, D. L. (2006). Publication bias: A case study of four test vendors. Personnel Psychology, 59(4), 927-953. http://doi.org/10.1111/j.1744-6570.2006.00059.x

McLachlan, G. J., & Peel, D. (2000). Finite mixture models. John Wiley & Sons.

Meece, D., & Mize, J. (2010). Multiple aspects of preschool children's social cognition: Relations with peer acceptance and peer interaction style. Early Child Development and Care, 180(5), 585-604. https://doi.org/10.1080/03004430802181452

Meeks-Gardner, J. M., Powell, C. A., & Grantham-McGregor, S. M. (2007). Determinants of aggressive and prosocial behaviour among Jamaican schoolboys. The West Indian medical journal, 56(1), 34-31. https://doi.org/10.1590/s0043-31442007000100007

Merk, W., De Castro, B. O., Koops, W., & Matthys, W. (2005). The distinction between reactive and proactive aggression: Utility for theory, diagnosis and treatment? The European Journal of Developmental Psychology, 2(2), 197-220. http://doi.org/10.1080/17405620444000300

Mesa-Gresa, P., Gil-Gómez, H., Lozano-Quilis, J. A., & Gil-Gómez, J. A. (2018). Effectiveness of virtual reality for children and adolescents with autism spectrum disorder: An evidence-based systematic review. Sensors, 18(8), 1-15. https://doi.org/10.3390/s18082486

Miadich, S., Shrewsbury, A.M., Doane, L., Davis, M., Clifford, S., & Lemery-Chalfant, K. (2020). Children's sleep, impulsivity, and anger: shared genetic etiology and implications for developmental psychopathology. Journal of Child Psychology and Psychiatry, 61(10), 1070-1079. http://doi.org/10.1111/jcpp.13328

Mikami, A. Y., Lee, S. S., Hinshaw, S. P., & Mullin, B. C. (2008). Relationships between social information processing and aggression among adolescent girls with and without ADHD. Journal of Youth and Adolescence, 37(7), 761-771. https://doi.org/10.1007/s10964-007-9237-8

Milich, R., & Dodge, K. A. (1984). Social information processing in child psychiatric populations. Journal of Abnormal Child Psychology, 12(3), 471-490. https://doi.org/10.1007/bf00910660

Moore, C. C., Hubbard, J. A., & Bookhout, M. K. (2018). Temperament and aggression. In T. Malti & K. H. Rubin (Eds.), Handbook of Child and Adolescent Aggression (pp. 107-126). The Guilford Press.

Moore, C. C., Hubbard, J., Morrow, M. T., Barhight, L. R., Lines, M. M., Sallee, M., & Hyde, C. T. (2018). The simultaneous assessment of and relations between children's sympathetic and parasympathetic psychophysiology and their reactive and proactive aggression. Aggressive Behavior, 44(2), 614–623. https://doi.org/10.1002/ab.21786

Morgan, A., & Lilienfeld, S. (2000). A meta-analytic review of the relation between antisocial behavior and neuropsychological measures of executive function. Clinical Psychology Review, 20(1), 113-136. http://doi.org/10.1016/S0272-7358(98)00096-8

Munoz, L. C., Frick, P. J., Kimonis, E. R., & Aucoin, K. J. (2008). Types of aggression, responsiveness to provocation, and callous-unemotional traits in detained adolescents. Journal of Abnormal Child Psychology, 36(1), 15–28. https://doi.org/10.1007/s10802-007-9137-0

Muris, P., Merckelbach, H., & Walczak, S. (2002). Aggression and threat perception abnormalities in children with learning and behavior problems. Child Psychiatry and Human Development, 33(2), 147-63. https://doi.org/10.1023/a:1020782208977

Nas, C. N., Brugman, D., & Koops, W. (2008). Measuring self-serving cognitive distortions with the "How I Think" Questionnaire. European Journal of Psychological Assessment, 24(3), 181–189. https://doi.org/10.1027/1015-5759.24.3.181

Nas, C. N., De Castro, B. O., & Koops, W. (2005). Social information processing in delinquent male adolescents. Psychology, Crime & Law, 11(4), 363-375. https://doi.org/10.1080/10683160500255307

Nelson, D. A., Cramer, C. M., Coyne, S. M., & Olsen, J. A. (2017). Children's hostile intent attributions and emotional distress: What do parents perceive? Aggressive Behavior, 44(1), 98-108. https://doi.org/10.1002/ab.21734

Nelson, D. A., Mitchell, C., & Yang, C. (2008). Intent attributions and aggression: A study of children and their parents. Journal of Abnormal Child Psychology, 36(6), 793-806. https://doi.org/10.1007/s10802-007-9211-7

Newman, J. P., Patterson, C. M., & Kosson, D. S. (1987). Response perseveration in psychopaths. Journal of Abnormal Psychology, 96(2), 145-148. https://doi.org/10.1037/0021-843X.96.2.145

Nigg, J. T. (1999). The ADHD response-inhibition deficit as measured by the stop task: Replication with DSM–IV combined type, extension, and qualification. Journal of Abnormal Child Psychology, 27(5), 393–402. https://doi.org/10.1023/a:1021980002473

Obradović, J. (2016). Physiological responsivity and executive functioning: Implications for adaptation and resilience in early childhood. Child Development Perspectives, 10(1), 65-70. http://dx.doi.org/10.1111/cdep.12164

Obradović, J., & Finch, J. E. (2016). Linking executive function skills and physiological challenge response: Piecewise latent growth curve modeling. Developmental Science, 20(3), 1-16. http://dx.doi.org/10.1111/desc.12476

Ogelman, H. G., & Seven, S. (2012). The effect social information processing in six-year-old children has on their social competence and peer relationships. Early Child Development and Care, 182(12), 1623-1643. http://doi.org/10.1080/03004430.2011.636810

Ogilvie, J. M., Stewart, A. L., Chan, R. C. K., & Shum, D. H. K. (2011). Neuropsychological measures of executive function and antisocial behavior: A meta-analysis. Criminology, 49(4), 1063-1107. http://doi.org/10.1111/j.1745-9125.2011.00252.x

Oostermeijer, S., Nieuwenhuijzen, M., van de Ven, P. M., Popma, A., & Jansen, L. M. (2016). Social information processing problems related to reactive and proactive aggression of adolescents in residential treatment. Personality and Individual Differences, 90, 54-60. http://doi.org/10.1016/j.paid.2015.10.035

Orue, I., & Calvete, E. (2009). Adaptation and validation of the What do You Think questionnaire to measure social information processing in children. Estudios de Psicologia, 30(3), 317-329. https://doi.org/10.1174/021093909789618503

Ostrov, J. M., & Godleski, S. A. (2010). Toward an integrated gender-linked model of aggression subtypes in early and middle childhood. Psychological Review, 117(1), 233–242. https://doi.org/10.1037/a0018070

Pechtel, P., & Pizzagalli, D. A. (2011). Effects of early life stress on cognitive and affective function: An integrated review of human literature. Psychopharmacology, 214(1), 55-70. http://doi.org/10.1007/s00213-010-2009-2

Pederson, C.A., Fite, P. J., & Bortolato, M. (2018). The role of functions of aggression in associations between behavior inhibition and activation and mental health outcomes. Journal of Aggression, Maltreatment & Trauma, 27(8), 811-830, http://doi.org/10.1080/10926771.2017.1370053

Peets, K., Hodges, E. V. E., Kikas, E., & Salmivalli, C. (2007). Hostile attributions and behavioral strategies in children: Does relationship type matter? Developmental Psychology, 43(4), 889-900. https://doi.org/10.1037/0012-1649.43.4.889

Peets, K., Hodges, E. V. E., & Salmivalli, C. (2008). Affect-congruent social-cognitive evaluations and behaviors. Child Development, 79(1), 170-185. https://doi.org/10.1111/j.1467-8624.2007.01118.x

Penton-Voak, I. S., Thomas, J., Gage, S. H., McMurran, 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. http://doi.org/10.1177/0956797612459657

Perren, S., Ettekal, I., & Ladd, G. (2013). The impact of peer victimization on later maladjustment: Mediating and moderating effects of hostile and self-blaming attributions. Journal of Child Psychology and Psychiatry, 54(1), 46-55. http://doi.org/10.1111/j.1469-7610.2012.02618.x

Peterson, E., & Welsh, M. (2014). The development of hot and cool executive functions: Are we getting warmer? In S. Goldstein & J.A. Naglieri (Eds.), Handbook of executive functions (pp. 45-68). Springer. http://doi.org/10.1007/978-1-4614-8106-5_4

Pinderhughes, E. E., Nix, R., Foster, E. M., Jones, D., & The Conduct Problems Prevention Research Group. (2001). Parenting in context: Impact of neighborhood poverty, residential stability, public services, social networks, and danger on parental behaviors. Journal of Marriage and Family, 63(4), 941-953. http://doi.org/10.1111/j.1741-3737.2001.00941.x

Polanczyk, G. V., Salum, G. A., Sugaya, L. S., Caye, A., & Rohde, L. A. (2015). Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. Journal of Child Psychology and Psychiatry and Allied Disciplines, 56(3), 345-365. http://doi.org/10.1111/jcpp.12381

Polman, H., de Castro, B. O., Koops, W., van Boxtel, 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(4), 522-535. https://doi.org/10.1007/s10802-007-9109-4

Polman, H., de Castro, B. O., 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(2), 183-193. http://doi.org/10.1007/s10802-008-9266-0

Poulin, F., & Boivin, M. (2000). Reactive and proactive aggression: Evidence of a two-factor model. Psychological Assessment, 12(2), 115-122. http://doi.org/10.1037/1040-3590.12.2.115

Quiggle, N. L., Garber, J., Panak, W. F., & Dodge, K. A. (1992). Social information processing in aggressive and depressed children. Child Development, 63(6), 1305-1320. http://doi.org/10.2307/1131557

Rabiner, D. L., Lenhart, L., & Lochman, J. E. (1990). Automatic versus reflective social problem solving in relation to children's sociometric status. Developmental Psychology, 26(6), 1010-1016. https://doi.org/10.1037/0012-1649.26.6.1010

Raine, A., Dodge, K. A., Loeber, R., Gatzke-Kopp, L., Lynam, D., Reynolds, C., et al. (2006). The reactive–proactive aggression questionnaire: Differential correlates of reactive and proactive aggression in adolescent boys. Aggressive Behavior, 32(2), 159-171. http://doi.org/10.1002/ab.20115

Reid, S. C., Salmon, K., & Lovibond P. F. (2006). Cognitive biases in childhood anxiety, depression, and aggression: Are they pervasive or specific? Cognitive Therapy and Research, 30(5), 531-549. http://doi.org/10.1007/s10608-006-9077-y

Reijntjes, A., Thomaes, S., Kamphuis, J., Bushman, B., De Castro, B. O., & Teich, M. (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. http://doi.org/10.1177/0146167211410247

Rijkeboer, M. M., & de Boo, G. M. (2010). Early maladaptive schemas in children: Development and validation of the schema inventory for children. Journal of Behaviour Therapy and Experimental Psychiatry, 41(2), 102–109 https://doi.org/10.1016/j.jbtep.2009.11.001

Ritter, D., & Eslea, M. (2005). Hot Sauce, Toy Guns, and Graffiti: A critical account of current laboratory aggression paradigms. Aggressive Behavior, 31(5), 407-419. https://doi.org/10.1002/ab.20066

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. Psychological Bulletin, 128(6), 934-960. http://doi.org/10.1037/0033-2909.128.6.934

Roeyers, H., Thys, M., Druart, C., De Schryver, M., & Schittekatte, M. (2011). Screeninglijst voor autismespectrumstoornissen. Handleiding. Hogrefe Uitgevers.

Romeo, R., Knapp, M., & Scott, S. (2006). Economic cost of severe antisocial behaviour in children–and who pays it. British Journal of Psychiatry 188(6), 547-553. http://doi.org/10.1192/bjp.bp.104.007625

Rothbart, M. K., & Derryberry, D. (1981). Development of individual difference in temperament. In M. E. Lamb, & A. L. Brown (Eds.), Advances in developmental psychology (pp. 37-86). Lawrence Erlbaum Associates.

Rubio-Garay, F., Carrasco, M. A., & Amor, P. J. (2016). Aggression, anger and hostility: evaluation of moral disengagement as a meditational process. Scandinavian Journal of Psychology, 57(2), 129-135. http://doi.org/10.1111/sjop.12270

Runions, K. C., & Keating, K. C. (2007). Young children's social information processing: Family antecedents and behavioral correlates. Developmental Psychology, 43(4), 838–849. http://doi.org/10.1037/0012-1649.43.4.838

Russo-Ponsaran, N. M., McKown, C., Johnson, J., Russo, J., Crossman, J., & Reife, I. (2018). Virtual environment for social information processing: Assessment of children with and without autism spectrum disorders. Autism Research, 11(2), 305-317. https://doi.org/10.1002/aur.1889

Sancilio, M. F., Plumert, J. M., & Hartup, W. W. (1989). Friendship and aggressiveness as determinants of conflict outcomes in middle childhood. Developmental Psychology, 25(5), 812-819. https://doi.org/10.1037/0012-1649.25.5.812

Schippell, P. L., Vasey, M. W., Cravens-Brown, L. M., & Bretveld, R. A. (2003). Suppressed attention to rejection, ridicule, and failure cues: A unique correlate of reactive but not proactive aggression in youth. Journal of Clinical Child and Adolescent Psychology, 32(1), 40-55. https://doi.org/10.1207/S15374424JCCP3201_05

Schniering, C. A., & Rapee, R. M. (2002). Development and validation of a measure of children's automatic thoughts: The Children's Automatic Thoughts Scale. Behaviour Research and Therapy, 40(9), 1091-1109. https://doi.org/10.1016/S0005-7967(02)00022-0

Scholte, E. M., & van der Ploeg, J. D. (2007). Handleiding Sociaal-Emotionele Vragenlijst (SEV). Bohn Stafleu van Loghum

Schoorl, J., Van Rijn, S., de Wied, M., Van Goozen, S., & Swaab, H. (2018). Boys with oppositional defiant disorder/conduct disorder show impaired adaptation during stress: An executive functioning study. Child Psychiatry and Human Development, 49(2), 298-307. https://doi.org/10.1007/s10578-017-0749-5

Schubert, T., Friedmann, F., & Regenbrecht, H. (1999). Embodied presence in virtual environments. In R. Paton, & I. Neilson (Eds.), Visual representations and interpretations (pp. 268-278). Springer-Verlag.

Schultz, D., Grodack, A., & Izard, C. E. (2010). State and trait anger, fear, and social information processing. In M. Potegal, G. Stemmler, & C. Spielberger (Eds.), International handbook of anger (pp. 311-328). New York: Springer. doi:10.1007/978-0-387-89676-2_18.

Shema-Shiratzky, S., Brozgol, M., Cornejo-Thumm, P., Geva-Dayan, K., Rotstein, M., Leitner, Y., Hausdorff, J.M., & (2018). Virtual reality training to enhance behavior and cognitive function among children with attention-deficit/hyperactivity disorder: Brief report. Developmental Neurorehabilitation, 22(6), 431-436https://doi.org/10.1080/17518423.2018.1476602

Shu, Y., & Luo, Z. (2021). Peer victimization and reactive aggression in junior high school students: A moderated mediation model of retaliatory normative beliefs and self-perspective. Aggressive Behavior. Advanced online publication. https://doi.org/10.1002/ab.21982

Siegel, D. (1999). The developing mind. The Guilford Press.

Simonds, J. & Rothbart, M. K. (2004). The Temperament in Middle Childhood Questionnaire (TMCQ): A computerized self-report measure of temperament for ages 7–10. Poster session presented at the Occasional Temperament Conference, October, Athens, GA.

Singh, P. (2017). Altering the way adolescents attribute negative ambiguous social encounters. Asia Pacific Journal of Counselling and Psychotherapy, 8(1), 15-28. https://doi.org/10.1080/21507686.2016.1256903

Smeets, K. C., Leeijen, A. A. M., van der Molen, M. J., Scheepers, F. E., Buitelaar, J. K., & Rommelse, N. N. J. (2015). Treatment moderators of cognitive behavior therapy to reduce aggressive behavior: A meta-analysis. European Child & Adolescent Psychiatry, 24(3), 255–264. https://doi.org/10.1007/s00787-014-0592-1

Smeets, K. C., Oostermeijer, S., Lappenschaar, M., Cohn, M., Van der Meer, J. M. J., Popma, A., Jansen, L. M. C., Rommelse, N. N. J., Scheepers, F. E., & Buitelaar, J. K. (2017). Are proactive and reactive aggression meaningful distinctions in adolescents? A variable-and person-based approach. Journal of Abnormal Child Psychology, 45(1), 1–14. https://doi.org/10.1007/s10802-016-0149-5

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. Personality and Social Psychology Review, 4(2), 108-131. http://doi.org/10.1207/S15327957PSPR0402_01Statistics Netherlands (2018, 2019). StatLine. Retrieved June 24, 2021, from https://opendata.cbs.nl

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. Psychological Bulletin, 87(2), 245-251. http://doi.org/10.1037/0033-2909.87

Steinberg, M. S., & Dodge, K. A. (1983). Attributional bias in aggressive adolescent boys and girls. Journal of Social and Clinical Psychology, 1(4), 312-321. https://dx.doi.org/10.1521/jscp.1983.1.4.312

Stickle, T. R., Kirkpatrick, N. M., & Brush, L. N. (2009). Callous-unemotional traits and social information processing: Multiple risk-factor models for understanding aggressive behavior in antisocial youth. Law and Human Behavior, 33(6), 515-529. http://doi.org/10.1007/s10979-008-9171-7

Stoltz, S., Deković, M., van Londen, M., De Castro, B. O., & Prinzie, P. (2013). What works for whom, how and under what circumstances? testing moderated mediation of intervention effects on externalizing behavior in children. Social Development, 22(2), 406-425. http://doi.org/10.1111/sode.12017

Stoltz, S., van Londen, M., Dekovic, M., De Castro, B. O., Prinzie, P., & Lochman, J. E. (2013). Effectiveness of an individual school-based intervention for children with aggressive behaviour: A randomized controlled trial. Behavioural and Cognitive Psychotherapy, 41(5), 525-548. https://doi.org/10.1017/s1352465812000525

Stoltz, S., van Londen, M., Dekovic, M., Prinzie, P., De Castro, B. O., & Lochman, J. E. (2013). Simultaneously testing parenting and social cognitions in children at-risk for aggressive behavior problems: Sex differences and ethnic similarities. Journal of Child and Family Studies, 22(7), 922-931. https://doi.org/10.1007/s10826-012-9651-8

Teisl, M., & Cicchetti, D. (2008). Physical abuse, cognitive and emotional processes, and aggressive/disruptive behavior problems. Social Development, 17(1), 1-23. https://doi.org/10.1111/j.1467-9507.2007.00412.x

Thomson, N. D., & Centifanti, L. C. (2018). Proactive and reactive aggression subgroups in typically developing children: the role of executive functioning, psychophysiology, and psychopathy. Child Psychiatry & Human Development, 49(2), 197-208. https://doi.org/10.1007/s10578-017-0741-0

Thomson, D. N., Gillespie, S. M., & Centafanti, L. C. M. (2020). Callous-unemotional traits and fearlessness: A cardiovascular psychophysiological perspective in two adolescent samples using virtual reality. Development and Psychopathology, 32(3), 803-815. http://doi.org/10.1017/S0954579419001196

Todorov, A., & Bargh, J. A. (2002). Automatic sources of aggression. Aggression and Violent Behavior, 7(1), 53-68. https://doi.org/10.1016/S1359-1789(00)00036-7

Underwood, M. K. (2005). Observing anger and aggression among preadolescent girls and boys: Ethical dilemmas and practical solutions. Ethics and Behavior, 15(3), 235-245.https://doi.org/10.1207/s15327019eb1503_4 U.S. Central Intelligence Agency. (2009). The World Factbook 2009. Retrieved from https://www.cia.gov/library/publications/the-world-factbook/index.html

Van Assen, M. A. L. M., Van Aert, R. C. M., Nuijten, M. B., & Wicherts, J. M. (2014). Why publishing everything is more effective than selective publishing of statistically significant results. PLoS ONE, 9(1), e84896. https://doi.org/10.1371/journal.pone.0084896

Van Baardewijk, Y., Andershed, H., Stegge, H., Nilsson, K. W., Scholte, E., & Vermeiren, R. (2010). Development and tests of short versions of the Youth Psychopathic Traits Inventory and the Youth Psychopathic Traits Inventory-Child Version. European Journal of Psychological Assessment, 26(2), 122–128. https://doi.org/10.1027/1015-5759/a000017

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three-level meta-analysis of dependent effect sizes. Behavior Research Methods, 45(2), 576-594. https://doi.org/10.3758/s13428-012-0261-6

Van Dijk, A., Hubbard, J. A., Deschamps, P. K. H., Hiemstra, W., & Polman, H. (2021). Do distinct groups of reactively and proactively aggressive children exist? A confirmatory latent profile approach. Research on Child and Adolescent Psychopathology, 49(10), 1303-1317. https://doi.org/10.1007/s10802-021-00813-0

Van Dijk, A., Poorthuis, A. M. G., & Malti, T. (2017). Psychological processes in young bullies versus bully-victims. Aggressive Behavior, 43(5), 430-439. https://doi.org/10.1002/ab.21701

Van Dijk, A., Poorthuis, A. M. G., Thomaes, S., & De Castro, B.O. (2018). Does parent-child discussion of peer provocations reduce young children's hostile attributional bias. Child development, 89(5), 1908-1920. https://doi.org/10.1111/cdev.13087

Van Dijk, A., Thomaes, S., Poorthuis, A. M. G., & de Castro, B. O. (2019). Can self-persuasion reduce hostile attribution bias in young children. Journal of Abnormal Child Psychology, 47(6), 989-1000. https://doi.org/10.1007/s10802-018-0499-2

Van Goozen, S.H.M., Fairchild, G., Snoek, H., & Harold, G.T. (2007). The evidence for a neurobiological model of childhood antisocial behavior. Psychological Bulletin, 133(1), 149-182. http://doi.org/10.1037/0033-2909.133.1.149

Van Nieuwenhuijzen, M., De Castro, B. O., Valk, V.D., Wijnroks, L., Vermeer, A., & Matthys, W. (2006). Do social information processing models explain aggressive behavior by children with mild intellectual disabilities in residential care? Journal of Intellectual Disability Research, 50(11), 801-812. https://doi.org/10.1111/j.1365-2788.2005.00773.x

Van Nieuwenhuijzen, M., Van Rest, M. M., Embregts, P. J. C. M., Vriens, A., Oostermeijer, S., Van Bokoven, I., & Matthys, W. (2017). E ecutive functions and social information processing in adolescents with severe behavior problems. Child Neuropsychology, 23(2), 228-241. http://doi.org/10.1080/09297049.2015.1108396

Van Nieuwenhuijzen, M., Vriens, A., Scheepmaker, M., Smit, M., & Porton, E. (2011). The development of a diagnostic instrument to measure social information processing in children with mild to borderline intellectual disabilities. Research in Developmental Disabilities, 32(1), 358-370. https://doi.org/10.1016/j.ridd.2010.10.012

Van Rest, M. M., Matthys, W., Van Nieuwenhuijzen, M., De Moor, M. H. M., Vriens, A., & Schuengel, C. (2019). Social information processing skills link executive functions to aggression in adolescents with mild to borderline intellectual disability. Child Neuropsychology, 25(5), 573-598. http://doi.org//10.1080/09297049.2018.1495186

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. https://doi.org/10.1017/s1352465814000149

Vassilopoulos, S. P., Brouzos, A., Rentzios, C. (2014). Evaluation of a universal social information-processing group program aimed at preventing anger and aggressive behaviour in primary school children. Hellenic Journal of Psychology, 11(4), 208-222.

Verhoef, R. E. J., Alsem, S. C., Verhulp, E. E., & De Castro, B. O. (2019). Hostile intent attribution and aggressive behavior in children revisited: A meta-analysis. Child Development, 90(5), 525-547. http://doi.org/10.1111/cdev.13255

Verhoef, R. E. J., Van Dijk, A., & De Castro, B. O. (2021).A dual-mode social-information-processing model to explain individual differences in children's aggressive behavior. Clinical Psychological Science, 10(1), 41-57.https://doi.org/10.1177/21677026211016396

Verhoef, R. E. J., Van Dijk. A., Thomaes, S., Verhulp, E. E., Van Rest, M. M., & De Castro, B. O. (2022). Detecting social information processing profiles of boys with aggressive behavior problems: An interactive virtual reality approach. Manuscript submitted for publication.

Verhoef, R. E. J., Van Dijk, A., Verhulp, E. E., & De Castro, B. O. (2021). Interactive virtual reality assessment of aggressive social information processing in boys withbehaviour problems: A pilot study. Clinical Psychology & Psychotherapy, 28(3), 489-499. https://doi.org/10.1002/cpp.2620

Verhoef, R. E. J., Verhulp, E. E., Van Dijk, A., & De Castro, B. O. (2021). Interactive virtual reality versus vignette-based assessment of children's aggressive social information processing. Research on Child and Adolescent Psychopathology. Advance Online Publication. https://doi.org/10.1007/s10802-021-00879-w

Verhulst, F.C., Ende, J. van der & Koot, H.M. (1997). Handleiding voor de Teachers's Report Form (TRF). Rotterdam: afd. Kinder- en jeugdpsychiatrie, Sophia Kinderziekenhuis/ AZR/ EUR.

Vlerick, P. (1994). The development of socially incompetent behavior in provocative situations. Psychologica Belgica, 34(1), 33-55.

Viechtbauer, W. (2010). Conducting meta¬analyses in R with the metafor package. Journal of Statistical Software. 36(3), 1-48. https://doi.org/10.18637/jss.v036.i03

Vitaro, F., Brendgen, M., & Barker, E. D. (2006). Subtypes of aggressive behaviors: A developmental perspective. International Journal of Behavioral Development, 30(1), 12-19. https://doi.org/10.1177/0165025406059968

Waldman, I. D. (1996). Aggressive boys' hostile perceptual and response biases: The role of attention and impulsivity. Child Development, 67(3), 1015-1033. https://doi.org/10.2307/1131877

Wang, Y., & Dix, T. (2017). Mothers' Depressive Symptoms and Children's Externalizing Behavior: Children's Negative Emotionality in the Development of Hostile Attributions. Journal of Family Psychology, 31(2), 214-223. https://doi.org/10.1037/fam0000241

Waschbusch, D. A. (2002). A meta-analytic examination of comorbid hyperactive impulsive attention problems and conduct problems. Psychological Bulletin, 128(1), 118-150. https://doi.org/10.1037/0033-2909.128.1.118

Webster-Stratton, C., & Lindsay, D.W (1999). Social competence and conduct problems in young children: issues in assessment. Journal of Clinical Chid Psychology, 28(1), 25-43. https://doi.org/10.1207/s15374424jccp2801_3

Weeland, J., Overbeek, G., Orobio, de, Castro, B., & Matthys, W. (2015). Underlying mechanisms of gene-environment interactions in externalizing behavior: A systematic review and search for theoretical mechanisms. Clinical Child and Family Psychology Review, 18(4), 413-441. http://doi.org/10.1007/s10567-015-0196-4

Weiss, B., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1992). Some consequences of early harsh discipline: Child aggression and a maladaptive social information processing style. Child Development, 63(6), 1321-1335. https://doi.org/10.2307/1131558.

Werner, R. S., Cassidy, K. W., & Juliano, M. (2006). The role of social-cognitive abilities in preschoolers' aggressive behaviour. British Journal of Developmental Psychology, 24(4), 775-799. https://doi.org/10.1348/026151005x78799

Werner, N. E. (2012). Do hostile attribution biases in children and parents predict relationally aggressive behavior? The Journal of Genetic Psychology: Research and Theory on Human Development, 173(3), 221-245. https://doi.org/10.1080/00221325.2011.600357

White, D. Y. (1984). Attributional bias in aggressive girls. Unpublished doctoral dissertation, American University, Washington, DC.

Wibbelink, C. J. M. & Assink, M. (2015). Handleiding voor het uitvoeren van een drie-level meta-analyse in R (Versie 3.0). Afdeling Forensische Orthopedagogiek, Universiteit van Amsterdam. Unpublished manual.

Williams, R. L. (1998). Social information-processing in lower and higher aggressive black adolescents. Unpublished doctoral dissertation, Hofstra University, Hempstead, NY.

Williams, S. C., Lochman, J. E., Phillips, N. C., & Barry, T. D.(2003). Aggressive and nonaggressive boys' physiological and cognitive processes in response to peer provocations. Journal of Clinical Child and Adolescent Psychology, 32(4), 568-576. https://doi.org/10.1207/s15374424jccp3204_9

Wilson, S. J., & Lipsey, M. (2006). The effects of school-based social information processing interventions on aggressive behavior: Part I: Universal programs: A systematic review. Campbell Systematic Reviews, 2(1), 1-42. https://doi.org/10.4073/csr.2006.5

Yaros, A., Lochman, J. E., Rosenbaum, J., & Jimenex-Camargo, L. A. (2014). Real-time hostile attribution measurement and aggression in children. Aggressive Behavior, 40(5), 409-420. https://doi.org/10.1002/ab.21532

Yaros, A., Lochman, J. E., & Wells, K. (2016). Parental aggression as a predictor of boys' hostile attribution across the transition to middle school. International Journal of Behavioral Development, 40(5), 452-458. https://doi.org/10.1177/0165025415607085

Yerkes, R.M., & Dodson, J.D. (1908). The relation of strength of stimulus to rapidity of habit formation. Journal of Comparative Neurology & Psychology, 18(5), 459-482. http://doi.org/10.1002/cne.920180503

Yeung, R. S., & Leadbeater, B. J. (2007). Does hostile attributional bias for relational provocations mediate the short-term association between relational victimization and aggression in preadolescence? Journal of Youth and Adolescence, 36(8), 973-983. https://doi.org/10.1007/s10964-006-9162-2

Yoon, J. S. (1998). Social cognitive differences between aggressive-rejected and aggressive-nonrejected children. Journal of School Psychology, 38(6), 551-570. https://doi.org/10.1016/s0022-4405(00)00052-2

Zelazo, P. D., & Müller, U. (2002). Executive function in typical and atypical development. In U. Goswami (Ed.), Blackwell handbook of childhood cognitive development (pp. 445-469). Blackwell Publishers Ltd.

Zelli, A., Cervone, D., & Huesmann, L. R. (1996). Behavioral experience and social inference: Individual differences in aggressive experience and spontaneous versus deliberate trait inference. Social Cognition, 14(2), 165-190. https://doi.org/10.1521/soco.1996.14.2.165

Zelli, A., Dodge, K. A., Lochman, J. E., Laird, R. D., & Conduct Problems Prevention Research Group. (1999). The distinction between beliefs legitimizing aggression and deviant processing of social cues: Testing measurement validity and the hypothesis that biased processing mediates the effects of beliefs on aggression. Journal of Personality and Social Psychology, 77(1), 150-166. https://doi.org/10.1037/0022-3514.77.1.150

Ziv, Y. (2012). Exposure to violence, social information processing, and problem behavior in preschool children. Aggressive Behavior, 38(6), 429-441. https://doi.org/10.1002/ab.21452

Ziv, Y., Kupermintz, H., & Aviezer, H. (2016). The associations among maternal negative control, children's social information processing patterns, and teachers' perceptions of children's behavior in preschool. Journal of Experimental Child Psychology, 142, 18-35. https://doi.org/10.1016/j.jecp.2015.09.004

Ziv, Y., & Sorongon, A. (2011). Social information processing in preschool children: Relations to sociodemographic risk and problem behavior. Journal of Experimental Child Psychology, 109(4), 412–429. https://doi.org/10.1016/j.jecp.2011.02.009

SUMMARY IN DUTCH (SAMENVATTING IN HET NEDERLANDS)

Het overkoepelende doel van deze dissertatie was om de sociale informatie verwerkingsprocessen (SIV) onderliggend aan agressief gedrag van kinderen beter te leren begrijpen. SIV bestaat uit een aantal mentale processen, ook wel stappen genoemd, die kinderen doorlopen voordat zij uiteindelijk tot bepaald gedrag komen in een sociale situatie. Het meest toonaangevende SIV model (Crick & Dodge, 1994) onderscheidt zes opeenvolgende verwerkingsstappen tijdens sociale interacties: (1) het coderen van sociale informatie, (2) het interpreteren van sociale informatie, (3) het bepalen van interactiedoelen, (4) het bedenken van mogelijke gedragsopties, (5) het evalueren van mogelijke gedragsopties, en (6) het selecteren van een gedragsoptie. Een breed scala aan onderzoek laat zien dat agressief gedrag van kinderen een gevolg zijn van afwijkingen in elk van deze SIV stappen, zoals het vaker waarnemen van vijandige signalen bij anderen, het vaker toeschrijven van vijandige intenties aan anderen, het nastreven van doelen gericht op wraak of instrumenteel gewin, het generen van agressieve gedragsopties en het positief beoordelen van agressieve gedragsopties (voor een overzicht, zie: De Castro & Van Dijk, 2017; Dodge, 2011).

Desondanks verklaarden eerdere SIV onderzoeken doorgaans slechts een relatief klein deel van de verschillen tussen kinderen in zowel hun huidige als toekomstige agressieve gedrag in het echte leven (De Castro et al., 2002; Dodge et al., 1986; Lansford et al., 2006). Dit komt mogelijk door drie hiaten in de huidige literatuur over SIV: Ten eerste hebben de huidige SIV meetmethoden een gebrekkige ecologische validiteit. Dat wil zeggen dat de gebruikte methoden in SIV onderzoek ver af lijken te staan van de alledaagse praktijk. Zo dienen kinderen zich vaak in te leven in verhaaltjes waarin een sociale gebeurtenis beschreven wordt. Een sociale gebeurtenis inbeelden is echter iets anders dan deze daadwerkelijk in het echt meemaken. Ten tweede zijn in eerder SIV onderzoek statistische technieken gebruikt die de individuele verschillen in SIV tussen kinderen niet goed in kaart brengen. Dit komt omdat eerder onderzoek louter gekeken heeft naar gemiddelde SIV scores van groepen kinderen, en niet naar unieke patronen van afwijkende SIV. Ten derde bieden de huidige theoretische SIV modellen louter een verklaring voor hoe agressie kan voortkomen uit SIV wanneer kinderen kalm zijn en weloverwogen sociale informatie verwerken. Deze SIV modellen specificeerden echter niet hoe agressie kan voortkomen uit SIV wanneer kinderen hevige emoties ervaren en vluchtig sociale informatie verwerken—terwijl juist veel agressie emotioneel van aard is.

De huidige dissertatie richtte zich op deze hiaten in de huidige SIV literatuur door drie onderzoeksvragen te beantwoorden:

1. Biedt interactieve Virtual Reality (VR) een meer ecologisch valide meting van agressie-gerelateerde SIV bij kinderen?
2. Is het mogelijk om unieke profielen van afwijkende SIV te onderscheiden bij kinderen met agressieve gedragsproblemen?
3. Kunnen we een verbeterd theoretisch SIV model specificeren om zowel kalme als emotionele agressie te verklaren?

Om deze onderzoeksvragen te beantwoorden, zijn vier empirische studies uitgevoerd en is een theoretisch artikel geschreven waarin we een nieuw theoretisch SIV model voorstellen. De antwoorden op deze onderzoeksvragen kunnen ons begrip vergroten van hoe SIV bij kinderen bijdraagt aan agressief gedrag en zo mogelijkheden bieden om behandelingen effectiever af te stemmen op de behoeften van individuele kinderen.

Deze dissertatie begint met een multi-level meta-analyse die de relatie onderzocht tussen agressief gedrag bij kinderen (leeftijd: 3-17 jaar) en één aspect van SIV, namelijk de neiging om vijandige intenties aan anderen toe te schrijven (**Hoofdstuk 2**). De resultaten lieten zien dat kinderen die geneigd zijn om vijandige intenties aan anderen toe te schrijven, vaker agressief gedrag vertonen in het dagelijks leven. De sterkte van dit verband varieerde echter sterk tussen studies. Een sleutelbevinding uit de meta-analyse was dat dit verband sterker bleek voor studies die emotioneel-geladen situaties gebruikten om SIV te meten. Dit suggereert dat emotioneel-geladen SIV meetmethoden een meer ecologisch valide meting opleveren van agressie-gerelateerde SIV bij kinderen dan traditionele vragenlijsten of vignetten.

Biedt interactieve Virtual Reality (VR) een meer ecologisch valide meting van agressie-gerelateerde SIV bij kinderen?

Om agressie-gerelateerde SIV te kunnen meten in een ecologisch valide, gestandaardiseerde, en emotioneel-geladen omgeving hebben wij een interactief VR klaslokaal ontwikkeld. In deze VR omgeving kunnen kinderen vrij rondlopen, praten met virtuele leeftijdsgenoten en spelletjes spelen (zoals een toren van blokken bouwen, of blikken van een tafel gooien met vijf ballen). Voor een betere meting van individuele verschillen in agressie en SIV hebben we, in tegenstelling tot eerder onderzoek, verschillende typen sociale situaties gebruikt om zowel reactieve als proactieve agressie en onderliggende SIV te kunnen meten. Reactieve agressie is defensieve, emotionele, impulsieve agressie in reactie op een provocatie, dreiging of frustratie (Dodge, 1991). Proactieve agressie daarentegen is offensieve, koelbloedige, gecontroleerde agressie gericht op het vergaren van instrumenteel gewin (Dodge, 1991). Om reactieve agressie en onderliggende SIV te meten hebben we twee VR scenario's ontwikkeld waarin kinderen geprovoceerd werden (hun spelletje werd bijvoorbeeld verpest). Om proactieve agressie en onderliggende SIV te meten hebben we twee VR scenario's ontwikkeld waarin kinderen de mogelijkheid hadden om instrumenteel gewin te vergaren (bijvoorbeeld winnen door vals te spelen). Omdat de triggers en uitingen van agressie kunnen verschillen tussen jongens en meisjes, hebben we de VR omgeving voor deze dissertatie gericht op jongens (waarbij agressie in de kindertijd vaker voorkomt).

Hoofdstuk 3 beschrijft een pilotstudie waarin we onderzochten of onze nieuwe VR omgeving een valide meting van SIV bij kinderen oplevert. Hieraan deden 32 jongens (leeftijd: 8-13 jaar) mee uit het regulier onderwijs en speciaal onderwijs voor kinderen met gedragsproblemen. Onze VR meting bleek over een goede convergente validiteit te beschikken: onze metingen correleerden matig tot hoog met een bestaand meetinstrument voor agressie-gerelateerde SIV (dit gold voor alle aspecten van SIV, behalve boosheid). Interactieve VR leverde ook een verbeterde meetsensitiviteit op: we vonden grotere verschillen tussen kinderen in SIV dan het bestaande meetinstrument (dit gold voor agressieve responsen, maar niet voor andere SIV variabelen). Tenslotte vonden we gedeeltelijke ondersteuning voor de discriminante validiteit van onze interactieve VR meting: onze provocatieve scenario's riepen meer reactief-agressieve SIV op dan andere scenario's, maar onze instrumentele scenario's riepen niet meer proactief-agressieve SIV op. Samengevat suggereert onze pilotstudie dat interactieve VR een veelbelovend instrument is om agressie-gerelateerde SIV en agressief gedrag bij kinderen te meten.

Een belangrijke volgende stap was om te onderzoeken of onze interactieve VR meting van SIV en gedrag beter kan voorspellen hoe agressief kinderen zich gedragen in het echte leven, vergeleken met een traditioneel SIV instrument dat verhaaltjes gebruikt in plaats van VR scenario's. **Hoofdstuk 4** testte daarom direct of interactieve VR een meer ecologisch valide meting van SIV bij kinderen oplevert dan een traditioneel SIV instrument. We gebruikten hiervoor VR scenario's en verhaaltjes die qua inhoud precies hetzelfde waren als in het traditionele SIV instrument (bijvoorbeeld een toren bouwen, die vervolgens wordt omgegooid). Aan deze studie deden 184 jongens mee (leeftijd: 7-13 jaar) uit het regulier onderwijs en speciaal onderwijs voor kinderen met gedragsproblemen. De resultaten lieten zien dat kinderen meer opgingen in de VR omgeving en zich meer emotioneel betrokken voelden bij de sociale situaties in VR dan bij de verhaaltjes. Daarnaast bleek dat onze VR meting inderdaad beter in staat was om agressie in het dagelijks leven te voorspellen dan de verhaaltjesmeting. Dit was met name het geval voor provocerende scenario's en minder voor de instrumentele scenario's. Echter bleek interactieve VR niet voor alle SIV stappen een verbeterde meting op te leveren: dit was wel zo voor de meting van agressief gedrag en onderliggende interactiedoelen, maar niet voor boosheid en de neiging om vijandige attributies te maken. Samengevat lijkt interactieve VR een meer ecologisch valide meting van agressief gedrag en onderliggende SIV bij kinderen op te leveren dan een traditionele verhaaltjesmeting.

Is het mogelijk om unieke profielen van afwijkende SIV te onderscheiden bij kinderen met agressieve gedragsproblemen?

Aangezien kinderen aanzienlijk verschillen in de SIV patronen die voorafgaan aan hun agressieve gedrag (De Castro & Van Dijk, 2017; Dodge, 2011), lijkt het logisch dat er unieke SIV profielen onderscheiden kunnen worden bij kinderen met agressieve gedragsproblemen. **Hoofdstuk 5** gebruikte daarom de SIV metingen in VR uit Hoofdstuk 4 om SIV profielen te onderscheiden bij jongens met agressieve gedragsproblemen. Op basis van latente profiel analyses werden vier verschillende SIV profielen onderscheiden. We vonden: 1) een algemeen reactief-agressief SIV profiel van kinderen die in beide provocerende scenario's agressieve gedragingen en agressie-gerelateerde SIV lieten zien; 2) een situatie-specifiek reactief-agressief SIV profiel van kinderen die alleen in één van de twee provocerende scenario's agressieve gedragingen en agressie-gerelateerde SIV lieten zien (het scenario waarbij hun spelletje werd verpest door een virtuele leeftijdsgenoot); 3) een gemengd reactief-proactief-agressief SIV profiel van kinderen die in zowel beide provocerende als instrumentele scenario's agressieve gedragingen en agressie-gerelateerde SIV lieten zien; 4) een niet-agressief SIV profiel van kinderen die in geen enkel scenario agressieve gedragingen en agressie-gerelateerde SIV lieten zien. Vergelijkingen tussen de profielen lieten daarnaast zien dat kinderen uit de verschillende SIV profielen, ook verschilden in onderliggend temperament, agressieve schema's en strafgevoeligheid, maar niet in executief functioneren en spanningsbehoefte. Samengevat is dit de eerste studie die laat zien dat unieke SIV profielen kunnen worden onderscheiden bij kinderen met agressieve gedragsproblemen. Dit biedt mogelijkheden voor de klinische praktijk om cognitieve-gedragsinterventies meer effectief af te stemmen op de behoeften van individuele kinderen.

Kunnen we een verbeterd theoretisch SIV model specificeren om zowel kalme als emotionele agressie te verklaren?

De bevindingen uit deze dissertatie suggereren dat hoe emotioneel kinderen zijn in een sociale situatie, een grote invloed heeft op hun SIV in deze situatie. Echter hoe dit precies in zijn werk gaat is niet gespecificeerd in de huidige SIV modellen. In **Hoofdstuk 6** poneren wij daarom een nieuw theoretisch SIV model dat specificeert hoe en waarom de emotionele staat van kinderen een dergelijke diepgaande invloed heeft op de aard van hun SIV. Ons nieuwe SIV model onderscheidt twee modi: 1) automatische SIV, waarbij kinderen vluchtig de situatie verwerken en direct met een reactie komen; en 2) reflectieve SIV, waarbij kinderen weloverwogen alle informatie in een situatie verwerken en daarna doordacht een reactie kiezen. Hierbij zou automatische SIV bijdragen aan snelle, emotie-gedreven agressie, terwijl reflectieve SIV zou bijdragen aan weloverwogen, gecontroleerde agressie. We stellen voor dat kinderen de automatische modus gebruiken als ze erg emotioneel zijn, en de reflectieve modus als ze kalm maar wel betrokken zijn. Ons model stelt daarnaast dat de emotionele staat van kinderen in deze situaties wordt bepaald door een interactie tussen kind-specifieke factoren (namelijk emotionele disposities, executief functioneren en motivationele disposities) en situatie-specifieke factoren (namelijk de huidige interne staat en type situatie). Ons theoretische SIV model heeft als doel om een meer ecologische verklaring te geven voor de unieke SIV stijlen van kinderen en zodoende behandelingen effectiever af te kunnen stemmen op individuele kinderen.

Discussie van bevindingen

Ons onderzoek naar interactieve VR als nieuwe SIV meetmethode heeft veel nieuwe inzichten opgeleverd. Ten eerste vonden we dat interactieve VR een meer ecologisch valide meting biedt van agressie-gerelateerde SIV bij kinderen, vergeleken met de traditionele verhaaltjesmeting van SIV. Ten tweede vonden we drie unieke profielen van afwijkende, agressie-gerelateerde SIV, die onderling verschilden in de aspecten van SIV die afweken, de situaties waarin afwijkende SIV optrad, en in de onderliggende kenmerken van kinderen. Ten derde hebben we een theoretisch model ontwikkeld dat zowel kalme als emotionele agressie en onderliggende SIV verklaart. Naast deze bevindingen, leverde ons onderzoek ook verschillende onverwachte bevindingen op.

Ten eerste bleek onze interactieve VR meting met name een verbeterde meting van kinderen hun agressieve gedrag en interactiedoelen op te leveren maar minder van hun boosheid en de neiging om vijandige intenties aan anderen toe te schrijven. Voor boosheid kan dit eraan gelegen hebben dat we kinderen vroegen om zelf aan te geven hoe boos ze zich voelden, zoals gebruikelijk in eerder SIV onderzoek. Echter laat onderzoek zien dat mensen het lastig vinden om nauwkeurig te rapporteren over de intensiteit van hun emoties (Robinson & Clore, 2002). Toekomstig onderzoek zou daarom een combinatie kunnen gebruiken van zelfrapportage en meer objectieve emotiematen zoals fysiologie (e.g., huidgeleiding, hartslag; Hubbard et al., 2002, 2004). Voor de neiging om vijandige attributies te maken kan de uitblijvende meerwaarde van VR een gevolg zijn van hoe de VR scenario's verliepen. Voor een goede meting van de neiging om vijandige attributies te maken, zijn scenario's nodig waarin het onduidelijk blijft wat de werkelijke intenties van de ander zijn (ambigue intenties). Aangezien de gemiddelde scores van kinderen hun vijandige attributies relatief hoog waren—ook in het regulier onderwijs—kan dit erop wijzen dat de intenties van virtuele karakters in onze VR scenario's eerder gemeen dan ambigu leken. Dit kan ervoor gezorgd hebben dat kinderen vijandige attributies toeschreven aan de virtuele karakters onafhankelijk van hun neiging tot vijandigheid of agressie in het echte leven. Om te zorgen dat VR scenario's inderdaad ambigu zijn, kunnen toekomstige studies de handelingen en reacties van de virtuele karakters stapsgewijs aanpassen om te ontdekken wanneer deze daadwerkelijk als ambigue worden ervaren.

Ten tweede liet onze interactieve VR meting van SIV bij kinderen de grootste toegevoegde waarde zien bij provocerende scenario's vergeleken met instrumentele scenario's. Onze interactieve VR meting leverde zodoende met name een verbeterde meting op van reactieve agressie en onderliggende SIV. Dit is bijzonder bruikbaar omdat reactieve agressie het meest voorkomt bij kinderen (e.g., Dodge et al., 1997; Van Dijk et al., 2021). Daarentegen bleek de toegevoegde waarde van interactieve VR om proactieve agressie en onderliggende SIV te meten minder sterk. Een verklaring kan zijn dat onze instrumentele scenario's minder emotionele betrokkenheid opwekten bij kinderen dan de provocerende scenario's. Dit zou betekenen dat iets stelen of valsspelen tijdens de spelletje in onze VR minder sterke emoties opwekte dan geprovoceerd worden. Aangezien de veronderstelde meerwaarde van VR juist is dat deze methode de emotionele betrokkenheid vergroot (kinderen maken de situaties immers echt mee), zou de toegevoegde waarde van instrumentele scenario's ten opzichte van een verhaaltjesmeting allicht minder zijn. Een alternatieve verklaring kan zijn dat interactieve VR simpelweg ook met betere stimuli minder toegevoegde waarde heeft om proactieve agressie en onderliggende SIV te meten. Sommige kinderen die agressie gebruiken om instrumenteel gewin te vergaren zullen dit mogelijk ook rapporteren bij een verhaaltjesafname omdat zij het gebruik van agressie altijd een acceptabele en bruikbare manier vinden om hun doelen te gebruiken.

Samengevat lijkt interactieve VR dus een belangrijke bijdrage te kunnen leveren aan een meer ecologisch valide meting van agressief gedrag en onderliggende SIV bij kinderen, en dan met name voor daadwerkelijk gedrag en achterliggende interactiedoelen, en met name in provocerende situaties die reactieve agressie meten.

Praktische tips voor VR gebruik in onderzoek

Hoe SIV bijdraagt aan agressief gedrag van kinderen als hun emoties oplopen, is lang te weinig onderzocht. Met interactieve VR als nieuwe SIV meetmethode en een nieuw theoretisch SIV model zijn er belangrijke eerste stappen genomen om ons begrip van de SIV voorafgaand aan het agressieve gedrag van kinderen beter te leren begrijpen. We hopen zodoende dat deze dissertatie wetenschappers aanmoedigt om emotioneel-geladen, ecologisch valide meetmethoden zoals interactieve VR te gebruiken om menselijk gedrag te bestuderen. Verschillende overwegingen verdienen aandacht wanneer onderzoekers inderdaad besluiten om interactieve VR voor onderzoeksdoeleinden te gebruiken.

Ten eerste is het bij het ontwikkelen van een interactieve VR omgeving belangrijk om een balans te vinden tussen drie aspecten: 1) de interactiviteit van de VR omgeving (welke bijdraagt aan het realisme en de emotionele betrokkenheid); 2) de ambiguïteit van de interacties in VR (welke belangrijk is bij het meten van de neiging om vijandige attributies te maken); en 3) standaardisatie van de interacties in VR (welke belangrijk is voor voldoende interne validiteit). Om dit te bewerkstelligen kunnen onderzoekers bijvoorbeeld vooraf een zorgvuldige selectie maken van voorgeprogrammeerde opties betreft verbale en fysieke responsen van de virtuele karakters. Dit om zo adequaat mogelijk te kunnen reageren op de uiteenlopende gedragingen van de deelnemers en tegelijkertijd de ambiguïteit en standaardisatie te waarborgen.

Ten tweede is het belangrijk dat deelnemers in interactieve VR zich zo veel mogelijk gedragen zoals in het echte leven. Hiervoor is het onder andere belangrijk dat kinderen, voordat zij de experimentele scenario's aangeboden krijgen, vertrouwd zijn geraakt met hoe zij kunnen reageren in de VR omgeving en hoe de VR omgeving op henzelf reageert. Dit kan bewerkstelligd worden door de VR omgeving zo veel mogelijk op de echte wereld te laten lijken en deelnemers de VR omgeving te laten ontdekken in meerdere oefenscenario's. Daarnaast is het belangrijk dat de gedragingen van de virtuele karakters adequaat getimed worden door de proefleider. Enige vertraging in bijvoorbeeld de responsen van de virtuele karakters kan de deelnemers er namelijk onbewust aan doen herinneren dat zij interacteren in een gesimuleerde computeromgeving in plaats van een realistische, natuurlijke omgeving die overeenkomt met de echte wereld.

Ten derde is het voor het meten van agressie in VR belangrijk dat kinderen het idee hebben dat zij interacteren met daadwerkelijke leeftijdsgenoten die echt schade ondervinden van hun agressieve gedrag. Dit is niet vanzelfsprekend, aangezien veel kinderen gewend zijn om agressief gedrag te vertonen tijdens digitale spelletjes omdat zij weten dat de computergestuurde karakters niet daadwerkelijk bestaan. In onze studies presenteerden we daarom de VR omgeving als een echt klaslokaal waarin deelnemers echte andere kinderen konden tegenkomen die ook deelnamen aan het onderzoek en op hetzelfde moment ingelogd waren op onze VR omgeving.

Klinische implicaties

Voor de klinische praktijk kan interactieve VR, wanneer dit verder gevalideerd is in toekomstig onderzoek, een meer aantrekkelijke, flexibele en ecologisch valide methode zijn om agressie-gerelateerde SIV van kinderen te diagnosticeren en te behandelen. Clinici kunnen de flexibiliteit van interactieve VR gebruiken om emotioneel-geladen oefeningen aan te bieden en deze aan te passen aan de kenmerken en behoeften van individuele cliënten door de inhoud, moeilijkheidsgraad en complexiteit van de sociale situaties gedurende de interventie te variëren. Kinderen die met name reactief-agressieve SIV laten zien zouden bijvoorbeeld stapsgewijs kunnen oefenen met steeds provocerender situaties om zo te leren hun boosheid te reguleren en hun vijandige attributies te onderzoeken. Kinderen die met name proactief-agressieve SIV laten zien, zouden stapsgewijs kunnen oefenen met een breed scala aan situaties waarin instrumenteel gewin vergaard kan worden, om daarbinnen met pro-sociale gedragsopties te oefenen. De toepassing van interactieve VR binnen cognitieve-gedragsinterventies wordt momenteel getest, waarbij is voortgebouwd op de bevindingen uit de huidige dissertatie (Alsem et al., 2021).

Daarnaast doet onze bevinding dat er verschillende SIV profielen onderscheiden kunnen worden bij kinderen met agressieve gedragsproblemen een nieuw licht schijnen op de huidige gestandaardiseerde diagnostiek. Aangezien kinderen met agressieve gedragsproblemen erg uiteenlopende SIV afwijkingen in verschillende situaties laten zien, lijkt het inzetten van louter één gestandaardiseerd SIV instrument dat identiek is voor elk kind niet het meest informatief. Zo zouden clinici ervoor kunnen kiezen om daarnaast een functieanalyse te gebruiken om de unieke situationele triggers en bijbehorende SIV- en gedragspatronen van kinderen in kaart te brengen (en deze functieanalyse vervolgens in VR te toetsen). Een dergelijke op-maat-gemaakte diagnostische benadering zou mogelijk een meer nauwkeurige weergave geven van de unieke SIV stijlen van kinderen met agressieve gedragsproblemen, waardoor cognitieve-gedragsinterventies nog beter afgestemd kunnen worden op de behoeften van het individuele kind.

Sterktes en beperkingen

De huidige dissertatie heeft verschillende sterke punten en beperkingen. De sterktes van deze dissertatie zitten vooral in het innovatieve karakter. Zo bevat deze dissertatie de eerste empirische studies naar interactieve VR als meetmethode voor agressief gedrag en onderliggende SIV bij kinderen. Dit leverde niet alleen een meer ecologisch valide meting op dan een traditionele verhaaltjesafname, maar ook belangrijke inzichten in de voor- en nadelen van het gebruik van interactieve VR bij het meten van agressief gedrag en onderliggende SIV. Een ander sterk punt van deze dissertatie is dat er verschillende VR scenario's gebruikt zijn om reactief- en proactief-agressieve SIV te meten. Dit om individuele verschillen in agressie en onderliggende SIV tussen kinderen nauwkeuriger in kaart te brengen—die inderdaad naar voren kwamen in onze studie naar unieke SIV profielen bij kinderen met agressieve gedragsproblemen Als laatste hebben we een nieuw theoretisch SIV model ontwikkeld dat mogelijk een meer ecologisch valide, gedetailleerde en testbare verklaring biedt voor de SIV patronen die voorafgaan aan agressief gedrag van kinderen.

De huidige dissertatie bevat ook verschillende beperkingen. Zo richtte deze zich op een relatief specifieke doelgroep (jongens tussen de 7-13 jaar, met relatief weinig variatie in sociaaleconomische en culturele achtergrond). Hierdoor kunnen de bevindingen niet zomaar gegeneraliseerd worden naar andere populaties. Daarnaast hebben we, gezien de hoge kosten voor het ontwikkelen van interactieve VR scenario's, slechts vier verschillende sociale situaties gebruikt om de SIV van kinderen te meten. Aangezien verschillende onderzoeken laten zien dat kinderen in uiteenlopende situaties agressief gedrag kunnen vertonen, hebben we uiteraard niet alle mogelijke situaties waarin kinderen agressie laten zien kunnen opnemen (De Castro & Van Dijk, 2017). Een andere beperking heeft betrekking op onze manier om SIV te meten. Wij kinderen na afloop van elk VR scenario naar iedere SIV stap. Dit leverde een verbeterde SIV meting op vergeleken met de traditionele verhaaltjesmethode, waarschijnlijk doordat kinderen over hun daadwerkelijke SIV rapporteerden in plaats van de SIV die ze zich voorstelden te hebben in een bepaalde situatie. Echter was onze methode nog steeds reflectief. Wanneer kinderen sterke emoties ervaren in interactieve VR, zullen velen van hen direct met agressie reageren zonder te reflecteren op elke SIV stap (Hoofdstuk 6; De Castro 2005; 2012; Helmsen et al., 2012). Het is dus de vraag of onze expliciete meting met het stellen van vragen achteraf wel gemeten heeft hoe sommige kinderen in echte situaties in het moment zelf zouden denken. Dit zou tevens kunnen verklaren dat, ondanks dat interactieve VR een verbeterde SIV meting opleverde vergeleken met de verhaaltjesmethode, ook VR nog lang niet alle verschillen kon verklaren in hoe agressief kinderen in het dagelijks leven zijn (er waren nog steeds relatief grote proporties onverklaarde variantie). Een laatste beperking is dat wij onvoldoende gegevens konden verzamelen betreft kinderen hun uitkomstverwachtingen en evaluaties van agressie. Wij besloten om kinderen te laten rapporteren over hun verwachte uitkomsten en evaluaties van hun daadwerkelijke gedrag in VR. Doordat kinderen vaak geen agressie vertoonden in de instrumentele scenario's (23.2%-23.8%, afhankelijk van het scenario), hadden we te weinig data om hierover statistische analyses uit te voeren. Onze conclusies hebben dus betrekking op 4 van de 6 aspecten van agressie-gerelateerde SIV.

Conclusie

Deze dissertatie laat zien dat het belangrijk is om agressie-gerelateerde SIV te meten in een emotionele context, zodat we het agressieve gedrag van kinderen in het dagelijks leven beter kunnen voorspellen. Interactieve VR blijkt hiervoor een geschikt instrument, aangezien het kinderen blootstelt aan emotioneel-geladen, realistische, en toch gestandaardiseerde sociale interacties. Interactieve VR biedt een breed scala aan nieuwe mogelijkheden voor SIV onderzoek en op-maat-gemaakte interventies die niet eerder mogelijk waren. Daarnaast kan ons nieuwe theoretische SIV model wetenschappers helpen om beter individuele verschillen in agressief gedrag en onderliggende SIV bij kinderen in kaart te brengen. Ik hoop dat deze dissertatie wetenschappers en clinici zal inspireren om interactieve VR te gebruiken en door te gaan met het vergroten van onze kennis van de SIV onderliggend aan het agressieve gedrag van kinderen.

ACKNOWLEDGEMENTS (DANKWOORD)

Ten eerste wil ik mijn promotoren bedanken. Ik heb werkelijk van het begin tot eind zo ontzettend genoten van dit project. Daar is geen woord (zelfs geen letter) aan gelogen. Zowel van het samen vormgeven van het project als de leuke sfeer die er heerste tijdens onze afspraken. **Bram**, het klinkt erg cliché en feel good maar ik had me geen fijnere promotor kunnen wensen. Ik denk dat er weinig promotoren zijn die zo goed een balans weten te vinden in hun begeleiding en deze weten af te stemmen op de promovendus. Jouw begeleiding is zeker tailored. Maar ook het feit dat je wekelijks tijd vrij maakte om het project te bespreken is iets wat ik heel erg waardeer en waarvan ik me goed besef dat dit niet de standaard is. En dan vergeet ik nog bijna te benoemen hoe indrukwekkend ik het vind hoe snel je tot de kern van iets complex kan komen. Ik ben vanaf dag één zo nieuwsgierig wat er te zien is als jouw hersenactiviteit gemeten wordt tijdens bijvoorbeeld een vergadering. Wel een emotionally engaging vergadering dan natuurlijk. **Sander**, ook van de samenwerking met jou heb ik intens genoten. Daarnaast ben ik blij dat je het vooroordeel van mij onderuit heb weten te halen door de eerste topwetenschapper te zijn die wel van het volkse voetbal houdt en er ook nog eens veel verstand van heeft. Ik weet dat naar jouw idee je misschien net iets minder betrokken bent geweest dan je had gewild. Zo heb ik dat zelf echter totaal niet ervaren. Ook jouw deur stond altijd open. Ik vond het heel fijn dat ik je altijd om advies kon vragen als ik ergens tegenaan liep. En ook alle tips over wat er allemaal komt kijken bij het nastreven van een carrière in de wetenschap waren zeer waardevol. Daarnaast heb ik tijdens het schrijven van de latente profielen paper ontzettend veel opgestoken van jouw indrukwekkende wetenschappelijk schrijfvaardigheden.

Anouk, ik gebruik niet graag clichés maar in jouw geval moet ik toch echt een uitzondering maken. Ik meen het oprecht dat ik mijn promotietraject nooit zonder al jouw grenzeloze steun, energie en kunde had kunnen volbrengen. Ik denk dat ik kan zeggen dat ik van jou het meest geleerd heb op elk vlak van het wetenschappelijk onderzoek doen. Of het nou het wetenschappelijk schrijven betrof, het nadenken over theoretische en methodologische vraagstukken of praktische obstakels, op jou kon ik altijd rekenen. Mijn proefschrift was echt nooit zo mooi geworden zonder jou. Daar ben ik je voor eeuwig grenzeloos dankbaar voor. Ik kijk met ontzettend veel plezier terug op onze samenwerking waarbij we met het theoretische paper denk ik ons hoogtepunt kenden. Ik lieg niet als ik zeg dat ik het zo allemaal over zou willen doen voor louter het proces. It is not the destination but the journey. **Esmée**, ook jou ben ik natuurlijk tot in het einde der tijden dank verschuldigd. Niet alleen voor al jouw inhoudelijke adviezen in de afgelopen jaren maar ook zeker voor het me telkens weer bij de hand te nemen bij het aanpakken van praktische obstakels. Dat heeft me heel veel stress gescheeld en daardoor levensjaren opgeleverd. Daar zijn zowel ik, als mijn directe sociale omgeving je zeer erkentelijk voor. Daarnaast waardeer ik het enorm dat je altijd paraat stond om een luisterend oor te bieden. Je was voor mij dan ook de stilte in de storm en daarom goud waard in de hectische periodes gedurende mijn promotietraject.

Ook onmisbaar in dit project waren Elise en Jessica. **Elise**, ik heb nog steeds geen woorden kunnen vinden om jouw onschatbare waarde binnen het project te beschrijven. Hoe ontzettend knap je de organisatie en coördinatie van de werving en dataverzameling volbracht hebt is werkelijk beyond imagination. Alleen dankzij jou hebben zoveel scholen deelgenomen, hebben we zoveel kinderen kunnen testen en verliep de werving en dataverzameling vlekkeloos. Naast de organisatie en coördinatie van de werving en dataverzameling heb je ook nog eens zelf heel veel data verzameld. Of het nou een Virtual Reality-, vignetten- of vragenlijstafname betrof, ik wist dat je dit elke keer weer tot in de perfectie zou volbrengen. Je hebt zodoende echt een sleutelrol gespeeld bij het slagen van het project. Mijn dank is daarom echt oneindig groot.

Jessica, ook jij verdient een welgemeend eigen dankwoord. Ook jij bent van onschatbare geweest bij de dataverzameling. Ik zou echt niet meer weten hoeveel afnames je voor je rekening hebt genomen maar ik weet nog wel dat ik er duizelig van werd. Daarnaast heb jij een Guinness Book of Records prestatie geleverd met de afname van de online vragenlijsten bij de leerkrachten en ouders. Naar mijn weten is het nog nooit vertoond binnen een onderzoek dat 95% van de ouders/leerkrachten alle online vragenlijsten hebben ingevuld. Ik snap nog steeds niet hoe je dat voor elkaar hebt weten te krijgen. Ook jij hebt daarmee echt een sleutelrol gespeeld bij het slagen van het project. Elise en Jessica, er is geen woord aan gelogen als ik zeg dat ik dit jullie in geen 100 jaar na had gedaan en dat ik jullie essentiële bijdrage aan het project nooit zal vergeten. Tevens wil ik natuurlijk **alle bachelor- en masterthesisstudenten** ontzettend bedanken voor hun tomeloze inzet bij de dataverzameling. Jullie stonden op wanneer het nog donker was en kwamen terug wanneer het wederom donker was geworden om data te verzamelen in alle uithoeken die Nederland kent en deden dit meerdere dagen in de week maandenlang. Dat volbrengen is op zich al een wereldprestatie. En om dit te doen met al jullie enthousiasme en secure manier van werken is buitenaards.

Daarnaast wil ik **CleVR** ontzettend bedanken voor de fijne samenwerking. Een speciaal dankwoord hierbij wil ik wijden aan **Steve**. Ik heb heel veel plezier beleefd aan onze middagen samen in Delft om de Virtual Reality te ontwikkelen. Daar kijk ik met een grote glimlach op terug. Daarnaast ben ook jij van onschatbare waarde geweest tijdens de dataverzameling. Tijdens de dataverzameling stond je werkelijk ieder moment paraat voor technische ondersteuning. Ik weet nog dat ik de eerste dagen tijdens de dataverzameling vrees had dat er misschien technische problemen zouden optreden en dat ik daardoor een testdag zou moeten afbreken. Maar de enkele keer dat ik een probleem had (of dacht te hebben) wist je dit binnen enkele minuten op te lossen. Sindsdien heb ik me nooit een seconde zorgen meer gemaakt. Nogmaals ontzettend bedankt!

Daarnaast wil ik alle **scholen, leerkrachten, ouders en natuurlijk de kinderen zelf** bedanken. En in het bijzonder de contactpersonen op de verschillende scholen. Het meedoen aan wetenschappelijk onderzoek vergt veel tijd en energie qua organisatie. Het verspreiden van de brieven aan de ouders, het afstemmen van het onderzoek met de schoolroosters van de kinderen en het beschikbaar maken van een leeg lokaal voor de afnames hebben jullie allemaal met volle overgave gedaan. Ik ben jullie echt heel erg dankbaar dat jullie dit allemaal, naast jullie dagelijkse taken op de scholen, er nog bij hebben willen doen. Zonder jullie medewerking was er immers geen project geweest.

Ook wil ik **alle collega's van de afdeling ontwikkelingspsychologie** bedanken voor de fantastische werksfeer. Ik heb me vanaf dag één enorm welkom gevoeld op de afdeling. Jullie deur stond, niet altijd letterlijk maar voor mijn gevoel wel figuurlijk, altijd open voor een vakkundig advies of gezellig praatje. Helaas heb ik jullie door de corona situatie in de laatste fase van mijn project minder gezien maar jullie betrokkenheid en warmte zal ik nooit vergeten. Enkele collega's wil ik in het bijzonder bedanken. **Jan**, wat heb ik geboft om jouw kamergenoot te mogen zijn en wat heb ik genoten van al onze gesprekken en vrijdagmiddag borrels. Je bracht elke werkdag weer een grote grijns van oor tot oor op mijn gezicht. **Lysanne**, ook jou wil ik natuurlijk enorm bedanken. Je was mijn grote steun en toeverlaat als medepromovendus. Ik heb echt ontzettend veel aan je gehad. Je hebt me onder andere wegwijs gemaakt op de afdeling maar ook in het nachtleven van Utrecht. Wat een prachtige tijd was het om tijdens een klein drankje te discussiëren over onderzoek maar ook gewoon als goede vrienden over het leven. Ik zal de herinneringen aan die tijd voor altijd koesteren. **Sophie**, ook jij verdient natuurlijk een welgemeend eigen dankwoord. Ook aan jou heb ik grenzeloos veel gehad. Je was niet alleen mijn gids op de afdeling, de stad Utrecht en San Francisco (ik bedoel: Washington) maar ook vaak de eerste hulplijn. Ik kijk met een grote glimlach terug op alle avonturen. Daarnaast bent jij degene met wie ik de eerste stappen heb gezet bij het doen van een eigen onderzoek in de vorm van een meta-analyse. En tevens degene die mijn eerste grijze haren publiekelijk bekend maakte terwijl ik nog nauwelijks als promovendus begonnen was. Ik ben je voor al het voorgaande heel dankbaar.

Tot slot wil ik **mijn familie** bedanken. **Mina**, jouw onvoorwaardelijke liefde, grenzeloze steun en oneindige begrip tijdens het traject betekenen heel veel voor me. Ik kan me voorstellen dat het soms lastig voor je was als ik in mijn eigen (werk)wereld opgesloten zat. Je klaagde echter nooit op die momenten. Je gaf me met een warme knuffel weer de veerkracht om nog eens goed de schouders eronder te zetten. Je zegt steeds dat je vanwege mijn promotie zo trots op mij bent maar ik ben pas trots op jou. De manier waarop je me hebt bijgestaan is in mijn ogen veel bijzonderder dan een proefschrift schrijven. En **Koa**, mijn kleine dappere strijder, ook jij hebt mij ontzettend geholpen bij de laatste loodjes. Jij zorgde er persoonlijk voor dat ik niks meer kon uitstellen omdat dit ten koste zou gaan van mijn tijd met jou en daar wilde ik niks van missen. Dank, grote kameraad! **Papa**, last but not least, jij verdient ook zeker een eigen speciaal dankwoord. Uiteindelijk ben jij degene die me de mogelijkheid hebt gegeven om een research master te gaan doen en dat zie ik toch als het startpunt van mijn studie tot wetenschapper. Ik vind het een heel bijzonder gebaar dat je toen besloot om mij de ruimte te geven om mijn passie na te jagen omdat je dit het allerbelangrijkste in het leven vindt. Jouw steun is grenzeloos gebleken. Dus ik sta hier ook absoluut dankzij jou.

ABOUT THE AUTHOR

Rogier Verhoef (1985) obtained both a research master Psychology and clinical master Clinical Psychology at the University of Amsterdam. During his studies, he worked as a teaching- and research assistant and completed a clinical internship at the UvA PsyPoli. In 2016, Rogier started his PhD project at the department of Developmental Psychology at Utrecht University. During his PhD project, he received an honourable mention (2nd place) for one of his publications included in this dissertation (Chapter 2) as part of the Child & Adolescent Studies (CAS) Research Article Award. At the moment, Rogier works as a post-doctoral researcher at the department of Clinical Child & Family studies at Utrecht University.

SCIENTIFIC PUBLICATIONS

*The manuscripts noted with an asterisk are included in this dissertation.

Oud, M., Arntz, A., Hermens, M. L. M., Verhoef, R., & Kendall, T. (2018). Specialized psychotherapies for adults with borderline personality disorder: A systematic review and meta-analysis. *The Australian and New Zealand Journal of Psychiatry*, *52*(10), 949-961. https://doi.org/10.1177/0004867418791257

Rameckers, S. A., Verhoef, R. E. J., Grasman, R. P. P. P., Cox, W. R., Van Emmerik, A. A. P., Engelmoer, I. M., & Arntz, A. (2021). Effectiveness of psychological treatments for Borderline Personality Disorder and predictors of treatment outcomes: A multivariate multilevel meta-analysis of data from all design types. *Journal of Clinical Medicine*, 10(23), 5622. https://doi.org/10.3390/jcm10235622

*Verhoef, R. E. J., Alsem, S. C., Verhulp, E. E., & De Castro, B. O. (2019). Hostile intent attribution and aggressive behavior in children revisited: A meta-analysis. *Child Development, 90*(5), 525-547. http://doi.org/10.1111/cdev.13255

*Verhoef, R. E. J., Van Dijk, A., & De Castro, B. O. (2021).A dual-mode social-information- processing model to explain individual differences in children's aggressive behavior. *Clinical Psychological Science, 10*(1), 41-57. https://doi.org/10.1177/21677026211016396

*Verhoef, R. E. J., Van Dijk. A., Thomaes, S., Verhulp, E. E., Van Rest, M. M., & De Castro, B.O. (2022). Detecting social information processing profiles of boys with aggressive behavior problems: An interactive virtual reality approach. *Manuscript submitted for publication*.

*Verhoef, R. E. J., Van Dijk, A., Verhulp, E. E., & De Castro, B. O. (2021). Interactivevirtual reality assessment of aggressive social information processing in boys withbehaviour problems: A pilot study. *Clinical Psychology & Psychotherapy, 28*(3), 489-499. https://doi.org/10.1002/cpp.2620

*Verhoef, R. E. J., Verhulp, E. E., Van Dijk, A., & De Castro, B. O. (2021). Interactivevirtual reality versus vignette-based assessment of children's aggressive social information processing. *Research on Child and Adolescent Psychopathology.* Advance Online Publication. https://doi.org/10.1007/s10802-021-00879-w